TWFN: An Architectural Framework for IoMT-Enabled Smart Healthcare System by Functional Heart Rate Variability Anomaly Detection Based on a Novel Optical Fiber Sensor

***Abstract*— Human vital signs are essential to the healthcare industry and applications of Internet of Medical Things (IoMT). As a significant vital sign signal, heart rate variability (HRV) provides into the functioning of the human stress levels and overall well-being. However, it is inconvenient for daily long-term HRV monitoring with conventional monitoring methods under the uncomfortable feelings. Moreover, it is hard to detect anomalies in HRV data to provide real-time previews about vital signs. To address these problems, a novel fiber interferometer-based optical fiber sensor is proposed to monitor human vital signs, and we propose a novel deep learning model (TWFN). Firstly, a novel module (ADSN) is proposed to apply graph modeling for obtaining spatial and temporal characteristics of HRV. Subsequently, an unsupervised generative adversarial network (VS-GAN) is proposed for the effective overcoming mode collapse and generator failure to converge to the target distribution better. The outcomes of the experiment might encourage the use of HRV-based healthcare application in IoMT-enabled healthcare sectors.**

***Index Terms*— Heart Rate Variability (HRV), Anomaly Detection, Optical Fiber Interferometer, Internet of Medical Things (IoMT), Healthcare Industry.**

I. Introduction

In an era defined by the convergence of the Internet of Medical Things (IoMT) technology and healthcare industry, the concept of heart rate variability (HRV) has emerged as a crucial metric for monitoring and understanding human cardiovascular health [1], [2]. HRV, as the variation in time intervals between successive heartbeats, reflects the influence of autonomic nervous system on the heart and can provide valuable insights into overall well-being of individuals [3]. It is one of the important vital sign parameters, and it can be extracted from the raw vital signs signals, such as ballistocardiography (BCG) and electrocardiogram (ECG) [4]. Recently, there is also a close connection between HRV, smart city development, public health policies, and image-based pollution estimation [5]. As an important indicator of individual health, HRV data can be collected through smart wearable devices and uploaded to the data center of smart cities. The big data platform of smart cities can analyze this data, identify possible health problem areas, and provide a basis for the formulation of public health policies [6]. Secondly, the environmental monitoring system of smart cities can use image-based pollution estimation technology to monitor environmental factors such as air quality and water quality in real time [7]. These data can help city managers understand the environmental conditions in a timely manner and formulate or adjust public health policies to reduce the impact of pollution on human health [8]. Smart city intelligent transportation systems and energy management systems can reduce pollution emissions and improve environmental quality, thereby indirectly improving HRV levels of human and promoting public health [9].

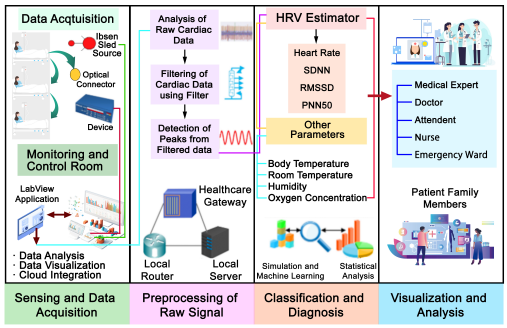


Fig. 1. Architectural diagram of proposed IoMT-based vital signs monitoring system with optical fiber sensor.

With the development of the healthcare industry requirements and related technology, the convergence of optical fiber sensors and the IoMT has sparked a revolution in healthcare monitoring [10]. Optical fiber sensors, harnessed for their precision and versatility, are seamlessly integrated into the fabrication of IoMT, creating a dynamic synergy that offers unparalleled advantages for the healthcare industry [11]. This transformative union empowers us to monitor health parameters with unprecedented accuracy, providing real-time insights into well-being while revolutionizing patient care. As shown in Fig. 1, as for the optical fiber sensor in healthcare monitoring, it has the advantages of high sensitivity, non-invasiveness, real-time, and continuous monitoring [12]. And optical fiber sensors are versatile and can simultaneously measure multiple parameters, offering a comprehensive view of human health. Traditional optical fiber sensors for vital

signs monitoring often rely on bending loss [13]. However, there are the following problems of this optical fiber sensor in practical application.

1) Bending loss may lead to signal loss when fibers are bent beyond a certain radius. This limits their application, especially in tight spaces. Optical fibers are fragile and prone to damage if bent or twisted excessively, resulting in costly repairs and downtime. Bending loss can distort signals, affecting data accuracy and reliability [14], such as in human blood pressure assessment and building safety monitoring.

2) There is a limit to how much a fiber can be bent before signal loss becomes unacceptable, posing challenges in confined spaces or around sharp edges [15].

Recent advancements in optical fiber sensors for human health monitoring, particularly fatigue condition monitoring, have been notable [16]. These sensors utilize light interference, scattering, and absorption, offering advantages over traditional electrical sensors, such as resistance to electromagnetic interference and enhanced electrical safety. Wearable fiber sensors are categorized into silicon and polymer fibers. Silicon fibers offer ultra-low-loss optical transmission but require structural design improvements for better mechanical properties [17]. Polymer fibers, on the other hand, provide excellent biocompatibility and stretchability, making them suitable for embedded sensing [18]. Optical fiber sensors exhibit high sensitivity and a good linear response, capable of monitoring body temperature, limb movement, and deformation with stability and speed. They are used in applications like joint posture, respiratory rate, and heart rate monitoring, and even in remote robotic arm control, highlighting their potential in human-computer interaction (HCI). However, challenges remain in system integration and ensuring long-term optical mechanical stability for reliable signal performance and monitoring [19]. As for the HRV anomaly detection, these sensors are usually difficult to meet the needs of long-term efficient and accurate for data acquisition and processing. In terms of data processing methods, some existing methods use the current popular graph modeling or spatial-temporal modeling mechanisms, which have also achieved better results [20]. However, in order to detect anomaly of HRV data with high quality and efficiency, the following problems need to be solved urgently.

1) These methods ignore the role of external knowledge for build sparse connection to prevent the over-fitting problem [21], such as in protein structure prediction and gear failure detection.

2) Mode collapse with generator failure to converge in the target distribution during the time-series data processing may affect the performance and result [22].

In this paper, to address the above-described important issues of bending loss, a novel optical fiber sensor system is proposed based on fiber interferometer, which utilizes the interference of light waves within an optical fiber to measure a wide variety of physical quantities. The proposed optical fiber sensor system is capable to exhibit great sensitivity, such that precise and accurate measurements can be carried out. In addition, a novel framework, termed TWFN, which consists of a novel module (ADSN) and a new generative adversarial network (VS-GAN) module, is proposed by us. It can be capable to detect anomalies of HRV data during human sleep better combined with the proposed optical fiber sensor system and tackle the above problems of data processing effectively.

The main contributions of this study are summarized as follows:

(1) A novel optical fiber sensor system based on fiber interferometer is proposed to monitor the raw HRV data accurately and effectively compared with other existing fiber sensors.

(2) We propose a novel deep learning framework (TWFN) for HRV anomaly detection, consisting of proposed ADSN and VS-GAN modules. It effectively extracts trend and spatial-temporal features using external knowledge, enhancing generalization and overcoming mode collapse in generator convergence better, outperforming state-of-the-art methods.

(3) To the best of our knowledge, we propose the first HRV dataset with the proposed optical fiber sensor, which is constructed during sleep, which enables anomaly detection framework making the most of human healthcare and possible IoMT applications.

In summary, optical fiber sensor system can accurately collect physiological data and provide accurate data sources for HRV anomaly detection. Deep learning framework can efficiently analyze the data acquired by optical fiber sensors, and achieve precise anomaly detection of HRV. IoMT integrates these technologies to build the intelligent medical monitoring system. Meanwhile, optical fiber sensors have high sensitivity and anti-interference, ensuring that the data is reliable. HRV anomaly detection can detect diseases such as cardiovascular and other diseases in the early stage. Deep learning techniques improve the accuracy and efficiency of detection. IoMT realizes remote medical and real-time monitoring, which is convenient for patients and doctors. These technologies are interrelated to promote progress in the field of medical and health. On the one hand, it can increase the early diagnosis rate of disease and reduce medical costs. On the other hand, it can also provide support for personalized medical care and formulate treatment plans according to the specific situation of the patient; promote the transformation of medical models to intelligence and remoteness; and improve the availability and quality of medical services.

II. Methodology

A. Proposed Optical Fiber Sensor System

The optical fiber sensor system in this work uses a 1550 nm distributed-feedback (DFB) laser as its light source. Because the optical fiber sensor operates on the concept of interference, variations in the external environment may result in random phase drift, which will cause the optical fiber interferometer's working point to stray from the linear region's center. The ability of interferometer to function can be significantly impacted by the working point's random drift. In order to acquire the BCG and ECG signal, a 3×3 coupler was used to split the interference light into three beams alongside a 120° phase variation. This technique is known as 3×3 demodulation. Through a 1×2 coupler, the DFB laser's light is sent into the reference and sensing fibers. Subsequently, at the 3×3

coupler, the light beams transmitting in the reference as well as sensing arms will interfere. The following equation can be used to determine the output light intensity of interferometer:

$$I_k = D + I_0 \cos[\varphi(t) - (r-1)(2\pi/3)] \tag{1}$$

where $\varphi(t)$ represents the phase difference of light signals, D represents the mean of output light intensity, I_0 represents the peak intensity of interference fringes, r represents the number of the output light path, and $r = 1, 2, 3$.

When used practically, $\varphi(t)$ comprises phase shifts brought on by the environment's fluctuation and the observed data, which are represented as follow

$$\varphi(t) = \phi(t) + \psi(t) \tag{2}$$

where $\varphi(t)$ reflects the signal to be determined, and $\psi(t)$ is the phase variance resulting from modifications in the environment.

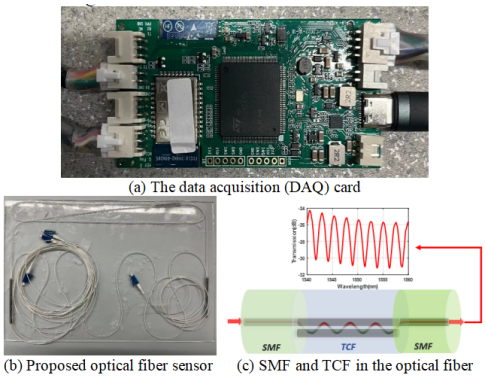
The interferometric optical fiber sensor receives the signal light emitted from a broadband source (BBS, SLED-1488/1650-10-FA-B, Shanghai Mai Xuan Laser CO., LTD). An optical spectrum analyzer (YOKOGAWA AQ6370D) alongside a resolution of 0.01 nm detects the transmission spectrum, as illustrated in Fig. 2. The spectra exhibit four interference dips at wavelengths of 1549.4 nm , $1549.72\ \mathrm{nm}, 1550.03\ \mathrm{nm}$, and 1550.34 nm , in that order. These dips are all useful for monitoring BCG signals.

As shown in Fig. 3, the 3×3 coupler demodulation technique can be used to demodulate the three optical intensity signals, which is illustrated in equation (1).

Following the process of distinction and cross-multiplication, the signal that has been demodulated is as follow

$$N = a(e-f) + b(f-d) + c(d-e) = \frac{3\sqrt{3}}{2} I_0^2 \varphi'(t) \tag{3}$$

variations in the polarization state and variations in light intensity will affect the value of I_0 in the real world. The following mathematical procedure is used to remove the influencing factors.



(a) The data acquisition (DAQ) card

(b) Proposed optical fiber sensor

(c) SMF and TCF in the optical fiber

Fig. 2. An optical fiber sensor-based system for monitoring human vital signs.

Initial, square the three input signals to gain

$$M = a^2 + b^2 + c^2 = \frac{3I_0^2}{2} \tag{4}$$

Subsequently split N through M and eliminate ${I_0}^2$ to obtain

$$P = \frac{N}{M} = \sqrt{3}\varphi'(t) \tag{5}$$

Following integral operation, the result is explained as follow

$$\sqrt{3}\varphi(t) = \sqrt{3}\phi(t) + \psi(t) \tag{6}$$

where $\psi(t)$ is thought of as a gradual changing quantity that a high-pass filter may remove. I_{out} represents the output signal of the 3×3 demodulation scheme, which is recognized as the BCG signal transmitted to the computer.

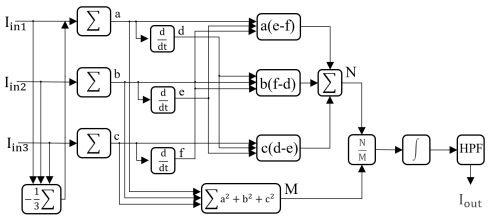


Fig. 3. Demodulation principle of the $\mathbf{3 \times 3}$ coupler.

The proposed optical fiber sensor system integrated into a smart mat is shown in Fig. 4. Fig. 5 shows special optical fiber manufacture industry in the production process of the proposed optical fiber sensor and the applications in human vital signs monitoring. Overall, compared with banding loss-based optical fiber sensors, the proposed optical fiber interference-based optical fiber sensor has the characteristics of light weight, compact structure, flexible geometry, explosion-proof, and anti-electromagnetic interference. It has higher sensitivity in vital signs monitoring.

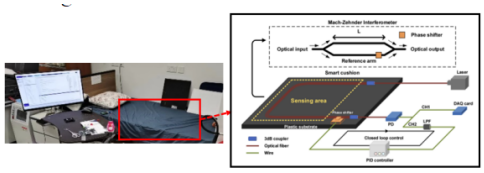


Fig. 4. The proposed smart mat-integrated optical fiber sensor system.

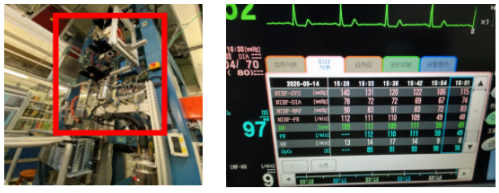


Fig. 5. Special optical fiber manufacture industry in the production process of the proposed optical fiber sensor and its application in vital signs monitoring.

B. Data Acquisition and Processing

HRV is a method used to assess the function of the autonomic nervous system, usually measured by the examination of the intervals during heartbeats. HRV can be employed to assess the stress levels of human body, mental health and

general physiological conditions. The following are the general steps for extracting HRV:

(1) Acquire raw data: HRV data is usually collected through BCG instruments, heart rate monitors, sports watches or other biosensors. It is necessary to obtain a set of continuous heartbeat timestamp data. Visualization of a sample of heart rate plot that extracted from the raw acquired signal is shown in Supplementary Fig. 1.

(2) Time domain analysis: time domain analysis method is usually used, which involves calculating a series of statistical parameters, such as mean, standard deviation, mean square error, etc., to describe changes in heartbeat intervals. Common time domain parameters include standard deviation of adjacent R-R intervals (RMSSD) and standard deviation of all R-R intervals (SDNN).

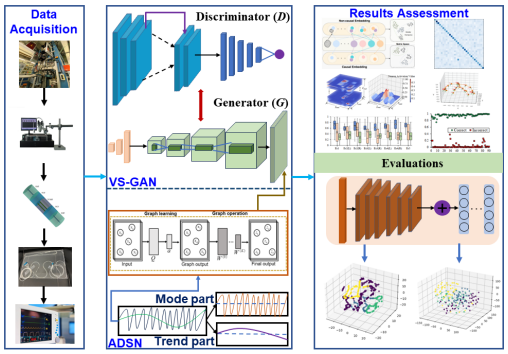


Fig. 6. Overall architecture of the proposed framework.

(3) Frequency domain analysis and spatial-temporal feature visualization: frequency domain analysis decomposes HRV into components in different frequency ranges, usually including low frequency (LF), high frequency (HF) and very low frequency (VLF) components. LF is associated with sympathetic nervous activity, and HF is associated with parasympathetic nervous activity. This can be achieved through methods such as Fourier transform. Visualization of three components and changes with different frequencies is shown in Supplementary Fig. 2.

(4) Nonlinear analysis: nonlinear analysis methods are used to explore complex dynamic characteristics in HRV signals, such as signal complexity and nonlinear relationships. These methods include Poincaré diagrams, complexity analysis, etc.

(5) Calculation of HRV: select appropriate HRV parameters and calculation methods based on application goals. Different parameters can provide different levels of information, including the autonomic regulation function of heart.

(6) Interpretation of results: explain the physiological significance of HRV parameters obtained through analysis. Fig. 6 shows the overall architecture of the proposed framework.

C. Spatial-Temporal Modeling and Module (ADSN)

In the proposed ADSN module, we define $G(V,E)$ as the graph structure of nodes, where V represents the set of all nodes, E represents the set of all adjacent node edges. When the distance $d(x_i,x_j)$ within adjacent nodes v_i and v_j is less compared to the radius l of the network, the two adjacent nodes are regarded to have connected edges in the graph $G(V,E)$. For $G(V,E), v_i \in V$, and $1 \leqslant i \leqslant n$. As for v_i, set of neighbor nodes with each node could be articulated by $N(v_i) = \{v_j \in V : (v_i,v_j) \in E\}$ based on external knowledge of graph structure.

The external knowledge, such as in the diffusion model for time series and spatial-temporal data, have been used in a branch of research [21]. The proposed ADSN module integrated the graph modeling into a deep neural network architecture and achieved a sparse connection between the layers of the proposed framework to prevent over-fitting effectively.

While identifying anomalies for node v_i, F_i represents the state of node, $F_i = 0$ indicates that v_i is a faultless node, and $F_i = 1$ indicates that v_i node has a fault. Based on the principle of spatial similarity in the measured values of adjacent nodes, the measured values of v_i are compared with all its adjacent nodes, and the state of the node is diagnosed based on the comparison results. Two adjacent nodes are given $(v_i,v_j) \in E, v_i \in N(v_i)$, defined as $c(v_i,v_j)$ is the comparison function between v_i and its neighboring nodes, abbreviated as c_{ij}, then

$$c_{ij} = c(v_i,v_j) = \begin{cases} 1, if\ |x_i - x_j| > \xi \\ 0, otherwise \end{cases} \tag{7}$$

when v_i and its neighboring nodes have similar results, it returns a result of 0 , otherwise it returns a result of 1 . If c_i is the amount of adjacent nodes with $c_{ij} = 0, 0 \leqslant |c_i| \leqslant |N(v_i)|$, then c_i refers to the degree of similarity between node v_i and adjacent nodes; if $|c_i| \geqslant \theta_1$ (θ_1 is the preset threshold), which can diagnose v_i as a faultless node and set $F_i = 0$.

The following is an improvement to the node state $G(V, E)$ is the graph of all nodes, assuming that node $v_i \in V$, and the number of neighbor nodes is $n.N(v_i) = \{v_1, v_2, \ldots, v_r\}$, where $n = |N(v_i)|$. It is necessary to address the issue that v_i may have an instantaneous failure, compare its neighbor nodes q-times at $t_1, t_2, \ldots, t_q$. At time $t_m, 1 \leqslant m \leqslant q$, record the measurement value of node v_i as x_i^m, and record the measurement values of r neighboring nodes of the node as ${x_1}^m, {x_2}^m, \ldots, {x_r}^m$. Therefore, at the time $t_m, c^m(v_i,v_j)$ can be defined as the comparison function between node v_i and neighboring nodes, abbreviated as ${c_{ij}}^m$, then

$$c_{ij}^m = \begin{cases} 0, \text{ if } \left|x_i^m - x_j^m\right| \leq \xi \\ 1, otherwise \end{cases} \tag{8}$$

For the comparison results of node v_i and all neighboring nodes at different times, can be represented by an $r \times q$ matrix $\boldsymbol{C}$. And the comparison result c_{ij} between node v_i and neighbor node v_j can be expressed using the time redundancy detection result in the matrix $\boldsymbol{C}$, that is

$$c_{ij} = c(v_i,v_j) = \begin{cases} 0, if \sum_{m=1}^{q} c_{ij}^m \leq \theta_2 \\ 1, otherwise \end{cases} \tag{9}$$

where θ_2 is a pre-given threshold. Finally, determine the F_i state value of node v_i according to the number of $c_{ij} = 0$ in adjacent nodes and the θ_1 threshold.

Meanwhile, we extract key features or patterns from the mode decomposition of HRV data, decrease the number of

dimensions in the data, retain important information, as well as thus simplify the complexity of the data. In addition, since the data of spatial-temporal modeling contains long-term trends, we can separate the long-term trends in the data by trend extraction on the HRV data, so that the basic trends of the data can be analyzed more clearly. The mode decomposition and trend extraction can be illustrated as follows

$$\rho_i = \text{SeriesDecomp}\left(L_{t_0:t_n}\right) \tag{10}$$

$$\vartheta_i = \text{SeriesTrendExtract}\left(L_{t_0:t_n}\right) \tag{11}$$

$$T = \sum_{i=0}^{n} \text{Roll}\left(\rho_i, \vartheta_i\right) \tag{12}$$

where ρ_i represents the mode components obtained after decomposition, ϑ_i refers to the extracted trend from raw signal. Fig. 7 shows the basic mechanism of the proposed ADSN module.

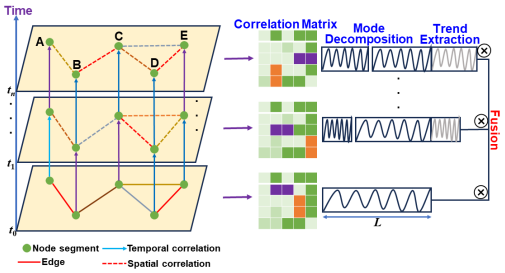


Fig. 7. Structure of the proposed ADSN module.

For the diagnosed node, according to the given q and θ_2 thresholds, the comparison results of the node and all neighboring nodes under time redundancy detection are calculated; then based on the given threshold θ_1, performing a threshold test on each node v_i; when $|c_i| \geqslant \theta$, let $F_i = 0$; if $|c_i| < \theta_1$, it means that the node fails the threshold test, and no operation is performed, and exit threshold testing. Then for the fault-free node $v_i\,(F_i = 0)$ that passes the threshold test, if the neighbor node v_i meets $c_{ij} = c_{ji} = 0$, then let $F_i = 0$, until propagate node status to all neighbor nodes. The algorithm for node failure detection and spatial-temporal feature extraction is described as Algorithm 1 in Supplemental Material.

D. Proposed Unsupervised Module (VS-GAN)

The Wasserstein loss, also referred to as Wasserstein distance or Earth Mover's distance (EMD), is a metric employed to calculate the separation across two probability distributions over a region χ in some space. In the context of generative adversarial network (GAN), it is used to train the generator (G) and discriminator (D).

There are some drawbacks in traditional GAN model, especially in high-dimensional spaces. The VS-GAN uses the Wasserstein distance instead. The main idea behind VS-GAN is to have a more stable and meaningful loss function, which can lead to better training dynamics and generate higher-quality samples. In the VS-GAN, the discriminator is constrained to be a 1-Lipschitz function. This means that its output should not vary too much within the space of inputs. This constraint is enforced by adding a penalty term to the loss function, which is proportional to the gradient norm of the discriminator's output with respect to its input. This is usually accomplished through weight clipping or, more effectively, by using a gradient penalty. The loss function for the VS-GAN can be formulated as follow

$$L_1 = E_{x\sim P_{data}}[D(x)] - E_{z\sim P_z}[D(G(z))] \tag{13}$$

where $D(x)$ refers to the discriminator's output for real data $x, G(z)$ reflects the generated data from noise z, and $D(G(z))$ is the output of discriminator for generated data. The generator makes an effort to reduce this loss, while the discriminator tries to maximize it.

In addition, VS-GAN offers a penalty term for the Lipschitz constraint. The loss function with the gradient penalty is as follow

$$\begin{aligned} L_2 =& E_{x\sim P_{data}}[D(x)] - E_{z\sim P_z}[D(G(z))] \\ &+ \lambda E_{\sigma\sim P_\sigma}\left[\left(\|\nabla_\sigma D(\sigma)\|_2 - 1\right)^2\right] \end{aligned} \tag{14}$$

where σ is a randomly interpolated point between real data and generated data, and λ represents a hyperparameter controlling the strength of the gradient penalty.

In VS-GAN model, the aim is to find a mapping across two domains, that is to say, domain X and domain Y, without the need for paired data, for example, there is no direct correspondence between individual samples in X and Y. For example, this could involve translating images of horses to zebras or vice versa. The key idea is to enforce cycle consistency. This implies that if we use the learnt mappings to transform a picture from domain X to domain Y and back to domain X, we should get back something similar to the original sample. The cycle consistency loss is used to enforce this property. The forward cycle consistency can be defined as follow

$$\begin{aligned} &L_c\left(G_{X\to Y}, G_{Y\to X}\right) = \\ &E_{x\sim P_{data}(X)}\left[\|G_{Y\to X}\left(G_{X\to Y}(x)\right) - x\|_1\right] \end{aligned} \tag{15}$$

This term measures the variation within the original input x and the signal that has been translated from X to Y and then back to X. The backward cycle consistency can be defined as follow

$$\begin{aligned} &L_c\left(G_{Y\to X}, G_{X\to Y}\right) = \\ &E_{y\sim P_{data}(Y)}\left[\|G_{X\to Y}\left(G_{Y\to X}(y)\right) - y\|_1\right] \end{aligned} \tag{16}$$

This term does the same, but in the reverse direction: from Y to X and back to Y. The total cycle consistency loss is the sum of the following two items:

$$L_t = L_c\left(G_{X\to Y}, G_{Y\to X}\right) + L_c\left(G_{Y\to X}, G_{X\to Y}\right) \tag{17}$$

This loss encourages the generator networks $G_{X\to Y}$ and $G_{Y\to X}$ to learn mappings that are consistent in both directions. It helps in producing realistic and coherent translations between the two domains.

Overall, the cycle consistency loss plays a crucial role in training VS-GAN by enforcing that the translation process is bidirectional and that the original information is preserved across the transformation process. The algorithm for the

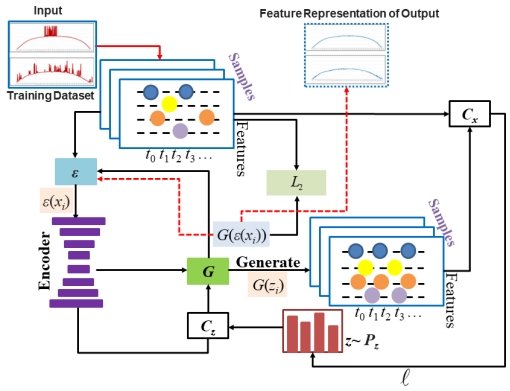


Fig. 8. Basic structure of the proposed VS-GAN module.

TABLE I
IMPORTANT HYPERPARAMETER VALUES OF ALL METHODS IN EXPERIMENT.

Method	Batch Size	Learning Rate	Iteration Step
SVM [24]	40	0.00001	100
KNN-TL [23]	40	0.00001	100
AMNet [25]	180	0.000001	200
DAEMON [27]	140	0.000001	200
BeatGAN [26]	60	0.000001	100
GHRN [20]	55	0.000001	150
GAAD [28]	50	0.000001	100
CCGR [29]	52	0.000001	100
PLSTM [30]	60	0.000001	100
TKF [31]	60	0.000001	100
TWFN (Ours)	64	0.000001	100

proposed VS-GAN module is described as Algorithm 2 in Supplemental Material.

Basic structure of the proposed VS-GAN module is shown in Fig. 8. To collect training samples, we implement a sliding window approach alongside a window size of t and a step size of s to partition the original time series into N subsequences denoted as $X = \left\{\left(x_i^{1.t}\right)\right\}_{i=1}^{N}$, where T represents the total length of the time series. In practical scenarios, determining the ground truth can be challenging, and instances of anomalous data points are relatively rare. As a result, we make the assumption that all data points within the training samples are considered normal. In addition, we create another set $Z = \left\{\left(z_i^{1\ldots k}\right)\right\}_{i=1}^{N}$ by generating data points z_i from a random space that follows a normal distribution. Here, k represents the dimension of this latent space. Subsequently, we utilize both X and Z as input data for the proposed VS-GAN module. The training process driven by the following objective function is as follow

$$F = \min_{\{\varepsilon,G\}} \max_{\{C_x \in X, C_z \in Z\}} L_X\left(C_x, G\right) + L_z\left(C_z, \varepsilon\right) + L_c\left(G_{X\to Y}, G_{Y\to X}\right) \tag{18}$$

III. Experiments and Analysis

A. Implementation Details

In this study, the framework was implemented through a computer that included graphics processing unit (GPU) and NVIDIA GeForce RTX4060 graphics card. Python was used to implement the model. Graph Heterophily Resistant Network (GHRN) [20], K-Nearest Neighbor Transfer Learning (KNN-TL) [23], support vector machine (SVM) [24], Adaptive Multi-frequency Graph Neural Network (AMNet) [25], BeatGAN [26], DAEMON [27], GAAD [28], CCGR [29], PLSTM [30] and TKF [31], are set as baseline models, respectively. We also listed important hyperparameter values of all methods in Table I, including batch size, learning rate and iteration step values.

Herein, there are four commonly used evaluation metrics applied, which includes F1-score, Accuracy, Precision and Recall. Some evaluation metrics are defined as follows

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

$$Recall = \frac{TP}{TP + FN} \tag{20}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{21}$$

where TN reflects true negative, FN denotes false negative, TP reflects true positive, and FP stands for false positive, respectively.

In experiment, we collected HRV data from 600 subjects, including 300 males and 300 females. We also used different optical fiber sensors for comparative experiments and research. There are many types of optical fiber sensors, and different types of optical fiber sensors have different requirements and responses to input light intensity and condition. For example, intensity-type optical fiber sensors mainly rely on changes in input light intensity for measurement. Changes in input light intensity will directly affect its output signal; phase-type optical fiber sensors are relatively insensitive to changes in input light intensity, but are more concerned with changes in the phase of the input light. It is usually used for high-precision measurements; distributed optical fiber sensors can measure parameters at multiple points along the length of the optical fiber, and have high requirements for the stability and uniformity of the input lighting condition and intensity.

Detail information of data preprocessing can be found at Supplemental Material.

B. Experimental Results and Analysis

Optical fiber sensors are particularly well-suited for vital sign monitoring due to their distributed measurement capabilities, cost-effectiveness, high sensitivity, and inconspicuous nature, yielding satisfactory outcomes. However, factors such as manufacturing process, size, cost, and feasibility significantly impact their performance. The manufacturing process determines the quality and sensitivity of the sensors; high-quality sensors offer better accuracy but can be costly due to complex processes. Sensor size also affects performance; smaller sensors are ideal for high-precision, space-constrained applications but may compromise sensitivity and stability. Easy installation and maintenance improve reliability and efficiency, whereas complex procedures increase costs and limit use. As shown in Table II, we compare the proposed

TABLE II
COMPARISON OF PERFORMANCE WITH DIFFERENT OPTICAL FIBER SENSORS.

Type	Manufacturing	Size	Feasibility
Mach-Zehnder interferometer	Easy	Large	Feasible
Sagnac interferometer	Medium	Large	Feasible
Michelson interferometer	Medium	Large	Feasible
Fabry-Perot interferometers	Difficult	Small	Feasible
Banding loss	Medium	Small	Feasible
Ours	Easy	Small	Feasible

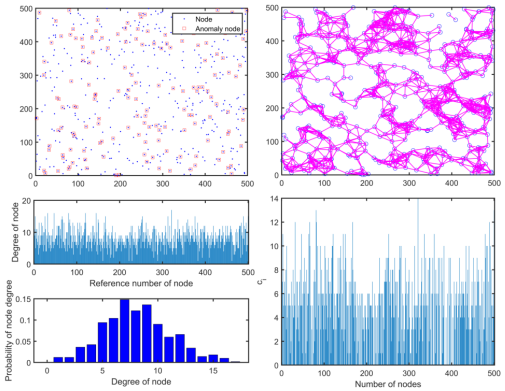


Fig. 9. The similarity degree of nodes and its spatial-temporal feature distribution.

optical fiber sensor with other important sensors based on manufacturing, size, and feasibility.

In order to fully extract the spatial characteristics of nodes and study the impact of the spatial similarity between nodes and the spatial-temporal features of nodes on the detection effect, in the experiment, we analyzed the correct detection rate (CDR) under different average node degrees of the network. Changes in CDR and false alarm rate (FAR) as well as the impact of the mean node degree and threshold θ_1 on CDR and FAR. When the probability of node failure is p, the fault diagnosis accuracy can be expressed by α/n_p, where n_p represents the amount of all faulty nodes and α is the number of correctly diagnosed nodes among all faulty nodes; enhancing the node defect detection technique is essential. Performance has to increase the pace of problem diagnosis while simultaneously bringing the rate of false fault alarms down to a low level.

In the experiment, every node has an equal transmission range. And replicate the node detection situation when the network node failure rate p is set as $p = \{0.05, 0.10, 0.15, 0.20, 0.25, 0.30\}$, respectively. Fig. 9 shows the changes in node CDR as the failure rate p and threshold θ_1 change, c_i represents the similarity degree of similar nodes.

It can be found from Table III that during the fault detection process, CDR will decrease as θ_1 increases. Therefore, both in experiments and model design, as the node failure rate p increases, in order to achieve better performance in assessment and anomaly detection, the threshold θ_1 should decrease accordingly. For example, when the failure rate $p = 0.05$, the threshold $\theta_1 = 8$ can make the CDR reaches 94.94%, which is better than the performance of other thresholds. Meanwhile, as shown in Fig. 10, we performed modal decomposition on the raw signal and extracted its changing trend. The decomposed modes and extracted trends make the source signal more visual and help to intuitively understand the structure and changes of the signal.

TABLE III
CHANGE OF CDR WITH DIFFERENT FAILURE RATES.

Failure Rate	CDR
0.30	70.05%
0.25	74.76%
0.20	80.12%
0.15	85.36%
0.10	89.83%
0.05	**94.94%**

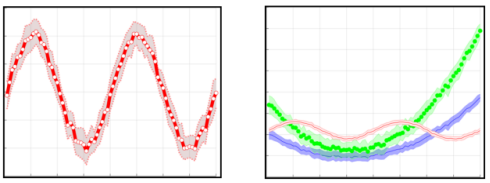


Fig. 10. Modal decomposition and trend extraction from the raw signal.

To investigate how the temporal redundancy detection approach affects the immediate fault diagnosis of nodes, the number of valid nodes (excluding instantaneous faulty nodes) is defined as n_g, the number of nodes with instantaneous faults is n_t, and the number of failed nodes is n_f.n_{gf} represents the number of valid nodes misdiagnosed as failed nodes, and n_{tf} represents the number of nodes with instantaneous faults misdiagnosed as failed nodes. Then the FAR of node fault diagnosis can be defined as follow

$$FAR = \frac{n_{gf} + n_{tf}}{n_g + n_t} \tag{22}$$

Meanwhile, the probability of an instantaneous failure due to sensor sensing errors is defined as p_s, and the probability of an instantaneous failure due to node problems is defined as p_c. Under the same experimental environment with the given node failure rate $p = \{0.05, 0.10, 0.15, 0.20, 0.25, 0.30\}$. When p_s and p_c change, it can affect the CDR and FAR. Fig. 11 illustrates the change of CDR with the change of $p_s(\theta = 1, 2,$ and 3 , respectively) and the change of FAR with the change of $p_c(\theta = 1, 2,$ and 3 , respectively). It can be found that, no matter how the instantaneous failure rate p_s of the node increases, the CDR of the node always remains at a value close to 1 , which shows that the instantaneous failure of the node has little impact on the assessment and anomaly detection accuracy.

We also used the proposed model in a number of tests and evaluated it against other baseline models with different subjects. The experimental results are shown in Table IV and Table V. The proposed TWFN framework achieves the best performance. It can be observed that most machine learning models, such as SVM and KNN-TL, which can achieve better results for the collected HRV data in experiments. The reason is that the structure of these models for processing sequence

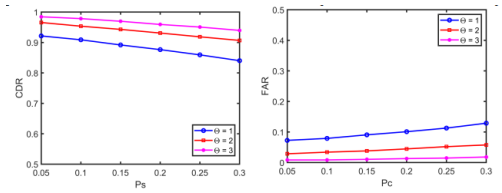


Fig. 11. The change of CDR and FAR with different θ values.

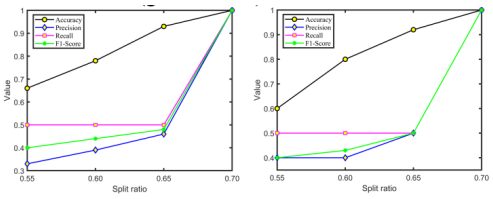


Fig. 12. Experimental results of different split ratios with different subjects.

abnormal patterns is relatively simple, mostly for timing-dependent anomalies, and obvious numerical changes in individual dimensions will not significantly affect the model. Other deep learning models as baselines often overlook external knowledge, unsupervised learning, model collapse and improve the convergence of generator, which are crucial for performance. External knowledge provides additional information and priority insights, helping models understand data features and patterns better, thus enhancing anomaly identification. For example, domain-specific knowledge can guide models to focus on particular anomalies, improving detection accuracy. Unsupervised learning is vital because anomalies are rare and hard to label. It allows models to learn patterns from large amounts of normal data, identifying exceptions when new data deviates from these patterns. This approach does not require extensive data, making it effective for real-world applications. Model collapse can lead to errors in anomaly detection. By integrating external knowledge and enhancing generators, this issue can be mitigated. Improved generators diversify data, increasing mode coverage and enhancing anomaly detection capabilities. The combination of these factors boosts the accuracy, robustness, and generalization of deep learning models, enabling them to better handle complex data. Together, they enhance the ability of models to process and adapt to intricate HRV data, ensuring more reliable and precise anomaly detection. And in statistics, the p-value by the T-test is generally $p < 0.05$ as significant, which means that the probability of the difference between samples caused by sampling error is less than 0.05 [32]. In experiments, we fine-tuned the split ratio of the HRV data as $\{0.55, 0.60, 0.65, 0.70\}$, respectively. Fig. 12 shows the experimental results of different split ratios with different subjects.

The fitness value is applied to evaluate the quality or fitness of a method. When the fitness value curve converges, it usually means that the model has found the optimal solution to the problem. Fast convergence means that the algorithm can reach a stable fitness value in a relatively short time or a small number of iterations, indicating that the model has high efficiency and optimization capabilities. Fig. 13 shows the change of objective function values and the best performance of the TWFN framework (green curve).

As shown in Table VI, we also conduct experiments to evaluate the average variance and GPU memory usage of the proposed TWFN framework and other baseline methods during model training. The average variance of the TWFN framework is lower than the baseline method. However, the performance in GPU memory usage is slightly inferior to PLSTM because PLSTM is an improvement based on long-short term memory (LSTM), and the parameter volume of

TABLE IV
AVERAGE EXPERIMENTAL RESULTS BASED ON MALE SUBJECTS.

Methods	Accuracy	Precision	Recall	F1-score	*p-value*
SVM [24]	0.91	0.92	0.97	0.95	0.0473
KNN-TL [23]	0.90	0.93	0.95	0.94	0.0421
AMNet [25]	0.91	0.91	1.00	0.95	0.0366
DAEMON [27]	0.88	0.92	0.94	0.93	0.0352
BeatGAN [26]	0.86	0.93	0.91	0.92	0.0360
GHRN [20]	0.91	0.91	1.00	0.95	0.0344
GAAD [28]	0.91	0.92	0.95	0.96	0.0315
CCGR [29]	0.93	0.94	0.97	0.97	0.0279
PLSTM [30]	**0.94**	0.94	0.97	0.96	0.0265
TKF [31]	0.91	0.92	0.94	0.94	0.0235
TWFN (Ours)	0.93	**0.95**	**1.00**	**1.00**	**0.0213**

the model is relatively low. According to all the above experimental results, compared with these models, the proposed TWFN framework is relatively more robust and the overall performance of the model is better. Some other baseline methods, such as DAEMON and TKF, can also extract the spatial-temporal features of data and analyze them, but they ignore the effective extraction of the changing trend of data, or are only based on thresholds or simple distance metrics, which makes it difficult to fully and automatically learn the latent feature representation of data, and it is also difficult to learn the complex latent distribution in the data. The proposed TWFN framework introduces ADSN and VS-GAN modules to fully obtain the changing trend of HRV data. It is based on the GAN architecture and can learn the complex latent distribution in the data. And compared with other methods based on manually designed feature extraction, the features learned by TWFN are often more representative and discriminative. Compared with other baseline methods, it can better obtain nonlinear and multimodal features in HRV data, and can effectively solve the problems of fuzzy anomaly definition, lack of data labels and complex time correlation in HRV data, and has better generalization ability.

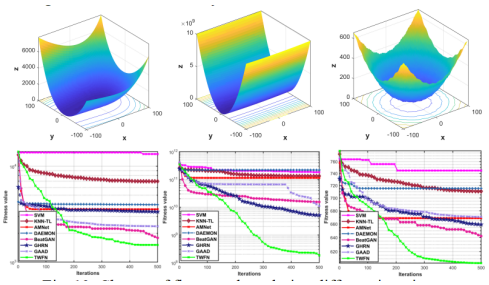


Fig. 13. Change of fitness values during different iterations.

We also conduct the statistical analysis on the relation-

TABLE V
AVERAGE EXPERIMENTAL RESULTS BASED ON FEMALE SUBJECTS.

Methods	Accuracy	Precision	Recall	F1-score	*p-value*
SVM [24]	0.94	0.96	1.00	0.94	0.0489
KNN-TL [23]	0.95	0.96	1.00	0.95	0.0455
AMNet [25]	0.96	0.97	0.97	0.96	0.0411
DAEMON [27]	0.96	0.96	1.00	0.97	0.0402
BeatGAN [26]	0.96	0.98	0.95	0.97	0.0395
GHRN [20]	0.96	0.96	0.98	0.97	0.0353
GAAD [28]	0.94	0.95	0.96	0.95	0.0348
CCGR [29]	0.98	0.97	0.98	0.98	0.0284
PLSTM [30]	0.97	0.98	0.98	0.97	0.0282
TKF [31]	0.95	0.95	0.98	0.96	0.0299
TWFN (Ours)	**0.98**	**1.00**	**1.00**	**0.99**	**0.0275**

TABLE VI
COMPARISON OF AVERAGE VARIANCE AND GPU MEMORY USAGE WITH DIFFERENT METHODS.

Methods	Male Variance	Female Variance	GPU
SVM [24]	16.67	16.95	3648MB
KNN-TL [23]	15.65	16.64	3552MB
AMNet [25]	14.68	15.75	3215MB
DAEMON [27]	14.86	15.53	4123MB
BeatGAN [26]	14.47	14.95	3166MB
GHRN [20]	14.59	15.02	3718MB
GAAD [28]	14.75	15.66	4126MB
CCGR [29]	13.25	14.22	2865MB
PLSTM [30]	12.98	13.85	**2810MB**
TKF [31]	14.81	13.94	2843MB
TWFN (Ours)	**12.96**	**13.52**	2824MB

ship between the sampling points of the average HRV data and anomaly detection for subjects of different genders. The p-value represents the probability of observing the current data or more extreme cases if the null hypothesis is true. The confidence interval is an estimate range of the overall parameter, giving a possible value range with a certain confidence level. p-value mainly provides evidence from the perspective of hypothesis testing, while confidence interval gives a specific range of parameter estimates. A smaller p-value means stronger evidence to reject the null hypothesis, which also corresponds to a narrower confidence interval width. If the p-value is small, it indicates that the result is statistically significant, which means that the estimate of the population parameter is more accurate, making the width of the confidence interval relatively narrow. On the contrary, if the p-value is large, the result is not significant, the uncertainty in the estimate of the population parameter increases, and the width of the confidence interval tends to be wider. Fig. 14 shows the changes between the p-value, estimated value, ground truth, and the widths of different confidence intervals.

Moreover, vital signs monitoring of drivers has been a popular research topic. HRV anomaly of driver can easily lead to some traffic accidents, endangering the lives of drivers and passengers. Therefore, HRV anomaly detection helps assess the driver's health and prevent accidents. Among them, HRV monitoring is particularly important because cardiovascular diseases (CVDs) have become the leading cause of death among various fatal diseases in the world. As shown in Fig. 15, we also focus on phase-sensitive fiber interferometer solutions, and also achieve the better performance of driver's HRV anomaly detection in practical.

Discussions of our work can be found at Supplemental Material.

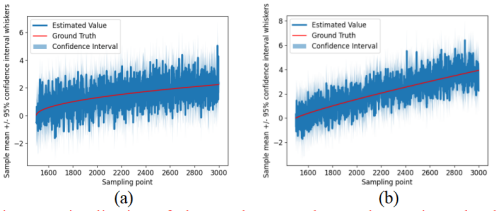


Fig. 14. Visualization of changes between the p-value, estimated value, ground truth, and the width of different confidence intervals. (a): Male subjects; (b): Female subjects.

Fig. 15. Application of the proposed optical fiber sensor for HRV monitoring and analysis in vehicle driving.

C. Ablation Study

To validate the importance of spatial-temporal modeling and unsupervised learning mechanism in the proposed model, we replaced the spatial-temporal modeling module (w/o ST), unsupervised learning module (w/o US), and replaced both spatial-temporal modeling module and unsupervised learning module (w/o USST). By comparing the Accuracy, Precision, Recall and F1-score, TWFN exhibits the best performance compared to another two variant models. Table VII shows the average experimental results of different variant models.

1) Impact of different components : To demonstrate the effectiveness of different modules in TWFN, we separately use the two modules (No. 1 and No. 2). Compared with the baseline results, these tests confirm each module can improve the detection accuracy, which is attributed to US enhancing potential structure and pattern representation capabilities, and ST effectively extracting and modeling spatial-temporal feature information. In addition, we also conduct experiments on combination of two modules (No. 3) to evaluate the impact when two modules are used together. Compared with using a single module, the results demonstrate that each combination can improve the performance, and the best detection results achieve when using ST and US modules together.

2) Requirements for computation: In variant models, due to the removal of some modules, the GPU memory usage will inevitably decrease. By comparison the GPU memory usage of the proposed TWFN framework with variant models (No. 1, No. 2, and No. 3), it may also lead to a relatively long simulation time. Long simulation time will reduce work efficiency, especially in practical scenarios that need to be responded quickly. Therefore, in future work, we will use parameter optimization technology, such as pruning and quantification

to reduce the number of parameters. Meanwhile, it will also try to use more efficient hardware equipment or optimization algorithms to shorten the simulation time.

TABLE VII
AVERAGE EXPERIMENTAL RESULTS OF VARIANT MODELS.

No.	Model	GPU	Accuracy	Precision	Recall	F1-score
1	w/o ST	2765MB	0.91	0.94	0.94	0.92
2	w/o US	2774MB	0.90	0.92	0.91	0.90
3	w/o USST	2715MB	0.88	0.90	0.90	0.89
4	TWFN (Ours)	**2824MB**	**0.94**	**0.96**	**0.95**	**0.95**

IV. CONCLUSION

In this paper, in the healthcare industry and IoMT, assessment the quality of HRV plays a pivotal role. It enables chronic disease management. HRV assessment and anomaly detection supports product development, enhances clinical trials, and ensures data security. In the preventive health, it helps identify individuals at risk, reducing healthcare industry costs. Moreover, it facilitates telehealth services and healthcare industry analytics, improving patient care and advancing healthcare innovation in real world. In this study, following fiber interferometer, a novel optical fiber sensor is proposed. It is capable of acquiring the raw human vital sign signal accurately and effectively. Furthermore, we introduce an architectural framework termed TWFN, comprising a novel module known as ADSN and a new generative adversarial network called VS-GAN. This framework is designed for enhanced processing of HRV data and improved detection of anomalies in HRV data during human sleep for medical quality monitoring when combined with the proposed optical fiber sensor. In summary, the quality of data can be improved and the lack of optical fiber sensors can be addressed to capture vital signs signals, which is a convenient and efficient solution to medical monitoring in both IoMT and healthcare industrial applications.

REFERENCES

[1] V. Catrambone and G. Valenza, "Nervous–system–wise functional estimation of directed brain–heart interplay through microstate occurrences," *IEEE Transactions on Biomedical Engineering*, 2023.

[2] S. Murase *et al.*, "Interstitial-fluid shear stresses induced by vertically oscillating head motion lower blood pressure in hypertensive rats and humans," *Nature Biomedical Engineering*, pp. 1–24, 2023.

[3] J. He and N. Jiang, "Fast multi-level mental stress identification from bispectrum-based heart rate variability feature," *IEEE Transactions on Industrial Informatics*, 2023.

[4] K. Wang *et al.*, "Forward privacy preservation in iot-enabled healthcare systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1991–1999, 2021.

[5] H. Liu *et al.*, "Opera: Optional dimensional privacy-preserving data aggregation for smart healthcare systems," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 857–866, 2022.

[6] T. J. Steiner and L. J. Stovner, "Global epidemiology of migraine and its implications for public health and health policy," *Nature Reviews Neurology*, vol. 19, no. 2, pp. 109–117, 2023.

[7] T.-C. Lin *et al.*, "Unmasking air quality: A novel image-based approach to align public perception with pollution levels," *Environment International*, vol. 181, p. 108289, 2023.

[8] M. Wang *et al.*, "Perceptually quasi-lossless compression of screen content data via visibility modeling and deep forecasting," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 6865–6875, 2022.

[9] D. Kim *et al.*, "Association between health-related quality of life and heart rate variability in elderly individuals with cognitive impairment in korea: cross-sectional study," *BMC Geriatrics*, vol. 23, no. 1, p. 847, 2023.

[10] F. Khan *et al.*, "A secure ensemble learning-based fog-cloud approach for cyberattack detection in iomt," *IEEE Transactions on Industrial Informatics*, 2023.

[11] S. Qahtan *et al.*, "Novel multi security and privacy benchmarking framework for blockchain-based iot healthcare industry 4.0 systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6415–6423, 2022.

[12] Q. Wang *et al.*, "Assessment of heart rate and respiratory rate for perioperative infants based on elc model," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13685–13694, 2021.

[13] H. Yang *et al.*, "Computational design of ultra-robust strain sensors for soft robot perception and autonomy," *Nature Communications*, vol. 15, no. 1, p. 1636, 2024.

[14] Z. C. Yi Zhang and H. I. Hee, "Noninvasive measurement of heart rate and respiratory rate for perioperative infants," *Journal of Lightwave Technology*, vol. 37, no. 11, pp. 2807–2814, 2019.

[15] Y. Li *et al.*, "An all-optical multidirectional mechano-sensor inspired by biologically mechano-sensitive hair sensilla," *Nature Communications*, vol. 15, no. 1, p. 2906, 2024.

[16] R. Jha, P. Mishra, and S. Kumar, "Advancements in optical fiber-based wearable sensors for smart health monitoring," *Biosensors and Bioelectronics*, p. 116232, 2024.

[17] Z. Gao *et al.*, "Advances in wearable strain sensors based on electrospun fibers," *Advanced Functional Materials*, vol. 33, no. 18, p. 2214265, 2023.

[18] T. Zheng *et al.*, "Mechanoluminescence and photoluminescence heterojunction for superior multimode sensing platform of friction, force, pressure, and temperature in fibers and 3d-printed polymers," *Advanced Materials*, vol. 35, no. 40, p. 2304140, 2023.

[19] Y. Zhao *et al.*, "Review of wearable optical fiber sensors: Drawing a blueprint for human health monitoring," *Optics & Laser Technology*, vol. 161, p. 109227, 2023.

[20] Y. Gao *et al.*, "Addressing heterophily in graph anomaly detection: A perspective of graph spectrum," in *Proceedings of the ACM International World Wide Web Conference (WWW)*, 2023.

[21] J. Deng *et al.*, "Disentangling structured components: Towards adaptive, interpretable and scalable time series forecasting," *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[22] D. Dera *et al.*, "Trustworthy uncertainty propagation for sequential time-series analysis in rnns," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 2, pp. 882–896, 2023.

[23] S. Liu *et al.*, "Knn-tl: k-nearest-neighbor transfer learning for low-resource neural machine translation," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.

[24] X. P. Jingxuan Pang and C. Li, "A hybrid algorithm incorporating vector quantization and one-class support vector machine for industrial anomaly detection," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8786–8796, 2022.

[25] F. Chen *et al.*, "Multivariate, multi-frequency and multimodal: Rethinking graph neural networks for emotion recognition in conversation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[26] Z. Chen *et al.*, "Supervised anomaly detection via conditional generative adversarial network and ensemble active learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7781–7798, 2022.

[27] X. Chen *et al.*, "Adversarial autoencoder for unsupervised time series anomaly detection and interpretation," in *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining (WSDM)*, 2023.

[28] N. M. Shahrbanoo Rezaei and A. Khojandi, "Gaad: Gan-enabled autoencoder for real-time sensor anomaly detection and recovery in autonomous driving," *IEEE Sensors Journal*, 2024.

[29] Z. Zhang *et al.*, "Complex-valued convolutional gated recurrent neural network for ultrasound beamforming," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[30] T. Liu and J. Bao, "A novel period-sensitive lstm for laser welding quality prediction," *IEEE Transactions on Industrial Informatics*, 2024.

[31] M. Ma *et al.*, "Transformer based kalman filter with em algorithm for time series prediction and anomaly detection of complex systems," *Measurement*, vol. 229, p. 114378, 2024.

[32] G. Francis and V. Jakicic, "Equivalent statistics for a one-sample t-test," *Behavior Research Methods*, vol. 55, no. 1, pp. 77–84, 2023.