Unsupervised Detection of Postoperative Complications in Home-Monitored Patients Preliminary Results

Djibrillah, Fatime Oumar; Waanders, Ilse; Lips, Daan; Nane, Gabriela F.; Van Keulen, Maurice; Witteveen, Annemieke; John, Arlene

DOI
10.1109/MeMeA65319.2025.11068044

Publication date
2025

Document Version
Final published version

Published in
20th Edition of the IEEE International Symposium on Medical Measurements and Applications, MeMeA 2025 - Proceedings

Unsupervised Detection of Postoperative Complications in Home-Monitored Patients: Preliminary Results

Fatime Oumar Djibrillah[*, ‡,1], Ilse Waanders[‡,2], Daan Lips[3], Gabriela F. Nane[4], Maurice van Keulen[5], Annemieke Witteveen [1], Arlene John [1]

***Abstract*—Wearable sensors enable remote, continuous patient monitoring at home, offering a promising approach for early detection of postoperative complications. However, analyzing continuous long-term physiological data remains challenging, particularly in the absence of precisely labeled deterioration events. Unsupervised change point detection methods can address this issue by identifying physiological deviations without requiring predefined event labels. This study investigates the feasibility of using a Long-Short-Term Memory (LSTM) autoencoder for detecting postoperative complications from continuous heart rate and respiration rate data using a wearable patch sensor while monitoring patients in their homes. The autoencoder was applied to identify physiological deviations that may indicate potential complications after major abdominal oncological surgeries in ten patients. The model was trained on data from seven patients to recognize deviations from normal physiological patterns and evaluated on three patients. The proposed model detected change points preceding the clinically documented complication time in two test patients, identifying these deteriorations an average of 3.25 hours earlier than the standard Remote Early Warning Score (REWS) alarm system. These findings suggest that LSTM autoencoder-based change point detection could be a valuable tool for identifying postoperative complications early in remote patient monitoring settings, to support timely intervention and potentially improving patient outcomes.**

***Index Terms*—Remote Patient Monitoring, Wearables, Postoperative complications, LSTM autoencoder, Change point detection.**

I. Introduction

Early detection of postoperative complications is critical for timely intervention and improved patient outcomes [1]. While in-hospital monitoring often provides periodic assessments, these may fail to capture early signs of physiological deterioration occurring between scheduled evaluations [2]. After discharge, the absence of monitoring further increases the risk of delayed complication detection, which may lead to adverse events [3]. Continuous remote monitoring with wearable sensors offers a promising solution by enabling continuous tracking of physiological signals, also at home and potentially detecting complications earlier [4]–[6].

Effectively analyzing data from wearable sensors requires advanced analytics to detect early signs of deterioration that may indicate potential complications. Change points in time-series data refer to instances where the statistical properties of the signal undergo significant deviations [7]. Change point detection (CPD) is a widely used technique to identify abrupt changes in time-series data and has been applied in various domains [7], including finance [8], environment [9], medicine [10], [11]. In the context of physiological monitoring, CPD can help detect abrupt changes in vital signs that may indicate potential complications. However, often only the moment of diagnosis of a complication is known and not the moment of onset, making it challenging to apply supervised learning approaches for detecting complications. Unsupervised CPD provides a promising solution by detecting change points without requiring predefined labels [12]. Long Short-Term Memory (LSTM) autoencoders have been widely studied for unsupervised change point or anomaly detection [13]–[15].

In this study, we investigated an LSTM autoencoder-based approach for unsupervised change point detection in physiological data to identify postoperative complications after major abdominal cancer surgery in patients. The contributions of this study are developing an LSTM autoencoder-based CPD model for postoperative monitoring; implementing dynamic thresholding, which adapts detection sensitivity based on individual patient variability to improve CPD performance, and comparing the proposed model's performance using the standard Remote Early Warning Score (REWS) alarm system.

II. Method

A. Data collection

This prospective observational study included thirty patients who underwent elective oncologic small bowel, colorectal or pancreatic resections between June and October 2024 at the Medisch Spectrum Twente (MST) Hospital in the Netherlands. Eligible patients were $\geq$ 18 years old, undergoing elective surgery (small bowel, colorectal or pancreatic) under general anesthesia, and provided written informed consent. Exclusion

* Corresponding author: o.d.fatime@utwente.nl
‡ The authors contributed equally
[1]Biomedical Signals and Systems group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, the Netherlands
[2] Sensing and Interaction Lab, Medisch Spectrum Twente, Enschede, the Netherlands
[3]Department of Surgery, Medisch Spectrum Twente, Enschede, the Netherlands
[4]Applied Probability group, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, the Netherlands
[5]Data Management & Biometrics group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, the Netherlands

2025 IEEE Medical Measurements & Applications (MeMeA) | 979-8-3315-2347-3/25/$31.00 ©2025 IEEE | DOI: 10.1109/MEMEA65319.2025.11068044

criteria were recent reoperations, having antibiotic-resistant infections, psychiatric hospitalization, having implantable devices, or skin conditions affecting sensor placement. Approval for this study was granted by the MST Board of Directors and the Local Advisory Committee on Feasibility on April 2, 2024, and the informed consent was obtained from all participants.

The patient data were continuously recorded using the Philips HealthDot, a wearable patch sensor, for a period of 14 days postoperatively at 5-minute intervals. The Healthdot is an accelerometer-based wearable patch sensor and is placed on the left lower rib of the patient. The sensor records heart rate (HR), respiratory rate (RR), posture and activity level, which were transmitted via the LoRa network [16]. The Philips HealthDot was placed on the patient following oncological abdominal surgery, including colorectal, pancreatic, and small bowel resections. All patients included in the study followed the same-day discharge protocol and the Healthdot-based monitoring was carried out in the patient's home.

The mean patient age was 70.0 years (standard deviation [SD] = 8.78, range: 54-85), with 63% male and 37% female.

B. Data preprocessing

To increase the robustness of our model, we applied several data preprocessing steps. For the current analysis, only HR and RR signals were used as input to the model. Outlier detection is crucial in time-series analysis, as anomalous data points can significantly distort model performance [17]. We used physiologically guided thresholding by defining HR values below 40 or above 200 bpm and RR values below 4 or above 40 bpm as outliers and replaced them with Not a Number (NaN) values. Missing data, which is common in wearable sensor recordings [18], was handled using linear interpolation followed by backward filling to maintain temporal consistency and avoid disruptions in sequential learning. Physiological signals such as HR and RR often have different ranges, which can negatively impact model training by causing certain features to dominate the learning process. Data normalization or transformation is a common step in machine learning models to ensure that all features contribute equally to model's learning process [19]. To achieve this, we applied z-score normalization to each patient data, which transforms each feature to have zero mean and unit variance [19]. This process ensures that the model effectively captures patterns without being biased toward features with larger numerical values. Following the normalization, the time-series data were segmented into overlapping 3-hour sequences (36-time steps per sequence) to provide temporal context for training the model.

C. Model architecture and training

In this study we investigated LSTM autoencoder-based unsupervised approach to identify change points that might indicate potential deterioration in patient vital signs. Autoencoders are neural networks designed to compress and reconstruct input data [20]. LSTM, a type of recurrent neural network (RNN), is well-suited for capturing long-term dependencies in sequential data. It enables effective anomaly or CPD by leveraging reconstruction loss to identify deviations in times series [21].

The proposed LSTM autoencoder architecture, illustrated in Figure 1, consists of an encoder-decoder structure. The encoder maps the input sequences into a compressed latent space, and the decoder reconstructs the original sequence. The architecture was designed as follows:

- Encoder: Two stacked LSTM layers with 64 and 32 units, respectively.
- Latent space representation: Encodes the time-series sequence into a reduced-dimensional feature space.
- Decoder: Two LSTM layers mirroring the encoder (32 and 64 units), followed by a dense output layer to reconstruct the input.

Prior to training, the dataset was divided into training and testing sets with patient stratification, with 70% (n = 7) of the patients used for training and 30% (n = 3) for testing. The model was trained to minimize the Mean Squared Error (MSE) reconstruction loss between the original input sequence X_i and its reconstructed output sequence $\hat{X}_i$. This loss function ensures that the model captures patterns and deviations in the physiological time series.

The model was trained using the Adam optimizer with a learning rate of 0.001. Early stopping was applied to prevent overfitting by monitoring validation loss and stopping training when no further improvement was observed. To ensure stability of the results, training was repeated with different random seed initializations, and results from one representative run using a fixed seed are reported.

D. Change point detection

In physiological monitoring, these shifts may indicate potential clinical deterioration. To identify physiological deviations which could indicate of potential complications, we analyzed the reconstruction errors of the LSTM autoencoder model on the test set.

The reconstruction error was computed using the MSE (1), which calculates the difference between the original input sequence X_i and its reconstructed output sequence $\hat{X}_i$:

$$e_t = \frac{1}{w}\sum_{i=1}^{w}\left(X_{t,i} - \hat{X}_{t,i}\right)^2 \tag{1}$$

where w is the sequence length,X_t, i, represents the original input values at time stept, and $\hat{X}_t, i$, represents the corresponding reconstructed values. Higher reconstruction errors indicate deviations from learned patterns and might potentially be due to a change point. In order to determine whether a given data point represents a change point, a threshold must be set on the reconstruction errors. A fixed global threshold may not be suitable due to inter-patient variability. Therefore, we implemented a patient-level dynamic thresholding approach [12]. This method allows us to adaptively define a threshold based on each patient's distribution of reconstruction errors. This ensures that the detection process accounts for individual differences rather than applying a one-size-fits-all threshold.

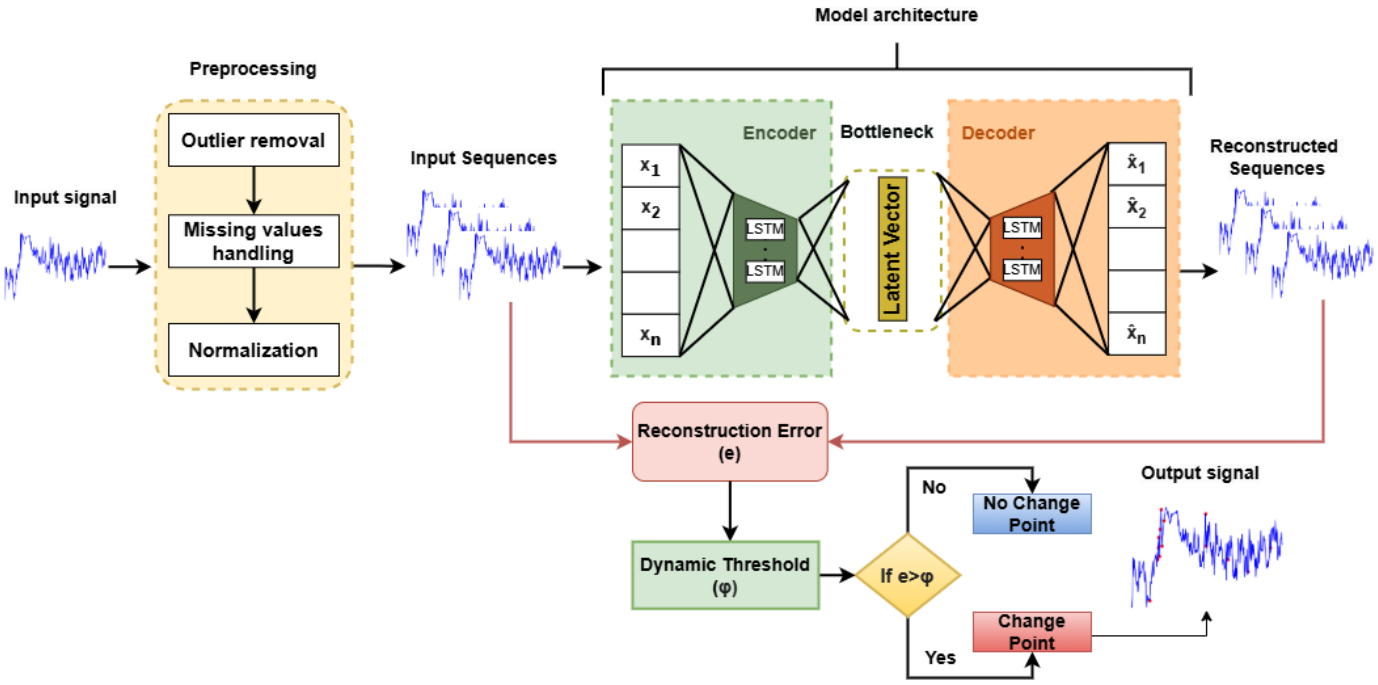


Fig. 1. General flow diagram of the proposed approach.

The dynamic threshold (2) for patient p is computed as follows:

$$\theta_p = \mu(e_p) + z\sigma(e_p) \tag{2}$$

where $\mu(e_p)$ and $\sigma(e_p)$ are the mean and standard deviation of reconstruction errors for patient p, and z is an adjustable parameter that controls sensitivity and determines how many standard deviations above the mean are considered indicative of change point. A previous study [12] suggests that selecting z between 2 and 5 provides a balance between detecting meaningful deviations and reducing false positives. In this study, we selected z=2.

In order to reduce false positives, we filtered consecutive detections within 12-time steps (1 hour) and kept only the first-detected change point, thereby maintaining only significant changes.

To evaluate the proposed LSTM autoencoder-based CPD, we defined Time-to-Detection (TTD) as difference between the first detected change point and the first clinically recorded complication timings.

E. Baseline method: Golden standard REWS features

In order to evaluate the effectiveness of our proposed CPD method, we compared it against a clinically established baseline feature, the Remote Early Warning Score (REWS). The REWS is calculated by summing the weight scores of 30-minute averages of HR and RR . The assigned weights are shown in Table I. A REWS alarm is generated when the sum of weights from HR and RR ≥ 3.

TABLE I
REWS SCORE WEIGHTS FOR HR AND RR

Vital Signs \ Score	2	1	0	1	2	3
HR (bpm)	<40	40-50	51-100	101-110	111-130	>130
RR (bpm)	<9		9-14	15-20	21-30	≥ 30

III. RESULTS AND DISCUSSION

We used a subset of ten patients (five with complications and five without complications) (Table II) for development and evaluation of the proposed methodology. These ten patients were selected to ensure a balanced comparison of both groups while maintaining computational feasibility for this preliminary analysis. The analysis focused on detecting change points indicative of potential complications and comparing the TTD of our approach with the time-to-alarm (TTA), which represents the interval between the first REWS-generated alarm and the clinically documented complication timing.

TABLE II
SUMMARY OF INCLUDED PATIENTS' COMPLICATIONS AND SURGERY TYPES

Category	Complication (n=5)	Non-Complication (n=5)
Clavien-Dindo Classification	CD = 1 (n=1), CD = 2 (n=1) CD = 3.5 (n=3)	CD = 0 (n=5)
Type of Complication	Respiratory (n=1), Surgical (n=3) Other (Constipation) (n=2)	None
Surgery Type	Laparoscopic (n=1), Open (n=1) Robotic (n=3)	Laparoscopic (n=3), Open (n=1), Robotic (n=1)

A. Model performance and reconstruction errors

The LSTM autoencoder was trained using data from seven patients and evaluated on the remaining three patients. The

reconstruction errors were computed at each time step and used as indicators of deviation in physiological patterns. The mean reconstruction errors across all test patients were 0.2748 (SD ± 0.0887), with variability observed among individual patients (Table III). The dynamic threshold, calculated per patient from the mean and SD of reconstructions errors, varied across patients. Higher error variability led to higher thresholds, as seen in Test Patient 3 (Table III).

TABLE III
MEAN RECONSTRUCTION ERRORS AND DYNAMIC THRESHOLDS FOR TEST PATIENTS

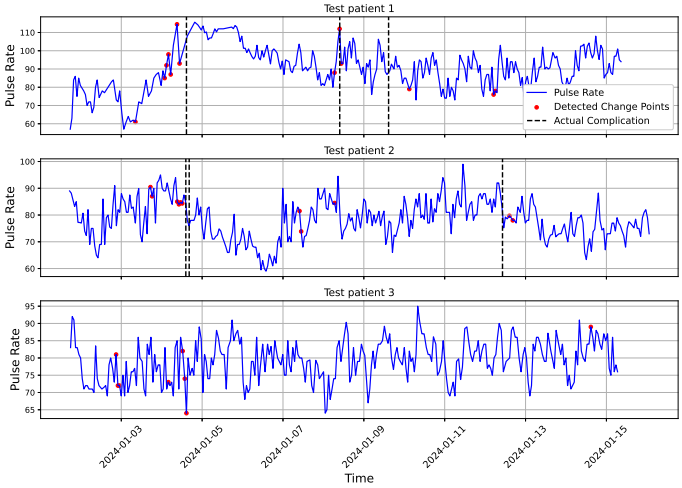
Patient ID	Mean Error	Standard Dev	Dynamic Threshold	Complication Status
Test Patient 1	0.2190	0.0581	0.3352	Yes (Surgical)
Test Patient 2	0.3023	0.0721	0.4465	Yes (Respiratory)
Test Patient 3	0.3033	0.1360	0.5753	No

B. Change point detection

The proposed model detected multiple change points in all test patients based on deviations in reconstruction errors. In patients who developed complications, a higher concentration of change points in period preceding the clinically recorded complication time was observed in the period preceding the clinically recorded complication time.

Figure 2 illustrates the detected change points timing (with red dots) in comparison to the clinically documented complication timing (black dashed line) across the three test patients. For patients who developed complications, the detected change points were more clustered around the clinically documented complication times (black dashed lines), which indicate early physiological deviations before clinical recognition.

This clustering of change points suggests that the model was able to identify early signs of potential deterioration. However, change points were also detected in the patient without complications (Figure 2). This highlights the challenge of distinguishing between clinically significant changes and normal physiological fluctuations without complications.



(a) Detected change points for heart rate (HR)

(b) Detected change points for respiratory rate (RR)

Fig. 2. Detected change points (red) and clinically documented complications (black dashed line) for Test Patients.

C. Time-to-Detection and comparison with REWS

The LSTM autoencoder detected signs of complications earlier than the REWS alarm system. As shown in Table IV, the proposed method identified deviations an average of 6.6 hours before the first REWS alarm. In both patients who had complications in the test set, the LSTM autoencoder provided an earlier indication of potential deterioration. The average TTD was 26.35 hours, while the average TTA was 19.73 hours, each calculated independently across patients.

These results demonstrate that while the model was able to detect early changes in physiological patterns related to complications, further refinement may be needed to improve its ability to differentiate between clinically relevant changes and normal fluctuations in cases with no complications.

TABLE IV
TIME-TO-DETECTION (TTD) AND TIME-TO-ALARM (TTA) FOR TEST PATIENTS

Patient ID	TTD (Hours)	TTA (Hours)	Complication Status
Test Patient 1	30.29	18.53	Yes (Surgical)
Test Patient 2	22.41	20.92	Yes (Respiratory)
Test Patient 3	N/A	N/A	No

D. Limitations

This study presented a preliminary analysis of LSTM autoencoder-based CPD for complication monitoring after major abdominal cancer surgery. While the results demonstrated the potential of this approach, several limitations must be considered. First, the analysis was conducted on a small subset of patients, which may limit the generalizability of the findings. Additionally, implementing cross-validation or leave-one-out strategies would provide more robust estimates of model performance, particularly for small datasets. Second, the dynamic thresholding method accounts for inter-patient variability, but further optimization of the adjustable parameter (z) is needed to balance sensitivity. Finally, change points were also detected in patients with no clinically recorded

complication, which highlights the challenge of distinguishing between normal physiological variations and clinically relevant deviations.

IV. Conclusion and Future Work

This preliminary study explored an LSTM autoencoder-based unsupervised CPD approach for postoperative complications. The proposed model effectively identified deviations in physiological signals with change points prior to the clinically documented complications. Compared to the REWS alarm system, our model demonstrated earlier detection of complications as well. However, the presence of change points in non-complication case indicates the need for further refinement to minimize false positive alarms. Future work will focus on expanding the dataset to improve model generalizability, perform additional preprocessing steps like smoothening the physiological data before processing, and incorporate cross-validation to further evaluate the model's robustness.

Acknowledgment

We would like to thank the patients who participated in this study to the advancement of postoperative monitoring research.

This study is part of the 4TU-programme RECENTRE (Risk-based lifEstyle Change: daily-lifE moNiToring and REcommendations). RECENTRE is funded by the 4TU-programme High Tech for a Sustainable Future (HTSF). 4TU is the federation of the four technical universities in The Netherlands (Delft University of Technology, DUT; Eindhoven University of Technology, TU/e; University of Twente, UT and Wageningen University and Research, WUR).

References

[1] W. R. Al-Kubati, "Optimizing Early Detection of Post-Operative Complications in Colorectal Surgery: A Meta-Analysis of Combined Inflammatory Markers," *World Journal of Colorectal Surgery*, vol. 13, no. 4, pp. 105–116, 10 2024.

[2] A. K. Khanna, S. Ahuja, R. S. Weller, and T. N. Harwood, "Postoperative ward monitoring – Why and what now?" *Best Practice & Research Clinical Anaesthesiology*, vol. 33, no. 2, pp. 229–245, 6 2019.

[3] L. T. Jonker, M. Plas, G. H. de Bock, E. Buskens, B. L. van Leeuwen, and M. M. Lahr, "Remote Home Monitoring of Older Surgical Cancer Patients: Perspective on Study Implementation and Feasibility," *Annals of Surgical Oncology*, vol. 28, no. 1, pp. 67–78, 1 2021.

[4] A. Y. A. Amer, F. Wouters, J. Vranken, D. de Korte-De Boer, V. Smit-Fun, P. Duflot, M. H. Beaupain, P. Vandervoort, S. Luca, J. M. Aerts, and B. Vanrumste, "Vital Signs Prediction and Early Warning Score Calculation Based on Continuous Monitoring of Hospitalised Patients Using Wearable Technology," *Sensors 2020, Vol. 20, Page 6593*, vol. 20, no. 22, p. 6593, 11 2020.

[5] L. T. Jonker, M. M. Lahr, M. H. Oonk, G. H. de Bock, and B. L. van Leeuwen, "Post-discharge Telemonitoring of Physical Activity, Vital Signs, and Patient-Reported Symptoms in Older Patients Undergoing Cancer Surgery," *Annals of Surgical Oncology*, vol. 28, no. 11, pp. 6512–6522, 10 2021.

[6] T. Amin, R. J. Mobbs, N. Mostafa, L. W. Sy, and W. J. Choy, "Wearable devices for patient monitoring in the early postoperative period: a literature review," *mHealth*, vol. 7, p. 50, 7 2021.

[7] S. Aminikhanghahi and D. J. Cook, "A Survey of Methods for Time Series Change Point Detection," *Knowledge and information systems*, vol. 51, no. 2, p. 339, 5 2016. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5464762/

[8] K. J. Oh and I. Han, "An intelligent clustering forecasting system based on change-point detection and artificial neural networks: Application to financial economics," *Proceedings of the Hawaii International Conference on System Sciences*, p. 66, 2001.

[9] J. F. Ducré-Robitaille, L. A. Vincent, and G. Boulet, "Comparison of techniques for detection of discontinuities in temperature series," *International Journal of Climatology*, vol. 23, no. 9, pp. 1087–1101, 7 2003.

[10] M. Bosc, F. Heitz, J. P. Armspach, I. Namer, D. Gounot, and L. Rumbach, "Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution," *NeuroImage*, vol. 20, no. 2, pp. 643–656, 10 2003.

[11] R. Malladi, G. P. Kalamangalam, and B. Aazhang, "Online Bayesian change point detection algorithms for segmentation of epileptic activity," *Conference Record - Asilomar Conference on Signals, Systems and Computers*, pp. 1833–1837, 2013.

[12] T. Tayeh, S. Aburakhia, R. Myers, and A. Shami, "An Attention-based ConvLSTM Autoencoder with Dynamic Thresholding for Unsupervised Anomaly Detection in Multivariate Time Series."

[13] S. Maleki, S. Maleki, and N. R. Jennings, "Unsupervised anomaly detection with LSTM autoencoders using statistical data-filtering," *Applied Soft Computing*, vol. 108, p. 107443, 9 2021.

[14] O. I. Provotar, Y. M. Linder, and M. M. Veres, "Unsupervised Anomaly Detection in Time Series Using LSTM-Based Autoencoders," *2019 IEEE International Conference on Advanced Trends in Information Theory, ATIT 2019 - Proceedings*, pp. 513–517, 12 2019.

[15] A. Sagheer and M. Kotb, "Unsupervised Pre-training of a Deep LSTM-based Stacked Autoencoder for Multivariate Time Series Forecasting Problems," *Scientific Reports 2019 9:1*, vol. 9, no. 1, pp. 1–16, 12 2019.

[16] "Healthdot — Philips." [Online]. Available: https://www.philips.com/a-w/about/dutch-design-week-2020/healthdot.cs

[17] V. J. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review 2004 22:2*, vol. 22, no. 2, pp. 85–126, 10 2004.

[18] J. Darji, N. Biswas, L. D. Jones, S. Ashili, J. Darji, N. Biswas, L. D. Jones, and S. Ashili, "Handling Missing Data in the Time-Series Data from Wearables," *Time Series Analysis - Recent Advances, New Perspectives and Applications*, 8 2023.

[19] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Applied Soft Computing*, vol. 97, p. 105524, 12 2020.

[20] M. Gupta, R. Wadhvani, and A. Rasool, "Real-time Change-Point Detection: A deep neural network-based adaptive approach for detecting changes in multivariate time series data," *Expert Systems with Applications*, vol. 209, p. 118260, 12 2022.

[21] Z. Atashgahi, D. C. Mocanu, R. Veldhuis, and M. Pechenizkiy, "Memory-free Online Change Point Detection: A Novel Neural Network Approach," 2023.