International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 12 - Issue 05 || May 2026 || PP. 01-10

An Interactive Metaverse-Based Digital Twin Framework for Organ-Oriented Healthcare Monitoring

Sevval Ural[1], Hasan Yetis[2], Mehmet Karakose[3]
[1,2,3]*Firat University, Computer Engineering Department, Elazig, Türkiye*

Abstract: This study proposes a metaverse-based digital twin system architecture for organ-based physiological anomaly detection and multi-class anomaly classification in the field of healthcare. The proposed system architecture was created by combining synthetic scenario generation, six-dimensional feature extraction, LSTM-based prediction model training, a threshold determination mechanism based on organ functions, a semantic inference-based decision-making module, and Unity-based interactive visualization components. In the proposed system, anomaly detection is performed for the left lung, right lung, and heart using heart rate and respiratory data, and the results are classified as normal, warning, lung anomaly, heart anomaly, and critical condition. Experimental evaluation was performed on a synthetic dataset of 24,000 sample data (each containing 20 time steps). The training model achieved 89.39% accuracy, 82.41% Micro-F1 score, 82.00% Macro-F1 score, 80.34% precision, 84.59% recall and 72.17% Exact Match Accuracy in the test dataset. Consequently, the findings indicate that the proposed approach provides a technically balanced and evaluable solution for multi-label physiological monitoring processes.

Keywords: Digital twin, metaverse, LSTM, physiological time-series, multi-label anomaly detection, healthcare monitoring, Unity visualization

I. Introduction

In recent years, it has been attracting a lot of attention that the intelligent systems developed in the field of healthcare have gained popularity rapidly. In particular, sensor technologies, wearable devices, edge-connected monitoring systems and data-driven medical platforms have caused a radical transformation in the collection, transmission and interpretation of physiological information. This transformation has not only led to the storage of patient data but has also contributed to the increasing need for advanced computational frameworks that enable real-time modeling, analysis, and interpretation of physiological processes. This situation has positioned digital twin technology (DT) as a powerful paradigm that enables continuous interaction between physical and digital assets by creating synchronized virtual counterparts of physical systems [1], [2]though this concept was initially developed for industrial and cyber-physical production environments, its fundamental principles such as data integration, simulation, prediction, and feedback have increasingly been adapted to healthcare applications [1], [2], [3], [4] the healthcare sector offer an extensible, flexible, and customizable structure that represents patients, organs, physiological processes, and medical equipment within an evolvable computational environment through real-time observations [5], [6]se systems can produce adaptable and interpretable solutions for diagnosis, monitoring, and prognosis processes by integrating physiological data, intelligent analysis methods, simulation capabilities, and decision support mechanisms beyond static medical records [7], [8], [9]. Current studies demonstrate that digital twins provide significant advancements in areas such as personalized medicine, patient-specific organ modeling, and real-time health monitoring, particularly when integrated with interoperable smart environments and sensor-enabled infrastructures [6], [7], [8], [10]e developments demonstrate that digital twins have evolved from being merely conceptual models to becoming practical components of next-generation intelligent healthcare systems.

One of the main challenges in the field of healthcare, particularly in intelligent healthcare systems, is the accurate interpretation of physiological time series signal data such as heart rate and respiratory rate. Although these signals contain clinically meaningful information regarding changes in health status, their noisy, nonlinear, and time-dependent nature limits the reliability of traditional threshold-based monitoring approaches. Detecting small deviations that occur before clinical deterioration emerges is quite challenging without temporal pattern learning mechanisms [11], [12] refore, the detection of anomalies stands out as a critical analysis component, particularly in remote patient monitoring, smart hospital environments, and IoT-based healthcare systems [7], [12].Additionally, predictive outputs remain limited when presented solely as raw model scores; however, when supported by interpretable decision layers, they can become much more meaningful and actionable [3], [4] situation highlights Long Short-Term Memory (LSTM) networks for physiological monitoring due to their ability to capture sequential dependencies and latent temporal structures in signal flows.

On the other hand, the development of metaverse technology has expanded the digital transformation in the healthcare field beyond mere numerical analysis, extending it toward immersive and interactive virtual environments [13], [14], [15]nt studies in the Metaverse field demonstrate that Metaverse-enabled healthcare systems strengthen areas such as telemedicine, collaborative diagnosis, rehabilitation, medical education, and patient engagement by combining immersive visualization with real-time digital interaction [13], [14], [15]combined with digital twin technology, this approach enables not only the monitoring of physiological states but also the visually intuitive presentation of predicted changes, semantic risk levels, and organ-specific outcomes. [16], [17] feature is of great importance in health monitoring scenarios, where interpretability is not only a preference but also an operational necessity. In this study conducted by us, this visual dimension was supported to produce organ-focused outputs through the adaptation of the open-source Z-Anatomy model in the Unity environment [19].

Another significant limitation in this field is the scarcity of large, well-annotated physiological data sets that can adequately represent both normal and abnormal temporal patterns under different clinical conditions. This limitation directly constrains the development and evaluation of machine learning models, particularly in anomaly detection problems where anomalous events are rare and heterogeneous. Therefore, synthetic data generation has become an effective and increasingly common approach for creating realistic yet controllable data sets that incorporate temporal variability, noise, and clinically meaningful scenario diversity [5], [11] itionally, studies on anomaly-focused digital twin systems have shown that simulated or semi-simulated data supports the learning of abnormal patterns and ensures analytical robustness in data-scarce situations. [11].

Based on this situation, this study proposes an interactive metaverse-based digital twin system for the analysis, monitoring, and anomaly detection of physiological time series data. The proposed system brings together synthetic data generation, feature engineering, an LSTM-based multi-label deep learning model, organ-specific threshold optimization, a rule-based decision layer, and a Unity-based real-time visualization environment within an integrated architecture. Through the analysis of heart rate and respiratory data, the system generates organ-level anomaly outputs for the left lung, right lung, and heart, and interprets these outputs via semantic system states such as normal, warning, lung anomaly, heart anomaly, and critical condition.

This study addresses the problem of monitoring physiological abnormalities at the organ level as a multi-label temporal learning problem and proposes a physiologically constrained synthetic scenario generation approach for healthcare settings with limited data. The proposed method enhances the interpretability of the system by combining probabilistic deep learning outputs with organ-specific threshold optimization and rule-based semantic inference. Additionally, the analytical findings obtained are transformed into real-time digital twin interactions within the Unity-based metaverse layer and presented interactively to the user. In this context, the primary contribution of the work lies not in a single algorithmic innovation, but in the integration of temporal modeling, interpretable decision support mechanisms, and organ-focused immersive visualization components within a single, holistic digital health system.

II. Methodology

This section describes the methodological structure of the metaverse-based digital twin framework developed for analyzing physiological time series data, performing multi-label anomaly detection, and enabling real-time visual interpretation. The proposed system is based on an integrated pipeline from processing raw physiological data to obtaining decision outputs. The system consists of the following components: synthetic data generation, feature engineering, data preprocessing, LSTM-based modeling, threshold optimization, and a rule-based decision mechanism. Through this approach, physiological data is presented to the user in an interactive format, both analytically and visually.

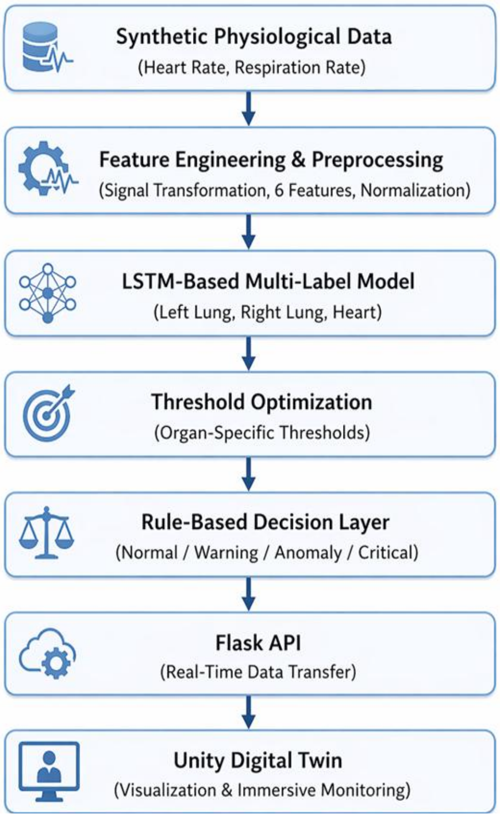

Fig. 1: Block diagram of the proposed system.

The overall architecture of the proposed system is shown in Figure 1. As seen in the figure, the system begins with synthetic physiological data generation and then proceeds through feature engineering, LSTM-based multi-label modeling, threshold optimization, and a rule-based decision mechanism. This system features a sequential pipeline structure. The obtained analysis results are transferred to the Unity environment via a Flask-based API and visualized in real time on the digital twin model. This system combines analytical processing steps and interactive visual interpretation within a single framework.

2.1 Synthetic Data Generation and Scenario-Based Modeling

Due to restrictions on access to original patient health data and the requirements for its confidentiality protection, the study adopted a synthetic data generation approach. In this context, a dataset consisting of 24,000 heart rate (HR) and respiratory rate (Resp) signals was created. Each sample was represented as a time series consisting of 20 time steps, resulting in an initial data tensor of size (24000, 20, 2). In the data generation process, five different scenarios have been defined: normal, warning, lung, heart, and critical. These scenarios correspond to probabilities of 0.34, 0.22, 0.18, 0.13, and 0, respectively, and were randomly sampled from 13 possible options.

To enhance temporal realism, the data were not generated entirely from random values. Instead, sinusoidal fluctuations, Gaussian noise, sudden changes, and slow changes were added to the process of generating data to make it look like how physiological signals naturally behave. Additionally, heart rate and respiratory values are restricted within physiologically meaningful ranges. Thus, the generated dataset better reflects realistic monitoring conditions and can represent different levels of physiological impairment. Therefore, the synthetic data generation process aims not only to increase the number of samples but also to provide a controlled working environment that reproducibly models inter-scenario variability, temporal continuity, abrupt disruption patterns, and organ-specific anomaly characteristics.

2.2 Multi-Label Anomaly Detection and Labeling Mechanism

In the proposed system, the anomaly detection process is treated as a multi-label classification problem instead of the conventional single-label decision problem. The proposed system produces simultaneous predictions for three target variables: the left lung, the right lung, and the heart. The label mechanism is constructed by taking into account the average heart rate (HR) and respiratory rate (Resp) values observed in the last five time intervals, along with the temporal trend information of these values.

Specifically, the probability of left lung anomaly is increased if the average respiratory rate over the last five steps exceeds 22, 24, and 27 breaths per minute, and further increases occur if the respiratory trend exceeds 2.5. The probability of right lung anomaly is increased if the average respiratory rate exceeds 25, 27, and 30 breaths per minute, with an additional increase if the respiratory trend exceeds 3.5. The probability of heart anomaly is increased if the average heart rate in the last five intervals exceeds 103, 110, and 118 values or falls below 58, 54, and 50 values, respectively. In addition, there is an additional increase if the absolute HR trend exceeds 7.

Instead of an approach based entirely on fixed thresholds, a probabilistic approach is adopted in the labeling mechanism. After constraining organ-specific abnormality probabilities to the range [0, 0.95], they are converted into binary labels through probabilistic sampling. About 2% of the labels in the dataset are noise labels, and one of the three target labels is changed at random using this method. This method makes the training data more varied and lowers the chance that the model will memorize it.

2.3 Feature Engineering and Data Preprocessing

To improve the system's prediction performance, the model is trained not only on raw physiological signals but also on derived features. In this context, a total of six features are used: heart rate, respiratory rate, HR difference, Respiratory difference, HR deviation from moving average, and Respiratory deviation from moving average. This feature design enables the network to learn signal magnitude, local temporal variations, and short-term deviations simultaneously. Thus, the model can capture not only instantaneous physiological values but also the trends and fluctuation characteristics of the signal dynamics. After feature engineering, the model input tensor reaches the dimension (24000, 20, 6).

During preprocessing, the dataset was divided into 80% training, 10% validation, and 10% test sets. Subsequently, all features were standardized using Z-score normalization to obtain a suitable and stable scale for model training. The mean and standard deviation values used for normalization were calculated solely on the training data and consistently reused across the validation, test, and real-time prediction processes.

2.4 LSTM-Based Deep Learning Model

We used a multi-layer LSTM-based deep learning model to look at temporal physiological data. The model has an input dimension of (20, 6) and three LSTM layers that are connected to fully connected layers. The first LSTM layer has 128 neurons and is set up to return sequences. It is followed by batch normalization and a 0.25 dropout layer. The second LSTM layer has 64 neurons and is again set up to return sequences. It is followed by batch normalization and a 0.20 dropout layer. The third LSTM layer has 32 neurons and is followed by a 0.15 dropout layer. After the repeated layers, there are two dense layers with 64 and 32 neurons each that use ReLU activation functions. After the first dense layer, an additional dropout layer with a rate of 0.15 was applied. The final output layer consists of three neurons with a sigmoid activation function. From an architectural perspective, the layered recurrent structure aims to separate short-term signal fluctuations from more persistent temporal dependencies and to gradually compress the latent representation. Because of this design, the model can generate independent anomaly probabilities for each target organ, thus solving the problem as a multi-label classification task.

2.5 Model Training and Optimization

During model training, the Adam optimization algorithm was used with a learning rate of 0.001, and gradient clipping was applied with the clipnorm parameter set to 1.0.CourseDuring model training, the Adam optimization algorithm was used with a learning rate of 0.001, and gradient clipping was applied with the clipnorm = 1.0 parameter. The model was trained for 50 epochs, with each epoch containing 64 samples in a mini-batch. To reduce the effects of class imbalance, the cross-entropy loss function was used. Class-specific weights were determined based on the positive sample ratio of each target To monitor the model's performance and prevent overfitting, Early Stopping, Reduce LR On Plateau, and Model Checkpoint procedures were used. In this way, the optimization process stops when the best validation performance is achieved, the learning rate is appropriately reduced, and the model weights that yield the best performance are preserved. reserved.

2.6 Threshold Optimization

The model outputs were not directly converted to binary labels using a fixed global threshold value. Instead, optimal threshold values were determined separately for each organ. For this purpose, threshold values between 0.30 and 0.70 are tested in steps of 0.02, and the F1 score is calculated for each value. The threshold value that provides the highest F1-score is selected as the optimal decision limit for the relevant body. This approach allows to improve the balance between sensitivity and specificity, especially in cases where there is a class imbalance. Dec.

2.7 Flask-Based API and Real-Time Data Flow

To enable real-time prediction, a Flask-based API was developed. This API loads the trained model into memory, performs predictions on the generated physiological data, and returns the results in JSON format. The system uses a floating (rolling) data buffer containing the last 20 time stages, and this buffer is constantly updated to ensure that the model remains compatible with the input structure. API; The service provides the endpoints /, /health_data, /refresh, and /reset_buffer respectively to perform the functions of service definition, current state retrieval, single-step buffer update, and buffer restart.

During the real-time prediction process, the simulator maintains a scenario state that includes controlled transitions between normal, warning, lung, heart, and critical modes. Heart rate and respiratory rate are updated through a first-order transition process toward scenario-specific target values; additionally, Gaussian noise and occasional random deviations are applied. The obtained HR and Resp values are converted into the six derived features used during the training phase, and then subjected to Z-score normalization using stored statistics. The model outputs are converted into binary anomaly labels using organ-specific threshold values loaded from the same file. In addition, confidence values are calculated based on the normalized Deciency between the estimated probabilities and the threshold values. The JSON response transmitted to the Unity environment contains physiological variables, organ-based anomaly labels, confidence scores, class probabilities, threshold values, operating mode, system status, sequence length, risk level, description, last update time, buffer size, and scenario information.

2.8 Rule-Based Decision Layer and Metaverse Integration

Instead of presenting model outputs directly to the user, they are interpreted through a rule-based decision layer. This layer analyzes the heart rate (HR), respiratory rate (Resp), and organ-level anomaly outputs together within clearly defined physiological ranges. In the applied decision logic, the normal heart rate range is 60–100 beats per minute, and the normal respiratory rate is 12–20 breaths per minute. Critical physiological conditions are defined as heart rate values being 110 beats per minute or above, or 52 beats per minute or below; and respiratory rate values being 28 breaths per minute or above, or 9 breaths per minute or below.

According to these thresholds, the system is classified as normal when no anomaly is detected and physiological values remain within the normal range; and as a warning when HR or Resp values exceed the normal range but no significant organ-level anomaly is observed. When an anomaly is detected in the left or right lung, the system is set to lung mode; in case of an anomaly originating from the heart, it is set to heart mode. If heart and lung anomalies are observed together or if HR and Resp values simultaneously meet critical thresholds, the system is classified as being in critical mode. Additionally, the system generates status levels (NORMAL, WARNING, ANOMALY) and risk levels (LOW, MEDIUM, HIGH) in addition to the mode label.

The analyzed results are transferred from the Flask API to the Unity environment and visualized on the digital twin model. The Anomaly Manager component on the Unity side periodically queries the API, parses the returned JSON data, updates physiological values and anomaly statuses, and gradually changes organ colors to normal or anomaly states. The same component also manages a textual interface that includes connection status, operation mode, system status, risk level, and the latest description information. In this way, the current state of the system is presented to the user as a complete entity, both visually and semantically.

Thanks to this approach, the proposed system transcends being merely an analytical pipeline and evolves into an interactive, user-centric digital twin platform that combines real-time analysis with immersive visual feedback.

III. Simulation Results

The performance of the proposed metaverse-based digital twin system was evaluated using synthetically generated physiological time series data. The dataset consists of heart rate (HR) and respiratory rate (Resp) signals, containing a total of 24,000 samples, each represented by 20 time steps. Each time step contains numerous features derived from the raw signals, thereby forming a multidimensional temporal input structure for the dataset.

The dataset is designed to simulate different physiological conditions, such as normal, warning, organ-level anomaly, and critical scenarios involving multiple organs. The data was split into 19,200, 2,400, and 2,400 samples. We used 80% of the samples to train the proposed LSTM-based model, 10% to validate it, and 10% to test it. The model was trained using an adaptive optimization method so that it would converge steadily across different time patterns and be able to generalize well.

3.1 General Performance Analysis

The overall performance of the proposed system was evaluated using various metrics suitable for multi-label classification problems. These metrics include Exact Match Accuracy (EMA), Micro F1-score, Macro F1-score, accuracy (accuracy), precision (precision), sensitivity (recall) and F1-score Dec.

Table 1. Overall model performance

Metric	**Value**
ExactMatchAccuracy (EMA)	0.7217
Micro F1-score	0.8241
Macro F1-score	0.8200
Accuracy	0.8939
Precision	0.8034
Recall	0.8459
F1-score	0.8241

The results show that the proposed model works well and consistently even when the simulations are very complicated and realistic. The 72.17% full match accuracy value shows that the model can accurately and simultaneously predict all organ-level anomaly labels in a large number of test samples. This metric is considered a highly challenging success criterion for multi-label classification problems.

The Micro and Macro F1-scores confirm that the model exhibits balanced performance across different anomaly classes. The overall accuracy of 89.39% also shows that the model is very good at classifying individual labels. From all of these results, we can say that the LSTM-based architecture can accurately capture temporal dependencies and work well in many different physiological situations.

3.2 Training Process Analysis

The training dynamics of the proposed model were analyzed using the accuracy and loss curves shown in Fig. 2 and Fig. 3. As shown in Fig. 2, both training and validation accuracy increase rapidly during the initial epochs and stabilize within the range of 0.89–0.90. This indicates that the model effectively learns the dominant temporal patterns. The close agreement between the training and validation curves suggests that the model maintains strong generalization capability without significant overfitting.

A similar trend is observed in the loss curves presented in Fig. 3. The training and validation loss values decrease steadily and converge to approximately 0.29–0.30, demonstrating stable optimization behavior throughout the learning process

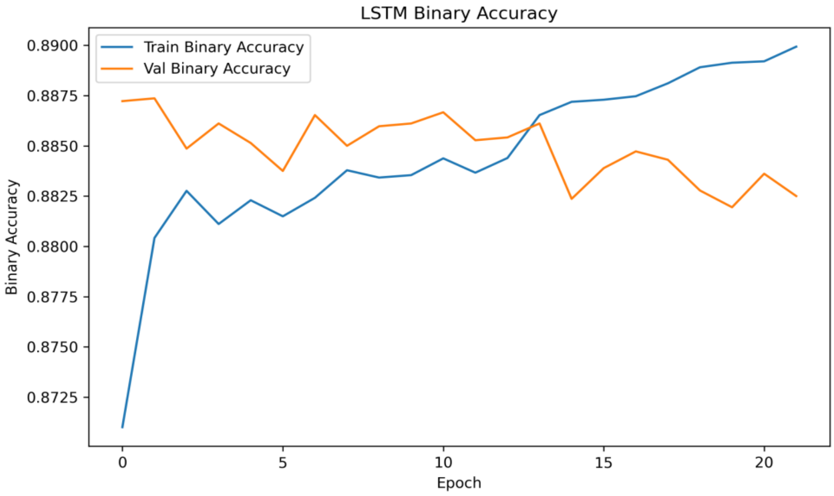

Fig. 2: Training and validation accuracy curves of the proposed LSTM-based model.

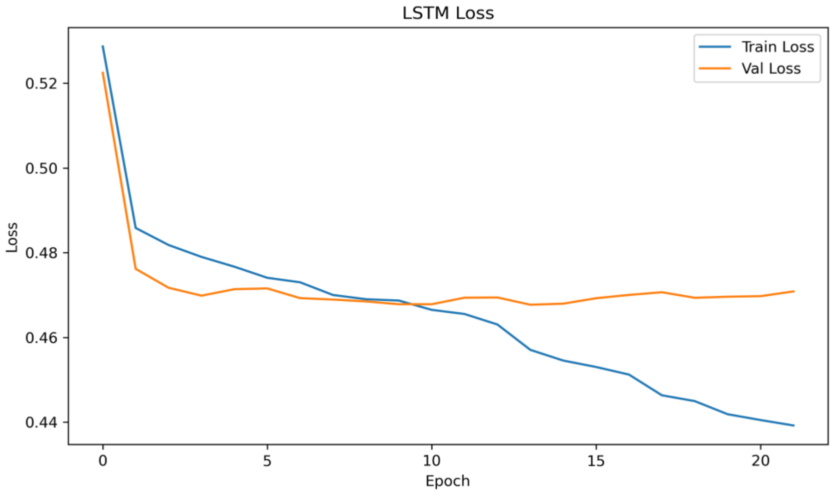

Fig. 3: Training and validation loss curves of the proposed LSTM-based model.

3.3 Organ-Based Performance Analysis

In order to evaluate the proposed model in more detail, organ-level performance metrics were analyzed separately for each target output.

Table 2: Organ-BasedPerformance

Organ	**Precision**	**Recall**	**F1-score**	**Accuracy**
LeftLung	0.8201	0.8601	0.8396	0.8738
RightLung	0.7863	0.8756	0.8286	0.9029
Heart	0.7963	0.7877	0.7920	0.9050

The results at the organ level show that the proposed model works well for different types of anomaly detection tasks. The left and right lung classes have high F1-scores, which means that it's easier to tell the difference between respiratory-based temporal patterns in the generated dataset. On the other hand, the heart class has lower recall and F1-score values, which means that it's harder to find cardiac anomalies with the same level of confidence. This is because heart rate signals have more variability and patterns that overlap with each other. However, the overall results show that the model can effectively generalize across different physiological signals within the unified multi-label framework The results at the organ level show that the proposed model works well for different types of anomaly detection tasks. The left and right lung classes have high F1-scores, which means that it's easier to tell the difference between respiratory-based temporal patterns in the generated dataset. On the other hand, the heart class has lower recall and F1-score values, which means that it's harder to find cardiac anomalies with the same level of confidence. This is because heart rate signals have more variability and patterns that overlap with each other. However, the overall results show that the model can effectively generalize across different physiological signals within the unified multi-label framework.

3.4 Performance Interpretation and Error Analysis

To better analyze the classification behavior of the proposed method, model performance was evaluated using precision, recall, and F1-score values for different organs.

The results for the left lung show that the model can correctly classify a large number of normal and abnormal samples and has a good balance between false positives and false negatives. The same is true for the right lung class, where the model is very good at telling normal examples apart while still being very sensitive to unusual cases.

In contrast, the heart class exhibits relatively lower recall and F1-score values. This indicates that cardiac anomalies are more difficult to detect with the same level of confidence. This difficulty can be explained by the high variability and partially overlapping patterns observed in heart rate signals.

Overall, it has been observed that the proposed model exhibits balanced performance in terms of precision and recall across different organs and provides stable classification behavior.

3.5 Visualization Results

The proposed system integrates real-time anomaly detection with a Unity-based digital twin environment to provide interactive and intuitive visualization of physiological states.

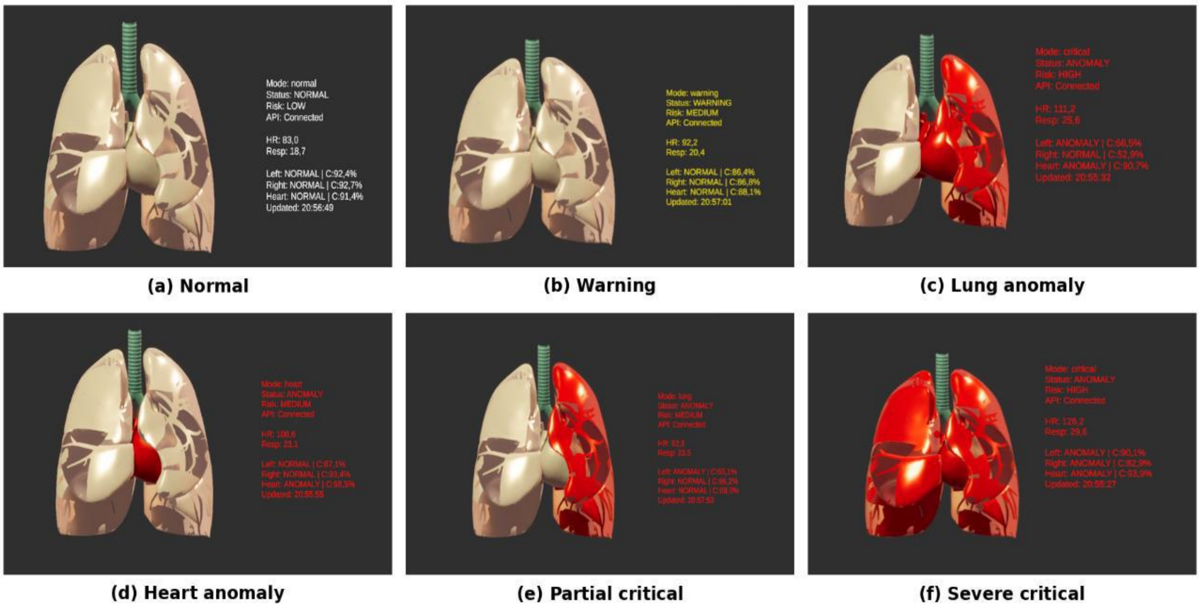

Fig. 4: Visualization results of the proposed system under different physiological conditions: (a) normal state, (b) warning state, (c) lung anomaly, (d) heart anomaly, (e) partial critical condition, and (f) severe critical condition.

Under normal conditions, physiological parameters remain within acceptable ranges and no abnormalities are detected. In warning states, moderate deviations produce early warning signals before a dominant organ-level abnormality occurs. In the lung anomaly scenario, affected areas are visually indicated in the digital twin environment. In critical cases, multiple anomalies occur simultaneously, and high-risk alerts are generated along with distinct visual changes.

The visualization layer enhances interpretability by transforming numerical model outputs into intuitive semantic and visual representations. This enables users to understand system behavior more rapidly and use the platform more effectively for real-time monitoring and decision support.

3.6 Comparison with Related Studies

To evaluate the effectiveness of the proposed method, the obtained results were compared with representative studies in the anomaly detection and digital twin literature.

Table 3: Indicative Comparison with Related Studies

Study	**Performance**
Ukil et al. [9]	~0.85 F1-score
Ahmed et al. [17], [18]	~0.90 accuracy
ProposedMethod	0.8241 F1-score / 0.8939 accuracy

This comparison demonstrates that the proposed system exhibits competitive performance when compared to representative studies. However, since the data sets used, target definitions, and evaluation metrics differ, these values should be interpreted with caution.

In this context, the primary purpose of the comparison is not to claim numerical superiority; rather, it is to demonstrate that the proposed framework, which integrates multi-label anomaly detection, rule-based interpretation, and real-time digital twin visualization within a single integrated architecture, is analytically valid.

IV. Conclusion

This study presents a metaverse-based digital twin framework that brings together synthetic data generation, six-dimensional feature engineering, an LSTM-based multi-label prediction model, organ-specific threshold optimization, rule-based semantic inference, and Unity-based visualization components into a single structure. The system predicts problems at the organ level in the left lung, right lung, and heart by looking at heart rate and respiratory signals at the same time. It then turns these predictions into system states that can be understood for real-time monitoring.

In the synthetic working environment used in this study, the proposed method showed good results, with accuracy rates of 89.39%, Micro-F1 of 82.41%, Macro-F1 of 82.00%, and Exact Match Accuracy of 72.17%. Also, at the organ level, the accuracy rates were 87.38% for the left lung, 90.29% for the right lung, and 90.50% for the heart. These results show that the proposed system is a useful and understandable way to analyze organ-level anomalies and monitor health in real time.

Nonetheless, the results of this study should be understood within its defined parameters. The LSTM model fits well with the time-based features of physiological signals. However, this study does not include a full ablation analysis or many comparisons with other architectures. The training and testing data were fabricated, thus the results should be regarded as preliminary indicators of potential rather than clinically validated outcomes. Later studies will use real patient data to test the proposed system, make systematic comparisons with other methods, include different biosignals, and test the system in real healthcare settings.

V. Acknowledgements

This study has been supported in part by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant 124E001.

References

[1] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, ‘Digital Twin in Industry: State-of-the-Art,’ IEEE Transactions on Industrial Informatics, vol. 15, no. 4, pp. 2405–2415, Apr. 2019, doi: 10.1109/TII.2018.2873186.

[2] M. Schluse and J. Rossmann, ‘From simulation to experimentable digital twins: Simulation-based development and operation of complex technical systems,’ in 2016 IEEE International Symposium on Systems Engineering (ISSE), Edinburgh, U.K., Oct. 2016, pp. 273–278, doi: 10.1109/SysEng.2016.7753162.

[3] T. Erol, A. F. Mendi, and D. Doğan, ‘The Digital Twin Revolution in Healthcare,’ in 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Oct. 2020, pp. 1–7, doi: 10.1109/ISMSIT50672.2020.9255249

[4] C. Angulo, L. Gonzalez-Abril, C. Raya, and J. A. Ortega, ‘A Proposal to Evolving Towards Digital Twins in Healthcare,’ in Bioinformatics and Biomedical Engineering, Cham, Switzerland: Springer, 2020, pp. 418–426, doi: 10.1007/978-3-030-45385-5_37.

[5] S. Adibi, A. Rajabifard, D. Shojaei, and N. Wickramasinghe, ‘Enhancing Healthcare through Sensor-Enabled Digital Twins in Smart Environments: A Comprehensive Analysis,’ Sensors, vol. 24, no. 9, Art. no. 2793, Apr. 2024, doi: 10.3390/s24092793.

[6] M. R. Kabir, F. S. Shishir, S. Shomaji, and S. Ray, ‘Digital twins in healthcare IoT: A systematic review,’High-Confidence Computing, vol. 5, no. 3, Art. no. 100340, Sep. 2025, doi: 10.1016/j.hcc.2025.100340.

[7] A. Musamih, I. Yaqoob, K. Salah, R. Jayaraman, Y. Al-Hammadi, M. Omar, and S. Ellahham, ‘Metaverse in Healthcare: Applications, Challenges, and Future Directions,’ IEEE Consumer Electronics Magazine, vol. 12, no. 4, pp. 33–46, Jul. 2023, doi: 10.1109/MCE.2022.3223522.

[8] Y. Yang, Z. Zhou, X. Li, X. Xue, P. C. K. Hung, and S. Yangui, ‘Metaverse for Healthcare: Technologies, Challenges, and Vision,’ International Journal of Crowd Science, vol. 7, no. 4, pp. 190–199, Dec. 2023, doi: 10.26599/IJCS.2023.9100020.

[9] A. Ukil, S. Bandyoapdhyay, C. Puri, and A. Pal, ‘IoT Healthcare Analytics: The Importance of Anomaly Detection,’ in 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), Crans-Montana, Switzerland, Mar. 2016, pp. 994–997, doi: 10.1109/AINA.2016.158.

[10] A. Castellani, S. Schmitt, and S. Squartini, ‘Real-World Anomaly Detection by Using Digital Twin Systems and Weakly Supervised Learning,’ IEEE Transactions on Industrial Informatics, vol. 17, no. 7, pp. 4733–4742, Jul. 2021, doi: 10.1109/TII.2020.3019788.

[11] Y. Santur, S. G. Santur, and M. Karaköse, ‘Architecture and implementation of a smart-pregnancy monitoring system using web-based application’, Expert Systems, vol. 37, no. 1, p. e12379, 2020, doi: 10.1111/exsy.12379.

[12] K. Arslanoğlu and M. Karaköse, ‘An Overview for Trustworthy and Explainable Artificial Intelligence in Healthcare’, in 2024 28th International Conference on Information Technology (IT), Feb. 2024, pp. 1–5.

doi: 10.1109/IT61232.2024.10475733.

[13] A. De Benedictis, N. Mazzocca, A. Somma, and C. Strigaro, 'Digital Twins in Healthcare: An Architectural Proposal and Its Application in a Social Distancing Case Study,' IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 10, pp. 5143–5154, Oct. 2023, doi: 10.1109/JBHI.2022.3205506.

[14] G. Bansal, K. Rajgopal, V. Chamola, Z. Xiong, and D. Niyato, 'Healthcare in Metaverse: A Survey on Current Metaverse Applications in Healthcare,' IEEE Access, vol. 10, pp. 119914–119946, 2022, doi: 10.1109/AC-CESS.2022.3219845.

[15] L. U. Khan, Z. Han, W. Saad, E. Hossain, M. Guizani, and C. S. Hong, 'Digital Twin of Wireless Systems: Overview, Taxonomy, Challenges, and Opportunities,' IEEE Communications Surveys & Tutorials, vol. 24, no. 4,pp. 2230–2254, 2022, doi: 10.1109/COMST.2022.3198273

[16] F. Laamarti, H. F. Badawi, Y. Ding, F. Arafsha, B. Hafidh, and A. E. Saddik, 'An ISO/IEEE 11073 Standardized Digital Twin Framework for Health and Well-Being in Smart Cities,' IEEE Access, vol. 8, pp. 105950–105961, 2020, doi: 10.1109/ACCESS.2020.2999871.

[17] S. Ahmed, J. S. Mertens, and L. Galluccio, 'A Digital Twin Framework for Human Organs Modeling and Forecasting in the Metaverse,' in 2025 3rd International Conference on Intelligent Metaverse Technologies & Applications (iMETA), Oct. 2025, pp. 39–44, doi: 10.1109/iMETA66706.2025.11306455.

[18] S. Ahmed, J. S. Mertens, and L. Galluccio, 'Metaverse-Enabled Digital Twins Employing Distributed Learning for Clinical Applications,' in 2025 IEEE Conference on Standards for Communications and Networking (CSCN), Sep. 2025, pp. 1–5, doi: 10.1109/CSCN67557.2025.11230713.

[19] LluisV, 'Z-Anatomy: Open-source 3D human anatomy model,' GitHub repository. [Online]. Available: https://github.com/LluisV/Z-Anatomy. [Accessed: 17-Apr-2026].