IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 19, 2021, accepted August 1, 2021, date of publication August 9, 2021, date of current version August 17, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3103211*

# Deep Learning for Human Activity Recognition Based on Causality Feature Extraction

**YU MIN HWANG, (Member, IEEE), SANGJUN PARK, (Member, IEEE), HYUNG OK LEE, (Member, IEEE), SEOK-KAP KO, (Member, IEEE), AND BYUNG-TAK LEE, (Member, IEEE)**

Honam Research Center, Electronics and Telecommunications Research Institute, Gwangju 61012, Republic of Korea

Corresponding author: Byung-Tak Lee (bytelee@etri.re.kr)

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant through by the Korean Government [Ministry of Science and ICT (MSIT)] (Solution for social problems using personalized care service for vulnerable social groups) under Grant 2020-0-01861.

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

**ABSTRACT** We propose a novel data-driven feature extraction approach based on direct causality and fuzzy temporal windows (FTWs) to improve the precision of human activity recognition and mitigate the problems of easily-confused activities and unlabeled data, which significantly degrade classification performance owing to the correlation of labeled data. In recognizing activities, the proposed approach not only considers the importance of oncoming short-term sensor data but also considers the continuity from past activities of the preceding long-term sensor data. In terms of the oncoming data, the causality feature is extracted using the direct transfer entropy to determine the unique pattern of an activity, which represents the quantified causal relationship between sensor activations. In terms of the preceding data, several hours of historical data are compressed to fuzzy features based on FTWs. Subsequently, the causality and fuzzy features are fused by matrix multiplication to express distinct features of activities. To effectively learn the spatiotemporal dependencies of the fused feature, deep long short-term memory (LSTM), two-dimensional convolutional neural network (2D-CNN), and hybrid models composed of a combination of LSTM and CNN were used. Leave-one-day-out cross-validation was performed based on the CASAS open datasets, including Aruba, Cairo, and Milan. The results showed that the macro-F1-scores were improved by 16.4, 37.5, and 18.5%, respectively, compared with those of the FTW-only environments. In addition, the proposed approach could improve the precision of activity recognition and mitigate the problems associated with the environments containing unlabeled data.

**INDEX TERMS** Activity of daily living, activity recognition, causality, convolutional neural network, deep learning, fuzzy feature, long short-term memory.

## I. INTRODUCTION

Owing to enhanced life expectancies, the number of elderly people has increased worldwide. Therefore, the social-level demand for caring for those living alone in terms of physical and mental health has increased [1]–[4]. Moreover, recognition of daily activities based on artificial intelligence and machine learning is a promising solution. By monitoring their health conditions, identifying deviations in their activities, and detecting emergencies, activity recognition (AR) can be used to assist the elderly that live on their own. AR can be classified into three main categories based on the type of data collection method: camera video [5]–[10], wearable devices [11]–[28], and binary sensors [29]–[36]. The camera video and wearable devices methods are not desirable because of their invasion of privacy and impracticality, such as discomfort while wearing a wearable device and maintenance burden in practical applications. Thus, this study focused on data-driven AR based on deep learning using a device-free and non-privacy-invading method. The binary sensor-based method was used because it provides a practical solution for real-life long-term activity monitoring.

The associate editor coordinating the review of this manuscript and approving it for publication was Haruna Chiroma.

In general, feature representation and extraction are crucial steps in the AR. To precisely classify and recognize easily-confused activities, such as meal preparation, dish washing, and eating in a kitchen area, this study focused on extracting a meta-action by quantifying the causal influence between a series of sensor activations. This is because human activity is a process variable comprising sequential activities formed based on the unique rules, habits, and lifestyles of a person [37]–[41]. Even if the areas of activities are identical, the unique orders or characteristics of sensor activation in the activities can differ depending on the habits and lifestyles and can be represented as a causality between sensors. These unique features can be extracted and learned to improve AR performance.

Studies on quantitative characterization while extracting causality on a given time series have been reported [42]–[46]. Among them, a well-known study analyzed Granger causality (G-causality) [42], [43], which represents the causality between two time series using an auto-regression model. However, G-causality is limited in dealing with nonlinear relationships because it is based on a linear model. To overcome the limitation of G-causality, transfer entropy (TE) [44], [45], which deals with nonlinear relationships, was proposed to measure the causal influence among multivariate time series. In [46], the authors proposed direct TE (DTE) by dividing it into direct and indirect quantitative causality between a pair of variables because the propagation of information between two variables can be affected by intermediate variables. In this study, to utilize the benefits of extracting the meta-action of current sensor activation and enhancing the recognition performance in AR, we adopted the concept of direct causality between oncoming short-term series of sensor activations, such as sensor activation for tens of seconds or a few minutes after query time. We implemented this concept by calculating the normalized DTE (NDTE) as a feature matrix to be learned.

Furthermore, processing of not only the oncoming short-term data but also of the preceding long-term data of sensor activation are crucial for recognizing the current activity because activities occur in succession based on the activity transition probability of residents. However, long-term historical data are significantly vast, for example, several hours or a day before query time, and they should be compressed and transformed into a single feature matrix to be effectively learned. In this study, fuzzy features extracted by fuzzy temporal windows (FTWs) [2], [47]–[51], which is a data representation method that effectively compresses large preceding data, were used.

Concurrently, there are still unresolved challenges in AR, such as unlabeled other activity data. Some of the challenges are irrelevant and routine activities, whereas others are related to a predefined target class. Unlabeled data can significantly reduce the AR performance [52]–[55] because of their correlation with labeled data. In real-time AR services that test all sensor data generated in real time in order, if the AR model is not trained on the other data, it will erroneously recognize queries as predefined activities even though there is no correspondence between them. In some studies, although the models were trained using such unlabeled data, queries were tested only with labeled data, excluding the queries from the unlabeled data in the performance evaluation. Model evaluation using this method is impractical because of the unreliability of the recognition accuracy.

By considering the above-mentioned limitations, the motivation and objectives of this study were to improve the precision of AR and mitigate the problems of easily-confused activities and unlabeled data for real-time reliable AR services. Therefore, the following steps were adopted:

*Step 1*: A novel sensor-based data-driven AR architecture based on causality and fuzzy feature learning was proposed. After data preprocessing for event-triggered feature extraction, the causality feature that describes the cause-and-effect information between sensor activation to express the unique patterns of resident activity was extracted. The extraction was based on NDTE and self-NDTE for all sensor pairs of oncoming short-term data, where self-NDTE is the causality between the same sensors.

*Step 2*: The fuzzy features were extracted from a large amount of historical data using FTWs for several hours to infer the current activity by determining the past activities. Thereafter, the two features (causality and fuzzy) were fused by the dot product to capture both the past and present features. Because the two-dimensional (2D) matrix of the fused feature contains spatiotemporal information, three deep learning models were trained: deep long short-term memory (LSTM), deep 2D convolutional neural network (2D-CNN), and deep 2D-CNN–LSTM to utilize the enhancing process of deep learning. *To the best of our knowledge, this is the first study to examine the approach of the extraction and learning of a fused feature based on direct causality in AR.*

*Step 3*: Experiments were conducted to investigate the recognition performance of the proposed AR architecture utilizing the three open datasets, Aruba, Cairo, and Milan, provided by a project of the Center for Advanced Studies in Adaptive Systems (CASAS) of Washington State University [56]. The experimental results were obtained using a well-known performance measure, precision, recall, F1-score, and confusion matrix in experimental environments. Three open datasets, three deep learning models, application/nonapplication of the causality features, and presence/absence of unlabeled data were utilized. The leave-one-day-out cross-validation results indicated that our proposed model outperformed the existing studies in terms of the macro-F1-scores, including the other activities.

The remainder of this paper is organized as follows. Related studies are discussed in Section II. In Section III, we describe the proposed model, including feature extraction and fusion. Section IV presents the experimental results. Finally, the conclusions of this study are summarized in Section V.

## II. RELATED STUDIES

This section discusses the existing studies in two main categories of human AR (HAR) subproblems: i) feature learning with temporal dependencies of binary sensor-based data and ii) interference with unlabeled data.

### A. TEMPORAL DEPENDENCY-BASED FEATURE EXTRACTION

Binary sensor-based AR is a useful and unobtrusive method for evaluating the conditions of daily living in a sensorized environment. Binary sensors, such as small and light devices, including passive infrared sensors, motion detectors, and contact switches, are suitable for obtaining data on sequential human activities and interactions. For high-precision AR, an effective method for extracting and expressing the hidden temporal features from sequential sensor activation data is essential, and the effectiveness of the extraction of the distinct temporal features leads to high recognition performance.

One efficient method for extracting temporal features is the using FTWs [2], [47], [48] and membership functions to implicitly express the features for several hours in preceding massive event data. In [2], an LSTM and CNN-based model that learns partial oncoming activities in addition to preceding sensor activations was proposed. In [47], an ensemble of activity-based classifiers, where each classifier is composed of LSTM networks, was proposed for balanced-based similarity training.

Feature extraction has been developed for dimensionality reduction methods without the loss of class separability. Unsupervised techniques, such as single k-means, fuzzy c-means, multi-view k-means, and multi-view fuzzy c-means clustering, are powerful techniques for finding hidden and unhidden labels considering feature reduction and unknown number cluster patterns in data. A novel algorithm was proposed in [57] for clustering multi-view data, termed feature-reduction multi-view K-means, which could automatically reduce unimportant features. In [58], a mechanism was presented for determining the number of clusters with feature-reduction behavior under an unknown number of clusters for k-means clustering, called Entropy-k-means. In [59], a novel multi-view fuzzy c-means clustering algorithm was proposed with view and feature weights based on collaborative learning for feature reduction to exclude redundant/irrelevant feature components. In [60], a robust learning-based fuzzy c-means framework was introduced to obtain free of the fuzziness, initializations without parameter selection, and to find the best number of clusters.

To develop intelligent systems from sensor data streams, a new approach was introduced for online AR with three temporal sub-windows [49], [50]. Moreover, an ensemble of activity-based classifiers was presented for balanced training and the selection of relevant sensors using FTWs [51]. Even though the fuzzy temporal feature extraction-based approaches in [2], [47]–[51] exhibited advanced performance, performance degradation owing to the unlabeled data for other activities that comprise half of the dataset and present high correlation to the labeled data were not investigated.

Mutual information and entropy-based methodologies quantifying information flow between multiple time series were studied for AR as alternative methods to extract features based on activation flow. Furthermore, a sensor-based method that recognizes a group of sensors as windows that are continuously activated together by their mutual information measure was proposed in [53]. In [35], the model presented in [53] was extended, and its performance was improved by altering the computation of mutual information. A causality-induced hierarchical Bayesian model was suggested to deal with interaction AR [40]. In [43], a G-causality-based framework was presented for the recognition of single-user activity. Additionally, another study proposed a human–object interaction model that integrates the causal relationship between humans and objects based on the TE for video AR [39].

### B. RECOGNITION OF UNLABELED OTHER ACTIVITIES

To improve recognition precision for realizing real-time reliable AR services, studies on the most problematic activity class of ''others'' of unlabeled sensor events are required. Some studies have discussed this problem in addition to using the corresponding data in training and not querying this class in performance analysis, which degrades the reliability of the AR architecture.

The problem of dealing with a large proportion of noncategorized data between predefined classes was investigated in [61], [62]. Moreover, patterns in the data were demonstrated by segmenting them into learnable targets, where the predefined classes denoted categories confirmed to be labeled among various human activities when generating a dataset. In [59], an architecture of feature extraction by merging the methods of natural language processing and time-series classification domains was proposed, and an unidentified class name, ''Other,'' was discussed for the Milan and Aruba datasets. In [52], the authors proposed an online application of the hierarchical hidden Markov model for detecting the current activity from live streaming of sensor events and identifying the activities that occur during another activity, called interrupted activities.

## III. METHODOLOGY

To improve the precision and sensitivity of AR and mitigate the problems of easily-confused activities and unlabeled data, a novel sensor-based data-driven AR architecture was proposed that extracts the causality and fuzzy features, as shown in Fig. 1. The proposed architecture was divided into feature extraction and AR.

In the feature extraction, fuzzy and causality features are extracted from the sensor data, and feature fusion is performed to transform these two features into one. In the extraction of the causality feature, sparsely recorded sensor activation data are restored to a uniform time series of 1 s, and the causal influence between the sensors is calculated

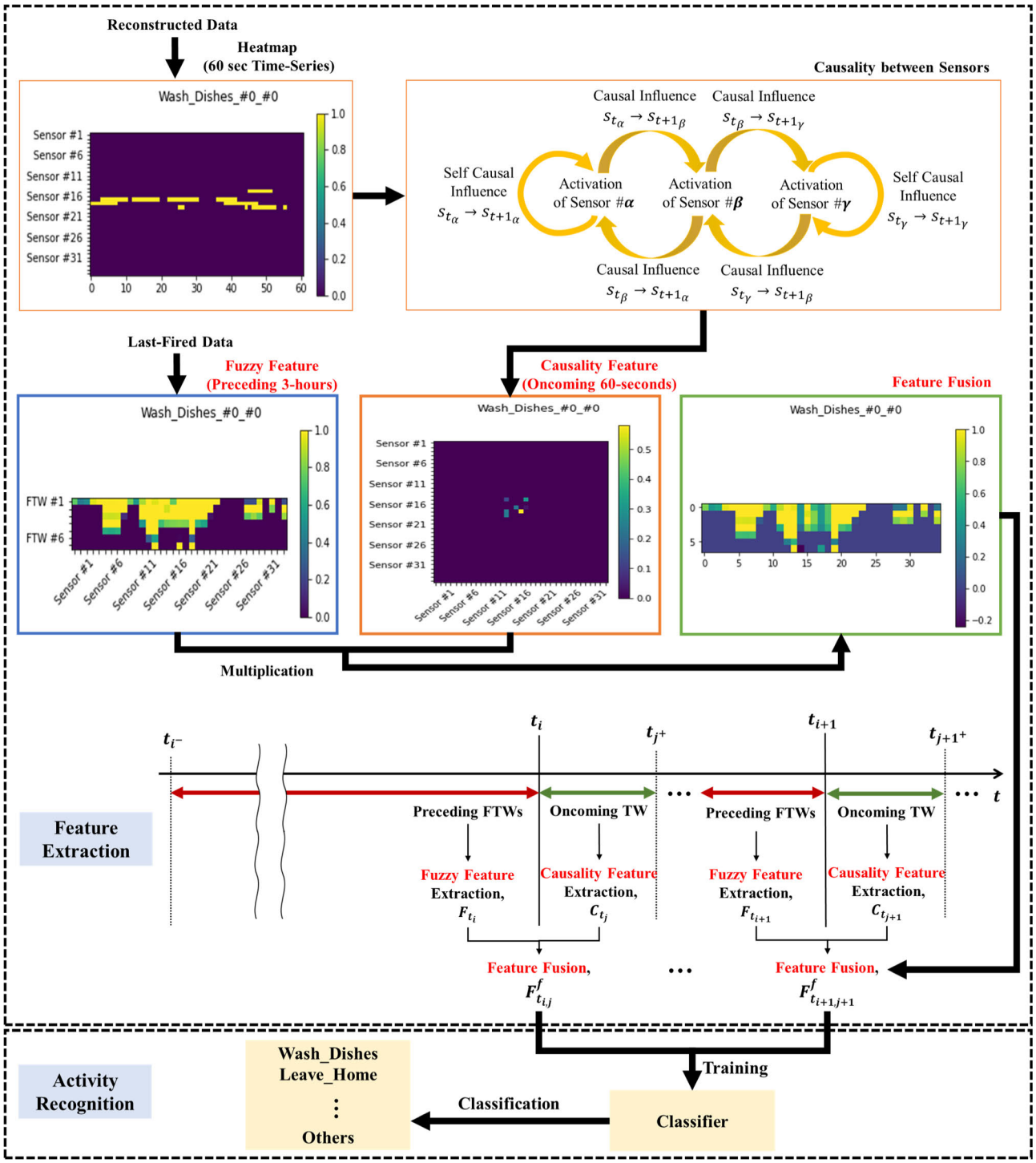
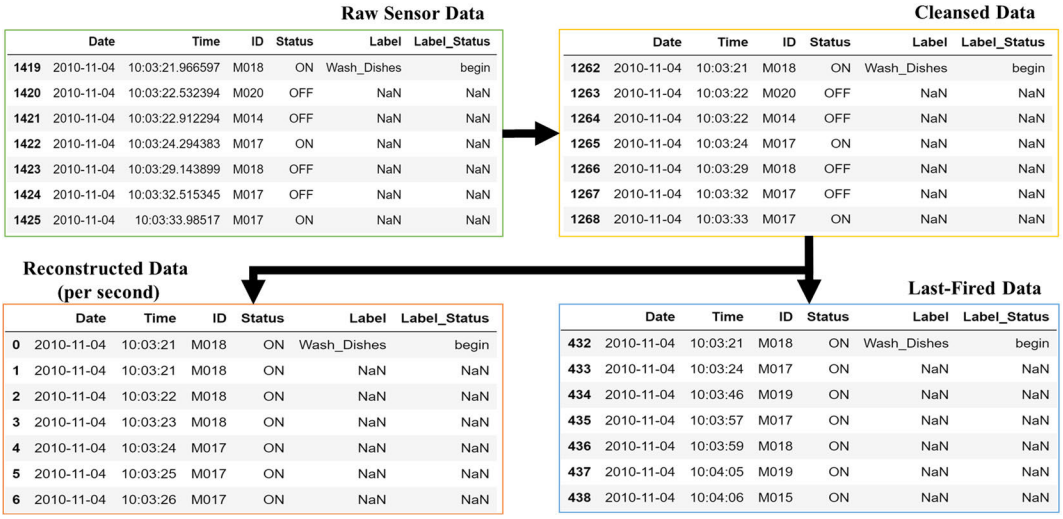



**FIGURE 1.** Flow chart of the proposed activity recognition architecture, where $j$ satisfies $t_i \leq t_j < t_{i+1}$.

using the restored data to form the causality feature matrix. In the extraction of the fuzzy feature, the FTWs composed of multiple membership functions are applied to calculate the fuzzy feature matrix for long-period preceding data.

In AR, deep neural network models designed based on LSTM, 2D-CNN, and hybrid models composed of a combination of LSTM and 2D-CNN were trained with the fused feature to effectively learn the dependencies of the feature. The trained model operates as a classifier for real-time AR, and the precision performance is evaluated using leave-one-day-out cross-validation with and without unlabeled data. In the architecture, we consider the problem of




**Raw Sensor Data**

| | Date | Time | ID | Status | Label | Label_Status |
|---|---|---|---|---|---|---|
| 1419 | 2010-11-04 | 10:03:21.966597 | M018 | ON | Wash_Dishes | begin |
| 1420 | 2010-11-04 | 10:03:22.532394 | M020 | OFF | NaN | NaN |
| 1421 | 2010-11-04 | 10:03:22.912294 | M014 | OFF | NaN | NaN |
| 1422 | 2010-11-04 | 10:03:24.294383 | M017 | ON | NaN | NaN |
| 1423 | 2010-11-04 | 10:03:29.143899 | M018 | OFF | NaN | NaN |
| 1424 | 2010-11-04 | 10:03:32.515345 | M017 | OFF | NaN | NaN |
| 1425 | 2010-11-04 | 10:03:33.98517 | M017 | ON | NaN | NaN |

**Cleansed Data**

| | Date | Time | ID | Status | Label | Label_Status |
|---|---|---|---|---|---|---|
| 1262 | 2010-11-04 | 10:03:21 | M018 | ON | Wash_Dishes | begin |
| 1263 | 2010-11-04 | 10:03:22 | M020 | OFF | NaN | NaN |
| 1264 | 2010-11-04 | 10:03:22 | M014 | OFF | NaN | NaN |
| 1265 | 2010-11-04 | 10:03:24 | M017 | ON | NaN | NaN |
| 1266 | 2010-11-04 | 10:03:29 | M018 | OFF | NaN | NaN |
| 1267 | 2010-11-04 | 10:03:32 | M017 | OFF | NaN | NaN |
| 1268 | 2010-11-04 | 10:03:33 | M017 | ON | NaN | NaN |

**Reconstructed Data (per second)**

| | Date | Time | ID | Status | Label | Label_Status |
|---|---|---|---|---|---|---|
| 0 | 2010-11-04 | 10:03:21 | M018 | ON | Wash_Dishes | begin |
| 1 | 2010-11-04 | 10:03:21 | M018 | ON | NaN | NaN |
| 2 | 2010-11-04 | 10:03:22 | M018 | ON | NaN | NaN |
| 3 | 2010-11-04 | 10:03:23 | M018 | ON | NaN | NaN |
| 4 | 2010-11-04 | 10:03:24 | M017 | ON | NaN | NaN |
| 5 | 2010-11-04 | 10:03:25 | M017 | ON | NaN | NaN |
| 6 | 2010-11-04 | 10:03:26 | M017 | ON | NaN | NaN |

**Last-Fired Data**

| | Date | Time | ID | Status | Label | Label_Status |
|---|---|---|---|---|---|---|
| 432 | 2010-11-04 | 10:03:21 | M018 | ON | Wash_Dishes | begin |
| 433 | 2010-11-04 | 10:03:24 | M017 | ON | NaN | NaN |
| 434 | 2010-11-04 | 10:03:46 | M019 | ON | NaN | NaN |
| 435 | 2010-11-04 | 10:03:57 | M017 | ON | NaN | NaN |
| 436 | 2010-11-04 | 10:03:59 | M018 | ON | NaN | NaN |
| 437 | 2010-11-04 | 10:04:05 | M019 | ON | NaN | NaN |
| 438 | 2010-11-04 | 10:04:06 | M015 | ON | NaN | NaN |

**FIGURE 2.** Data preprocessing for feature extraction.

dealing with unlabeled "Others" data, which comprise half of the dataset. The unlabeled data are correlated with the labeled data, which confuses the classifier and degrades its classification performance. To alleviate performance degradation problem, we introduce a method of extracting NDTE-based causality feature extraction along with the conventional FTW-based fuzzy feature extraction to express unique patterns of activities.

### A. DATA PREPROCESSING

For binary sensor-based data, three open datasets of Aruba, Cairo, and Milan [56] were utilized. These datasets capture daily activities, such as cooking, sleeping, and leaving home, using sensors installed in smart homes.

Before performing the feature extraction, 1,719,558, 433,665, and 726,534 binary sensor activation events were preprocessed from the raw datasets of Aruba, Cairo, and Milan, respectively, as depicted in Fig. 2. For each dataset, the distribution of the entire target class, including the other classes (unlabeled data), is depicted in Fig. 3. The raw sensor data in this data processing were cleansed by correcting some corrupted values (e.g., "OF" to "OFF"), mapping the time form to "hh:mm:ss," removing the temperature sensor data, and performing a single labeling operation on scarce multi-labeled data. To extract and train the other class data, unlabeled sensor events between the end of a labeled activity and beginning of the next activity were labeled with the target name, "Others," in the data cleansing. Subsequently, two data transformations were considered: last-fired and reconstructed data, which are described as follows:

*Last-fired data*: This data refer to the last implemented sensor. The sensor that last changed the state remains "ON," and it changes to "OFF" when another sensor changes its state. Last-fired data are transformed from the cleansed data.

*Reconstructed data*: These data refer to the reconstruction of the "ON" or "OFF" status of each sensor in increments of 1 s and conversion into a uniformly sampled time series instead of irregular sensor events. The reconstructed data are transformed from the cleansed data.

### B. FEATURE EXTRACTION

The main purpose of this study is to use the causality feature and the fuzzy feature simultaneously by fusing the two features for real-time AR because the advantages of the two features are different. We propose an evolved feature extraction approach based on NDTE and FTWs to take advantage of both the causality feature and the fuzzy feature as follows:

- Causality feature: It is used for single activity and has the advantage in that it contains meta-action information by calculating direct causal influence between sensors to express unique orders and characteristics of an activity.
- Fuzzy feature: It is for multiple activities and has the advantage in that it contains historical activity information by compressing long-term sensor event data; however, there should be a loss of information in extracting short-term and specific features of each activity.

This subsection describes the extraction of the two types of features: i) NDTE-based causality features and ii) FTW-based fuzzy features. To mine the meta-activity information during a predefined oncoming time interval (e.g., 60 s), the causality feature matrix is generated by calculating the NDTE [46] of the regular time series of each sensor from the reconstructed data. While considering the preceding time interval (e.g., 3 h), a large amount of sensor activation data is expressed as one fuzzy feature matrix based on the number of sensors and the FTWs; subsequently, a training set is prepared by combining the causality feature matrix with the fuzzy matrix.

A set of binary sensors is represented by $S = \{s_1, s_2, \ldots s_\alpha, \ldots, s_{|S|}\}$, where $|S|$ is the total number of sensors. The ON–OFF status of sensor event index $i \in \{1, 2, \ldots, I\}$ from the last-fired data and $j \in \{1, 2, \ldots, J\}$

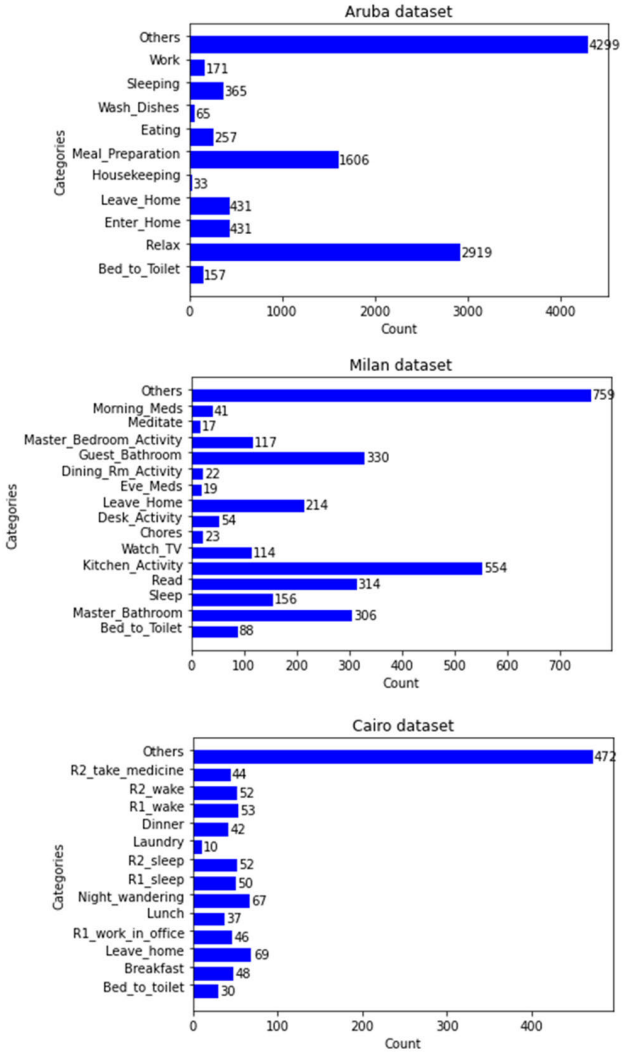



**FIGURE 3.** Human AR dataset of CASAS [56]; activity distributions of Aruba (top), Milan (middle), and Cairo (bottom) datasets and related targets.

from the cleansed data ($I < J$) can be expressed as follows:

$$S_i = \left\{s_{i_1}, s_{i_2}, \ldots s_{i_\alpha}, \ldots, s_{i_{|s|}}\right\}, \quad s_{i_\alpha} = \{0, 1\}, \tag{1}$$

$$S_j = \left\{s_{j_1}, s_{j_2}, \ldots s_{j_\alpha}, \ldots, s_{j_{|s|}}\right\}, \quad s_{j_\alpha} = \{0, 1\}, \tag{2}$$

where $s_{i_\alpha}$ and $s_{j_\alpha}$ are the activation statuses of irregular and regular sensor events $i$ and $j$ of the $\alpha^{th}$ sensor, respectively. It should be noted that in this study, observations are the sequence of sensor events, and features are extracted for each observation.

### 1) CAUSALITY FEATURE MATRIX

In an environment in which various sensors are installed, causality is the process of extracting the cause-and-effect information to express the unique patterns of resident activity for a predefined time interval (e.g., 60 s). As an asymmetric matrix with directional and dynamic information, the causality, $c_{\alpha\beta} = NDTE_{\alpha\to\beta} \in C$, denotes the causal relationship between a pair of sensors $(\alpha, \beta) \in \{1, 2, \ldots, |S|\}$, where $C \in \mathcal{R}^{|S|\times|S|}$ denotes the causality matrix. The quantitative causal influence, $c_{\alpha\beta} \neq c_{\beta\alpha}$, can be calculated using certain patterns of the joint probability of activation between sensors.

The causality from variable $s_{t_\alpha}$ to variable $s_{t_\beta}$ is composed of direct and indirect paths, owing to an intermediate variable. For example, $s_{t_\gamma}$, where variable sets $\left\{s_{t_\alpha}, s_{t_\beta}, s_{t_\gamma}\right\} \in S_t = \left\{s_{t_1}, s_{t_2}, \ldots s_{t_\alpha}, \ldots, s_{t_{|s|}}\right\}$ is a time series reconstructed from the sensor event set, $\left\{s_{j_\alpha}, s_{j_\beta}, s_{j_\gamma}\right\} \in S_j$, with time stamp $t$ in units of seconds from the *reconstructed data*, $\gamma \in \{1, 2, \ldots, |S|\}$, $\alpha \neq \gamma$ and $\beta \neq \gamma$. In [46], a TE-based methodology was used to detect and differentiate between direct and indirect causalities for process variables. In this study, only direct causality is used. For a binary sensor-based dataset for AR, direct causality can be considered as pure causality (or the information of the meta-action) in which the causal influence of the intermediate sensors is removed. Schreiber proposed a TE-based measure of causality to compute the deviation from the following generalized Markov condition [44]:

$$p\left(s_{t+1_\alpha}|s^n_{t_\alpha}, s^m_{t_\beta}\right) = p\left(s_{t+1_\alpha}|s^n_{t_\alpha}\right), \tag{3}$$

where $s^m_{t_\alpha} = \left(s_{t_\alpha}, \ldots, s_{t-m-1_\alpha}\right)$, $s^n_{t_\beta} = \left(s_{t_\beta}, \ldots, s_{t-n-1_\beta}\right)$, $m$ and $n$ represent the orders (memory) of Markov processes $s_{t_\alpha}$ and $s_{t_\beta}$, respectively, and $p\,(|)$ represents the transitional probability. The TE from $s_{t_\alpha}$ and $s_{t_\beta}$ [45] is formulated as

$$\begin{aligned} &TE_{\alpha\to\beta} \\ &= H\left(s_{t+1_\beta}|s^n_{t_\beta}\right) - H\left(s_{t+1_\beta}|s^n_{t_\beta}, s^m_{t_\alpha}\right) \\ &= \sum_{s_{t+1_\beta}, s^n_{t_\beta}, s^m_{t_\alpha}} p\left(s_{t+1_\beta}, s^n_{t_\beta}, s^m_{t_\alpha}\right) \log \frac{p\left(s_{t+1_\beta}|s^n_{t_\beta}, s^m_{t_\alpha}\right)}{p\left(s_{t+1_\beta}|s^n_{t_\beta}\right)}, \end{aligned} \tag{4}$$

where $H\left(s_{t+1_\beta}|s^n_{t_\beta}\right)$ denotes the conditional Shannon entropy of $s_{t_\beta}$ at time stamp $t+1$, provided its value at time $t$ is expressed as

$$\boldsymbol{H}(\boldsymbol{s_{t+1_\beta}}|\boldsymbol{s^n_{t_\beta}}) = -\sum_{\boldsymbol{s_{t+1_\beta}}, \boldsymbol{s^n_{t_\beta}}} \boldsymbol{p}(\boldsymbol{s_{t+1_\beta}}|\boldsymbol{s^n_{t_\beta}}) \log \frac{\boldsymbol{p}(\boldsymbol{s_{t+1_\beta}}|\boldsymbol{s^n_{t_\beta}})}{\boldsymbol{p}(\boldsymbol{s^n_{t_\beta}})}. \tag{5}$$

$TE_{\alpha\to\beta}$ can be regarded as the information about future observations $s_{t+1_\beta}$ gained from past observations of $s^n_{t_\beta}$, and it is subtracted from the information gained from future observations $s_{t+1_\beta}$ when the joint probability, $\left(s^n_{t_\beta}, s^m_{t_\alpha}\right)$, is known. If $s^n_{t_k}$ is independent of $s^m_{t_\alpha}$, then the second term in (3) is zero. The TE can be in the range of $0 \leq TE_{\alpha\to\beta} < \infty$. In [46], the DTE was calculated for variables $s_{t_\alpha}$, $s_{t_\beta}$, and $s_{t_\gamma}$, where $s^o_{t_\gamma} = \left(s_{t_\gamma}, \ldots, s_{t-o-1_\gamma}\right)$ and $o$ represent the orders.

$$\begin{aligned} &DTE_{\alpha\to\beta} \\ &= TE_{\alpha\to\beta|\gamma} \\ &= H\left(s_{t+1_\beta}|s^n_{t_\beta}, s^o_{t_\gamma}\right) - H\left(s_{t+1_\beta}|s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right) \end{aligned}$$

$$= \sum_{\substack{s_{t+1_\beta}, s^n_{t_\beta}, \\ s^o_{t_\gamma}, s^m_{t_\alpha}}} p\left(s_{t+1_\beta}, s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right) \log \frac{p\left(s_{t+1_\beta} | s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right)}{p\left(s_{t+1_\beta} | s^n_{t_\beta}, s^o_{t_\gamma}\right)}, \tag{6}$$

where

$$\frac{p\left(s_{t+1_\beta} | s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right)}{p\left(s_{t+1_\beta} | s^n_{t_\beta}, s^o_{t_\gamma}\right)} = \frac{p\left(s_{t+1_\beta}, s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right) p\left(s^n_{t_\beta}, s^o_{t_\gamma}\right)}{p\left(s_{t+1_\beta}, s^n_{t_\beta}, s^o_{t_\gamma}\right) p\left(s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right)}. \tag{7}$$

Equation (7) is an example of a DTE with an intermediate variable, $s_{t_\gamma}$. In the case of the Aruba dataset, a total of 35 sensors and 33 sensors exist, excluding the target sensor pair, which can be a set of intermediate sensors every time the DTE is calculated for all possible pairs of sensors. Furthermore, the self-DTE is calculated to obtain the causal influence between themselves, where the target sensor pair consists of the same sensors, as follows:

$$DTE_{\alpha \to \beta} = \begin{cases} DTE_{\alpha \to \beta} & \text{if } \alpha \neq \beta \\ Self - DTE_{\alpha \to \alpha} & \text{if } \alpha = \beta, \end{cases} \tag{8}$$

where

$$\begin{aligned} &Self\, DTE_{\alpha \to \alpha} \\ &= TE_{\alpha \to \alpha | \beta, \gamma} \\ &= H\left(s_{t+1_\alpha} | s^n_{t_\beta}, s^o_{t_\gamma}\right) - H\left(s_{t+1_\alpha} | s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right) \\ &= \sum_{\substack{s_{t+1_\alpha}, s^n_{t_\beta}, \\ s^o_{t_\gamma}, s^m_{t_\alpha}}} p\left(s_{t+1_\alpha}, s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right) \log \frac{p\left(s_{t+1_\alpha} | s^n_{t_\beta}, s^o_{t_\gamma}, s^m_{t_\alpha}\right)}{p\left(s_{t+1_\alpha} | s^n_{t_\beta}, s^o_{t_\gamma}\right)}. \end{aligned} \tag{9}$$

DTE matrices can contain large amounts of noise and bias-generating bias entropy. Therefore, to improve the recognition precision, this bias should be removed from the DTE separately estimate from the process of filtering out direct causality. Bias entropy can be removed by subtracting the average $DTE_{\alpha \to \beta}$ using the shuffled version of $s_{t_\alpha}$ denoted by $\langle DTE_{\alpha_{shuffle} \to \beta} \rangle$, over several shuffles. $\alpha_{shuffle}$ denotes the randomly rearranged $s_{t_\alpha}$ in a shuffled order to deteriorate time sensitivity, and the NDTE from $s_{t_\alpha}$ to $s_{t_\beta}$ can be calculated as follows [46]:

$$NDTE_{\alpha \to \beta} = \begin{cases} NDTE_{\alpha \to \beta} & \text{if } \alpha \neq \beta \\ Self - NDTE_{\alpha \to \alpha} & \text{if } \alpha = \beta, \end{cases} \tag{10}$$

where

$$\begin{aligned} &NDTE_{\alpha \to \beta} \\ &= \frac{DTE_{\alpha \to \beta} - \langle DTE_{\alpha_{shuffle} \to \beta} \rangle}{H\left(s_{t+1_\beta} | s_{t_\beta}\right)}, \end{aligned} \tag{11}$$

$$\begin{aligned} &SelfN - DTE_{\alpha \to \alpha} \\ &= \frac{Self - DTE_{\alpha \to \alpha} - \langle Self - DTE_{\alpha_{shuffle} \to \alpha} \rangle}{H\left(s_{t+1_\alpha}\right)}. \end{aligned} \tag{12}$$

The NDTE is in the range of $0 \leq NDTE_{\alpha \to \beta} \leq 1$. It is zero when $s_{t_\alpha}$ transfers no information to $s_{t_\beta}$, and is one when $s_{t_\alpha}$ transfers the maximum information to $s_{t_\beta}$. In this study, self-NDTE was first derived and applied. It is important to obtain self-NDTE because it prevents the diagonal component in the causality matrix from being null. If the self-NDTE is not calculated, feature information loss occurs during the fusion of causality and fuzzy features.

In this study, the causality matrix, $C_{t_j}\left(S_t, \Delta t_{j^+}\right) \in \mathcal{R}^{|S| \times |S|}$, is extracted for the set of reconstructed time series, $S_t$, for the predefined temporal window (or interval), $\Delta t_{j^+} = \left[t_j, t_{j^+}\right]$, where $t_j$ denotes the time at which event $j$ occurs, and $t_{j^+}$ denotes the oncoming point of time from $t_j (t_j < t_{j^+})$. An element of the causality matrix, $c_{\alpha\beta} \in C_{t_j}\left(S_t, \Delta t_{j^+}\right)$, can be calculated as follows:

$$\begin{aligned} c_{\alpha\beta} &= NDTE_{\alpha \to \beta}\left(S_t, \Delta t_{j^+}\right) \\ &= NDTE_{\alpha \to \beta}\left(S_t\right), \quad \forall t \in \Delta t_{j^+}. \end{aligned} \tag{13}$$

#### 2) FUZZY FEATURE MATRIX

This subsection describes the extraction of the fuzzy features of long-period preceding data using multiple and incremental FTWs [2], [47]. Each FTW $T_k$ is defined as a set of fuzzy membership functions whose shape corresponds to the trapezoidal function, $T_k(x)[l_1, l_2, l_3, l_4]$, as expressed in (14).

$$T_k(x)[l_1, l_2, l_3, l_4] = \begin{cases} 0 & x \leq l_1 \\ (x - l_1)/(l_2 - l_1) & l_1 < x \leq l_2 \\ 1 & l_2 < x \leq l_3 \\ (l_4 - x)/(l_4 - l_3) & l_3 < x \leq l_4 \\ 0 & l_4 < x. \end{cases} \tag{14}$$

The well-known trapezoidal membership functions are defined by a lower limit $l_1$, an upper limit $l_4$, a lower support limit $l_2$, and an upper support limit $l_3$. The values of $l_1$, $l_2$, $l_3$, and $l_4$ are defined based on the Fibonacci sequence. To extract features, the FTWs slide over the temporal window, $\Delta t_{i^-} = [t_{i^-}, t_i]$, between a particular time $t_i$ and a preceding point of time $t_{i^-}$, where $i^*$ denotes the corresponding provided event index for $t_{i^*}$. Once the FTW $T_k$ is defined, the activation degree of a binary activation $s_{i_\alpha}$ from the sensor activation $S_i$ at the evaluated time $t_{i^*}$ can be computed using (15).

$$T_k\left(s_{i_\alpha}, \Delta t_{i^-}\right) = \begin{cases} \max(T_k(s_{i_\alpha})), & \forall t \in \Delta t_{i^-} \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

Subsequently, the fuzzy feature $F_{t_i}\left(s_{i_\alpha}, \Delta t_{i^-}\right) \in \mathcal{R}^{k \times |S|}$, can be extracted and aggregated by calculating $f_{k\alpha} = T_k\left(s_{i_\alpha}, \Delta t_{i^-}\right) \in F$, applying $k$ FTWs for $|S|$ binary sensors based on the last-fired data. In Fig. 4, some examples of the activation degree from FTW $T_k$ and different binary activations $s_{i_\alpha}$ are presented.

### C. WINDOWING AND FEATURE FUSION

The most common windowing techniques include the TW and sensor event windowing (SEW) [63]. Each approach

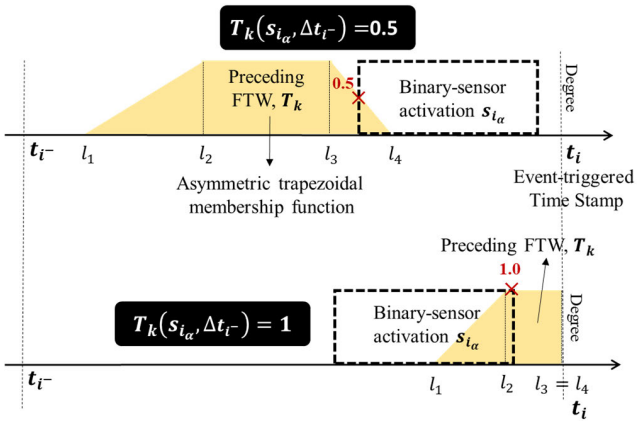



**FIGURE 4.** Activation degrees of FTW $T_k$ for different binary activations $s_{i_\alpha}$ and different limit sets $\{l_1, l_2, l_3, l_4\}$ by partial overlapping between $T_k$ and $s_{i_\alpha}$.

**Algorithm 1** Feature Extraction Based on Causality and FTWs

**Input:** cleansed data $S_j$, last-fired data $S_i$, sensors $S = \{s_1, s_2, \ldots s_\alpha, \ldots, s_{|S|}\}$, preceding window $\Delta t_{i^-} = [t_{i^-}, t_i]$, oncoming window $\Delta t_{j^+} = [t_j, t_{j^+}]$, membership function $(k, l_1, l_2, l_3, l_4)$, $t_i$ = time of event $i$.
**Result:** fused feature $F^f_{t_{i,j}}$.
**For** $S_j$, event index $j \in \{1, 2, \ldots, J\}$ **do**
  **Reconstruct** uniform time series $S_t$ from $S_j$
  **For** $s_\alpha, s_\beta \in S$ **do**
    **Calculate** causality feature matrix $C_{t_j}(S_t, \Delta t_{j^+})$ using (13)
  **end for**
**end for**
**For** $S_i$, event index $i \in \{1, 2, \ldots, I\}$ **do**
  **For** $s_\alpha \in S$ **do**
    **Calculate** fuzzy feature matrix $F_{t_i}(s_{i_\alpha}, \Delta t_{i^-})$ using (15)
  **end for**
**end for**
**For** $i \in \{1, 2, \ldots, I\}$ **do**
  **For** $j \in \{1, 2, \ldots, J\}$ **do**
    **If** $t_i \le t_j < t_{i+1}$ **do**
**Calculate**fused feature $F^f_{t_{i,j}}$ using (16)
    **end if**
  **end for**
**end for**

has its own advantages and disadvantages. TW breaks a dataset into equal periods, and is preferred for selecting the optimal period for streamed data because of the availability of extremely few or numerous features for representing an activity. SEW splits a dataset into equal segments of fired sensor events and has the advantage that the varying time windows owing to the fixed number of sensor events can appropriately capture the context. This is achievable because a wide time range is necessary to acquire the context in the case of sleeping, whereas a narrow time range is sufficient for frequent events, such as housekeeping.

In this study, a sensor event-triggered time window (STW) combining TW and SEW was applied. STW is a windowing technique in which the start time is set as the time an event (index) occurs, and the end time is set by adding (or subtracting) a predefined time interval to (from) the start time. When the start and end times are defined, the events occurring within that period are included in the window and used for feature extraction. Therefore, the number of SEW events may vary, and features are extracted for every event index to form a training dataset.

As depicted in the second block of Fig. 1, the fuzzy feature $F_{t_i}$ is extracted from an STW with a start time $t_i$ defined as the time at which the last-fired data-based event index $i$ occurred. As described in Subsection IIIB, only preceding sensor events up to 3 h from the start time were utilized to extract the fuzzy features. Moreover, the fuzzy membership function was subdivided into $k$ functions within the STW. Furthermore, to extract key patterns of the behavior that occur over a short period, the causality $C_{t_j}$ is extracted from the STW with the start time $t_j$ and a time interval of 60 s for the oncoming sensor events, where event index $j$ satisfies $t_i \le t_j < t_{i+1}$. Therefore, index $i$ can have multiple indices, $j$.

If the fuzzy features are extracted based on the cleansed data-based sensor event, $j$, several completely identical feature vectors that do not contain the behavioral pattern changes can be obtained. This is because of the FTWs-based feature extraction, which generates one feature matrix for the long-term TW for several hours. Moreover, it is difficult to contain relevant context information while extracting fuzzy features in an extremely short time interval (or low time resolution). However, it is important to extract the fuzzy feature for the event index of the last-fired data that expresses only the data in which the sensor status has changed clearly by describing the *last-fired data*. Concurrently, the causality expresses the changed behavior pattern for every event $j$, and it cannot have a smaller resolution than event $j$. In particular, the behavior pattern is extracted with maximum detail in both cases of fuzzy features and causality. The fused feature $F^f_{t_{i,j}}$ can be expressed as follows:

$$F^f_{t_{i,j}} = F_{t_i} \cdot \left(I_{|S|} - C_{t_j}\right), \quad t_i \le t_j < t_{i+1}, \tag{16}$$

where $I_{|S|}$ denotes an $|S| \times |S|$ identity matrix and "$\cdot$" represents multiplication of two matrices using the dot product. The pseudo-code algorithm for the proposed feature extraction method is described in Algorithm I.

The method of fusing two feature matrices is typically divided into physical fusion based on concatenation and chemical fusion via matrix multiplication. In this study, the above fusion was achieved experimentally via matrix multiplication, as expressed in (16).

In Fig. 5, we extracted the fuzzy and causality features for the preceding 3 h and oncoming 60 s, respectively, from each sensor event time of the Aruba dataset. The size of the

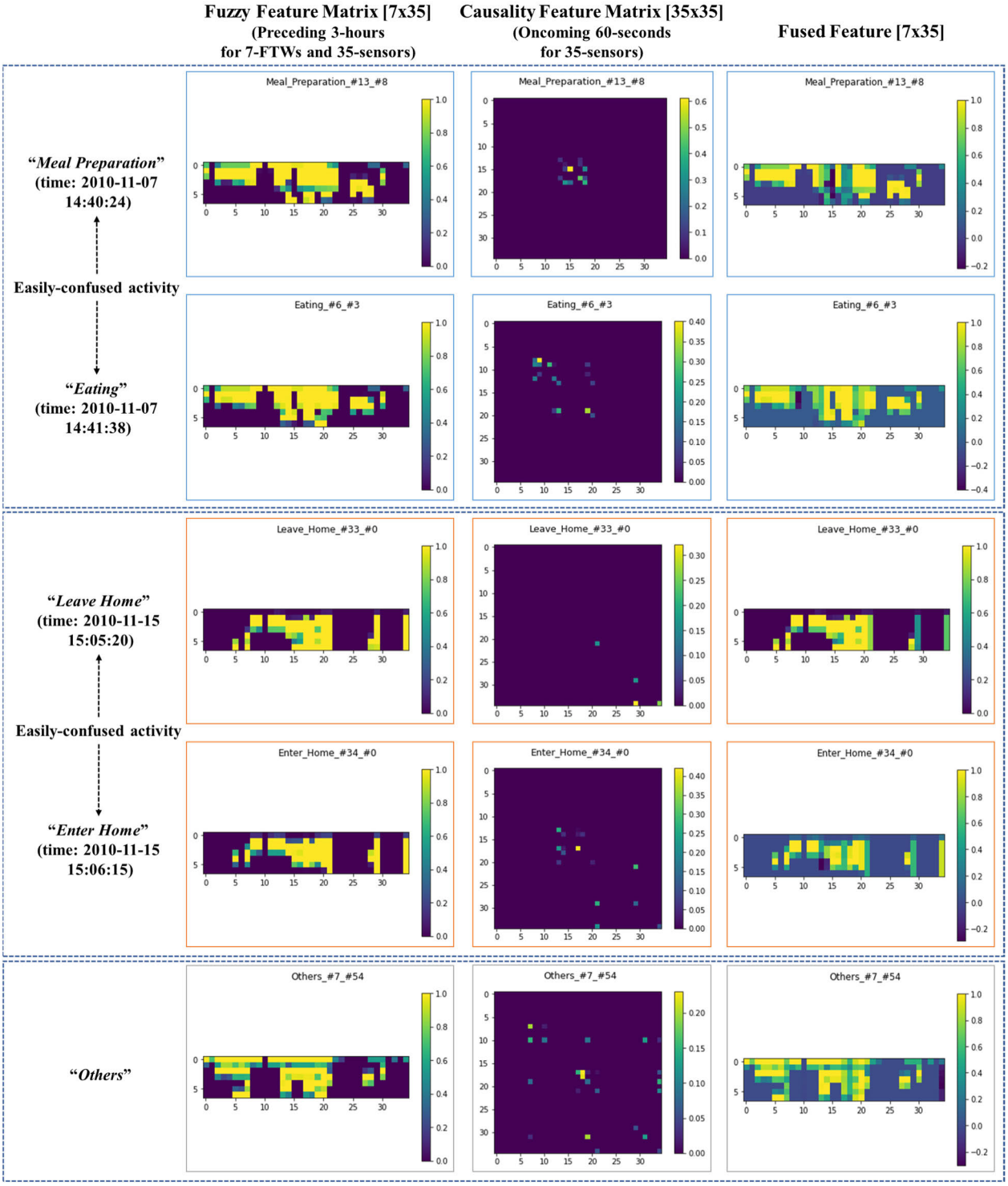



**FIGURE 5.** Example of extraction of fuzzy, causality, and fused features for easily-confused activities and "Others" class for Aruba dataset.

fuzzy and causality feature matrices becomes 7 × 35 and 35 × 35, respectively, when the 7-FTWs listed in Table 1 and 35-sensors are applied. The fused feature is generated using matrix multiplication of the fuzzy and causality features using (16), to combine the strengths of the two features.

The fuzzy features of the easily-confused activity pairs in Fig. 5 (Meal_Preparation, Eating) and (Leave_Home, Enter_Home), have similar patterns that confuse the classifier in classifying activities and cause misclassifications, which can be confirmed in the confusion matrix of Fig. 7 in Section IV. The confusion is inevitable because the event times of easily-confused activities are very close to each other, and the time interval for the FTW-based extraction is extremely long at 3 h, that is, the time period of the data to

**TABLE 1.** Shape of trapezoidal functions (min).

| $k$ | $l_1$ | $l_2$ | $l_3$ | $l_4$ |
|---|---|---|---|---|
| 1 | 180 | 60 | 30 | 15 |
| 2 | 60 | 30 | 15 | 5 |
| 3 | 30 | 15 | 5 | 3 |
| 4 | 15 | 5 | 3 | 2 |
| 5 | 5 | 3 | 2 | 1 |
| 6 | 3 | 2 | 1 | 0 |
| 7 | 2 | 1 | 0 | 0 |

extract the fuzzy feature approximately coincides with the two activities.

Moreover, the causality features of the easily-confused activities exhibit a distinct aspect and pattern of the activities, as shown in Fig. 5. The reason is because not only the feature extraction method differs from the FTWs but also overlapping time intervals are not considered as a problem. Fused features have the same matrix size as fuzzy features; however, their aspects differ from fuzzy features through the fusion of causality features.

The ''Others'' unlabeled data that occupy approximately half of the entire dataset cause correlation with other activities, which degrades the precision performance of the classifier. To mitigate this degradation and recognize ''Others'' as ''Others'' as much as possible, we generated the fused features of ''Others,'' and used them for model training.

### D. MODEL ARCHITECTURE

In the previous section, we extracted the 2D matrix of the fused features for each activity from the sensor data. In this subsection, we focus on effectively learning the dependencies of the fused features to classify the activities. Because the fused feature is generated through matrix multiplication of the fuzzy and causality features, it has both the temporal dependency of the fuzzy feature and spatial dependency of the causality feature.

Here, we consider that there is a temporal dependency between the elements of the fuzzy feature matrix because the fuzzy feature is calculated based on the FTWs having a temporal relationship between the windows. As for the causality feature, it can be seen that only the spatial dependency, in which the elements are arranged and located in the matrix, is provided in the causality feature because the elements of the causality feature matrix are independently calculated based on the NDTE.

To effectively learn the spatiotemporal dependencies of the fused feature, LSTM, 2D-CNN, and hybrid models composed of a combination of LSTM and CNN were used as classifiers in this study. LSTM is a type of recurrent neural network that includes memory to learn the temporal dependencies from sequences of observations over time. In the experiments, the LSTM model was designed by stacking three LSTM layers followed by a fully connected dense layer and a soft-max layer. For all models in this study, a learning rate of 0.001, batch size of 100, and 20 training epochs were used in the experiments. As a regularization technique, a dropout rate of 40% was applied to the LSTM network to prevent overfitting.

Furthermore, to learn the spatial dependencies of the fused feature, a 2D-CNN model was designed. CNN, which is a promising solution for image classification, speech recognition, and text analysis, has two advantages for HAR in learning local dependencies and scale invariance. In this study, a 2D-CNN architecture was adopted to extract and learn local 2D subsequences from sequence data. The model was designed by stacking three convolutional layers with dropout and learning rates of 40% and 0.001, respectively, followed by a max-pooling layer. This was further followed by three fully connected dense layers with a dropout rate of 40% and a soft-max layer.

The hybrid 2D-CNN–LSTM model learns spatiotemporal dependencies and is also employed to induce a local sensitivity-based 2D-CNN as a preprocessing step before LSTM. In this model, 2D-CNN is important for recognizing the spatial patterns, and the output of 2D-CNN becomes the input of LSTM to learn the temporal dependencies of the fused feature. The model was designed by stacking three convolutional layers with a dropout rate of 40%, followed by a max-pooling layer. This is followed by three LSTM layers with a dropout and learning rates of 40% and 0.001, respectively, followed by a fully connected dense layer and a soft-max layer.

## IV. EXPERIMENTS

### A. DATA DESCRIPTION

In this study, the most useful binary sensor open datasets, including Aruba, Milan, and Cairo from the CASAS [56], were used to identify human activities. Aruba is a single-occupancy dataset in which a volunteer woman interacts with 31 motion sensors, four door sensors, and five temperature sensors. In the Aruba dataset, ten annotated activity categories and the ''Others'' class as unlabeled sensor events were included in the model training, as depicted in Fig. 3. Milan is a dataset based on a volunteer woman with a dog, 28 motion sensors, and three door sensors with 16 categories, including the ''Others'' class. Cairo is a dataset of a volunteer adult couple and a dog. The 13 annotated categories and the ''Others'' class sensor events were generated from 27 motion sensors and five temperature sensors. All datasets contain five attributes (date, time, sensor ID, status, label, and label status), and the events from the temperature sensors were excluded while extracting the features and training the models. The unlabeled sensor events between the end of a labeled activity and beginning of the next activity are labeled as ''Others.''

### B. FEATURE VECTOR

The fuzzy features are computed by applying $k = 7$ FTWs with trapezoidal functions, $T_k(x)\,[l_1, l_2, l_3, l_4]$, for each sensor using (14). We set the shapes of the membership functions listed in Table 1 in this experiment because a larger number

| | Meal_Preparation | Wash_Dishes | Bed_to_Toilet | Eating | Enter_Home | Leave_Home | Relax | Work | Sleeping | Housekeeping |
|---|---|---|---|---|---|---|---|---|---|---|
| **Meal_Preparation** | 0.37492 | 0.00000 | 0.00000 | 0.12287 | 0.0 | 0.03466 | 0.44991 | 0.01575 | 0.00000 | 0.00189 |
| **Wash_Dishes** | 0.06154 | 0.01538 | 0.00000 | 0.01538 | 0.0 | 0.09231 | 0.69231 | 0.07692 | 0.00000 | 0.04615 |
| **Bed_to_Toilet** | 0.00637 | 0.00000 | 0.01274 | 0.00000 | 0.0 | 0.00000 | 0.00000 | 0.00000 | 0.98089 | 0.00000 |
| **Eating** | 0.14400 | 0.14800 | 0.00000 | 0.12000 | 0.0 | 0.18800 | 0.34400 | 0.04400 | 0.00800 | 0.00400 |
| **Enter_Home** | 0.37587 | 0.00000 | 0.00000 | 0.01392 | 0.0 | 0.23898 | 0.31555 | 0.04640 | 0.00464 | 0.00464 |
| **Leave_Home** | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 1.0 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| **Relax** | 0.18395 | 0.00861 | 0.00000 | 0.00551 | 0.0 | 0.06235 | 0.63314 | 0.02377 | 0.07613 | 0.00654 |
| **Work** | 0.23977 | 0.01170 | 0.00000 | 0.01170 | 0.0 | 0.15789 | 0.30994 | 0.22222 | 0.02339 | 0.02339 |
| **Sleeping** | 0.52500 | 0.00000 | 0.38750 | 0.00000 | 0.0 | 0.00250 | 0.04500 | 0.00000 | 0.04000 | 0.00000 |
| **Housekeeping** | 0.12500 | 0.00000 | 0.00000 | 0.00000 | 0.0 | 0.34375 | 0.43750 | 0.09375 | 0.00000 | 0.00000 |

**FIGURE 6.** Probability of activity transition for Aruba dataset; rows and columns are before and after transition, respectively.

**TABLE 2.** Macro-F1-scores for CASAS Aruba, Milan, and Cairo datasets with leave-one-day-out cross-validation.

| Dataset | Feature | Target | Deep LSTM | Deep 2D-CNN | Deep 2D-CNN–LSTM |
|---|---|---|---|---|---|
| Aruba | FTW only | Without Others | 0.823 | 0.839 | 0.851 |
| | Causality + FTW | Without Others | **0.903** | **0.908** | **0.919** |
| | FTW only | With Others | 0.610 | 0.591 | 0.634 |
| | Causality + FTW | With Others | **0.700** | **0.736** | **0.738** |
| Milan | FTW only | Without Others | 0.483 | 0.468 | 0.506 |
| | Causality + FTW | Without Others | **0.678** | **0.734** | **0.726** |
| | FTW only | With Others | 0.398 | 0.400 | 0.416 |
| | Causality + FTW | With Others | **0.552** | **0.571** | **0.572** |
| Cairo | FTW only | Without Others | 0.794 | 0.839 | 0.822 |
| | Causality + FTW | Without Others | **0.871** | **0.923** | **0.911** |
| | FTW only | With Others | 0.531 | 0.547 | 0.628 |
| | Causality + FTW | With Others | **0.646** | **0.628** | **0.744** |

of windows and a wider time interval did not improve the AR performance. The causality features were extracted as a core feature to improve AR performance and the FTWs-based fuzzy features were utilized as the base features of this experiment to see the improvement of the AR performance. Because the current activity has a pattern that is linked or derived from a past activity, as depicted in Fig. 6 (for the Aruba dataset), it is excellent to use the past information of several hours to perform AR. To extract the features of the oncoming data, causality was calculated for all possible sensor pairs using (12). The memory orders were set as $m = n = o = 58$, and the time windows for oncoming data $\Delta t_{j^+}$ and preceding data $\Delta t_{i^-}$ were set as 60 s and 180 min, respectively. The fused features as the training dataset were calculated using (16), and labeled based on the label information of the corresponding raw datasets. The resulting datasets were used for the real-time AR.

### C. PERFORMANCE MEASURES

The performance evaluation method commonly used in classification problems is utilized as follows:

*Accuracy* represents the proportion of all cases that is correctly classified.

$Precision = \frac{TP}{TP+FP}$ shows the proportion of cases correctly predicted to belong to a class, where true positive (TP) denotes the number of samples correctly classified as positive and false positive (FP) denotes the number of samples incorrectly classified as positive.

$Recall = \frac{TP}{TP+FN}$ represents the proportion of correctly classified cases of a class, where false negative (FN) denotes the number of samples incorrectly classified as negative.

$F1 - score = \frac{2}{1/Sens+1/Prec}$ yields the weighted average of precision and sensitivity.

### D. RESULTS

In this section, the results of the different models and datasets are discussed based on the different feature extraction approaches. The main objective of the study was to verify the effectiveness of using the causality feature in an environment in which the ''Others'' class is considered. To examine the capability of the deep LSTM, deep 2D-CNN, and deep 2D-CNN–LSTM models in predicting new data that were not used in estimating it, leave-one-day-out cross-validation was applied. In this validation method, one day of the activities was used for the test, and the remaining days were used for the training set.

#### 1) COMPARISON OF THE EXTRACTION METHODS

Table 2 lists the macro-F1-scores for the three CASAS datasets to compare the performance of the conventional method, which uses only FTWs, and the proposed feature fusion-based method. The latter uses both the causality and

**TABLE 3.** Precision, recall, and F1-score for Aruba dataset with leave-one-day-out cross-validation.

| Activity | Feature | Target | Precision | | | Recall | | | F1-score | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM |
| Bed_to_Toilet | FTW only | Without Others | 0.97 | 0.94 | 0.94 | 0.94 | 0.95 | 0.97 | 0.96 | 0.95 | 0.96 |
| | Causality-applied | Without Others | 0.92 | 0.97 | 0.88 | 0.83 | 0.77 | 0.88 | 0.87 | 0.86 | 0.88 |
| | FTW only | With Others | 0.76 | 0.77 | 0.78 | 0.88 | 0.71 | 0.85 | 0.82 | 0.74 | 0.81 |
| | Causality-applied | With Others | 0.63 | 0.66 | 0.70 | 0.81 | 0.47 | 0.81 | 0.71 | 0.55 | 0.75 |
| Relax | FTW only | Without Others | 0.84 | 0.86 | 0.86 | 0.86 | 0.86 | 0.89 | 0.85 | 0.86 | 0.87 |
| | Causality-applied | Without Others | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.98 | 0.99 | 0.99 |
| | FTW only | With Others | 0.47 | 0.61 | 0.54 | 0.54 | 0.54 | 0.60 | 0.50 | 0.58 | 0.56 |
| | Causality-applied | With Others | 0.66 | 0.70 | 0.66 | 0.87 | 0.94 | 0.91 | 0.75 | 0.80 | 0.76 |
| Enter_Home | FTW only | Without Others | 0.85 | 0.81 | 0.82 | 0.85 | 0.84 | 0.85 | 0.85 | 0.82 | 0.83 |
| | Causality-applied | Without Others | 0.92 | 0.95 | 0.93 | 0.90 | 0.86 | 0.91 | 0.91 | 0.90 | 0.92 |
| | FTW only | With Others | 0.57 | 0.79 | 0.56 | 0.69 | 0.61 | 0.72 | 0.62 | 0.69 | 0.63 |
| | Causality-applied | With Others | 0.45 | 0.73 | 0.54 | 0.69 | 0.58 | 0.75 | 0.54 | 0.65 | 0.62 |
| Leave_Home | FTW only | Without Others | 0.57 | 0.81 | 0.70 | 0.87 | 0.78 | 0.83 | 0.69 | 0.79 | 0.76 |
| | Causality-applied | Without Others | 0.82 | 0.92 | 0.91 | 0.96 | 0.94 | 0.95 | 0.88 | 0.93 | 0.93 |
| | FTW only | With Others | 0.16 | 0.47 | 0.20 | 0.45 | 0.03 | 0.36 | 0.24 | 0.05 | 0.26 |
| | Causality-applied | With Others | 0.41 | 0.75 | 0.58 | 0.86 | 0.67 | 0.86 | 0.55 | 0.71 | 0.69 |
| Housekeeping | FTW only | Without Others | 0.93 | 0.92 | 0.90 | 0.96 | 0.84 | 0.94 | 0.94 | 0.88 | 0.92 |
| | Causality-applied | Without Others | 0.92 | 0.95 | 0.93 | 0.91 | 0.87 | 0.95 | 0.91 | 0.91 | 0.94 |
| | FTW only | With Others | 0.82 | 0.96 | 0.80 | 0.92 | 0.54 | 0.84 | 0.87 | 0.69 | 0.82 |
| | Causality-applied | With Others | 0.77 | 0.92 | 0.79 | 0.92 | 0.76 | 0.93 | 0.84 | 0.83 | 0.86 |
| Meal_Preparation | FTW only | Without Others | 0.90 | 0.90 | 0.92 | 0.84 | 0.89 | 0.87 | 0.87 | 0.89 | 0.89 |
| | Causality-applied | Without Others | 0.99 | 0.99 | 0.99 | 0.95 | 0.96 | 0.96 | 0.97 | 0.98 | 0.98 |
| | FTW only | With Others | 0.57 | 0.63 | 0.57 | 0.52 | 0.60 | 0.53 | 0.55 | 0.61 | 0.55 |
| | Causality-applied | With Others | 0.80 | 0.83 | 0.80 | 0.74 | 0.79 | 0.76 | 0.77 | 0.81 | 0.78 |
| Eating | FTW only | Without Others | 0.30 | 0.43 | 0.40 | 0.56 | 0.46 | 0.60 | 0.39 | 0.44 | 0.48 |
| | Causality-applied | Without Others | 0.56 | 0.62 | 0.65 | 0.93 | 0.89 | 0.94 | 0.70 | 0.73 | 0.77 |
| | FTW only | With Others | 0.28 | 0.29 | 0.31 | 0.55 | 0.22 | 0.50 | 0.37 | 0.25 | 0.38 |
| | Causality-applied | With Others | 0.33 | 0.50 | 0.37 | 0.90 | 0.80 | 0.90 | 0.49 | 0.62 | 0.53 |
| Wash_Dishes | FTW only | Without Others | 0.62 | 0.79 | 0.75 | 0.94 | 0.88 | 0.94 | 0.75 | 0.83 | 0.84 |
| | Causality-applied | Without Others | 0.77 | 0.75 | 0.74 | 0.97 | 0.93 | 0.97 | 0.86 | 0.83 | 0.84 |
| | FTW only | With Others | 0.42 | 0.70 | 0.54 | 0.88 | 0.58 | 0.87 | 0.57 | 0.63 | 0.67 |
| | Causality-applied | With Others | 0.61 | 0.89 | 0.70 | 0.94 | 0.66 | 0.95 | 0.74 | 0.76 | 0.81 |
| Sleeping | FTW only | Without Others | 0.99 | 0.97 | 0.97 | 0.99 | 0.98 | 0.99 | 0.99 | 0.98 | 0.98 |
| | Causality-applied | Without Others | 0.97 | 0.96 | 0.97 | 0.99 | 1.00 | 0.99 | 0.98 | 0.98 | 0.98 |
| | FTW only | With Others | 0.74 | 0.89 | 0.82 | 0.87 | 0.75 | 0.89 | 0.80 | 0.81 | 0.85 |
| | Causality-applied | With Others | 0.76 | 0.78 | 0.77 | 0.99 | 0.97 | 0.99 | 0.86 | 0.86 | 0.87 |
| Work | FTW only | Without Others | 0.90 | 0.96 | 0.98 | 0.98 | 0.94 | 0.95 | 0.94 | 0.95 | 0.97 |
| | Causality-applied | Without Others | 0.89 | 0.95 | 0.92 | 1.00 | 1.00 | 0.99 | 0.94 | 0.97 | 0.96 |
| | FTW only | With Others | 0.58 | 0.79 | 0.58 | 0.70 | 0.54 | 0.78 | 0.64 | 0.64 | 0.67 |
| | Causality-applied | With Others | 0.50 | 0.54 | 0.51 | 0.91 | 0.84 | 0.93 | 0.65 | 0.66 | 0.66 |
| Others | FTW only | Without Others | - | - | - | - | - | - | - | - | - |
| | Causality-applied | Without Others | - | - | - | - | - | - | - | - | - |
| | FTW only | With Others | 0.76 | 0.78 | 0.78 | 0.73 | 0.83 | 0.76 | 0.74 | 0.81 | 0.77 |
| | Causality-applied | With Others | 0.86 | 0.88 | 0.87 | 0.77 | 0.82 | 0.76 | 0.81 | 0.85 | 0.81 |
| (Macro) | FTW only | Without Others | 0.79 | 0.84 | 0.82 | 0.88 | 0.84 | 0.88 | 0.82 | 0.84 | 0.85 |
| | Causality-applied | Without Others | 0.88 | 0.90 | 0.89 | 0.94 | 0.92 | 0.95 | 0.90 | 0.91 | 0.92 |
| | FTW only | With Others | 0.56 | 0.70 | 0.59 | 0.70 | 0.54 | 0.70 | 0.61 | 0.59 | 0.63 |
| | Causality-applied | With Others | 0.62 | 0.74 | 0.66 | 0.85 | 0.75 | 0.87 | 0.70 | 0.74 | 0.74 |

fuzzy features (indicated as ''Causality + FTW'') in an environment with or without the ''Others'' class from the unlabeled data. The macro-F1-scores listed in Table 2 were calculated based on the precision and recall data summarized in Tables 3 and 5 for each dataset. The F1-score is the average of ten implementations of the training model. The boldface values in Table 2 denote the macro-F1-scores of the ''Causality + FTW'' cases for comparison with those of the only FTW cases. The results corresponding to FTW only of deep LSTM in Table 2 refer to the FTW-based approach in [47].

A comparison of the models based on the macro-F1-scores shows that deep 2D-CNN–LSTM presents a superior overall performance among all the other methods. Moreover, learning spatiotemporal dependencies is more effective in recognizing activities than learning only the spatial or temporal dependencies of the fused feature.

Furthermore, we compared the AR performance with that of the delayed FTW method [2], which is also an FTW-based AR method that uses both the preceding and oncoming data. The experimental environment was set identical to the environment, unlabeled data was included in the feature extraction, and the time windows of the preceding and oncoming data were 180 and 1 min, respectively. Additionally, features were extracted from Aruba dataset, the deep 2D-CNN–LSTM

**TABLE 4.** Precision, recall, and F1-score for Milan dataset with leave-one-day-out cross-validation.

| Activity | Feature | Target | Precision | | | Recall | | | F1-score | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM |
| Bed_to_Toilet | FTW only | Without Others | 0.47 | 0.40 | 0.54 | 0.66 | 0.17 | 0.57 | 0.55 | 0.24 | 0.55 |
| | Causality-applied | Without Others | 0.31 | 0.39 | 0.25 | 0.78 | 0.54 | 0.79 | 0.45 | 0.45 | 0.38 |
| | FTW only | With Others | 0.50 | 0.45 | 0.53 | 0.70 | 0.67 | 0.55 | 0.58 | 0.54 | 0.54 |
| | Causality-applied | With Others | 0.23 | 0.43 | 0.24 | 0.76 | 0.50 | 0.72 | 0.35 | 0.47 | 0.36 |
| Master_Bathroom | FTW only | Without Others | 0.30 | 0.39 | 0.36 | 0.45 | 0.42 | 0.48 | 0.36 | 0.40 | 0.41 |
| | Causality-applied | Without Others | 0.46 | 0.44 | 0.49 | 0.89 | 0.89 | 0.960 | 0.61 | 0.59 | 0.63 |
| | FTW only | With Others | 0.22 | 0.20 | 0.25 | 0.37 | 0.33 | 0.36 | 0.28 | 0.25 | 0.29 |
| | Causality-applied | With Others | 0.35 | 0.34 | 0.35 | 0.81 | 0.74 | 0.78 | 0.49 | 0.46 | 0.48 |
| Sleep | FTW only | Without Others | 0.63 | 0.61 | 0.62 | 0.52 | 0.51 | 0.54 | 0.57 | 0.55 | 0.58 |
| | Causality-applied | Without Others | 0.92 | 0.96 | 0.95 | 0.80 | 0.86 | 0.83 | 0.86 | 0.90 | 0.89 |
| | FTW only | With Others | 0.54 | 0.55 | 0.53 | 0.47 | 0.45 | 0.46 | 0.50 | 0.50 | 0.49 |
| | Causality-applied | With Others | 0.83 | 0.80 | 0.78 | 0.79 | 0.80 | 0.80 | 0.81 | 0.80 | 0.79 |
| Read | FTW only | Without Others | 0.53 | 0.58 | 0.56 | 0.54 | 0.57 | 0.57 | 0.53 | 0.57 | 0.57 |
| | Causality-applied | Without Others | 0.78 | 0.78 | 0.82 | 0.95 | 0.95 | 0.95 | 0.86 | 0.85 | 0.88 |
| | FTW only | With Others | 0.34 | 0.37 | 0.38 | 0.43 | 0.45 | 0.45 | 0.38 | 0.50 | 0.41 |
| | Causality-applied | With Others | 0.44 | 0.42 | 0.43 | 0.82 | 0.75 | 0.83 | 0.57 | 0.54 | 0.57 |
| Kitchen_Activity | FTW only | Without Others | 0.81 | 0.80 | 0.82 | 0.74 | 0.84 | 0.79 | 0.77 | 0.82 | 0.81 |
| | Causality-applied | Without Others | 0.96 | 0.94 | 0.96 | 0.80 | 0.88 | 0.81 | 0.87 | 0.91 | 0.88 |
| | FTW only | With Others | 0.65 | 0.65 | 0.64 | 0.59 | 0.59 | 0.61 | 0.62 | 0.62 | 0.62 |
| | Causality-applied | With Others | 0.78 | 0.81 | 0.80 | 0.68 | 0.79 | 0.73 | 0.73 | 0.80 | 0.76 |
| Watch_TV | FTW only | Without Others | 0.53 | 0.56 | 0.55 | 0.60 | 0.59 | 0.61 | 0.56 | 0.57 | 0.58 |
| | Causality-applied | Without Others | 0.88 | 0.88 | 0.89 | 0.87 | 0.86 | 0.91 | 0.87 | 0.87 | 0.90 |
| | FTW only | With Others | 0.46 | 0.43 | 0.45 | 0.54 | 0.52 | 0.55 | 0.50 | 0.47 | 0.49 |
| | Causality-applied | With Others | 0.69 | 0.68 | 0.70 | 0.77 | 0.72 | 0.82 | 0.73 | 0.70 | 0.75 |
| Chores | FTW only | Without Others | 0.73 | 0.76 | 0.76 | 0.78 | 0.63 | 0.71 | 0.76 | 0.69 | 0.73 |
| | Causality-applied | Without Others | 0.81 | 0.89 | 0.81 | 0.85 | 0.71 | 0.85 | 0.83 | 0.79 | 0.83 |
| | FTW only | With Others | 0.68 | 0.68 | 0.74 | 0.73 | 0.78 | 0.75 | 0.70 | 0.73 | 0.75 |
| | Causality-applied | With Others | 0.71 | 0.91 | 0.85 | 0.79 | 0.66 | 0.85 | 0.75 | 0.76 | 0.85 |
| Desk_Activity | FTW only | Without Others | 0.62 | 0.76 | 0.67 | 0.81 | 0.70 | 0.77 | 0.70 | 0.73 | 0.72 |
| | Causality-applied | Without Others | 0.93 | 0.94 | 0.91 | 0.99 | 0.99 | 0.99 | 0.96 | 0.96 | 0.95 |
| | FTW only | With Others | 0.51 | 0.50 | 0.58 | 0.80 | 0.85 | 0.74 | 0.62 | 0.63 | 0.65 |
| | Causality-applied | With Others | 0.46 | 0.47 | 0.55 | 0.97 | 0.94 | 0.99 | 0.62 | 0.63 | 0.71 |
| Leave_Home | FTW only | Without Others | 0.36 | 0.46 | 0.45 | 0.50 | 0.34 | 0.52 | 0.42 | 0.39 | 0.48 |
| | Causality-applied | Without Others | 0.57 | 0.78 | 0.74 | 0.75 | 0.72 | 0.81 | 0.65 | 0.75 | 0.77 |
| | FTW only | With Others | 0.19 | 0.18 | 0.24 | 0.32 | 0.30 | 0.32 | 0.24 | 0.22 | 0.27 |
| | Causality-applied | With Others | 0.21 | 0.27 | 0.24 | 0.57 | 0.33 | 0.61 | 0.30 | 0.29 | 0.34 |
| Eve_Meds | FTW only | Without Others | 0.06 | 0.08 | 0.09 | 0.37 | 0.03 | 0.46 | 0.10 | 0.04 | 0.15 |
| | Causality-applied | Without Others | 0.06 | 0.14 | 0.07 | 0.43 | 0.14 | 0.43 | 0.10 | 0.14 | 0.12 |
| | FTW only | With Others | 0.05 | 0.08 | 0.08 | 0.26 | 0.43 | 0.37 | 0.09 | 0.13 | 0.14 |
| | Causality-applied | With Others | 0.08 | 0.00 | 0.09 | 0.49 | 0.00 | 0.51 | 0.14 | 0.00 | 0.15 |
| Dining_Rm_Activity | FTW only | Without Others | 0.40 | 0.53 | 0.51 | 0.71 | 0.56 | 0.66 | 0.51 | 0.55 | 0.57 |
| | Causality-applied | Without Others | 0.72 | 0.88 | 0.79 | 0.94 | 0.94 | 0.93 | 0.82 | 0.91 | 0.85 |
| | FTW only | With Others | 0.32 | 0.40 | 0.38 | 0.68 | 0.68 | 0.71 | 0.44 | 0.50 | 0.49 |
| | Causality-applied | With Others | 0.69 | 0.77 | 0.72 | 0.92 | 0.87 | 0.94 | 0.78 | 0.82 | 0.82 |
| Guest_Bathroom | FTW only | Without Others | 0.33 | 0.36 | 0.36 | 0.39 | 0.32 | 0.38 | 0.36 | 0.34 | 0.37 |
| | Causality-applied | Without Others | 0.66 | 0.71 | 0.64 | 0.77 | 0.77 | 0.86 | 0.71 | 0.74 | 0.73 |
| | FTW only | With Others | 0.21 | 0.20 | 0.22 | 0.30 | 0.28 | 0.29 | 0.25 | 0.24 | 0.25 |
| | Causality-applied | With Others | 0.40 | 0.47 | 0.42 | 0.72 | 0.79 | 0.76 | 0.52 | 0.58 | 0.54 |
| Master_Bedroom_Activity | FTW only | Without Others | 0.50 | 0.49 | 0.50 | 0.44 | 0.49 | 0.46 | 0.47 | 0.49 | 0.48 |
| | Causality-applied | Without Others | 0.86 | 0.87 | 0.89 | 0.58 | 0.60 | 0.65 | 0.69 | 0.71 | 0.75 |
| | FTW only | With Others | 0.34 | 0.34 | 0.34 | 0.31 | 0.34 | 0.31 | 0.32 | 0.34 | 0.33 |
| | Causality-applied | With Others | 0.68 | 0.68 | 0.68 | 0.56 | 0.49 | 0.55 | 0.61 | 0.57 | 0.61 |
| Meditate | FTW only | Without Others | 0.37 | 0.64 | 0.26 | 0.41 | 0.43 | 0.47 | 0.39 | 0.51 | 0.34 |
| | Causality-applied | Without Others | 0.39 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 0.56 | 1.00 | 0.99 |
| | FTW only | With Others | 0.18 | 0.14 | 0.21 | 0.57 | 0.53 | 0.47 | 0.27 | 0.22 | 0.29 |
| | Causality-applied | With Others | 0.47 | 0.64 | 0.38 | 1.00 | 1.00 | 1.00 | 0.64 | 0.78 | 0.55 |
| Morning_Meds | FTW only | Without Others | 0.12 | 0.11 | 0.18 | 0.46 | 0.14 | 0.49 | 0.19 | 0.12 | 0.26 |
| | Causality-applied | Without Others | 0.22 | 0.29 | 0.20 | 0.96 | 0.75 | 0.95 | 0.35 | 0.42 | 0.34 |
| | FTW only | With Others | 0.07 | 0.07 | 0.08 | 0.36 | 0.36 | 0.33 | 0.12 | 0.12 | 0.13 |
| | Causality-applied | With Others | 0.18 | 0.31 | 0.252 | 0.89 | 0.51 | 0.82 | 0.30 | 0.38 | 0.39 |
| Others | FTW only | Without Others | - | - | - | - | - | - | - | - | - |
| | Causality-applied | Without Others | - | - | - | - | - | - | - | - | - |
| | FTW only | With Others | 0.49 | 0.52 | 0.53 | 0.42 | 0.45 | 0.47 | 0.45 | 0.48 | 0.50 |
| | Causality-applied | With Others | 0.64 | 0.65 | 0.67 | 0.39 | 0.47 | 0.39 | 0.49 | 0.55 | 0.50 |
| (Macro) | FTW only | Without Others | 0.45 | 0.50 | 0.48 | 0.56 | 0.45 | 0.57 | 0.48 | 0.47 | 0.51 |
| | Causality-applied | Without Others | 0.63 | 0.73 | 0.69 | 0.82 | 0.77 | 0.84 | 0.68 | 0.73 | 0.73 |
| | FTW only | With Others | 0.36 | 0.36 | 0.39 | 0.49 | 0.50 | 0.48 | 0.40 | 0.40 | 0.42 |
| | Causality-applied | With Others | 0.49 | 0.54 | 0.51 | 0.75 | 0.65 | 0.76 | 0.55 | 0.57 | 0.57 |

**TABLE 5.** Precision, recall, and F1-score for Cairo dataset with leave-one-day-out cross-validation.

| Activity | Feature | Target | Precision | | | Recall | | | F1-score | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM | LSTM | 2D-CNN | 2D-CNN-LSTM |
| Bed_to_toilet | FTW only | Without Others | 0.81 | 0.87 | 0.88 | 0.78 | 0.93 | 0.96 | 0.79 | 0.90 | 0.92 |
| | Causality-applied | Without Others | 0.78 | 0.95 | 0.87 | 0.91 | 0.88 | 0.94 | 0.84 | 0.91 | 0.90 |
| | FTW only | With Others | 0.60 | 0.50 | 0.59 | 0.55 | 0.47 | 0.92 | 0.57 | 0.49 | 0.72 |
| | Causality-applied | With Others | 0.54 | 0.38 | 0.57 | 0.83 | 0.37 | 0.91 | 0.66 | 0.37 | 0.70 |
| Breakfast | FTW only | Without Others | 0.89 | 0.92 | 0.97 | 0.98 | 0.91 | 0.89 | 0.94 | 0.92 | 0.93 |
| | Causality-applied | Without Others | 0.96 | 0.99 | 1.00 | 0.95 | 0.91 | 0.87 | 0.95 | 0.95 | 0.93 |
| | FTW only | With Others | 0.49 | 0.56 | 0.61 | 0.92 | 0.66 | 0.74 | 0.64 | 0.60 | 0.67 |
| | Causality-applied | With Others | 0.59 | 0.71 | 0.71 | 0.89 | 0.32 | 0.51 | 0.71 | 0.44 | 0.60 |
| Leave_home | FTW only | Without Others | 0.78 | 0.57 | 0.50 | 0.71 | 0.94 | 0.92 | 0.74 | 0.71 | 0.65 |
| | Causality-applied | Without Others | 0.84 | 0.82 | 0.65 | 0.92 | 1.00 | 1.00 | 0.88 | 0.90 | 0.79 |
| | FTW only | With Others | 0.09 | 1.00 | 0.17 | 0.44 | 0.01 | 0.60 | 0.15 | 0.03 | 0.27 |
| | Causality-applied | With Others | 0.42 | 1.00 | 0.50 | 0.75 | 0.16 | 0.84 | 0.54 | 0.27 | 0.63 |
| R1_work_in_office | FTW only | Without Others | 0.63 | 0.70 | 0.64 | 0.69 | 0.81 | 0.87 | 0.66 | 0.75 | 0.74 |
| | Causality-applied | Without Others | 0.63 | 0.93 | 0.94 | 0.90 | 0.89 | 0.89 | 0.74 | 0.91 | 0.91 |
| | FTW only | With Others | 0.32 | 0.49 | 0.36 | 0.57 | 0.33 | 0.73 | 0.41 | 0.40 | 0.48 |
| | Causality-applied | With Others | 0.47 | 0.85 | 0.68 | 0.75 | 0.38 | 0.80 | 0.58 | 0.53 | 0.73 |
| Lunch | FTW only | Without Others | 0.67 | 0.78 | 0.89 | 1.00 | 0.78 | 0.84 | 0.80 | 0.78 | 0.86 |
| | Causality-applied | Without Others | 0.86 | 0.85 | 0.94 | 0.99 | 0.87 | 0.86 | 0.92 | 0.86 | 0.90 |
| | FTW only | With Others | 0.41 | 0.59 | 0.62 | 0.90 | 0.47 | 0.69 | 0.57 | 0.52 | 0.65 |
| | Causality-applied | With Others | 0.58 | 0.81 | 0.90 | 0.85 | 0.24 | 0.67 | 0.69 | 0.37 | 0.77 |
| Night_wandering | FTW only | Without Others | 0.83 | 0.96 | 0.95 | 0.67 | 0.92 | 0.94 | 0.74 | 0.94 | 0.95 |
| | Causality-applied | Without Others | 0.89 | 0.94 | 0.98 | 0.86 | 0.96 | 0.95 | 0.87 | 0.95 | 0.97 |
| | FTW only | With Others | 0.67 | 0.74 | 0.80 | 0.54 | 0.68 | 0.81 | 0.60 | 0.71 | 0.81 |
| | Causality-applied | With Others | 0.53 | 0.86 | 0.72 | 0.73 | 0.79 | 0.87 | 0.61 | 0.82 | 0.79 |
| R1_sleep | FTW only | Without Others | 0.63 | 0.68 | 0.74 | 0.71 | 0.77 | 0.80 | 0.67 | 0.72 | 0.77 |
| | Causality-applied | Without Others | 0.63 | 0.97 | 0.97 | 0.75 | 0.99 | 1.00 | 0.69 | 0.98 | 0.98 |
| | FTW only | With Others | 0.35 | 0.61 | 0.53 | 0.42 | 0.56 | 0.79 | 0.38 | 0.59 | 0.63 |
| | Causality-applied | With Others | 0.39 | 0.85 | 0.73 | 0.64 | 0.79 | 0.97 | 0.48 | 0.82 | 0.83 |
| R2_sleep | FTW only | Without Others | 0.86 | 0.80 | 0.86 | 0.74 | 0.81 | 0.85 | 0.80 | 0.80 | 0.85 |
| | Causality-applied | Without Others | 0.89 | 0.99 | 1.00 | 0.79 | 1.00 | 0.99 | 0.83 | 0.99 | 0.99 |
| | FTW only | With Others | 0.42 | 0.64 | 0.57 | 0.44 | 0.53 | 0.68 | 0.43 | 0.58 | 0.62 |
| | Causality-applied | With Others | 0.51 | 0.84 | 0.80 | 0.57 | 0.72 | 0.90 | 0.54 | 0.78 | 0.85 |
| Laundry | FTW only | Without Others | 0.58 | 1.00 | 0.62 | 0.95 | 0.99 | 0.99 | 0.72 | 0.99 | 0.76 |
| | Causality-applied | Without Others | 0.97 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 0.98 | 1.00 | 1.00 |
| | FTW only | With Others | 0.44 | 0.94 | 0.35 | 0.84 | 0.69 | 0.94 | 0.58 | 0.79 | 0.51 |
| | Causality-applied | With Others | 0.88 | 1.00 | 1.00 | 0.98 | 0.96 | 0.99 | 0.93 | 0.98 | 0.99 |
| Dinner | FTW only | Without Others | 0.99 | 0.91 | 0.94 | 0.81 | 0.87 | 0.96 | 0.89 | 0.89 | 0.95 |
| | Causality-applied | Without Others | 0.99 | 0.96 | 0.98 | 0.96 | 0.96 | 0.98 | 0.97 | 0.96 | 0.98 |
| | FTW only | With Others | 0.34 | 0.47 | 0.61 | 0.34 | 0.45 | 0.68 | 0.34 | 0.46 | 0.64 |
| | Causality-applied | With Others | 0.56 | 0.71 | 0.66 | 0.60 | 0.65 | 0.88 | 0.58 | 0.68 | 0.75 |
| R1_wake | FTW only | Without Others | 0.69 | 0.83 | 0.87 | 0.86 | 0.88 | 0.89 | 0.77 | 0.85 | 0.88 |
| | Causality-applied | Without Others | 0.87 | 0.75 | 0.76 | 0.88 | 0.94 | 0.93 | 0.88 | 0.83 | 0.84 |
| | FTW only | With Others | 0.62 | 0.73 | 0.74 | 0.78 | 0.54 | 0.76 | 0.69 | 0.62 | 0.75 |
| | Causality-applied | With Others | 0.67 | 0.48 | 0.52 | 0.80 | 0.71 | 0.91 | 0.73 | 0.57 | 0.66 |
| R2_wake | FTW only | Without Others | 0.94 | 0.95 | 0.97 | 0.94 | 0.93 | 0.89 | 0.94 | 0.94 | 0.92 |
| | Causality-applied | Without Others | 0.94 | 0.98 | 0.96 | 0.89 | 0.93 | 0.94 | 0.91 | 0.95 | 0.95 |
| | FTW only | With Others | 0.90 | 0.84 | 0.82 | 0.81 | 0.75 | 0.82 | 0.85 | 0.79 | 0.82 |
| | Causality-applied | With Others | 0.76 | 0.74 | 0.76 | 0.72 | 0.83 | 0.89 | 0.74 | 0.78 | 0.82 |
| R2_take_medicine | FTW only | Without Others | 0.81 | 0.58 | 0.38 | 0.91 | 0.92 | 0.83 | 0.86 | 0.71 | 0.52 |
| | Causality-applied | Without Others | 0.77 | 0.67 | 0.55 | 0.96 | 1.00 | 1.00 | 0.85 | 0.80 | 0.71 |
| | FTW only | With Others | 0.26 | 0.20 | 0.22 | 0.57 | 0.22 | 0.68 | 0.36 | 0.21 | 0.34 |
| | Causality-applied | With Others | 0.23 | 0.41 | 0.23 | 0.72 | 0.51 | 0.87 | 0.35 | 0.46 | 0.37 |
| Others | FTW only | Without Others | - | - | - | - | - | - | - | - | - |
| | Causality-applied | Without Others | - | - | - | - | - | - | - | - | - |
| | FTW only | With Others | 0.89 | 0.86 | 0.91 | 0.82 | 0.88 | 0.85 | 0.86 | 0.87 | 0.88 |
| | Causality-applied | With Others | 0.93 | 0.91 | 0.96 | 0.89 | 0.95 | 0.90 | 0.91 | 0.93 | 0.93 |
| (Macro) | FTW only | Without Others | 0.78 | 0.81 | 0.79 | 0.83 | 0.88 | 0.89 | 0.79 | 0.84 | 0.82 |
| | Causality-applied | Without Others | 0.85 | 0.91 | 0.89 | 0.90 | 0.95 | 0.95 | 0.87 | 0.92 | 0.91 |
| | FTW only | With Others | 0.49 | 0.66 | 0.57 | 0.64 | 0.52 | 0.76 | 0.53 | 0.55 | 0.63 |
| | Causality-applied | With Others | 0.58 | 0.75 | 0.70 | 0.77 | 0.60 | 0.85 | 0.65 | 0.63 | 0.74 |

model was trained, and leave-one-day-out cross-validation was applied to examine the AR performance of the delayed FTW method [2]. In the delayed FTW method, fuzzy features were extracted from both the preceding and oncoming data.

The experimental results showed that there was insignificant improvement compared to the macro-F1-scores of the Baseline in Table 7 when the unlabeled "Others" data were considered. Unlike [2], where performance is improved by

**TABLE 6.** Performance comparison with other methods.

| | Feature extraction method | | Target | Macro-F1-scores |
|---|---|---|---|---|
| | Preceding data (180 min) | Oncoming data (1 min) | | (Deep 2D-CNN-LSTM) |
| FTW [48] | FTW | N/A | with Others | 0.634 |
| Delayed FTW [2] | FTW | FTW | with Others | 0.651 |
| Ours | FTW | Causality | with Others | 0.738 |

**TABLE 7.** Welch's t-test values based on leave-one-day-out cross-validation.

| Dataset | Feature | Target | Model | t-value | p-value | $\mu$ |
|---|---|---|---|---|---|---|
| Aruba | FTW only | With Others | Deep 2D-CNN–LSTM | -6.067 | 2.451e-08 | 71.14 |
| | Causality + FTW | | | | | 78.66 |
| Milan | FTW only | With Others | Deep 2D-CNN–LSTM | -7.860 | 5.036e-12 | 53.26 |
| | Causality + FTW | | | | | 63.38 |
| Cairo | FTW only | With Others | Deep 2D-CNN–LSTM | -7.567 | 2.100e-11 | 70.66 |
| | Causality + FTW | | | | | 79.92 |

using up to 4 h of the oncoming data, this experiment has an extremely short time of 1 min of the oncoming data; thus, it does not yield a considerable improvement. However, because a long time delay is difficult to practically apply to real-time AR, improvements should be made with a short delay. Because the causality feature extracted from 1 min of the oncoming data differs from the fuzzy feature of the preceding data in the proposed method in this study, the performance is improved by generating a new feature using the synthesis of the two features.

#### 2) STATISTICAL SIGNIFICANCE TEST

We used hypothesis testing for the proposed AR architecture. Therefore, we performed a statistical significance test to determine whether the classification accuracy of the models differs with and without the causality feature, where the accuracy represents the proportion of the prediction that is correctly classified. The main idea of the proposed AR architecture was to extract causality features based on the NDTE in addition to extracting fuzzy features with FTWs to improve the precision of AR. The two models were set with the same architecture as deep-2D-CNN-LSTM, and Model 1 was trained with the fuzzy feature only, and Model 2 was trained with the fused feature in which the causality and fuzzy features were fused.

We performed a two-tailed Welch's unequal variances $t$-test [64] on the two sets of accuracies, which are the result of the two models through the leave-one-day-out cross-validation. Test data were generated with randomly selected 50 days for the Aruba, Milan, and Cairo datasets, and 50 accuracies for each model were sampled. The $t$-test was performed with the null hypothesis $H_0$:$\mu_1 = \mu_2$ versus the alternative hypothesis $H_a$:$\mu_1 < \mu_2$ and significance level $\alpha = 0.05$, where $\mu_1$ and $\mu_2$ denote the mean accuracies of Models 1 and 2, respectively.

The results are shown in Table 7, where it is clearly shown that the null hypothesis is rejected every time because the $p$-value is less than $\alpha$, and $\mu_2$ is calculated to be greater than $\mu_1$. Therefore, the results suggest that Model 2 predicts better than Model 1 does.

**TABLE 8.** Improvement of macro-F1-scores by applying the causality.

| Dataset | Feature | Target | Deep 2D-CNN-LSTM | Improvement |
|---|---|---|---|---|
| Aruba | FTW only | With Others | 0.634 | +0.104 |
| | Causality + FTW | | 0.738 | |
| Milan | FTW only | With Others | 0.416 | +0.156 |
| | Causality + FTW | | 0.572 | |
| Cairo | FTW only | With Others | 0.628 | +0.116 |
| | Causality + FTW | | 0.744 | |

### *E. DISCUSSION*

Herein, we discussed the proposed causality-based method by considering the unlabeled data.

In Table 2, it is observed that the AR performance degrades for all cases involving the ''Others'' class. In particular, the precision, recall, and F1-score of Leave_Home of the Aruba dataset, as shown in Table 3, is significantly reduced because of the large correlation with the ''Others'' class, which explains the reason for dealing with unlabeled data for real-time AR. If the unlabeled data are not learned as the ''Others'' class, the model can erroneously classify a query from the unlabeled data as the most probable predefined activity because the model does not know about the ''Others'' class.

To overcome the performance degradation caused by the ''Others'' class, an AR architecture using the causality between sensor activations was presented in this study. We confirmed that the macro-F1-scores was improved in all cases when the causality was applied. Among the datasets, the improvement in the macro-F1-scores of the Milan dataset was the most remarkable (see Table 2), and the proposed method of applying causality was effective in AR when the size of the dataset was not sufficiently large. Table 6 provides the improvement in the macro-F1-scores for the best model in the training dataset containing the ''Others'' class.

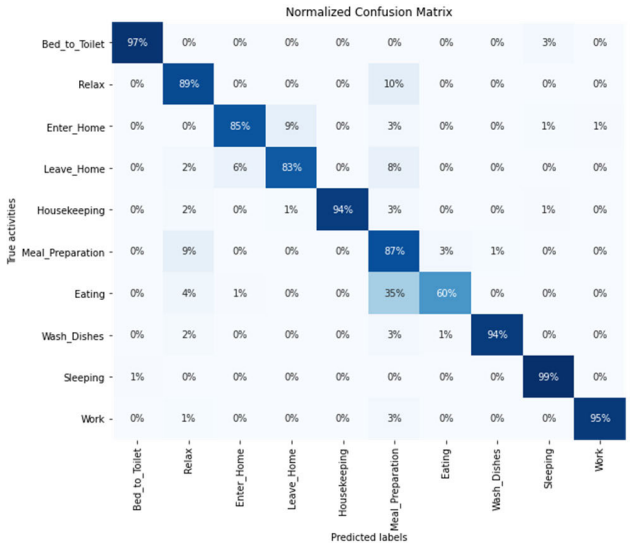



**FIGURE 7.** Normalized confusion matrix of the deep 2D-CNN–LSTM model for only FTW without "Others" class for Aruba dataset.

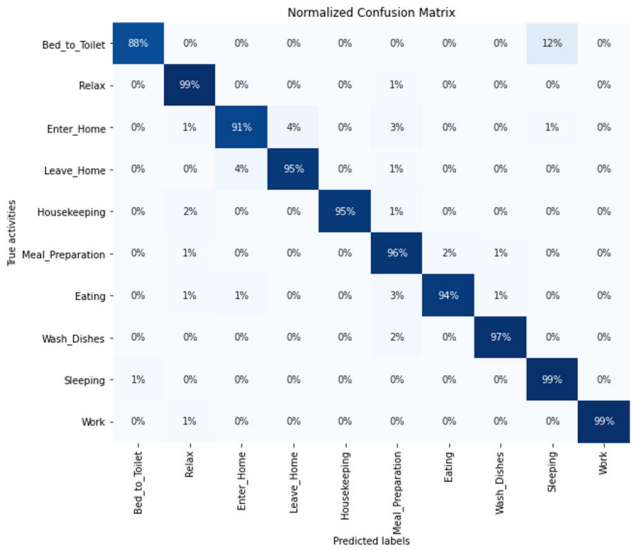



**FIGURE 8.** Normalized confusion matrix of the deep 2D-CNN–LSTM model for causality application without "Others" class for Aruba dataset.

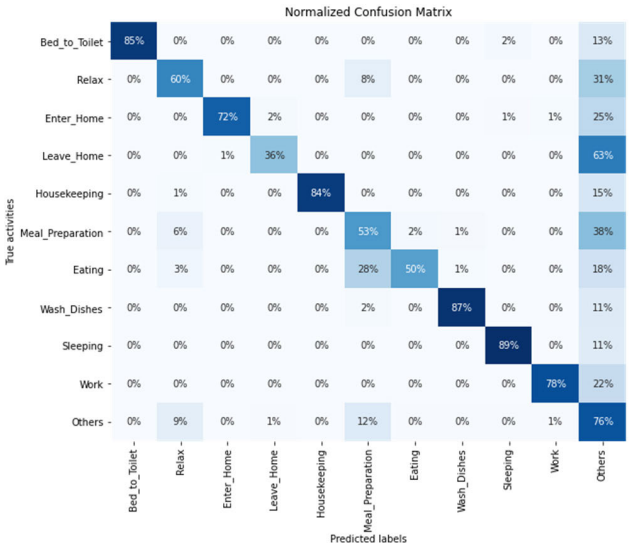



**FIGURE 9.** Normalized confusion matrix of the deep 2D-CNN–LSTM model for only FTW with "Others" class for Aruba dataset.

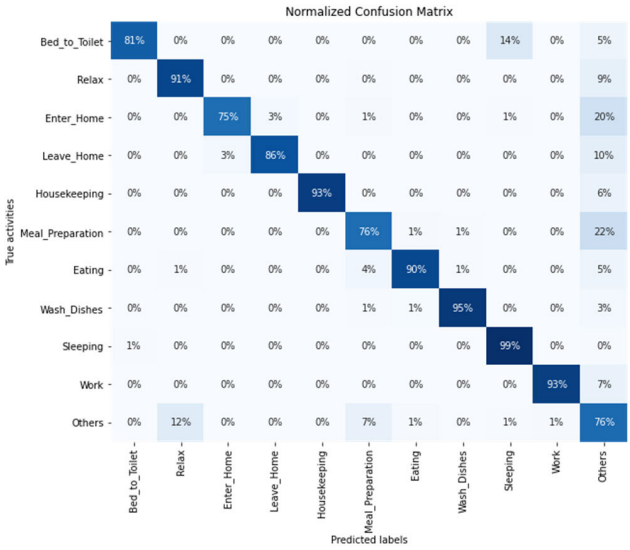



**FIGURE 10.** Normalized confusion matrix of the deep 2D-CNN–LSTM model for causality application with "Others" class for Aruba dataset.

Among the normalized confusion matrices depicted in Figs. 7–10, the ''Others'' class presented in Fig. 9 seems to absorb the predicted label from other categories as white noise. This is because the ''Others'' class data account for approximately half of the raw datasets, and some of the ''Others'' data are unlabeled even though they have to be labeled as predefined activities. Using the causality feature, the problems of easily-confused activities (Meal_Preparation, Eating, Enter_Home, and Leave_Home, as shown in Figs. 7–8), which absorb predicted labels into the ''Others'' class, are significantly mitigated, as depicted in Figs. 9–10.

This is possible because the causal relationship of each sensor activation uniquely indicates distinct dependencies for each activity. Therefore, we employed distinct features to recognize activities. Consequently, the proposed feature fusion can be used to complement the shortcomings of the FTW-based fuzzy features in improving AR performance.

## V. CONCLUSION

In this study, a novel feature extraction approach based on binary sensor data was proposed to improve classification performance by mitigating the problems of easily confused activities and unlabeled data. To clearly recognize an activity, the proposed method not only considers that the current activity is a continuation of the long-term preceding activity but also that it is important for predicting a short-term

oncoming activity. For the preceding activity, a large amount of historical data for several hours was used as the fuzzy features obtained by FTWs, and for the oncoming activity, the direct causal influence, serving as a spatial feature, was extracted for the relationship of sensor activation. Moreover, they were fused as a unique feature of the activity. The spatiotemporal dependencies of the fused feature were learned by deep LSTM, deep 2D-CNN, and deep 2D-CNN–LSTM. The experiments demonstrated that the F1-score improved when the causality feature was applied to recognize all activities performed in the CASAS open datasets, Aruba, Milan, and Cairo. In addition, it was confirmed that the precision and the recall performance were improved even in the case the training model included the unlabeled data in the others class. This result indicated that it is important to consider causality for enhancing the learning process of the models.

In future studies, we will explore a chain model for classifying current activities and predicting future activities based on the results of the classifier, further providing the probability of the next activities. For the chain model, a multi-labeled problem can be considered to express the transition of an activity. In addition, an AR service can be designed to deal with unexpected emergency scenarios prior to predicting future behavior.

## REFERENCES

[1] D. Chen, S. Yongchareon, E. M.-K. Lai, J. Yu, and Q. Z. Sheng, "Hybrid fuzzy c-means CPD-based segmentation for improving sensor-based multiresident activity recognition," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11193–11207, Jul. 2021, doi: 10.1109/JIOT.2021.3051574.

[2] R. A. Hamad, A. G. Salguero, M.-R. Bouguelia, M. Espinilla, and J. Medina-Quero, "Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 2, pp. 387–395, Feb. 2020.

[3] K. Bouchard, J. Hao, B. Bouchard, S. Gaboury, M. T. Moutacalli, C. Gouin-Vallerand, H. Kenfack, Ngankam, H. Pigot, and S. Giroux, "The cornerstones of smart home research for healthcare," *Smart Innov. Syst. Technol.*, vol. 93, pp. 185–200, Apr. 2018.

[4] F. Ordóñez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.

[5] M. Wang, C. Luo, B. Ni, J. Yuan, J. Wang, and S. Yan, "First-person daily activity recognition with manipulated object proposals and non-linear feature fusion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 10, pp. 2946–2955, Oct. 2017.

[6] K. Adhikari, H. Bouchachia, and H. Nait-Charif, "Activity recognition for indoor fall detection using convolutional neural network," in *Proc. 15th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2017, pp. 81–84.

[7] C.-D. Huang, C.-Y. Wang, and J.-C. Wang, "Human action recognition system for elderly and children care using three stream ConvNet," in *Proc. Int. Conf. Orange Technol. (ICOT)*, Dec. 2015, pp. 5–9.

[8] D. Chen, A. J. Bharucha, and H. D. Wactlar, "Intelligent video monitoring to improve safety of older persons," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 3814–3817.

[9] M. Papakostas, T. Giannakopoulos, F. Makedon, and V. Karkaletsis, "Short-term recognition of human activities using convolutional neural networks," in *Proc. 12th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, 2016, pp. 302–307.

[10] S. N. Gowda, "Human activity recognition using combinatorial deep belief networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1589–1594.

[11] F. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, pp. 1–25, 2016.

[12] D. Wagner, K. Kalischewski, J. Velten, and A. Kummert, "Activity recognition using inertial sensors and a 2-D convolutional neural network," in *Proc. 10th Int. Workshop Multidimensional (nD) Syst. (nDS)*, Sep. 2017, pp. 1–6.

[13] S. Matsui, N. Inoue, Y. Akagi, G. Nagino, and K. Shinoda, "User adaptation of convolutional neural network for human activity recognition," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 783–787.

[14] Y. Chen and Y. Xue, "A deep learning approach to human activity recognition based on single accelerometer," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 1488–1492.

[15] J. Li, R. Wu, J. Zhao, and Y. Ma, "Convolutional neural networks (CNN) for indoor human activity recognition using ubisense system," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 2068–2072.

[16] K.-J. Kim, M. M. Hassan, S.-H. Na, and E.-N. Huh, "Dementia wandering detection and activity recognition algorithm using tri-axial accelerometer sensors," in *Proc. 4th Int. Conf. Ubiquitous Inf. Technol. Appl.*, Dec. 2009, pp. 1–5.

[17] S.-M. Lee, S. Min Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2017, pp. 131–134.

[18] S. Abbate, M. Avvenuti, and J. Light, "MIMS: A minimally invasive monitoring sensor platform," *IEEE Sensors J.*, vol. 12, no. 3, pp. 677–684, Mar. 2012.

[19] N. K. Vuong, S. Chan, C. T. Lau, S. Y. W. Chan, P. L. K. Yap, and A. S. H. Chen, "Preliminary results of using inertial sensors to detect dementia- related wandering patterns," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 3703–3706.

[20] D. De Venuto, V. F. Annese, and G. Mezzina, "Remote neuro-cognitive impairment sensing based on P300 spatio-temporal monitoring," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8348–8356, Dec. 2016.

[21] K. Ohnishi, A. Kanehira, A. Kanezaki, and T. Harada, "Recognizing activities of daily living with a wrist-mounted camera," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3103–3111.

[22] T. Zebin, P. J. Scully, and K. B. Ozanyan, "Human activity recognition with inertial sensors using a deep learning approach," in *Proc. IEEE SENSORS*, vol. 1, Oct. 2016, pp. 1–3.

[23] M. Panwar, S. R. Dyuthi, K. C. Prakash, D. Biswas, A. Acharyya, K. Maharatna, A. Gautam, and G. R. Naik, "CNN based approach for activity recognition using a wrist-worn accelerometer," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2438–2441.

[24] S. Ha, J.-M. Yun, and S. Choi, "Multi-modal convolutional neural networks for activity recognition," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 3017–3022.

[25] W. D. Kearns, D. Algase, D. H. Moore, and S. Ahmed, "Ultra wideband radio: A novel method for measuring wandering in persons with dementia," *Gerontechnology*, vol. 7, no. 1, pp. 48–57, Jan. 2008.

[26] K. Makimoto, M. Yamakawa, N. Ashida, Y. Kang, and K.-R. Shin, "Japan-Korea joint project on monitoring people with dementia," in *Proc. 11th World Congr. Internet Med.*, 2006, pp. 1–5.

[27] C. Price, "Monitoring people with dementia—Controlling or liberating?" *Qual. Ageing*, vol. 8, no. 3, pp. 41–44, Apr. 2007.

[28] A. Kumar, C. T. Lau, S. Chan, M. Ma, and W. D. Kearns, "A unified grid-based wandering pattern detection algorithm," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 5401–5404.

[29] G. Liu, J. Liang, G. Lan, Q. Hao, and M. Chen, "Convolution neutral network enhanced binary sensor network for human activity recognition," in *Proc. IEEE SENSORS*, Oct. 2016, pp. 1–3.

[30] H. H. Dodge, N. C. Mattek, D. Austin, T. L. Hayes, and J. A. Kaye, "In-home walking speeds and variability trajectories associated with mild cognitive impairment," *Neurology*, vol. 78, pp. 1946–1952, Jun. 2012.

[31] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Automated cognitive health assessment using smart home monitoring of complex tasks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 43, no. 6, pp. 1302–1313, Nov. 2013.

[32] B. Das, D. J. Cook, N. C. Krishnan, and M. Schmitter-Edgecombe, "One-class classification-based real-time activity error detection in smart homes," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 5, pp. 914–923, Aug. 2016.

[33] J. Petersen, D. Austin, N. Mattek, and J. Kaye, "Time out-of-home and cognitive, physical, and emotional wellbeing of older adults: A longitudinal mixed effects model," *PLoS ONE*, vol. 10, no. 10, pp. 1–16, 2015.
[34] L. G. Fahad, S. F. Tahir, and M. Rajarajan, "Activity recognition in smart homes using clustering based classification," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 1348–1353.
[35] N. Yala, B. Fergani, and A. Fleury, "Feature extraction for human activity recognition on streaming data," in *Proc. Int. Symp. Innov. Intell. Syst. Appl. (INISTA)*, Sep. 2015, pp. 1–6.
[36] N. Yala, B. Fergani, and A. Fleury, "Towards improving feature extraction and classification for activity recognition on streaming data," *J. Ambient Intell. Humanized Comput.*, vol. 8, no. 2, pp. 177–189, 2017.
[37] Y. Zhou, B. Ni, S. Yan, and T. S. Huang, "Recognizing pair-activities by causality analysis," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 1, pp. 1–20, Jan. 2011.
[38] T. Bossomaier, L. Barnett, M. Harre, and J. T. Lizier, *An Introduction to Transfer Entropy: Information Flow in Complex Systems*. Berlin, Germany: Springer, 2016.
[39] L. Zhou, J. Liu, S. Nishimura, J. Antony, and C. Gurrin, "Causality inspired retrieval of human-object interactions from video," in *Proc. Int. Conf. Content-Based Multimedia Indexing (CBMI)*, Sep. 2019, pp. 1–6.
[40] Z. Xie, T. Wu, X. Yang, L. Zhang, and K. Wu, "Jointly social grouping and identification in visual dynamics with causality-induced hierarchical Bayesian model," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 62–75, Feb. 2019.
[41] R. Kuznets, L. Prosperi, U. Schmid, and K. Fruzsa, "Causality and epistemic reasoning in Byzantine multi-agent systems," *Electron. Proc. Theor. Comput. Sci.*, vol. 297, pp. 293–312, Jul. 2019, doi: 10.4204/EPTCS.297.19.
[42] C. W. J. Granger, "Investigating causal relationships by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.
[43] C. Aviles-Cruz, E. Rodriguez-Martinez, J. Villegas-Cortez, and A. Ferreyra-Ramirez, "Granger-causality: An efficient single user movement recognition using a smartphone accelerometer sensor," *Pattern Recognit. Lett.*, vol. 125, pp. 576–583, Jul. 2019.
[44] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, no. 2, pp. 461–464, 2000.
[45] A. Kaiser and T. Schreiber, "Information transfer in continuous processes," *Phys. D, Nonlinear Phenomena*, vol. 166, nos. 1–2, pp. 43–62, Jun. 2002.
[46] P. Duan, F. Yang, T. Chen, and S. L. Shah, "Direct causality detection via the transfer entropy approach," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 6, pp. 2052–2066, Nov. 2013.
[47] J. Medina-Quero, S. Zhang, C. Nugent, and M. Espinilla, "Ensemble classifier of long short-term memory with fuzzy temporal windows on binary sensors for activity recognition," *Expert Syst. Appl.*, vol. 114, pp. 441–453, Dec. 2018.
[48] J. Medina, L. Martinez, and M. Espinilla, "Subscribing to fuzzy temporal aggregation of heterogeneous sensor streams in real-time distributed environments," *Int. J. Commun. Syst.*, vol. 30, no. 5, pp. 1–17, 2017.
[49] M. Espinilla, J. Medina, J. Hallberg, and C. Nugent, "A new approach based on temporal sub-windows for online sensor-based activity recognition," *J. Ambient Intell. Humanized Comput.*, pp. 1–13, Mar. 2018. [Online]. Available: https://link.springer.com/article/10.1007%2Fs12652-018-0746-y
[50] M. Espinilla and C. Nugent, "Computational intelligence for smart environments," *Int. J. Comput. Intell. Syst.*, vol. 10, no. 1, pp. 1875–6883, 2017.
[51] J. Quero, C. Orr, S. Zang, C. Nugent, A. Salguero, and M. Espinilla, "Real-time recognition of interleaved activities based on ensemble classifier of long short-term memory with fuzzy temporal windows," *Proceedings*, vol. 2, no. 19, p. 1225, Oct. 2018.
[52] P. Asghari, E. Soleimani, and E. Nazerfard, "Online human activity recognition employing hierarchical hidden Markov models," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 3, pp. 1141–1152, Mar. 2020.
[53] N. C. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," *Pervasive Mobile Comput.*, vol. 10, pp. 138–154, Feb. 2014.
[54] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, p. 33, 2014.
[55] D. Bouchabou, S. Nguyen, C. Lohr, B. Leduc, and I. Kanellos, "Fully convolutional network bootstrapped by word encoding and embedding for activity recognition in smart homes," 2020, *arXiv:2012.02300*. [Online]. Available: http://arxiv.org/abs/2012.02300
[56] D. Cook, M. Schmitter-Edgecombe, A. Crandall, C. Sanders, and B. Thomas, "Collecting and disseminating smart home sensor data in the CASAS project," in *Proc. CHI Workshop Develop. Shared Home Behav. Datasets Adv. HCI Ubiquitous Comput. Res.*, 2009, pp. 1–7.
[57] M.-S. Yang and K. P. Sinaga, "A feature-reduction multi-view K-means clustering algorithm," *IEEE Access*, vol. 7, pp. 114472–114486, Aug. 2019.
[58] K. P. Sinaga, I. Hussain, and M.-S. Yang, "Entropy K-means clustering with feature reduction under unknown number of clusters," *IEEE Access*, vol. 9, pp. 67736–67751, May 2021.
[59] M.-S. Yang and K. P. Sinaga, "Collaborative feature-weighted multi-view fuzzy c-means clustering," *Pattern Recognit.*, vol. 119, pp. 1–15, May 2021.
[60] M.-S. Yang and Y. Nataliani, "Robust-learning fuzzy c-means clustering algorithm with unknown number of clusters," *Pattern Recognit.*, vol. 71, pp. 45–59, Nov. 2017.
[61] D. J. Cook, N. C. Krishnan, and P. Rashidi, "Activity discovery and activity recognition: A new partnership," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 820–828, Jun. 2013.
[62] H. Gjoreski and D. Roggen, "Unsupervised online activity discovery using temporal behaviour assumption," in *Proc. ACM Int. Symp. Wearable Comput.*, 2017, pp. 42–49.
[63] B. Quigley, M. Donnelly, G. Moore, and L. Galway, "A comparative analysis of windowing approaches in dense sensing environments," *Proceedings*, vol. 2, no. 19, p. 1245, Oct. 2018.
[64] G. D. Ruxton, "The unequal variance t-test is an underused alternative to student's t-test and the Mann–Whitney U test," *Behav. Ecol.*, vol. 17, no. 4, pp. 688–690, Jul. 2006.

**YU MIN HWANG** (Member, IEEE) received the B.S. and Ph.D. degrees from the Department of Wireless Communications Engineering, Kwangwoon University, Seoul, South Korea, in 2012 and 2018, respectively. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, Western University, Canada, from 2019 to 2020. He is currently with the Electronics and Telecommunications Research Institute (ETRI), South Korea. His research interests include the Internet of Energy, deep learning, activity recognition, and anomaly detection.

**SANGJUN PARK** (Member, IEEE) received the B.S. degree in computer engineering from Chungnam National University, Daejeon, South Korea, in 2009, and the Ph.D. degree in electrical engineering and computer science from Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2019. Since 2020, he has been working at the Electronics and Telecommunications Research Institute. His research interests include information theory, numerical optimization, compressed sensing, blockchain, deep-neural networks, and finite state machine.

**HYUNG OK LEE** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer engineering from Chonnam National University, Gwangju, South Korea, in 2006, 2008, and 2015, respectively. Since 2016, he has been a Senior Member of Engineering Staff with the Electronics and Telecommunications Research Institute (ETRI), South Korea. His current research interests include micro service, the energy IoT, and big data.

**SEOK-KAP KO** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in information telecommunication engineering from Soongsil University, South Korea, in 1997, 2002, and 2009, respectively. Since 2008, he has been the Principal Research Engineer with the Electronics and Telecommunications Research Institute (ETRI), South Korea. His research interest includes machine learning for energy management systems.

**BYUNG-TAK LEE** (Member, IEEE) received the B.S. degree from Yonsei University, Seoul, South Korea, in 1992, and the M.S. and Ph.D. degrees from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 1994 and 2000, respectively. From 2000 to 2004, he was the Principal Research Engineer with LG Electronics, where he was engaged in the area of Tbps optical transmission systems. Since 2005, he has been the Principal Research Engineer with the Electronics and Telecommunications Research Institute (ETRI), South Korea. His current research interests include the Internet of Things, deep anomaly detection, and system optimization.