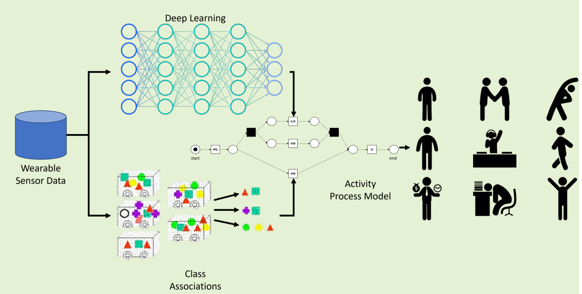
PROMPT: PROcess Mining and Paravector Tensor based Physical Health Monitoring Framework

Sunder Ali Khowaja* *Senior Member, IEEE*, Parus Khuwaja, Kapal Dev *Senior Member, IEEE*, Muhammad Aslam Jarwar *Senior Member, IEEE*

***Abstract*— The provision of physical healthcare services during the isolation phase is one of the major challenges associated with the current COVID-19 pandemic. Smart healthcare services face a major challenge in the form of human behavior, which is based on human activities, complex patterns, and subjective nature. Although the advancement in portable sensors and artificial intelligence has led to unobtrusive activity recognition systems but very few studies deal with behavior tracking for addressing the problem of variability and behavior dynamics. In this regard, we propose the fusion of PRocess mining and Paravector Tensor (PROMPT) based physical health monitoring framework that not only tracks subjective human behavior, but also deals with the intensity variations associated with inertial measurement units. Our experimental analysis on a publicly available dataset shows that the proposed method achieves 14.56% better accuracy in comparison to existing works. We also propose a generalized framework for healthcare applications using wearable sensors and the PROMPT method for its triage with physical health monitoring systems in the real world.**



***Index Terms*— Activity Recognition, Deep Learning, Process Mining, Smart Healthcare, Wearable Sensors**

I. Introduction

CURRENTLY we are facing a pandemic in the form of COVID-19. The situation requires self-isolation at home as the first line of defense. If an individual is asymptotic or has mild symptoms, physical inactivity becomes the risk factor for various diseases [1]. Researchers are trying to build a defense mechanism to slow the spread of COVID-19 through automated means that include physical health monitoring, physical action recognition, and contact tracking. Physical human activity has always been an epitome of a healthy lifestyle. In the context of the COVID-19 pandemic, that is, self-isolation, lifestyle, and health monitoring, can all be performed together through physical action recognition which is required for personal wellbeing. Monitoring of such sort is only possible if the monitoring devices and software are ubiquitous and ubiquitous. Thanks to wearable technology such as smartwatches, body-worn sensors, and smartphones that allow the collection of sensor data without affecting the activities of daily life [2]. The modalities that can be recorded automatically through the devices mentioned vary from inertial units to heart rate and blood pressure measurement.

A process in process mining can be referred to as a set of tasks that are carried out temporally to attain the desired goal [3]. With the recent advancements in micro-electromechanical systems (MEMS) the industries and organizations prefer to record the log of process-related events. The field of process mining combines the characteristics of process modeling, data mining, and computational intelligence to discover temporal events and transform them into a process model that could be observed and analyzed accordingly [3]. Process discovery, i.e., extracting the model from event log, is a vital component of process mining framework which can be realized through state-charts, UML activity diagrams, process trees, Petri Nets, and BPMN models [4]. The process mining techniques were originally proposed for analyzing and modeling business processes, but over time its scope has been broadened to other domains and therefore can be used for human activity recognition.

The semantic meaning of the event labels generated from business process event logs was clear in their terms, such as to request, mortgage, register, and so forth. However, the event logs generated from inertial sensors worn by the humans or motion sensors placed at different places in a home environment have much more complex event labels. This discriminates the use of process mining with traditional business

*Corresponding Author
Sunder Ali Khowaja with Department of Telecommunication Engineering, Faculty of Engineering and Technology, University of Sindh. (email: sandar.ali@usindh.edu.pk, sunder.ali@ieee.org)
Parus Khuwaja with Institute of Business Administration, University of Sindh, Pakistan. (email: parus.khuwaja@usindh.edu.pk)
Kapal Dev is with the Department of institute of intelligent systems, University of Johannesburg, South Africa.(e-mail:kapal.dev@ieee.org)
Muhammad Aslam Jarwar is with Department of Computing, Sheffield Hallam University, Sheffield, UK (e-mail: a.jarwar@shu.ac.uk)

processes and in smart environments such as smart homes, smart learning, and more. Now, process mining techniques are widely used to model human behavior [5]. The modeling of human behavior through generated process models is helpful in the independent living of elder people, healthy lifestyle, health monitoring, learning behaviors, gaming behaviors, and so forth. However, modeling behavior is a complicated task due to the generation of event logs on the sensor level, but the analyst might be interested in investigating the process of how the activities are performed by an individual while interacting with the sensors. In simple words, generating the event log based on sensor activations in smart homes produces non-informative behavior models while overgeneralizing events [6]. In addition to temporal characteristics, researchers have also tried to recognize human activities through various combinations of feature extraction and machine learning techniques. However, most of the features extracted are based on correlations and the intensity levels of sensor readings, which are sensitive to abrupt changes of hand or wrist movements. The inhomogeneity in the intensity levels of sensor readings at different scales makes it difficult for the machine learning algorithm (specifically shallow learning methods such as support vector machines and logistic regression) to recognize actions accurately. Furthermore, the changes also vary in terms of human behavior because performed activities depends on individual's style [7]. The use of paravectors was introduced for complex shape analysis using 3D vectors in Euclidean space. Considering that the wearable sensors have three-axial inertial measurement units, we can think if as plane fragments that can be used to define a new object, i.e., a paravector.

In this study, we propose a human behavior-based physical healthcare monitoring system. The underlying recognition relies on the fusion of process mining and paravector tensors based deep learning methods. The process mining helps to learn the temporal associations among the activities based on human behavior, while the paravector tensors would suppress the shortcomings regarding correlation-based features for learning activity patterns from 3-D sensor data. We show that the proposed approach achieves state-of-the-art results on a publicly available dataset in comparison to the existing works. Furthermore, we propose a generalized framework based on the fusion method for monitoring physical health using multiple modalities for a potential triage and integration method with emerging technologies. To the best of our knowledge, this is the first work to propose a paravector-based solution for sensor-based activity recognition and its fusion with the process mining approach. The contributions of this paper are summarized below.

- We propose a paravector tensor-based solution for human activity recognition using wearable sensors.
- We propose a fusion scheme to combine the process mining and the deep learning method based on a paravector tensor to improve the recognition process.
- We propose a generalized framework that can use multiple modalities for physical health monitoring during the COVID19 pandemic.
- We report the best quantitative results on a publicly available dataset in comparison to existing approaches.

The remainder of this paper is structured as follows. Section II provides a review of articles that focus on modeling the behavior of human activities. Section III presents the methodology for the proposed PROMPT framework to recognize human actions. Section IV presents an experimental setup, results, and analysis to show the importance of behavior modeling in the context of the activity recognition domain and its comparison with existing methods. Section V proposes a generalized framework for physical healthcare monitoring in the context of a pandemic situation using multiple modalities. Section VI concludes the study, followed by future research directions.

II. Related Works

In this section, we present a consolidated review of the studies using process mining to model human behavior in the context of human activities. As the field of process mining is quite recent, there are a few studies that consider these techniques to extract behavior patterns from daily activities. The underlying assumption for using process mining is that an activity can be considered as an event of the process instance, and the behavior can be explored by analyzing the series of events executed in that process. The study in [6] paved the way for analyzing user behavior for high-level activities using process mining techniques. The logs of activities were recorded through smartphones and smartwatches. The logs are then transformed into modeling macroactivities. This study captured the true essence of instrumental activities of daily life (IADL) using process mining techniques, as they focused on the interaction of user activity with respect to location, reflecting human behavior when performing different activities at different places.

Fernandez-Llatas et al. [8] focused on modeling the behavior of workers in the health care domain, which is the hospital area. The activity information was based on the location sensors deployed in the surgical areas of the hospital. The collected data from the indoor location was then used to recognize macro activities followed by the process discovery to analyze the staff behavior. The authors adapted the process mining tool: parallel activity-based log inference algorithm (PALIA) [9] to build a Petri Net on the principles of formal automaton, i.e., timed parallel automaton for representing the process. [10] proposed the use of a "knowledge layer" that acts as an intermediary to map raw sensor data, that is, location and time, from real-time location systems (RTLS) to event logs. The authors [10] refer to the knowledge layer as an interaction. Process mining approaches are then used to discover the process models from the transformed event logs, assuming that the interaction module has prior domain knowledge. [11] focused on modeling user habits in a smart home environment. Their work considered the case study of the smart home project from the center for advanced studies in adaptive systems (CASAS) laboratory of Washington State University. The smart home was equipped with temperature sensor, switch sensors on the main entrance door, and passive infrared (PIR) sensors; which were activated in a sequence when a user performed a specific activity. The user habit modeling was performed using Fuzzy miner for multiple inhabitants in the smart home. [12] proposed the event abstraction method, which derives high-level activity from the relationship of their patterns and low-level event logs. Domain knowledge is required to capture user behavior patterns to perform event abstraction. The advantage of using such abstraction is two-fold. First, process mining techniques will be able to get more insights from the abstracted event log compared to low-level events, and second, the robustness to handle the noise. Their proposed method was evaluated in a hospital setting with the help of experts in the Norwegian hospital domain. [13] exploited process discovery techniques from sensor activation logs to derive activity models. The activity models were then used to analyze user behavior. The major contribution of this work was to characterize event logs in macro- and microactivities generated from user annotations and sensor activations, respectively. The event logs from the respective source, i.e., macro- and micro-activities were pre-processed to remove the noise, and process discovery techniques were used to generate the macroactivity models for the user behavior analysis. [14] focused more on deriving an alternative method to conventional pattern recognition to infer activities and user behavior. Their study relied on knowledge representation and process mining techniques to infer a graphical visualization that depicts the action sequences performed daily. These visualizations were then used to infer the activities and user behavior patterns. [7] proposed a way to combine the characteristics of machine learning with process mining techniques to recognize personalized macro-activities, as well as to extract user behavior patterns to solve a cold-start problem. The contribution of this work mainly lies with the solution of activity handling problem suggesting that the user performs the same activity differently when performing them at different locations, and also they might perform different activities in comparison to other users. Their study proposed a similarity metric that can determine the closest available model in the existing pool of samples based on user behaviors. Once selected, the model will then be used to predict the macro-activities of the user. This work is also one of a kind as it can be applied to both sensor activations and inertial measurement units. The study [15] proposed the use of process mining to discover behavior patterns from activities carried out using sensor data. The process model was generated using flow patterns based on sensors, time, and users in a smart environment. Another study [16] used the pattern mining approach to model user behavior using gait pattern analysis. The study used various filters and cubic spline techniques to analyze the user's gait to recognize high-level activities.

Although the studied used process mining techniques to understand user behavior, some studies focused on improving the modeling process for mining activity patterns [17]. These studies performed filtering, abstraction, and other techniques to reduce the complexity of the generated process model, i.e., transforming a spaghetti model to a lasagna-like process.

III. PROMPT

The proposed PRocess Mining and Paravector Tensor (PROMPT) framework for physical healthcare monitoring using wearable sensors is shown in Figure 1. PROMPT leverages the characteristics of process mining to model human behavior while using paravector tensors to improve the variability in intensity of sensor readings for improving recognition performance. We first define the method for extracting 3D paravector representation, followed by the extraction of indicators and responses and their transformation to tensors. We then briefly define the adopted methodology for process mining based approach. Lastly, we describe our fusion method for improving the action recognition performance, accordingly.

A. Paravector Tensor Module

As shown in Figure 1, we first apply a separable quadrature filter on the alpha scale space in polar space. Quadrature filters have been successfully used for 2D image signals and have proved that the features are resistant to lighting and contrast. We intend to use the quadrature filter for making the features resistant to intensity levels, accordingly. We denote the angular frequency and the radial frequency by $\Im(r_k)$ and $\Im_\gamma(w)$. The filter kernel based on the said frequencies can be defined as $(\Im_\gamma(r_k, w) = \Im(r_k)\Im_\gamma(w))$. The notation r_k refers to orientation, and γ represents scale space, respectively. The frequency coordinates for the scale space are defined as $w = (w_1, w_2)$. We adopt the technique introduced in [18] to apply a fractional order derivative on scale spaces for constructing the radial function while considering the value of $\gamma = 0.25$, hence the name quarter-scale kernel. The quarter-scale derivative is then represented in frequency domain using the formulation shown in equation 1. The mathematical proofs for deriving the scale spaces are given in [18], [19].

$$\Im_\gamma(w) = norm_c\ w^j\ exp(-\sqrt{scale \cdot w}) \tag{1}$$

where $scale$ refers to the scaling factor, $norm_c$ represents the normalization constant and $j \in \mathbb{R}^+$. The formulation for the octave bandwidth $\mathbb{B}$, $norm_c$, and the peak tuning frequency w_0 are given in equation 2.

$$\begin{aligned} w_0 &= \frac{4j^2}{scale},\ norm_c = 2\frac{\sqrt{\pi/4}(2^{2j+1})(scale^{j+0.5})}{\sqrt{2(2j+1)}} \\ \mathbb{B} &= \frac{2\ln(\frac{\mathbb{W}(-1,\alpha)}{\mathbb{W}(0,\alpha)})}{\ln(2)},\ \alpha = -\frac{2^{-\frac{1}{2j}}}{e} \end{aligned} \tag{2}$$

The notation $\mathbb{W}$ represents the Lambert function in equation 2. The derived filter comprises both the imaginary (odd-symmetric) and real (even-symmetric) parts, respectively. The selectivity of responses to the orientations is determined by the angular component, that is, $(\hat{r_k} \cdot \hat{w})^2$. The orientation of the filter directing vector is given as $(k\frac{\pi}{3})$ suggesting that only three directions will be considered for the feature extraction, accordingly.

The second feature extraction technique employed for the paravector tensor module is the rotational paravector function using the Riesz transform [20]. The transform projects the summation of oriented symmetric patterns from time-series data in a 3D vector, i.e. two imaginary parts and one real part. We denote the rotational paravector as $\Re_T^{r_k}(w)$ and the

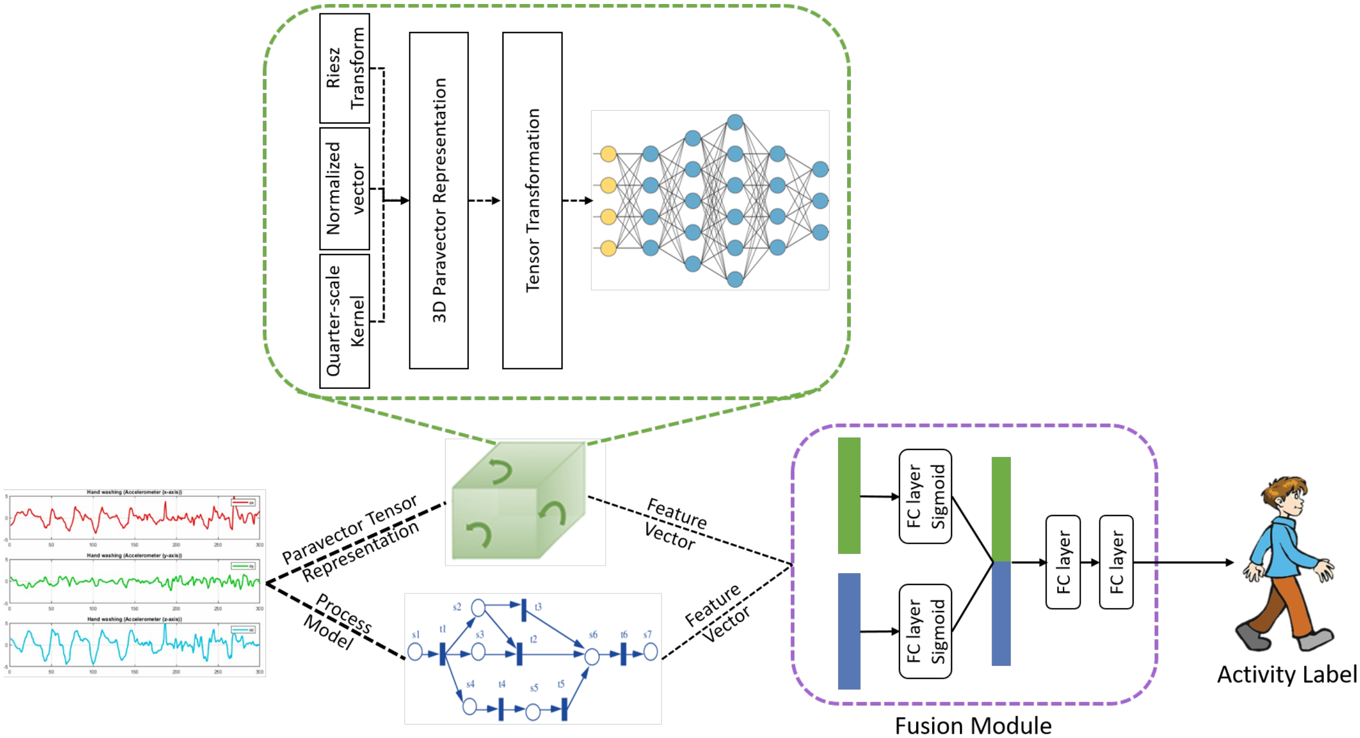


Fig. 1. Proposed PROMPT Architecture for Physical Healthcare Monitoring

formulation to calculate the rotational paravector is shown in equation 3.

$$\Re_T^{r_k}(w) = (\bar{\mathfrak{h}} * \mathfrak{U} * \Im_\gamma^{r_k})(w), when\ \bar{\mathfrak{h}} = (\bar{\mathfrak{h}}_1, \bar{\mathfrak{h}}_2) \tag{3}$$

The notation $\bar{\mathfrak{h}}$ and $\mathfrak{U}$ refers to the Fourier domain representation of the Riesz transform and the L_2 norm of the sensor modality readings, respectively. The representation of odd parts regarding 3D paravector is shown in Equation 4.

$$\begin{aligned} \bar{\mathfrak{h}} &= \sqrt{-1}(\bar{\mathfrak{h}}_1(w) + \sqrt{-1}\bar{\mathfrak{h}}_2(w)) \\ &= \sqrt{-1}\frac{w_1}{\|w\|_2} + \sqrt{-1}(\sqrt{-1}\frac{w_2}{\|w\|_2}) \\ &= \frac{\sqrt{-1}w_1 - w_2}{\|w\|_2} \end{aligned} \tag{4}$$

With reference to Equation 4, we can compute even and odd responses using Equations 5 - 7.

$$\mathbb{E}^{r_k}(x) = real(F^{-1}(\bar{\mathfrak{U}}(w) * \Im_\gamma^{r_k}(w))) \tag{5}$$

$$\mathfrak{h}_1^{r_k}(x) = real(F^{-1}(\bar{\mathfrak{U}}(w) * \Im_\gamma^{r_k}(w) * \bar{\mathfrak{h}}(w))) \tag{6}$$

$$\mathfrak{h}_2^{r_k}(x) = imag(F^{-1}(\bar{\mathfrak{U}}(w) * \Im_\gamma^{r_k}(w) * \bar{\mathfrak{h}}(w))) \tag{7}$$

where $\mathbb{E}^{r_k}$, $\mathfrak{h}_1^{r_k}$, $\mathfrak{h}_2^{r_k}$, represent the even and odd parts while F^{-1} refers to the inverse Fourier transform, respectively. The study [21] showed that the local features from 3D paravector function can be estimated using Clifford analysis. These local features exhibit the direction information in terms of phase through local orientation $\theta_T^{r_k}$ and local phase $\phi_T^{r_k}$. The aforementioned features are obtained using Riesz energy and paravector energy functions which are shown in equation 8 and 9.

$$\mathbb{E}_T^{r_k} = \sqrt{\mathbb{E}^{r_k\,2} + \mathbb{E}_{\mathfrak{h}}^{r_k}},\ \mathbb{E}_{\mathfrak{h}}^{r_k} = \sqrt{\mathfrak{h}_1^{r_k\,2} + \mathfrak{h}_2^{r_k\,2}} \tag{8}$$

$$\theta_T^{r_k} = arctan2(\mathfrak{h}_2^{r_k}, \mathfrak{h}_1^{r_k}), \phi_T^{r_k} = arctan(\frac{\mathbb{E}_{\mathfrak{h}}^{r_k}}{\mathbb{E}^{r_k}}) \tag{9}$$

where $\theta_T^{r_k}$ ranges between $[-\pi, \pi]$ and $\phi_T^{r_k}$ ranges between $[-\pi/2, \pi/2]$. Finally, we transform the paravector valued functions into tensor representation using the formulation shown in equation 10.

$$\tau = \sum_k \mathbb{M}^{r_k}(\hat{Z}_k \hat{Z}'_k - \rho\mathbb{I}) \tag{10}$$

where τ represents the paravector tensor, $\mathbb{M}$ represents the features, i.e., $\mathbb{E}_T^{r_k}$, $\theta_T^{r_k}$, and $\phi_T^{r_k}$. The notation $\hat{Z}_k$ refers to the normalized column vector defined as $\hat{Z}_k = [cos(norm_c); sin(norm_c); 1]$. An identity matrix of size 3x3 is denoted by $\mathbb{I}$ and ρ refers to the orientation scale, which is 0.25, respectively. The derived tensor representation is then processed through a deep learning architecture to extract the feature vectors. It should be noted that the paravector tensors derived above are for a single window and a single modality. For instance, τ_1^{acc} represents the tensor representation for 1st window of accelerometer sensor. Therefore, the tensor representation for all windows and modalities are provided as input to the deep learning architecture.

B. Process Mining Module

This section provides information regarding the basic notation used in process mining techniques in the context of

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3195613

activity recognition studies. The process instance can be formally characterized by three elements, i.e. event, trace, and event-log.

An event E can be represented in the form of a tuple $<\alpha, r, t>$, where α refers to an activity of a set of activities $\mathcal{A}$, r refers to the resource who performed the activity, and t represents the time stamp at which the activity is performed by a specific resource. Let us assume that a user named Oliver $\in \mathcal{R}$ performs an activity *Personal Grooming* $\in \mathcal{A}$ at 9.15 on June, 1^{st} 2016. The tuple for this event can be represented as ⟨Personal Grooming, Oliver,"2016-06-01 09:15:00"⟩. A trace $\mathcal{T}r$ refers to the sequence of events $\{\epsilon_1, \ldots, \epsilon_n\}$ in temporal ordering such that each event is unique, i.e. $\nexists \epsilon_i, \epsilon_j$ and $i \neq j | \epsilon_i = \epsilon_j \wedge \forall_{i,j}$ with $i < j, t_i \leq t_j$. An event log ϵl records all process instances in a dataset. An event log stores the trace ($\mathcal{T}r$) of a process instance according to its temporal order of occurrences. An ϵl can be represented as $\{\mathcal{T}r_1, \ldots, \mathcal{T}r_k\}$ such that every ϵ appears only once in the complete log, i.e $\forall \epsilon \in \epsilon l, \nexists \mathcal{T}r_i, \mathcal{T}r_j | \epsilon \in \mathcal{T}r_i \wedge \epsilon \in \mathcal{T}r_j$. An example of a trace for a set of activities from a publicly available dataset is shown in Table I.

TABLE I
SNIPPET OF TOY DATASET SHOWING MORNING ACTIVITIES

	$\mathcal{A}$	$\mathcal{R}$	$\mathcal{T}$
ϵ_1	Personal Grooming	Oliver	2016-06-01 09:15:00
ϵ_2	Eating/Drinking	Oliver	2016-06-01 09:57:03
ϵ_3	House Work	Oliver	2016-06-01 10:09:27
ϵ_4	Socializing	Oliver	2016-06-01 11:21:47
ϵ_5	Personal Grooming	Robert	2016-06-01 07:12:21
ϵ_6	Desk Work	Robert	2016-06-01 07:23:02
ϵ_7	Socializing	Robert	2016-06-01 07:52:28
ϵ_8	Personal Grooming	Oliver	2016-06-02 08:29:11
ϵ_9	House Work	Oliver	2016-06-02 09:02:03
ϵ_{10}	Eating/Drinking	Oliver	2016-06-02 09:52:07
ϵ_{11}	Socializing	Oliver	2016-06-02 10:33:29

With reference to the example shown in Table I, the set of activities and resources can be described as $\mathcal{A}$ ={*Personal Grooming, Eating/Drinking, House Work, Socializing*}, and $\mathcal{R} = \{Oliver, Robert\}$, respectively. The traces with respect to the particular sequences of events shown in Table I are represented as follows: $\mathcal{T}r_1 = \{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4\}$, $\mathcal{T}r_2 = \{\epsilon_5, \epsilon_6, \epsilon_7\}$, and $\mathcal{T}r_3 = \{\epsilon_8, \epsilon_9, \epsilon_{10}, \epsilon_{11}\}$. These event logs provide some insights regarding the user behavior, for instance each event starts with activity "Personal Grooming" and ends with an activity "Socializing". However, the ones performed between these activities vary from $\mathcal{T}r_1 - \mathcal{T}r_3$, with some abnormalities/commonalities of course, for instance, activities "House Work" and "Eating/Drinking" appear together in two of the traces. In reality, there are thousands of events in real-world event logs, and it is nearly impossible to perform a manual analysis. Therefore, generating process flows and looking for such events corresponds to the field of process modeling.

There are several process discovery techniques for deriving a process model from the event log if it is not available in the first place. The most common techniques for process discovery include infrequent inductive miner, Heuristic miner, Fuzzy miner, and Alpha Miner [3]. The infrequent inductive miner discovers the model based on the process tree principles, where the activity or the process is represented through a leaf node and an operator is represented as a branch node. The output of the infrequent inductive miner is represented by Petri Net. The miner filters the infrequent instances. The model discovered through inductive miner is considered to be one of the soundest ones (refer to the soundness [3]), i.e. the final marking is always reachable when the process is executed using Petri Net. The heuristic miner is similar to the infrequent inductive miner in the sense that they both filter the infrequent behaviors. The heuristic miner is also related to Alpha miner, where it extends the mapping of casual relationships between activities and applies some heuristics while considering the frequency of the relationships. This allows a heuristic miner to differentiate between the parallel and sequential activities, accordingly. The fuzzy miner provides a high-level abstraction of the process model by grouping the most infrequent instances. The output generated using a fuzzy miner aggregates the infrequent activities as a single activity, while displaying the most frequent ones with their prior relations. In this study, we employ the associative classifier method that uses process mining technique at its core as proposed in study [7]. However, instead of directly classifying activities and generating process model, we use the class association rules and transform them into a numeric matrix and flatten them to generate the feature vector, accordingly.

C. Fusion Module

It has been proved by several existing studies that fusing multiple modalities or diverse characteristics from different feature set improves the activity recognition performance [2], [22]. Most of the studies employ decision-level fusion, i.e., combining the classification scores using weighted averaging, due to its simplistic approach. Although improved results have been achieved using the said approach, it does not exploit correlations among heterogeneous feature sets or multiple modalities. In this regard, the PROMPT architecture performs feature-level fusion to model the correlations between features based on the paravector tensor and features based on the rules of the class association to enhance recognition performance. There are many techniques proposed to fuse the feature vectors from multiple streams such as hard fusion, soft fusion, and direct fusion [23]. Direct fusion simply concatenates the feature vectors, thus, it does not model the correlations quite well. Therefore, in this study, we investigate the latter two techniques for feature level fusion. The hard fusion strategy passes the feature vector through a fully connected layer followed by a sigmoid function. A Gumbel softmax method is used to compute the hardmask obtained from the sigmoid function [24] which consists of binary values. Let's denote the feature vector obtained using paravector tensor module as $feat_{pt}$ and the feature vector from the process mining module as $feat_{pm}$, the computation for hard fusion is given as:

$$\begin{aligned} &S_{pt} = \sigma(feat_{pt}),\ S_{pm} = \sigma(feat_{pm}) \\ &G_{pt} = gumbel_{sm}(S_{pt}),\ G_{pm} = gumbel_{sm}(S_{pm}) \\ &\psi_{hard}(feat_{pt}, feat_{pm}) = [feat_{pt} \odot G_{pt}; feat_{pm} \odot G_{pm}] \end{aligned} \tag{11}$$

The only difference between the hard and soft fusion method is that the Gumbel softmax method is not used, therefore, soft fusion only generates a softmask multiplied element wise with its corresponding modality. The computation for the soft fusion is shown in Equation 12

$$S_{pt} = \sigma(feat_{pt}),\ S_{pm} = \sigma(feat_{pm})$$
$$\psi_{soft}(feat_{pt}, feat_{pm}) = [feat_{pt} \odot S_{pt}; feat_{pm} \odot S_{pm}] \tag{12}$$

IV. Experiment and Analysis

In this section, we first provide the details regarding experimental setup, followed by the network and parameter details used for the proposed method. We then present experimental results and comparisons with existing work to prove the efficacy of the proposed approach.

A. Experimental Setup

The experimental analysis has been carried out on daily life log dataset [5]. The rationale for choosing the dataset is because of the availability of raw sensor data as well as high-level activity labels in relation to low-level actions. We employ recently proposed Transformer networks [25] as our deep learning architecture, and CAPHAR method [7] for computing class association rules. The process model is generated using Inductive miner in PRoM and decision trees with Python [1] were used for predicting the high-level activity label, accordingly. We also compare the performance in terms of accuracy and average execution time obtained using the proposed framework with some existing methods, as shown in Figures 2 and 3. The existing methods include, DeepConvLSTM [26], Deep Residual Bidirectional LSTM (DRBLSTM) [27], DeepSense [28], and Associative learning (AL) [7], respectively.

B. Network and Parameter Details

We consider 150 samples in a single window obtained using inertial sensor measurements, i.e., 50 Hz sampling rate while window size is of 3 seconds. The kernel size of Conv1D layers in the transformer architecture is set to 2. We use 7 residual blocks for the data extraction from paravector tensor module. A ReLU and drop out layer is placed after each Conv1D layer of each residual block. The initial layers use 64 filters while the last two layers use 128 filters. In the fusion module, two fully connected layers are used. The first layer considers the same number of units as that of the feature maps, followed by a ReLU and dropout layer having 0.4% dropout factor. The second fully connected layer has a number of units similar to that of activity classes. For stable training and fast convergence, we use Rectified ADAM optimizer [29]. The learning rate is set to 0.001 with a decay rate of 10 after every 10 epochs. The batch size was set to be 32, accordingly.

[1]https://pm4py.fit.fraunhofer.de/implemented-approaches

C. Experimental Results

We perform leave-one-subject-out (LOSO) analysis, which is a standard for human activity recognition studies. The results have been compared with state-of-the-art approaches, i.e., DeepConvLSTM, DRBLSTM, DeepSense, and Associative Learning. The results for accuracy are shown in Figure 2. The results represent the superiority of the proposed PROMPT method in comparison to existing approaches. The proposed work achieves 74.09% accuracy which is an improvement of 14.56% in comparison to DeepSense (achieved the least accuracy), and 6.24% better in comparison to Associative learning method (achieved second best accuracy). It should be noted that the associative learning method is trained for individual subjects, i.e., a separate test model for each subject, respectively, while the proposed work uses a single model to classify all the high-level activities. The results also show that the PROMPT method achieves relatively less execution time in comparison to the majority of existing methods. The execution time can further be lessened by reducing the number of trees employed for classification purposes, but there would be a trade-off in terms of achievable accuracy.

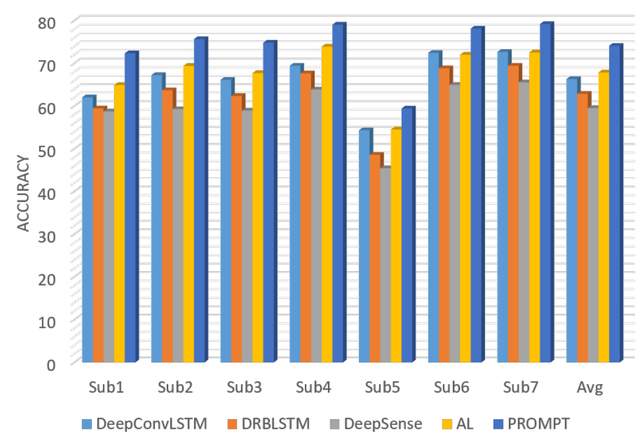


Fig. 2. Comparative Analysis of the proposed work with existing methods on Daily Life Log Dataset

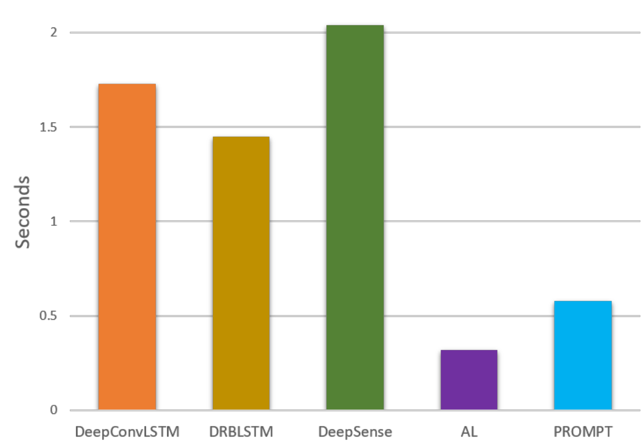


Fig. 3. Comparison of the average execution time of the proposed work with existing methods on Daily Life Log Dataset

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3195613

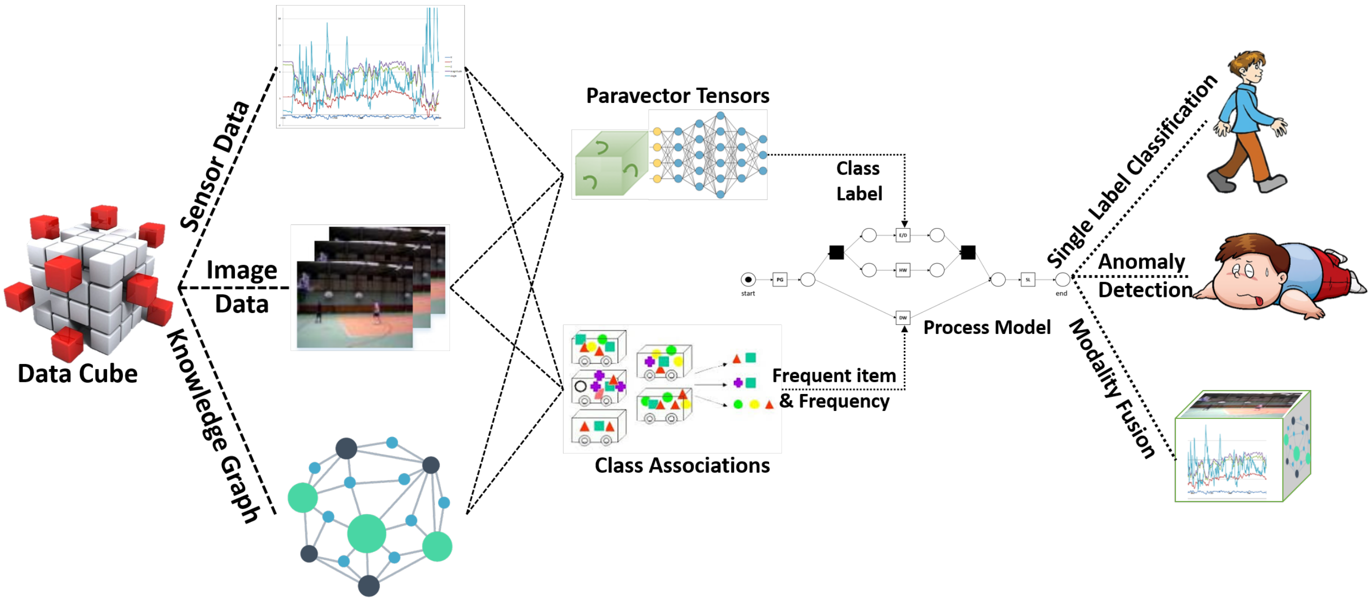


Fig. 4. Generalized PROMPT Framework for Triage in Pandemic Situations with Physical Healthcare Monitoring Systems using multiple modalities.

V. Triage of the PROMPT for Multiple Modalities and Physical Healthcare Applications

We propose a generalized PROMPT framework for its triage with multiple modalities and physical healthcare applications as shown in Figure 4. The framework is designed in the context of smart health concerning pandemic situation such as COVID-19. We explain the components of the proposed framework in the context of smart healthcare monitoring with respect to recognition of physical activity. Data cube is a generalized term that can store the data from multiple sensor modalities, such as sensor data, image data, and knowledge graphs. The framework can use the data from all the modalities or anyone modality, considering its availability. Now, most of the architectures include pre-processing and feature extraction layers. Considering the recent deep learning architectures, the need for pre-processing and feature extraction has been reduced significantly. However, if the system designer wants to use such techniques, it won't affect the overall process of the proposed architecture. The data is then sent to the paravector tensor module to extract informative features that are resilient to changes in intensity, which is a recurring problem in inertial measurement units associated with wearable sensors. The paravector tensors are then sent to a deep learning architecture such as convolutional neural networks, long-short-term memory networks, gated recurrent units, transformers, or others to extract feature representation for predicting the class label. Alternatively, the same data will be an input to the class association block, which will compute the class association rules based on the frequent item sets. For example, walking and standing are performed quite frequently together while performing "personal grooming" activity. In this regard, the class association rules will be constructed in order to provide a context later. The class labels from the deep learning architecture and the class association rules with frequency from the association block are then provided input to the process model. The process model will undergo conformance checking and the process will be iteratively performed until the maximum conformance checking is obtained. The process model is then used for predicting the high-level or anomalous activities. Furthermore, process model from each of the modality can be constructed, and the final results might be fused to increase the classification performance, accordingly.

VI. Conclusion

This work proposes a fusion of PRocess Mining and Paravector Tensor (PROMPT) for human action recognition based on their behavior. The experimental analysis on a publicly available dataset shows that the PROMPT achieves the best accuracy compared to existing work while utilizing less computation time. The work also proposes a generalized architecture for physical healthcare monitoring in pandemic situation using wearable sensors and PROMPT method. The underlying application is quite relevant in the current COVID-19 situation. Activity monitoring is important in a sense that people live in isolation, and user behavior patterns can provide us with insights regarding their anomalous behavior. Moreover, patterns can help us to distinguish between the COVID patient and the normal one. Furthermore, the behavior patterns can be extended to check the compliance of the isolation guidelines followed by the COVID patient. The process mining framework is currently one of the best resources for extracting such behavior patterns.

One of the limitations of the proposed approach is that it requires both the raw sensor data for low-level actions and contextual data for high-level activities. Due to this limitation, our experimental results are restricted to the analysis of only one dataset which provides both of the required information.

The architecture also highlights some potential future works for the integration of process mining techniques with machine learning and knowledge representations. As an extension, we would also like to apply the proposed framework for students learning behavior and work on a real-life case study by collecting a large volume of data from students watching online videos in the current pandemic crisis globally.

References

[1] K. Dev, S. A. Khowaja, A. S. Bist, V. Saini, and S. Bhatia, "Triage of potential COVID-19 patients from chest X-ray images using hierarchical convolutional networks," *Neural Computing and Applications*, pp. 1–16, 2021.

[2] S. A. Khowaja, A. G. Prabono, F. Setiawan, B. N. Yahya, and S.-L. Lee, "Contextual activity based Healthcare Internet of Things, Services, and People (HIoTSP): An architectural framework for healthcare monitoring using wearable sensors," *Computer Networks*, vol. 145, pp. 190–206, nov 2018.

[3] W. van der Aalst, "Data Science in Action," in *Process Mining*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 3–23.

[4] S. J. J. Leemans, D. Fahland, and W. M. P. van der Aalst, "Discovering Block-Structured Process Models from Event Logs - A Constructive Approach," in *International Conference on Applications and Theory of Petri Nets and Concurrency*, 2013, pp. 311–329.

[5] T. Sztyler, J. Carmona, J. Völker, and H. Stuckenschmidt, "Self-tracking Reloaded: Applying Process Mining to Personalized Health Care from Labeled Sensor Data," in *Transactions on Petri Nets and Other Models of Concurrency XI*, 2016, pp. 160–180.

[6] N. Tax, N. Sidorova, R. Haakma, and W. M. van der Aalst, "Log-based Evaluation of Label Splits for Process Models," *Procedia Computer Science*, vol. 96, pp. 63–72, 2016.

[7] S. A. Khowaja, B. N. Yahya, and S.-L. Lee, "CAPHAR: context-aware personalized human activity recognition using associative learning in smart environments," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, p. 35, dec 2020.

[8] C. Fernandez-Llatas, A. Lizondo, E. Monton, J.-M. Benedi, and V. Traver, "Process Mining Methodology for Health Process Tracking Using Real-Time Indoor Location Systems," *Sensors*, vol. 15, no. 12, pp. 29 821–29 840, nov 2015.

[9] C. Fernandez-Llatas, B. Valdivieso, V. Traver, and J. M. Benedi, "Using Process Mining for Automatic Support of Clinical Pathways Design," in *Data Mining in Clinical Medicine*, 2015, pp. 79–88.

[10] A. Senderovich, A. Rogge-Solti, A. Gal, J. Mendling, and A. Mandelbaum, "The ROAD from Sensor Data to Process Instances via Interaction Mining," in *International Conference on Advanced Information Systems Engineering*, 2016, pp. 257–273.

[11] M. Dimaggio, F. Leotta, M. Mecella, and D. Sora, "Process-Based Habit Mining: Experiments and Techniques," in *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*. IEEE, jul 2016, pp. 145–152.

[12] F. Mannhardt, M. de Leoni, H. A. Reijers, W. M. P. van der Aalst, and P. J. Toussaint, "From Low-Level Events to Activities - A Pattern-Based Approach," in *International Conference on Business Process Management*, 2016, pp. 125–141.

[13] M. Cameranesi, C. Diamantini, A. Mircoli, D. Potena, and E. Storti, "Extraction of User Daily Behavior From Home Sensors Through Process Discovery," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8440–8450, sep 2020.

[14] M. R. Ma'arif, "Revealing daily human activity pattern using process mining approach," in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. IEEE, sep 2017, pp. 1–5.

[15] F. Prathama, B. N. Yahya, and S.-L. Lee, "A Multi-case Perspective Analytical Framework for Discovering Human Daily Behavior from Sensors using Process Mining," in *IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2021, pp. 638–644.

[16] V. Bijalwan, V. B. Semwal, and V. Gupta, "Wearable sensor-based pattern mining for human activity recognition: deep learning approach," *Industrial Robot*, vol. 49, no. 1, pp. 21–33, 2021.

[17] N. Tax, N. Sidorova, R. Haakma, and W. M. P. van der Aalst, "Event Abstraction for Process Mining Using Supervised Learning Techniques," in *Proceedings of SAI Intelligent Systems Conference*, 2018, pp. 251–269.

[18] A. Belaid and Boukerroui, "A new generalised alpha scale spaces quadrature filters ," *Pattern Recognition*, vol. 47, no. 10, pp. 3209–3224, 2014.

[19] E. J. Pauwels, L. J. van Gool, P. Fiddelaers, and T. Moons, "An extended class of scale-invariant and recursive scale space filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 7, pp. 691–701, 1995.

[20] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001.

[21] S. Held, M. Storath, P. Massopust, and B. Forster, "Steerable Wavelet Frames Based on the Riesz Transform," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 653–667, 2010.

[22] S. A. Khowaja, B. N. Yahya, and S.-L. Lee, "Hierarchical classification method based on selective learning of slacked hierarchy for activity recognition systems," *Expert Systems with Applications*, vol. 88, pp. 165–177, dec 2017.

[23] C. Chen, S. Rosa, Y. Miao, C. X. Lu, W. Wu, A. Markham, and N. Trigoni, "Selective Sensor Fusion for Neural Visual-Inertial Odometry," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 10 542–10 551.

[24] E. Jang, S. Gu, and B. Poole, "Categorical Reparameterization with Gumbel-Softmax," 2016.

[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," jun 2017.

[26] F. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors*, vol. 16, no. 1, p. 115, jan 2016.

[27] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, "Deep Residual Bidir-LSTM for Human Activity Recognition Using Wearable Sensors," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–13, dec 2018.

[28] S. Yao, S. Hu, Y. Zhao, A. Zhang, and T. Abdelzaher, "DeepSense: A Unified Deep Learning Framework for Time-Series Mobile Sensing Data Processing," in *Proceedings of the 26th International Conference on World Wide Web*. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, apr 2017, pp. 351–360.

[29] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han, "On the Variance of the Adaptive Learning Rate and Beyond," 2019.