SRDA: Mobile Sensing based Fluid Overload Detection for End Stage Kidney Disease Patients using Sensor Relation Dual Autoencoder

Mingyue Tang* UTD8HJ@VIRGINIA.EDU
Department of Systems and Information Engineering, University of Virginia, US

Jiechao Gao* JG5YCN@VIRGINIA.EDU
Department of Computer Science, University of Virginia, US

Guimin Dong GUIMIND@AMAZON.COM
Amazon.com

Carl Yang J.CARLYANG@EMORY.EDU
Department of Computer Science, Emory University, US

Bradford Campbell BRADJC@VIRGINIA.EDU
Department of Computer Science, University of Virginia, US

Brendan Bowman BTB5K@HSCMAIL.MCC.VIRGINIA.EDU
School of Medicine, Division of Nephrology, University of Virginia, US

Jamie Marie Zoellner JZ9Q@VIRGINIA.EDU
Department of Public Health Sciences, University of Virginia, US

Emaad Abdel-Rahman EA6N@VIRGINIA.EDU
School of Medicine, Division of Nephrology, University of Virginia, US

Mehdi Boukhechba MBOUKHEC@ITS.JNJ.COM
The Janssen Pharmaceutical Companies of Johnson & Johnson

Abstract

Chronic kidney disease (CKD) is a life-threatening and prevalent disease. CKD patients, especially end-stage kidney disease (ESKD) patients on hemodialysis, suffer from kidney failures and are unable to remove excessive fluid, causing fluid overload and multiple morbidities including death. Current solutions for fluid overtake monitoring such as ultrasonography and biomarkers assessment are cumbersome, discontinuous, and can only be performed in the clinic. In this paper, we propose SRDA, a latent graph learning powered fluid overload detection system based on Sensor Relation Dual Autoencoder to detect excessive fluid consumption of EKSD patients based on passively collected bio-behavioral data from smartwatch sensors. Experiments using real-world mobile sensing data indicate that SRDA outperforms the state-of-the-art baselines in both F1 score and recall, and demonstrate the potential of ubiquitous sensing for ESKD fluid intake management.

* These authors contributed equally

Data and Code Availability Research data are not shared due to privacy and ethical concerns, the detailed data description, and hyper-parameter tuning are in Section 4 and 5.1. Code and the supplemental material are available at https://tinyurl.com/cm3vp974

1. Introduction

Chronic kidney disease (CKD) is a condition characterized by a gradual loss of kidney function over time, which impairs patients' capacity to filter wastes from their blood Levin et al. (2013). CKD is ranked as the 9^{th} biggest cause of mortality in America, and 15% of US adults, over 37 million individuals, are estimated to have CKD, with many more at risk. Of those diagnosed with CKD, 750,000 Americans have the most severe form of CKD—end stage kidney disease (ESKD), whose kidney function falls below 15% of normal Ozieh et al. (2021).

One of the most important medical requirements for ESKD patients is that they must follow unique dietary restrictions (32 oz or less total fluid intake per day) Liska et al. (2019) to reduce the risk of fluid overload. Any ex-

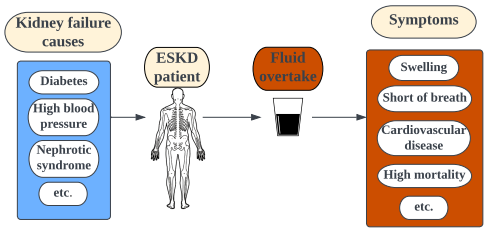

Figure 1: *ESKD patients' fluid overload causes and symptoms.*

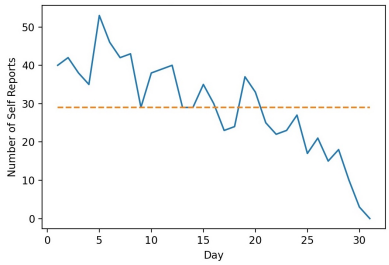

Figure 2: *Total self-reported fluid intake daily compliance in a 4 weeks study. Blue line represents trend of number of self-reported daily fluid intake. Yellow line represents average number of self-reported daily fluid intakes.*

cess fluid must be removed through dialysis, and drinking too much fluid may make a buildup between dialysis sessions, causing symptoms such as swelling, trouble breathing from fluid in the lungs, heart damage, and refractory hypertension (e.g., Figure 1). However, it is nearly impossible to maintain such low-level fluid intake every day in practice, since high sodium western diets, dialysis sessions, and hypotension can all induce patients to exceed the limitation of daily fluid intake. Current fluid monitoring techniques Morimoto et al. (1993); Efremov et al. (2020); Grant et al. (2003); Dasselaar et al. (2005) are expensive, require experts, and are only administered in clinics. Since clinicians lose sight of their patients outside of the clinic, most deaths happen at home and mainly because of fluid overload Pellicori et al. (2015). Therefore, it is critical to developing a ubiquitous fluid management system to monitor patients' fluid consumption and detect their fluid overload with high sensitivity, and to enable clinical intervention between dialysis sessions (2˜3 days) in a timely manner. However, self-report fluid management system fails to track fluid intake reliably because of human forgetfulness and lack of motivation until it is too late. Figure 2 shows the self-reported fluid intake daily compliance of our participants. We observe that the number of self-reports decreases dramatically over time, which makes it hard to track the patient's fluid intake and challenging to provide ground truth for supervised training of fluid overtake prediction model over a long period.

The development of mobile sensing technologies brings advantages in health monitoring, and the potential to implement bio-marker and contextual marker-based ubiquitous fluid intake management (FIM) systems. Existing works Adler et al. (2020); Dong et al. (2022, 2021b); Salem et al. (2014); Ukil et al. (2016) have shown that mobile sensing techniques utilizing ubiquitous devices can build bridges between people's health status and their behaviors as characterized by fine-grained human behavioral signatures extracted from the continuously collected mobile sensing data. Intuitively, the number of steps taken, the intensity of physical activity, the hand movements, the ambient environment, physiology and other variables that can be captured through mobile sensing may all be associated with fluid consumption. Meanwhile, the relationship changes between all these factors may indicate an abnormal situation occurred (e.g., step counts decreased, but heartrate increased). Compare to other latest deep learning architecture, graph neural networks (GNNs) are compelling tools to take all the factors (i.e., sensors data) and their correlations into account. Existing methods Zhang et al. (2022); Tang et al. (2022); Dong et al. (2022) show the great power of multi-sensor modeling and classification using GNNs. However, current solutions Tang et al. (2022); Dong et al. (2021a) are designed to handle supervised tasks, which have a large number of labeled data. Given the specificity of ESKD diseases, abnormal cases are hard to find in daily life, even though ESKD patients may have fluid overtake situations, they will not report it to their healthcare providers unless there is a serious discomfort Taber et al. (2015), which makes ESKD fluid overload detection unrealistic to use supervised learning. *Ipso facto*, developing a weakly-supervised fluid overload detection method with high F1 and recall is important to detect abnormal states of high-level fluid consumption through physiological and behavioral changes, since most abnormal situations of ESKD patients are closely associated with fluid excess Yilmaz et al. (2016).

To address the challenges of implementing a ubiquitous fluid overload detection system, such as physiological and bio-behavioral data collection, representation of multi-modal sensory data and their correlations, and label sparsity. 1) We deploy FluiSense Boukhechba et al. (2022), a smartwatch application to collect and process physiological and behavioral multi-modal sensory data for 14 patients in a total of 496 days. 2) We propose SRDA, a weakly-supervised latent graph learning based sensor relation dual autoencoder (SRDA) to detect abnormal fluid overload. We define abnormal fluid overload as the anomaly, which represents "excessive fluid consump-

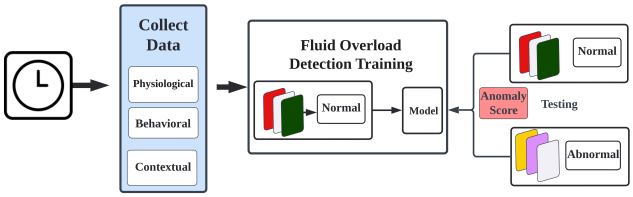

Figure 3: *Overview of our fluid overload monitoring system.*

tion" in a given time window (e.g., top 5% fluid consumption in 30 days). Figure 3 shows the overview of our proposed system. First, we utilize the denoising graph autoencoders to learn the latent interdependence within mobile sensor networks, since there is no pre-defined sensor relation structure prior to the ubiquitous sensor system. Second, we propose the dual autoencoder structure: relation autoencoder and feature autoencoder, to take inner relations between heterogeneous sensors and continuous on-body sensor data features into integrated consideration. Third, we only train SRDA on normal fluid intake data (which can be obtained easily in daily life) and detect the fluid overtake abnormals when testing by ranking the reconstruction error as the anomaly (overload) score. Our contributions can be summarized as follows:

(1) We present a novel weakly-supervised ubiquitous mobile sensing method for personal-level continuous fluid intake estimation. To the best of our knowledge, SRDA is the first attempt for modeling ESKD patients' fluid intake management in a weakly-supervised fashion, which is more practical for ESKD patients since labeled abnormal data is hard to collect.

(2) We propose a latent graph learning-based sensor relation dual autoencoder (SRDA) to automatically learn the relation among sensors and detect abnormal fluid intake signals by reconstructing the structural and feature information of multimodal input data on a normal fluid intake dataset.

(3) We test SRDA in a real-human subject study with 14 ESKD patients collected in a kidney center. SRDA achieves state-of-the-art performance, improves average 1.25% on F1 score, and average 1.22% on recall rate in comparison with over 10 baseline methods, including the state-of-the-art methods in graph anomaly detection.

2. Related work

2.1. Fluid Intake Management

Many researchers investigated the use of ubiquitous sensing devices to detect fluid intake using smart objects such as smart tables and smart containers Cohen et al. (2021). The solutions of FIM can be classified into portable methods such as smart containers and wearable devices or non-portable methods such as smart tables and cameras. Zhou et al. used a smart table equipped with pressure sensors and force sensor resistors to monitor container weight fluctuations Zhou et al. (2015). Cippitelli et al. monitored eating and drinking behaviors using a depth and RGB camera positioned on the ceiling Cippitelli et al. (2016). However, non-portable equipment for monitoring fluid intake is incapable of providing continuous monitoring of fluid intake because users may drink fluids in locations where smart tables and cameras are not accessible. Existing works of using wearable sensors for FIM are largely based on gesture recognition by using motion sensors in a controlled environment Huang et al. (2020). A recent work Tang et al. (2022) shows consumer-grade wearable products have the potential fluid intake monitoring ability in a supervised learning fashion. Different from the existing solutions, in this study, we deploy Fluisense in real-world scenarios and collect unlabeled multi-modal sensor data (physiological, behavioral bio-markers) in order to enable continuous fluid overload alerting and potential just-in-time interventions for ESKD patients. To the best of our knowledge, this is the first work using smartwatches to detect the fluid overtake abnormals of ESKD patients using only the normal data that can be collected in daily life.

2.2. Graph Neural Networks for Anomaly Detection

With the emerging advancement of Graph Neural Networks (GNNs), GNNs have been applied in graph structure/relation modeling widely. The core motivation for using GNNs to model multivariate data anomaly detection is that GNNs can model the interdependence and interactions between nodes and their neighbor nodes. For the multivariate data anomaly detection task, two major GNN anomaly detection approaches are presented in the latest literature. One type is forecasting-based detection, Cao et al. Cao et al. (2020) proposed a spectral temporal GNN based on self-attention latent correlation graph learning and time series forecasting to identify anomalies. Another type is auto-encoder based, in order to consider both structural information and node attributes of input graphs, Fan et al. presented a dual autoencoder anomaly detection that also addressed both relations and features reconstruction problems in a single framework Fan et al. (2020). Despite the state-of-art results in the anomaly detection achieved in existing works, the forecasting-based methods only rely on the power of graph neural networks, which may lead to over-smoothing problems when recov-

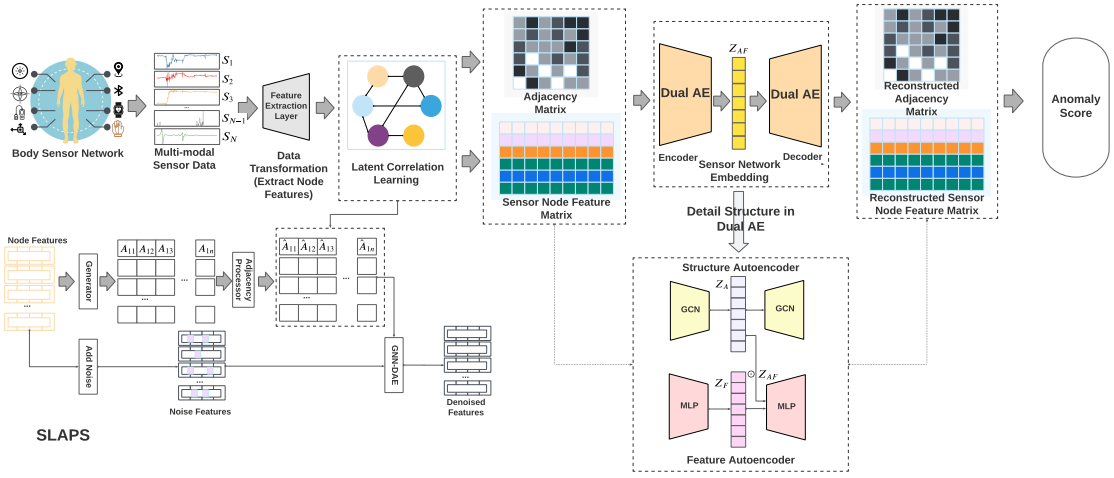

Figure 4: *The diagram of SRDA model with multivariate sensor data. The left part from the body sensor network to data transformation are the feature preprocessing steps; the latent correlation layer learns the graph relation using given features. The right part is the main model structure; it consists of two autoencoders, one structure autoencoder reconstructs the sensor relations, and one feature autoencoder decodes the original features.*

ering the features Li et al. (2019). Furthermore, most of the GNN-based methods aim to find node anomalies over graph/instance anomalies. In our study, we leverage the physiological and behavioral data collected from on-body sensor networks and train a model to profile the normal fluid consumption behaviors of ESKD patients. We use our proposed sensor-relation dual autoencoder framework to detect the graph anomalies, which automatically capture the sensor relation while reconstructing the structural and feature information in a semi-supervised fashion.

3. Methodology

3.1. Problem Formulation

Given a set of ubiquitous sensors S, the goal of our method is to detect the abnormal fluid overload according to top k% anomaly scores for the multimodal data R_S collected by S as input. In order to capture the inter-relation within a heterogeneous sensor network, we design a latent graph learning method to learn **multimodal sensor relation graph** (sensor network structure) $G = (X, A)$, representing the latent interdependence between different sensors S. $X^t(i) = h(R_S^t(i))$ presents the homogeneous transformed input features from raw multimodal sensor $S(i)$ input $R_S^t(i)$, where $i \in N$ is the number of sensors, $t \in T$ shows the number of inter-dialytic instances(data collected between each two dialysis sessions). $A = (u_{i,j})$ is the adjacency matrix, where $u_{i,j} > 0$ indicates there is an edge (weighted interaction) between sensor(node) i and sensor(node) j. Given T instances of multimodal sensor relation graphs $G = (X, A)$, our approach aims to calculate T anomaly scores $AS = F(G)$ for each instance graph, which is generated by our auto-encoder-based model F with parameters θ, from reconstructing both graph features X and structures A. The anomalies will be selected from the top k% AS. The length of T is the total number of days between the dialysis section for weights change, and days of collected reports for self-reported data during the data collection.

3.2. Overview of SRDA

Our proposed overload detection framework aims to detect abnormal fluid intake cases by reconstructing the features extracted from multivariate time series input and relation structures from the dense representations produced by encoders on normal fluid intake observations. SRDA considers both inner relations between heterogeneous sensors and on-body sensor data features. The framework consists of five parts: 1) Heterogeneous Data Transformation 2) Latent Correlation Layer 3) Structure Autoencoder 4) Feature Autoencoder 5) Overload (anomaly) Detection. Figure 4 illustrates the overall architecture of SRDA. SRDA feeds the sensor data streams into a latent correlation layer to learn the relationship between different sensor data as a relation graph, then combines the relation graph along with sensor features as inputs to feed into the relation and feature autoencoders. After that, the learned node embedding from the structure encoder

with the learned node embedding from the feature encoder generates latent variables which capture both structure and feature information by decoders during training. The final anomaly score is measured by the reconstruction errors based on both sensor relation and node attribute.

3.3. Heterogeneous Data Transformation

The data collected by on-body sensors S are in the different frequency domains (e.g, PPG 100HZ, GPS 10HZ, GPS node will collect fewer observations than PPG node), which means the raw data streams $R_S(i)$ cannot be directly fed to the model. Therefore, we first apply a heterogeneous-to-homogeneous transformation function h to extract features from raw data:

$$X(i) = h(R_S(i)), i \in \{1 \dots N\}. \tag{1}$$

$X \in \mathbb{R}^{N \times F}$, F is feature size. We adopt time-series features extraction as the transformation function, and we use tsfellibrary Barandas et al. (2020) to extract all time-series features in tsfel website with default settings[1], such as Min, Max, Variance in the statistic domain; Absolute energy, Entropy in temporal domain; and FFT mean coefficient, Fundamental frequency in spectral-domain.

3.4. Latent Correlation Layer

To better model the relationship between different sensors, it's essential to connect the highly related nodes (sensors) in a relation graph. In previous works, researchers utilize prior knowledge (e.g., the road network in traffic prediction) to construct a relation graph. However, pre-defined knowledge is not available all the time. In our case, we do not have the ground-truth sensors relation structure. An intuitive way of constructing a relation graph is to select a similarity metric and set the edge weight between two nodes to be their similarity Belkin et al. (2006). However, the quality of downstream tasks can be influenced by the choice of similarity metric(s). To learn the latent relation structure, we adopt the unsupervised **latent correlation layer** from SLAPS Fatemi and El Asri (2021), which is inspired by, and similar to, the pre-training strategies for GNNs Gomes et al. (2020); Hu et al. (2020). The latent correlation graph learning from SLAPS has been proven effective by various experiments in Fatemi and El Asri (2021). Therefore, we use the self-supervised denoising autoencoders from SLAPS to learn the latent correlations between sensors based on sensor features, which automatically emphasizes a task-specific correlation by real sensor data. The graph structure training is separate from the overload detection training, which practically decreased the training difficulty of an unsupervised task given a small dataset.

The latent correlation layer as shown in figure 4 bottom left consists of three main components to learning the latent graph structure. 1) **Generator**: generates an initial, non-symmetric correlation graph 2) **Adjacency processor**: produces symmetric, normalized adjacent graph matrix. 3) **Self-supervision** refines the correlation graph by predicting the node features. Given the X from the previous step as an input, the multi-layer perceptron (MLP) **generator** generates the initial graph $\hat{A}$ by:

$$\hat{A} = G_{MLP}(X; \theta_G) = kNN(MLP(X)) \tag{2}$$

where $\hat{A} \in \mathbb{R}^{N \times N}$, θ denotes the weights of MLP and the kNN produces a sparse matrix according to the updated node representations $MLP(X)$ and the top k neighbors according to the similarity between node features. The detailed implementation of kNN can be found in Fatemi and El Asri (2021).

Since the generated adjacent matrix $\hat{A}$ is not normalized nor symmetric, the **adjacency processor** function defined as follow:

$$A = \frac{1}{2} D^{-\frac{1}{2}} (P(\hat{A}) + P(\hat{A})^T) D^{-\frac{1}{2}} \tag{3}$$

where $A \in \mathbb{R}^{N \times N}$, D is the degree matrix $\in \mathbb{R}^{N \times N}$, and P is non-negative range function. Through this operation, the adjacent matrix A becomes symmetric and does not contain a negative value.

To refine the initialized relation structure, we utilize **self-supervision** to train a denoising autoencoder. In the self-training process, we predict the node features with learnable graph structure and noisy version node features $\hat{X}$, where we add Gaussian noise on X. The training loss function can be shown as:

$$\mathcal{L}_{SE} = L(X_i, GNN(\hat{X}, A; \theta_{GNN})_i) \tag{4}$$

where i indicates the noise addition indices of X, and the θ_{GNN} is the weights of a graph neural network model (i.e. the denoising autoencoder).

After minimizing the above loss function, we can get a self-learned adjacency matrix A for future relation reconstruction. The learned relation adjacency matrix is analyzed and validated by the ESKD expert in section 5.3.

3.5. Structure Autoencoder

As shown in Figure 4, SRDA consists of two parts of autoencoder, one is a structure autoencoder for graph relation structure reconstruction, and another one is a feature

1. https://tsfel.readthedocs.io/en/latest/descriptions/feature_list.html

autoencoder for sensor attributes reconstruction. For the structure autoencoder, the goal of this module is to learn the community node representations of the given graph G, where the nodes in the same community will share similar characteristics and have higher correlations which can be decoded as structures.

An intuitive solution is to apply Graph Autoencoder (GAE) Kipf (2016) to embed and reconstruct the relation graph structure, where use a Graph Convolutional Networks (GCN) Kipf and Welling (2017) as the encoder, and reconstruct the adjacency matrix by the latent variable Z_A encoded by GCN with parameter θ.

$$Z_A = (I + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}) X\theta \tag{5}$$

where $\theta \in \mathbb{R}^{F \times H}$, H is the output dimension of $Z_A \in \mathbb{R}^{N \times H}$. The structure decoder takes the latent variable Z_A to reconstruct the original adjacency matrix by: $\hat{A} = Sigmoid(Z \cdot Z^T)$. However, GAE does not contain the node (sensor) features reconstruction process, which is critical in anomaly detection-like tasks. Since features usually contain a significant amount of information and the node features may be over-mixed with their neighbors when applying GCN repeatedly, which is also known as the over-smoothing problem in GCN Li et al. (2019). Also, the feature reconstruction process provides the latent correlation in a data-driven fashion.

3.6. Feature Autoencoder

To address the above issues in the structure-only autoencoder, we apply a feature autoencoder to reconstruct feature information while also capturing and refining the structure learning by interacting with the structural latent variable Z_A. The feature autoencoder consists of two nonlinear feature transformation layers as traditional AE Aggarwal et al. (2015), the feature latent embedding Z_F can be formulated as:

$$X_{F_2} = sigmoid((X)^T W^{F_1} + b^{F_1}) \tag{6}$$

$$Z_F = X_{F_2} W^{F_2} + b^{F_2} \tag{7}$$

where the $W \in \mathbb{R}^{F_{n-1} \times F_n}$ and $b \in \mathbb{R}^{F_n \times F_n}$ (n indicates the n^{th} layer) are weights and biases of MLP functions. To get the final representation that covers both structure and feature information. We interact the structural latent embedding Z_A encoder with Z_F using element-wise multiplication:

$$Z_{AF} = Z_A \odot Z_F \tag{8}$$

$Z_{AF} \in \mathbb{R}^{N \times H}$. The feature reconstruction can be written as:

$$\hat{X} = MLP(Z_{AF}) \tag{9}$$

with MLP parameters W^X and b^X. Thus, the training objective of SRDA is to minimize the sum of both reconstruction errors:

$$\mathcal{L}_{rec} = \lambda \cdot BCE(A, \hat{A}) + (1 - \lambda) \cdot MSE(X, \hat{X}) \tag{10}$$

where the λ is a parameter to control the trade-off between structural and feature reconstruction loss, $BCE(A, \hat{A}) = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(A_{ij} \cdot log\hat{A_{ij}} + (1 - A_{ij}) \cdot log(1 - \hat{A_{ij}}))$ is a criterion that measures the Binary Cross Entropy between the target and input adjacency matrix, and MSE (Mean Squared Error), $MSE(X, \hat{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{X}_i)^2$ is a loss function that measures the average of the squares of the feature reconstruction errors.

3.7. Overload (anomaly) Detection

Since we train our SRDA framework on normal cases, the patterns of abnormal situation reconstruction will differ from the patterns we learned from the training dataset. Therefore, we calculate the anomaly score by the reconstruction error function that we use during the training phase:

$$AS_t = \lambda \cdot BCE(A, \hat{A}) + (1 - \lambda) \cdot MSE(X_t, \hat{X}_t) \tag{11}$$

where t stands for the t^{th} instance. Based on the anomaly score set AS, we select a threshold α that can distinguish the top $k\%$ anomaly scores.

4. Study and Data Description

With the approval of the Institutional Review Board (IRB), we recruit 14 ESKD patients from the Kidney Clinic to participate in a study investigating the use of wearable sensing to estimate fluid intake for four weeks. We choose the participants from different age groups, gender groups, and races from a fairness perspective. Figure 5 shows the statistics of our selected participants.

We give each participant an Android smartwatch (Fossil Gen 5) with the pre-installed Fluisense application (available in the Android play store now). Table 1 summarizes the bio-behavioral features gathered through Fluisense that are used for monitoring fluid intake. The five sensors are represented as the five nodes in the sensor relation graph. We ask the patients to log their fluid intake through the app by choosing from a list of predefined volumes each time they consume any liquid. We create the self-report daily volume intake feature to help patients monitor their fluid consumption. We also record the weight changes between dialysis sessions as our ground truth because ESKD patients don't have the

ability to remove fluid themselves. These two measurements represent current clinical practice and characterize both objective and subjective fluid consumption Flythe et al. (2017); Ku et al. (2018). In practice, the self-fluid intake amount correlated to the weight changes (with $P < 0.01$, $r > 0.25$). 14 patients record their fluid intake for a total of 496 days/samples, The average inter-dialytic weight change is 3.18 kg +/- 1.38 and the average self-reported interdyalic fluid consumption is 2.97 kg +/- 2.12. When the daily fluid intake level is higher than 32oz, patients may experience life-threatening danger Liska et al. (2019), we analyze the portion of daily fluid overload (fluid intake amount $>$ 32oz) as 20.36%.

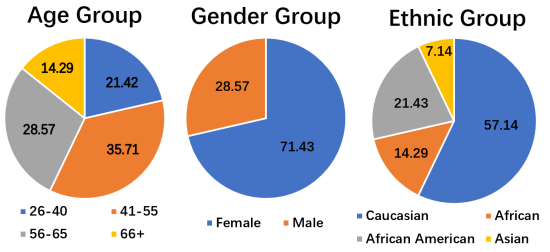

Figure 5: *Participants' statistics: we demonstrate demographic information for different groups including different age, gender and ethnic groups.*

Table 1: *Multi-modal on-body sensory data description.*

Modality	Description
Heart Rate	The number of heartbeats per minute
PPG	Detects volumetric changes in blood in peripheral circulation
Step Count	Measures the number of steps and distance traveled
GPS	Longitude and latitude of locations
Gyroscope	Senses the change in rotational angle per unit of time
Weight Changes	Converts an input mechanical force such as weight
Self-Reported Intake	Self-reported fluids consumed in a given period

5. Experiment

We design our experiments to evaluate SRDA on ESKD fluid overload anomaly detection problem, focusing on the following research questions:

- **RQ1**: How do SRDA and its major ablations compare to the baseline methods?
- **RQ2**: How to interpret the node/sensor correlation learned in fluid overload anomaly detection by using SRDA?
- **RQ3**: What are the impacts of the major model hyper-parameters — modules — modalities, including embedding size, running time, latent graph learning, and effectiveness of modalities on SRDA?

5.1. Experimental Setting

Datasets. We define our fluid-intake anomaly detection task as finding the potential dangerous fluid overload situations. In a real-world setting, when fluid intake level achieves 50% percentile of patients' long-term intake, patients need to start their fluid intake control. Therefore, we specify abnormality as a patient drinking 95%, 90%, or 80% percentile of his/her long-term (4-weeks) recorded water intake (i.e., 5%, 10%, 20% anomalies). Note, the anomaly detection process is trying to help patients start to control their fluid intake, therefore, there is no "hard line" for setting an anomaly. The reasons why we set percentage (i.e., 5%, 10%, 20%) instead of a specific number threshold (e.g., 32 oz per day) are because 1) both weight changes and fluid intake data cannot reflect the accurate fluid intake, since weight changes may involve eating solid food other than fluid consumption, and the self-reported result may not record all the fluid intake activity completely according to Fig. 2, 2) percentage thresholds are more practical in long-term unsupervised fluid intake anomaly alerting since we don't have the real ground-truth about how much fluid the patients consumed each day. This detection can be used to send alerts to patients aiming to reduce fluid intake and reduce the risks of health complications associated with fluid overload.

We study the anomaly detection performance on two types of acceptable ground truth used in real clinical practice and evaluate the model performance using Leave-one-subject out cross-validation (LOOCV). The LOOCV is the most common testing method in the clinical field, especially for a small sample size Wong (2015). Thus, we split N-2 participants as the training set, N=1 as validation set, and the last N=1 as test set, the hyper-parameters are selected using the validation set. We make predictions on the test set (with selected hyper-parameters on training sets and validation sets, N-1 folds) and calculate the performance metrics across all test sets without training data leakage. Then we take the average for 10 times repeat experiments (i.e, LOOCV on training and validation set first, then predict on test set).

Data Pre-processing and Data Segmentation. We collected a total of 496 days of data for 14 patients. We prepared the collected data for two types of tasks in two-time granularity - 1) Weight changes anomaly alerting between two dialysis sessions 2) Self-report fluid intake anomaly alerting for each day. For 1, ESKD patients usually have three dialysis sessions per week - Monday, Wednesday, and Friday. So we calculated the weight changes between two dialysis sessions to perform an inter-dialysis time granularity anomaly detection. For 2, patients recorded

Table 2: *Interdialytic fluid intake anomaly detection accuracy in terms of precision(%), recall(%), and F1-score, SRDA performed the best over all baselines especially on Recall and F1-score. All experiments are performed for 10 times on testsets and take the mean (with the hyper-parameter selected by LOOCV on training data).*

Model	Weights Change(5%/10%/20%)			Self Reported(5%/10%/20%)		
	Precision	Recall	F1	Precision	Recall	F1
PCA	87.94/79.75/65.97	74.19/71.17/64.72	80.43/75.10/65.33	87.52/79.35/65.63	73.99/70.97/64.52	80.14/74.81/65.06
KNN	87.68/79.15/68.06	74.19/70.77/67.14	80.38/74.66/67.59	87.26/78.75/67.72	73.99/70.56/66.94	80.08/74.37/67.32
FB	88.53/79.42/68.70	75.00/70.77/67.54	81.04/74.75/68.11	87.82/79.02/68.37	74.40/70.56/67.84	80.44/74.46/67.84
OC-SVM	90.98/82.60/71.42	51.81/51.21/52.02	63.64/60.37/57.22	90.82/82.48/71.37	52.02/51.41/52.22	63.66/60.42/57.32
Deep-SVDD	89.12/81.72/67.14	75.81/73.59/65.93	81.65/77.19/66.52	88.41/80.34/65.63	75.20/72.18/64.52	81.06/75.86/65.06
COPOD	87.52/79.35/66.02	73.99/70.97/64.92	80.14/74.81/65.46	89.00/83.47/67.83	76.01/75.20/67.34	81.68/78.77/67.58
SUOD	87.99/80.61/71.42	83.47/80.44/72.38	85.67/80.53/69.20	87.99/80.61/69.14	83.47/80.44/74.40	85.67/80.53/**71.26**
CoLA	88.81/80.35/69.37	89.72/80.85/68.75	89.26/80.59/69.06	88.37/81.05/67.66	89.31/81.05/67.54	88.84/81.05/67.60
SRDA w/o GAE	88.39/80.98/67.42	83.67/80.65/**72.98**	85.97/80.81/69.92	87.99/80.61/67.09	83.47/80.44/72.78	85.67/80.53/69.65
SRDA w/o AE	90.23/81.78/70.78	28.83/30.65/35.48	39.00/37.62/37.96	89.31/80.73/70.02	28.63/30.44/37.68	38.69/37.32/37.68
SRDA	89.25/81.54/72.45	**90.12/81.85/71.77**	**89.68/81.70/72.10**	90.14/82.22/67.55	**90.93/82.06/77.02**	**90.52/82.14/71.02**

their fluid intake right after they have a drink, soup or other types of fluid consumption. We calculated the total fluid intake amount for each day as ground-truth, to have a daily time granularity detection. The experiments performed on different time granularity can prove the system had the potential to use and alert fluid overtake in an out-of-clinic situation.

For the different modality sensor stream data we collected from smartwatches, all the data were recorded simultaneously. We checked there was no missing data in a single sensor. To have a reliable data input, we filtered and denoised each signal. For instance, we apply NeuroKit2 Makowski et al. (2021) to filter the PPG signals and take out all frequencies that definitely not coming from the heart. For heartrate (HR) we excluded data points below 20bpm or above 250pm. For motion data, like a gyroscope, we filter and process it through a low pass filter to remove the noises. For GPS, we leverage the distance information according to the GPS signals in every time window as a general-purpose input feature. After data pre-processing, we input the processed data into the SRDA pipeline. All the baseline functions are using the same input data.

Baseline. We compare SRDA with 10 popular and recent advanced anomaly detection baselines, representing two types of anomaly detection models; 1) Statistical machine learning: PCA Shyu and Chen (2003), KNN Angiulli (2002), FB Lazarevic and Kumar (2005), and OC-SVM Schölkopf et al. (1999), 2) Deep learning and GNNs: Deep-SVDD Ruff et al. (2018), COPOD Li et al. (2020), SUOD Zhao et al. (2021), CoLA Liu et al. (2021), AE Aggarwal et al. (2015), and GAE Aggarwal et al. (2015). The AE and GAE can be considered as two ablations of SRDA (i.e., SRDA w/o GAE and SRDA w/o AE), which represent SRDA without structure reconstruction and without feature reconstruction, respectively.

- **PCA**, Principal Component Analysis Shyu and Chen (2003), captures most of the variance in data by finding a low-dimensional projection and calculates the projection reconstruction error as anomaly scores.
- **KNN**, K Nearest Neighbors Angiulli (2002), calculates the anomaly score as the distance from coming instances to their kth nearest neighbors.
- **FB**, Feature Bagging Lazarevic and Kumar (2005), detects anomalies by fitting a number of detectors on subsamples of the dataset using a meta-estimator.
- **OC-SVM**, One Class SVM Schölkopf et al. (1999), estimates the distribution boundary of most observations and labels the instances which lie far from the boundary.
- **Deep-SVDD**, Deep-Support Vector Data Description Ruff et al. (2018), encloses instances of network representations in a minimized volume hypersphere, detects anomalies based on the distance to the center of the learned sphere.
- **COPOD**, Copula-Based Outlier Detection (COPOD) Li et al. (2020), is an outlier detection system based on empirical copula models that are parameter-free and highly interpretable.
- **SUOD**, is a large-scale unsupervised outlier detection training and prediction acceleration framework Zhao et al. (2021) with sets of regular outlier detection algorithms (e.g., LOF, IForest, COPOD, etc).
- **CoLA**, is a state-of-art contrastive self-supervised learning-based GNN method Liu et al. (2021) for graph anomaly detection.
- **AE**, Autoencoder Aggarwal et al. (2015), consists of an encoder and decoder, which reconstruct the data sample

from the low-dimensional latent variable. Considered as ablation of our method w/o structure reconstruction.

- **GAE**, Graph Autoencoder Kipf (2016), reconstructs graph structures via GCN encoders and decoders. Considered as ablation of our method w/o feature reconstruction.

Metrics. We choose precision(P), recall(R), and F1-Score(F1) as our evaluation metrics over our method and all the baseline models. Note that our anomaly detection tasks are performed on unbalanced datasets, so the choice of these metrics is fair and suitable for the unbalanced data setting. Specifically, $P = \frac{TP}{TP+FP}$, $R = \frac{TP}{TP+FN}$, and $F1 = \frac{2\times P\times R}{P+R}$, where TP, TN, FP, FN are true positives, false positives, false positives, and false negatives. To detect anomalies, we label top k% anomaly scores as anomalies, and the threshold α is set to $(100 - k\%)^{th}$ anomaly score.

Training setting. We perform a cross-validation grid search to select the best-required hype parameters for all models that involve in training (e.g., number of neighbors, different sets of detectors). For models with GCN, we set the number of GCN layers to 2. For SRDA, the learning rate is select from {5e-2, 5e-3, 5e-4, 5e-5}, epoch size {10, 20, 50, 100, 200, 300}, embedding size {20, 40, 60, 80, 100}, trade-off parameter λ {0, 0.2, 0.4, 0.6, 0.8, 1}. The training processing of latent graph learning models follows Fatemi and El Asri (2021). Most of the implementations of baseline functions are provided by Python libraries PyOD Zhao et al. (2019) and PyGOD Liu et al. (2022). For baselines that do not support multivariate time-series data (e.g., deep-SVDD), we concatenate different modalities into one. We use the same top k% anomaly percentage for all baselines. All experiments are performed on an 8GB NVIDIA RTX 3070 GPU. For baselines that involve GNNs, we use the same latent graph layer to learn the correlations in the sensor graph.

5.2. RQ1: Performance Analysis

We compare SRDA with the 10 most popular and recent advanced anomaly detection baselines in table 2, where AE and GAE can be considered as two ablations of our method (i.e., w/o structural AE and w/o feature AE). The experimental results show that the proposed SRDA outperforms all baselines in both **interdialytic weights change** and **self-reported fluid intake** anomaly detection tasks. Specifically, 1) In terms of **F1 score**, SRDA outperforms or is comparable to the baselines regarding all 5%, 10%, and 20% anomalies on both tasks. For weights change, SRDA has 0.47%, 1.01%, 3.12% higher F1 score than the next best baseline on 5%, 10%, 20% anomaly tasks respectively, and 1.89%, 1.34%, -0.33% higher F1 score on self-reported fluid intake. 2) While there are three baselines(OC-SVM, Deep-SVDD, SUDO, CoLA, and GAE) that have comparable performance to SRDA in **precision**, SRDA surpass all the baselines in **recall** perspective over all tasks (average 1.22% above). It shows that our model can find the anomalies more comprehensively, which is critical in our fluid overload anomaly finding tasks to provide interventions for more possible suspicious cases. 3) ColA has slightly worse performance than SRDA, CoLA is a state-of-art graph anomaly detection method, but not designed specifically for our task (it is trained in a node-wise fashion, not for normal/abnormal data points). 4) SRDA w/o GAE and SRDA w/o AE are two ablations of our study. Our model outperforms both ablations proving that both structural and feature autoencoders have positive effects on the final prediction results. Especially for GAE that only encodes and decodes the structural information, it shows extremely bad performance regarding recall rate (e.g., -68.00% on weights change 5% anomaly task). 5) When anomaly proportion increases, the performance decreases on all models. The rationale behind this is that when we label enlarged anomalies percentage in the dataset, normal indicators and symptoms become indistinguishable from abnormal ones since the amount of water consumption does not exceed much by the recommended values.

Note that, the essential problem we want to solve in this research is to identify as many and early as possible when the abnormalities possibly happen so the providers and dependents can receive alerts or provide intervention in time. Therefore, the **recall** and **F1 scores** are the most important metrics for the fluid overtake alert system. Based on the large amount of EKSD patients worldwide, We believe that even small enhancement in F1 and recall score is meaningful for the fluid overtake alert system. In this sense, mobile sensing with provided sensors clearly is a feasible solution for fluid intake monitoring. The labeled sample size is small, but the result shows that one benefit of applying our weakly-supervised learning approach here is it can learn the pattern of limited data.

5.3. RQ2: Sensor Correlation Analysis

Graph relation structure learned by the latent correlation layer highlights the relationship between sensor nodes and explains the structure of each of how a sensor's neighbors model the sensor's behavior. We draw a sensor relation plot depicted in Figure 6. Understanding the sensor relation plot increases the interpretability of the proposed model and validates our model on automatically constructing the relation graph.

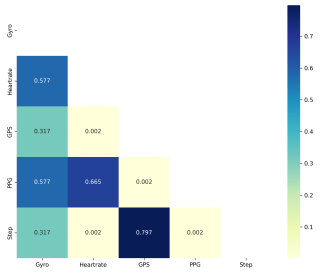

Figure 6: *Self-learned sensor correlation matrix via the trained latent correlation layer in SRDA, step count and GPS, PPG, and heart rate are highly correlated.*

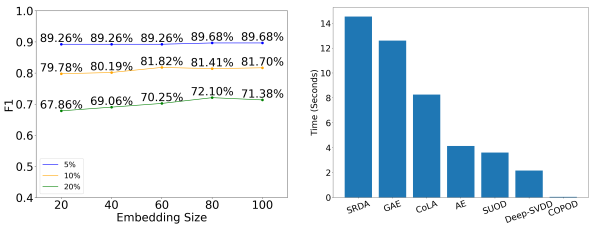

Figure 7: *Hyperparamter sensitivities. The left figure illustrated the impact of embedding size to F1 score given different weight changes; the right figure illustrates the empirical inference time comparison.*

From the sensor relation heatmap, physiological biomarkers like "PPG" and "heartrate" are highly correlated (0.665). This finding proves how validated our approach is because heart rate is known to be extracted from PPG (by estimating the beat-to-beat intervals from the PPG signal). Some behavior biomarkers like "StepCount" is related to the "gyroscope" (0.317) and "GPS" (0.797), which is intuitively reasonable since both the angle of the gyroscope will change periodically and the location will change when walking or doing other activities. In addition, the motion sensor "gyroscope" is highly correlated with almost all the other sensors since all the sensors may change when wearable devices detect a motion change. "Step" has a low correlation with "heartrate", which is a little bit surprising, but it's reasonable since our SLAPS only select top k (k=2 in our test) closest sensors in it's kNN operation. Therefore, the result only indicates the "heartrate" has a relatively lower correlation to "Step" compare to "gyroscope" and "GPS". The SRDA self-learned sensor relation graph was presented to three hospital experts and they verified the self-learned relation as valid with the discussion above.

5.4. RQ3: In-depth Studies

Hyperparameters. The embedding size is critical for autoencoder-based solutions since it reflects the effectiveness of information compression. We selected embedding sizes starting from 20 to 100 on the self-reported fluid-intake anomaly detection task. As shown in in Figure 7(a), our model is not very sensitive to embedding size change when predicting high-level fluid intake, and maintains the state-of-arts performance over all embedding dimensions. The model performs slightly better when the embedding size becomes larger.

To further understand the efficiency and feasibility of our proposed method, we compare the running time of SRDA with all the other deep learning-based anomaly detection approaches.

We measure the run time of all compared algorithms on Intel i7 7700k processors, 16 GiB of RAM and 8GB NVIDIA RTX 3070 GPU. We train all the models to converge on 20% anomaly weights to change datasets and recorded the training time. From Figure 7(b), although all other deep models are faster than SRDA since SRDA need to train an additional graph correlation layer (same for GAE and CoLA), our model has a comparable running speed with the normal GAE. Note, our SRDA framework does not rely on more hyper-parameters than the basic ones for classic GNN models like GCN and GAE. The inference time is similar for all the baseline models, which is a more practical metric for patients (near 0.3s).

Latent graph learning. The latent graph learning module is an essential part of SRDA model. Thus, the effect of different graph correlation learning modules is interesting ablation of our study. We compare four different latent correlation layers in the interdialytic weight change task: 1) SLAPS (i.e., the layer we used in SRDA) 2) Cosine similarity 3) Euclidean distance 4) All connected graph (i.e., all edge weights are equal to one). For the sake of fairness, all the *hyper-parameter settings* used in the dual autoencoder module part are the **same**.

Table 3: *Performance effects of different latent graph learning layers, SLAPS maintains the highest performance while Euclidean distance metrics performed worse than fully connected graphs.*

Layer	Precision	Recall	F1
Cosine	89.22/81.13/71.27	89.92/81.45/70.56	89.57/81.29/70.91
Euclidean	88.79/80.72/68.90	89.11/81.05/68.15	88.73/80.88/68.52
Connected	88.79/80.31/70.88	89.52/80.65/70.16	89.15/80.48/70.51
SLAPS	**89.25/81.54/72.45**	**90.12/81.85/71.77**	**89.68/81.70/72.10**

As shown in Table 3, SLAPS latent graph learning performed the best over all the other types of graph structure constructors. The performance of pair-wise similarity graph construction methods highly depends on the metrics it selected. For instance, the cosine similarity can show better results compared to the euclidean distance, and the euclidean distance sometimes even has worse performance than all connected graph settings. Therefore, a self-learning correlation layer is necessary in our method and can be generalized to other problem settings.

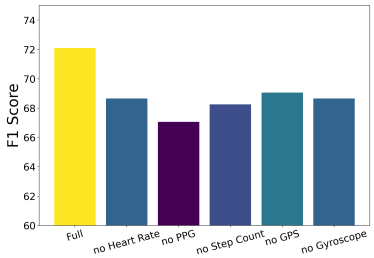

Figure 8: *F1 score of the modality ablation studies. Each bar presents the performance when one of the five sensor types (as shown in Table 3) is removed and the whole training process was applied on the remaining sensors.*

Effectiveness of different modalities. To explore how different sensor modalities contribute to the final overload detection result. We further conducted a modality/feature ablation study. For the five feature types - Heart Rate, PPG, Step Count, GPS, and Gyroscope (as presented in table 3), we removed one of them and re-ran the whole model using the remaining four sensor types on 20% anomaly weight change task. Figure 8 summarized the results above. We found that all modalities are important in the fluid overload anomaly detection task, among them, PPG and Step Count sensors were the most important, and removing them lead to the biggest performance drop. While removing the GPS sensor has the least effect. This finding indicates that both behavior and physiological data are essential in fluid monitoring.

6. Conclusion

In this work, we present a novel weakly-supervised wearable sensing fluid intake anomaly detection framework to address the health problem in the growing kidney patient population. The proposed solution leverages the power of low-cost ubiquitous mobile sensing and graph neural networks to effectively model the collected multi-modal sensory data and detect potential fluid intake anomalies. Our findings show that mobile sensing techniques (e.g., GPS, PPG, Gyroscope, etc.) can be used to design accurate, ubiquitous, and unobtrusive fluid management systems, and they can be used in addition to the clinical examination, hemodynamic monitoring, and weight assessment that dialysis patients undergo three times weekly. But still, long-term experiments are needed to determine the clinical importance of the mobile sensing monitoring system. Besides, in section 5.4, we explore which sensor modalities are more useful in fluid overtake anomaly detection. It can inspire future fluid management research on these sensors.

From the performance analysis, the experiment results show that our proposed solution improved the detection of fluid intake anomalies by 1.25% on F1 on average. In addition, SRDA achieves a higher recall rate of 1.22% than the state-of-the-art baseline methods, which indicates that SRDA has the capability of capturing more abnormal situations, which is critical for ESKD patients' life-saving fluid control regimes. The in-depth sensor relation and ablation analysis provide a view of how physiological and behavioral sensors correlate with each other and validate the feasibility of our solution. We acknowledge our current solution still has limitations due to 1) limited patient *sample size* and *sensors used*, 2) ignoring *temporal information*, and 3) lack of ground-truth *anomaly proportion*. We believe the data we collected and the proposed method are meaningful in end-stage kidney patients' fluid control and will address the limitations in future work. In addition, our fluid intake anomaly detection solution can provide continuous and unobtrusive fluid management solutions to many health conditions in which fluid intake and potentially other types of vital monitoring is essential.

Institutional Review Board (IRB) The study is approved by the Institutional Review Board (IRB): IRB-HSR/UVA Study Tracking #: HSR200294.

References

Daniel A Adler, Dror Ben-Zeev, Vincent WS Tseng, John M Kane, Rachel Brian, Andrew T Campbell, Marta Hauser, Emily A Scherer, and Tanzeem Choudhury. Predicting early warning signs of psychotic relapse from passive sensing data: an approach using encoder-decoder neural networks. *JMIR mHealth and uHealth*, 8(8):e19962, 2020.

Charu C Aggarwal et al. *Data mining: the textbook*. Springer, 2015.

Fabrizio Angiulli. Fast outlier detection in high dimensional spaces. In *European conference on principles of data mining xand knowledge discovery*, 2002.

Marília Barandas, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, and Hugo Gamboa. Tsfel: Time series feature extraction library. *SoftwareX*, 2020.

Mikhail Belkin, Partha Niyogi, and Vikas Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of machine learning research*, 2006.

Mehdi Boukhechba, Mingyue Tang, Zoellner, Jamie Bowman, Brendan, and Emaad Rahman. A smartwatch based system for monitoring fluid consumption of end-stage kidney patients. In *AHFE*, 2022.

Defu Cao, Yujing Wang, Juanyong Duan, Ce Zhang, Xia Zhu, Congrui Huang, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, et al. Spectral temporal graph neural network for multivariate time-series forecasting. *NeurIPS*, 2020.

Enea Cippitelli, Samuele Gasparrini, and Ennio Gambi. Unobtrusive intake actions monitoring through rgb and depth information fusion. In *ICCP*, 2016.

Rachel Cohen, Geoff Fernie, and Atena Roshan Fekr. Fluid intake monitoring systems for the elderly: A review of the literature. *Nutrients*, 2021.

Judith J Dasselaar, Roel M Huisman, Paul E de Jong, and Casper FM Franssen. Measurement of relative blood volume changes during haemodialysis: merits and limitations. *Nephrology Dialysis Transplantation*, 2005.

Guimin Dong, Lihua Cai, Debajyoti Datta, Shashwat Kumar, Laura E Barnes, and Mehdi Boukhechba. Influenza-like symptom recognition using mobile sensing and graph neural networks. In *CHIL*, 2021a.

Guimin Dong, Mingyue Tang, Lihua Cai, Laura E Barnes, and Mehdi Boukhechba. Semi-supervised graph instance transformer for mental health inference. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1221–1228. IEEE, 2021b.

Guimin Dong, Mingyue Tang, Zhiyuan Wang, Jiechao Gao, Sikun Guo, Lihua Cai, Robert Gutierrez, Bradford Campbell, Laura E Barnes, and Mehdi Boukhechba. Graph neural networks in iot: A survey. *ACM Transactions on Sensor Networks (TOSN)*, 2022.

Sergey M Efremov, Vsevolod V Kuzkov, Evgenia V Fot, Mikhail Y Kirov, Dmitry N Ponomarev, Roman E Lakhin, and Evgenii A Kokarev. Lung ultrasonography and cardiac surgery: a narrative review. *Journal of cardiothoracic and vascular anesthesia*, 2020.

Haoyi Fan, Fengbin Zhang, and Zuoyong Li. Anomalydae: Dual autoencoder for anomaly detection on attributed networks. In *Proc. of ICASSP*, 2020.

Bahare Fatemi and Layla El Asri. Slaps: Self-supervision improves structure learning for graph neural networks. *Proc. of NeurIPS*, 2021.

Jennifer E Flythe, Magdalene M Assimon, and Robert A Overman. Target weight achievement and ultrafiltration rate thresholds: potential patient implications. *BMC nephrology*, 18(1):1–13, 2017.

Weihua Hu Bowen Liu Joseph Gomes, Marinka Zitnik Percy Liang Vijay Pande, and Jure Leskovec. Strategies for pre-training graph neural networks. *ICLR. Weihua Hu Bowen Liu Joseph Gomes Marinka Zitnik Percy Liang Vijay Pande and Jure Leskovec*, 2020.

Edward G Grant et al. Carotid artery stenosis: gray-scale and doppler us diagnosis—society of radiologists in ultrasound consensus conference. *Radiology*, 2003.

Ziniu Hu, Yuxiao Dong, Kuansan Wang, Kai-Wei Chang, and Yizhou Sun. Gpt-gnn: Generative pre-training of graph neural networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1857–1867, 2020.

Hsiang-Yun Huang, Chia-Yeh Hsieh, Kai-Chun Liu, Steen Jun-Ping Hsu, and Chia-Tai Chan. Fluid intake monitoring system using a wearable inertial sensor for fluid intake management. *Sensors*, 2020.

Thomas N Kipf. Variational graph auto-encoders. *NeurIPS*, 2016.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *ICLR*, 2017.

Elaine Ku, Joel D Kopple, Kirsten L Johansen, Charles E McCulloch, Alan S Go, Dawei Xie, Feng Lin, L Lee Hamm, Jiang He, John W Kusek, et al. Longitudinal weight change during ckd progression and its association with subsequent mortality. *American Journal of Kidney Diseases*, 71(5):657–665, 2018.

Aleksandar Lazarevic and Vipin Kumar. Feature bagging for outlier detection. In *Proc. of KDD*, 2005.

Adeera Levin et al. Kidney disease: Improving global outcomes (kdigo) ckd work group. kdigo 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney international supplements*, 2013.

Yuening Li, Xiao Huang, Jundong Li, Mengnan Du, and Na Zou. Specae: Spectral autoencoder for anomaly detection in attributed networks. In *CIKM*, 2019.

Zheng Li, Yue Zhao, Nicola Botta, Cezar Ionescu, and Xiyang Hu. Copod: copula-based outlier detection. In *Proc. of ICDM*, 2020.

DeAnn Liska, Eunice Mah, Tristin Brisbois, Pamela L Barrios, Lindsay B Baker, and Lawrence L Spriet. Narrative review of hydration and selected health outcomes in the general population. *Nutrients*, 2019.

Kay Liu, Yingtong Dou, Yue Zhao, Xueying Ding, Xiyang Hu, Ruitong Zhang, Kaize Ding, and Philip S. Chen, Yu. Pygod: A python library for graph outlier detection. *arXiv preprint arXiv:2204.12095*, 2022.

Yixin Liu, Zhao Li, Shirui Pan, Chen Gong, Chuan Zhou, and George Karypis. Anomaly detection on attributed networks via contrastive self-supervised learning. *IEEE transactions on neural networks and learning systems*, 2021.

Dominique Makowski, Tam Pham, Zen J. Lau, Jan C. Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and S. H. Annabel Chen. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4):1689–1696, feb 2021. doi: 10.3758/s13428-020-01516-y. URL https://doi.org/10.3758%2Fs13428-020-01516-y.

Tadaoki Morimoto, Sugura Kimura, Yasunori Konishi, Kansei Komaki, Tadashi Uyama, and Yasumasa Monden. A study of the electrical bio-impedance of tumors. *Journal of Investigative Surgery*, 1993.

Mukoso N Ozieh, Emma Garacci, and Rebekah J Walker. The cumulative impact of social determinants of health factors on mortality in adults with diabetes and chronic kidney disease. *BMC nephrology*, 2021.

Pierpaolo Pellicori, Kuldeep Kaur, and Andrew L Clark. Fluid management in patients with chronic heart failure. *Cardiac Failure Review*, 1(2):90, 2015.

Lukas Ruff, Robert Vandermeulen, Nico Goernitz, Lucas Deecke, Shoaib Ahmed Siddiqui, Alexander Binder, Emmanuel Müller, and Marius Kloft. Deep one-class classification. In *ICML*, 2018.

Osman Salem, Yaning Liu, Ahmed Mehaoua, and Raouf Boutaba. Online anomaly detection in wireless body area networks for reliable healthcare monitoring. *IEEE journal of biomedical and health informatics*, 2014.

Bernhard Schölkopf, Robert C Williamson, Alex Smola, John Shawe-Taylor, and John Platt. Support vector method for novelty detection. *NeurIPS*, 1999.

Mei-Ling Shyu and Shu-Ching Chen. A novel anomaly detection scheme based on principal component classifier. Technical report, 2003.

Jennifer M Taber, Bryan Leyva, and Alexander Persoskie. Why do people avoid medical care? a qualitative study using national data. *Journal of general internal medicine*, 30(3):290–297, 2015.

Mingyue Tang, Guimin Dong, Jamie Zoellner, Brendan Bowman, Emaad Abel-Rahman, and Mehdi Boukhechba. Using ubiquitous mobile sensing and temporal sensor-relation graph neural network to predict fluid intake of end stage kidney patients. *IPSN*, 2022.

Arijit Ukil, Soma Bandyoapdhyay, and Chetanya Puri. Iot healthcare analytics: The importance of anomaly detection. In *2016 IEEE 30th international conference on advanced information networking and applications (AINA)*, 2016.

Tzu-Tsung Wong. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9):2839–2846, 2015.

Süreyya Yilmaz, Yasar Yildirim, Zülfükar Yilmaz, Ali Veysel Kara, Mahsuk Taylan, Melike Demir, Mehmet Coskunsel, Ali Kemal Kadiroglu, and Mehmet Emin Yilmaz. Pulmonary function in patients with end-stage renal disease: effects of hemodialysis and fluid overload. *Medical science monitor: international medical journal of experimental and clinical research*, 22:2779, 2016.

Xiang Zhang, Marko Zeman, Theodoros Tsiligkaridis, and Marinka Zitnik. Graph-guided network for irregularly sampled multivariate time series. *ICLR*, 2022.

Yue Zhao, Zain Nasrullah, and Zheng Li. Pyod: A python toolbox for scalable outlier detection. *Journal of Machine Learning Research*, 2019.

Yue Zhao, Xiyang Hu, Cheng Cheng, Cong Wang, Changlin Wan, Wen Wang, Jianing Yang, Haoping Bai, Zheng Li, Cao Xiao, et al. Suod: Accelerating large-scale unsupervised heterogeneous outlier detection. 2021.

Bo Zhou, Jingyuan Cheng, Mathias Sundholm, Attila Reiss, Wuhuang Huang, Oliver Amft, and Paul Lukowicz. Smart table surface: A novel approach to pervasive dining monitoring. In *PerCom*, 2015.