Inferring Work Routines and Behavior Deviations with Life-logging Sensor Data

Shohreh Deldari
RMIT University
Melbourne, Australia
shohreh.deldari@student.rmit.edu.au

Jonathan Liono
RMIT University
Melbourne, Australia
jonathan.liono@rmit.edu.au

Flora D. Salim
RMIT University
Melbourne, Australia
flora.salim@rmit.edu.au

Daniel V. Smith
Data61, CSIRO
Hobart, Australia
Daniel.V.Smith@data61.csiro.au

ABSTRACT

Recently human activity recognition has encouraged a great deal of interest due to its impact on various areas of application. As human's brain own ability to recognize the actions relies upon obtaining information from a number of senses, not only our vision system; we propose to enhance vision-based activity recognition systems by integrating additional contextual temporal data being sensed from ubiquitous sensors as well. In this paper, we focus upon exploiting time series to extract a user's daily life patterns and identify any deviations in the temporal patterns. First, we apply Information Gain Temporal Segmentation (IGTS) method, a generic and robust temporal segmentation approach for heterogeneous time series data. From the output of temporal segmentation, daily working patterns are extracted for which unusual user behavior is identified automatically based on high deviation scores of days from the reference transition times. Finally, we evaluate the accuracy of identified days that have such behavior deviations by comparing them with the visual dataset as the ground truth.

CCS CONCEPTS

• **Computing methodologies → Classification and regression trees**; Multi-task learning; Semi-supervised learning settings;

KEYWORDS

Task Intelligence; Temporal Segmentation; Activity Recognition; Behavior Deviations; Work Routines;

ACM Reference format:
Shohreh Deldari, Jonathan Liono, Flora D. Salim, and Daniel V. Smith. 2019. Inferring Work Routines and Behavior Deviations with Life-logging Sensor Data. In *Proceedings of ACM International Conference on Web Search Data Mining, 2019, ACM International Conference on Web Search Data MiningTask Intelligence Workshop @ WSDM (Task Intelligence Workshop @ WSDM),* 6 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Task Intelligence Workshop @ WSDM, 2019

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Human activity recognition is the problem of classifying the movement, actions and tasks of an individual using only the data collected from sensors, which could include cameras and other wearable devices such as motion sensors, biometric sensors, global positioning systems and microphones. Such knowledge can be utilized to develop personalized and interactive technologies that assist (i.e. home automation systems), inform (i.e. health care alerts) or persuade (i.e. recommendation systems) users in their daily life. The activity could be low-level (e.g., sleeping, walking or running) or high-level, which is an aggregate of several low-level activities with a more complex semantic meaning (e.g., cooking, working at the office, shopping or commuting)[7]. We may call these high-level activities as task since each task is a combination of several low-level activities. Although there are lots of research on recognizing low-level activities, task recognition is still a challenging issue. The problem is that these high-level activities are represented by complex motions that can only be adequately captured by employing multiple sensors. Consequently, high-level activities often need to be represented as multiple time series where each time series represents a simple but different motion task.

Therefore, due to the heterogeneity of the input time series, the inherent noisy nature of the input, the high sampling rate and the lack of a clear way to relate different input sources to known movements, task recognition is a challenging problem [2].

The other issue is processing videos and images taken by cameras mounted on the agent's body to track and understand their behavior. Vision-based activity recognition has found many usecases. However, despite remarkable progress of image-based activity recognition, its usage for most applications remains a distant aspiration [8]. In contrast, the human brain's ability to recognize actions mainly relies on extracting information from a number of senses not only our vision system. Based on this observation, we propose to enhance vision-based activity recognition systems by integrating additional time series data that is being sensed from ubiquitous sensors on the body[3]. Computer vision based recognition approaches are typically applicable to image-processing techniques to provide activity segmentation based on the semantic correlation between consecutive images or video frames[9].

In this study, we focus using the time series to extract daily lifestyle patterns of the user based upon temporal deviations. Furthermore, we investigate whether integrating image and time series data can produce more accurate task segmentation results than using images alone. We focus upon utilizing body metric sensors to find routine patterns for regular working and non-working days. In this study, we exploit one of the state of the art segmentation methods proposed by A. Sadri, et al [7]. They have proposed an Information Gain based Temporal Segmentation (IGTS), an unsupervised technique to extract the transition times associated with different daily tasks from multi-variant time series.

Furthermore, by annotating the images and categorizing them into five different task categories, we compare to our segmented results and validate our hypothesis that aggregating image and video data with other contextual knowledge can not only accurately label high-level activities but also can provide a reliable intuition about low-level activities.

The next sections are organized as follows. Section 2 defines the problem and we explore the data-set in section 3. Section 4 discusses the applied multiple time series temporal segmentation method (IGTS), patterns extracted from the biometric data and the effectiveness of aggregating body metric time series along with the cameras. In section 5, we provide a brief conclusion and outline our suggestions for future work.

2 PROBLEM DEFINITION

The main goal of our experiments is to extract specific activity patterns across a day or part of the day in order to classify each day as either a working, non-working, weekend, or break day with the aim of determining the user's working behavior. In order to compare the accuracy of our results, we labeled each image of the **LSC2018** dataset with a tag.

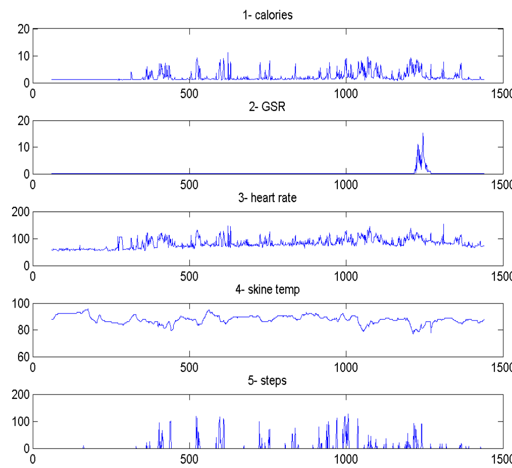
Since we are only interested in working behavior, we only consider the following user tasks: walking, driving, working, and eating. The images that do not belong to any of these four task categories are considered as "other", which may include both low-level and high-level activities, such as running, socializing with a friend, sleeping, watching television, and etc. The annotation process is fully described in section 3. It should be noted that we annotate these sequences of activity when the user is not at home, and hence, we are focusing upon the tasks that could relate to the user's working routines.

3 DATASET AND PRELIMINARY ANALYSIS

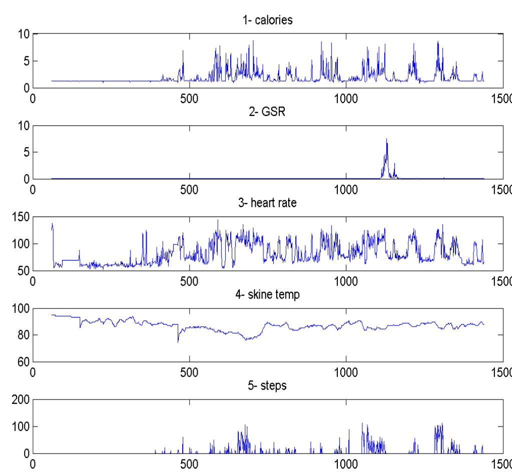
In this study, we are using the existing lifelogging dataset from the NTCIR-13 Lifelogging track [4], comprising 27 days of logs, with over 1,500 images per day. The dataset consists of wearable camera data, physical activity data, semantic locations, and human biometrics. There is also an associated visual concept detector that accompanies the images[1].

At first, we focused on the biometric data that has been collected from 5 different sensors measuring heart rate, skin temperature, galvanic skin response, calorie burn, and number of steps. Moreover, there are readings of Blood Pressure and Blood Sugar levels. Unfortunately, these readings are only available for the daily summary. Therefore, we exclude them from our experiment as we only consider sensor data that can be projected in a time series manner within a day.

We created a representative array for each individual day based on these five sensors extracted from the associated metadata file. Consequently, there are five different time series that we can plot and observe per day. It should be noted that the granularity of the sensor readings are limited to minutes precision, based on the given metadata in this life-logging data challenge. Figure 1 shows the aforementioned parameters over almost 1500 minutes from waking up in the morning to sleep time across two sample days.



(a) Biometrics captured on 15/8/2016



(b) Biometrics captured on3/10/2016

Figure 1: Time series of five biometric sensors for two random sample days.

As shown in Figure 1, the GSR values have been captured for a limited time duration each day. Therefore, we omit this sensor data from our experiment as it is not a continuous time series that can directly characterize the work routines. In addition, by exploring all of the parameters over the 27 days, we realized that there were some missing values for each of the sensors. We ignore the time points which had no values across all of the sensors. When particular sensors had missing values, we filled those values with the preceding values in its time series.

In order to understand the working routine of the user in the life-logging dataset, we define the following task categorizations that can be identified from the image data (captured via the camera held by the user during the data collection):

(1) *Driving*: Images are considered as driving if the user is sitting in the car.
(2) *Walking*: Images are annotated as *Walking* if it is inferred that the user is walking considering the preceding and proceeding images.
(3) *Working*: It is a more difficult category to annotate given it is a high-level task that can happen in a wide variety of situations, places and times. Images would be categorized as working if the user is in a place other than home (i.e. cafe, airport and other obvious non-working areas) and is working with computer/laptop or is having a meeting with colleagues. In this phase, the correctness of the labeling is subject to the annotator's perception.
(4) *Eating*: Identifying this category is more straightforward. Any images that specifically contain a kind of food, is classified as *Eating*.
(5) *Other*: Every other image that is not classified as one of the above categories, gets this label.

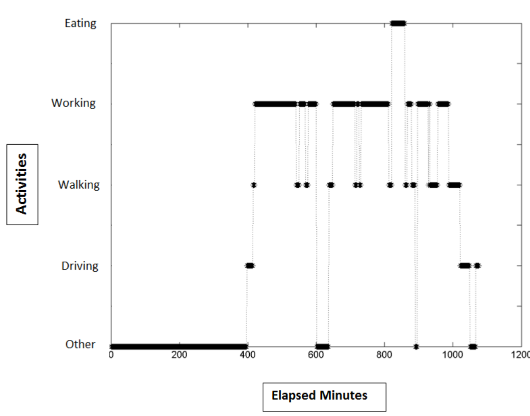


Figure 2: Annotating a random day based on the five different task categories (Other, Driving, Walking, Working, and Eating).

In order to construct the ground-truth of task labelling, three independent volunteers manually annotated all of the images from the 27 days. Figure 2 represents the division of a randomly selected day based on the different task categories mentioned above. Categories *Other*, *Driving*, *Walking*, *Working*, and *Eating* for a random day are shown in the diagram in the first to the fifth level respectively.

By exploring the data set, we found that there were seven off-days (including weekends and public holidays). Consequently, these days were used to model the weekend patterns.

4 INFERRING BEHAVIOUR DEVIATIONS OF WORK ROUTINES

4.1 Information Gain-based segmentation

In order to observe the working patterns, a robust temporal segmentation method is required to deal with the heterogeneous time series data. Hence, we used a state-of-the-art segmentation technique based on Information Gain theory. According to information theory, entropy reflects uncertainty. In our case, the entropy refers to the predictability of the user's bio-metric condition. The IGTS algorithm finds low-entropy segments by using a cost function that identifies the coherent segments in the time series. As a result, the best segmentation is the one with highest information gain value[7].

In this paper, we apply the IGTS method to heterogeneous input channels recorded by mounted biometric sensors to detect the transition times of the human activities. According to our intuition, this method should be able to help in capturing the following activities of *Walking*, *Driving*, *Working*, Eating and *Other*. Nevertheless, we aim to extract a reliable user-specific daily pattern, in order to distinguish between working and non-working behaviour.

4.2 Extracting the Routine Patterns

As shown in Figure 1, the GSR parameters have been recorded in a sparse manner. Joint analysis GSR time series with the other four metrics do not provide an informative result, especially for temporal segmentation. Therefore, we excluded the readings of GSR from our segmentation process.

In short, we consider only following sensor data, that are sampled every minute, as input for our segmentation method:

- Heart rate
- Calories burnt
- Number of steps
- Skin temperature

According to [7], it is necessary to specify whether we are interested in finding the daily transitions in high-level activities or low-level activities. Hence, the dynamic programming implementation of the IGTS method is applied to the **LSC2018** dataset, which contains both high-level and low-level activity labels. This lifelogging dataset captures both sensing data and images from when the user wakes up until when before going to sleep. For the IGTS method, the granularity of recognized activity segments is highly dependent upon the given parameter **K** (the number of segments that are specified to partition the daily time series).

In our experiment, we observe the patterns of time series by performing sensitivity analysis on various values of parameter **K**, ranging from 4 to 11. Considering the limited number of task categorizations, a set of smaller value of **K** would be reasonable for our analysis.

Figure 3 shows the average extracted transition times for four to eleven segments per day for both weekdays (3a) and weekends (3b). As we increase the number of segments in a day, it means we will have more granularity in the detected tasks.

There are several distinct differences in transition time patterns between the weekdays and weekends. In a typical day, the data captured from this user could start at around 4:00 am. It should be noted that this start time is not necessarily the time that user wakes up in the morning, but the time that salient changes in the values of biometric parameters have been detected.

By exploring the raw images along with the metadata[1], we could identify that these pre-wake up segments are produced due to reasons of the user's biological clock (in which it gradually awakes the person). Moreover, higher values of heart rate or even calories burnt are noticeable in an early morning in contrast to the segments that include mid-night time slot. The first transition time (where $K = 4$) in a typical working day occurs at around 4:00 am, whereas the first transition time is shifted to around 5:00 AM on the weekends. In this case, the user is likely to wake up later in the weekends due to no obligation to come to the office, which may result in a deeper sleep state (from Friday night to Saturday morning).

As shown in Figure 3, the last transition time (for $8 \leq K \leq 11$) on weekdays and weekends are around 17:30 and 20:00, respectively. Therefore, the user usually sleep later on weekends.

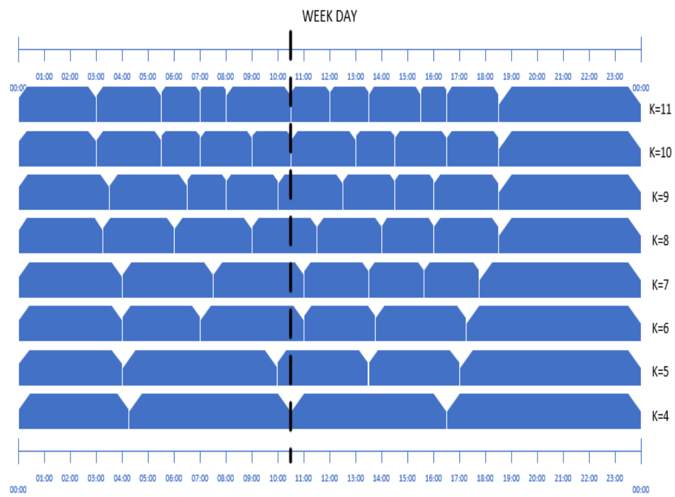
Alternatively, as a distinguishing highlight, there is another important difference that we could observe based on the distribution of the suggested transition times on weekdays and weekends. The vertical black line in both diagrams represents the point that the average number of task transitions before this line is almost equal to the number of transition times occur after that. It means we expect a higher density of transition times before 11:00 AM and after 13:00 PM for weekdays and weekends respectively. Consequently, this can lead us to several conclusions on the daily routine pattern for this specific user. First of all, the user may have a tighter schedule in the morning during the weekdays. On the other hand, the user could be busier in the afternoon of the weekends. He may also prefer to have a more flexible schedule at the weekends and relax.

4.3 Deviation From Reference Pattern

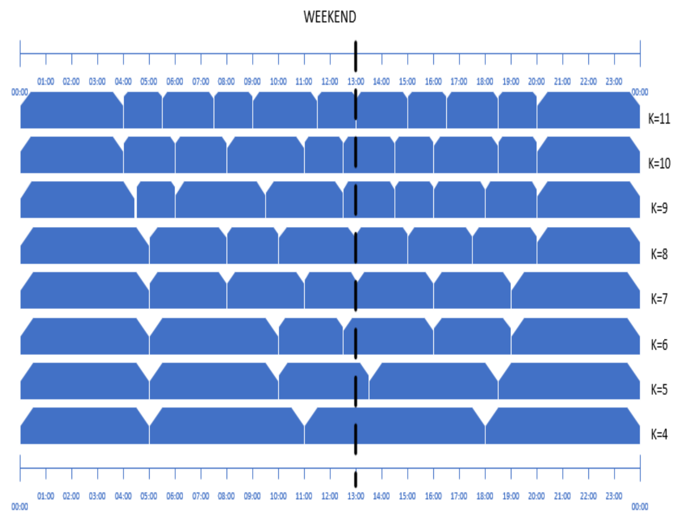
In order to derive the behavior deviations from the work routines of the user in a lifelogging scenario, we present a threshold-based technique to identify unusual patterns automatically, by leveraging a certain dissimilarity score function. In this paper, we designed a simple principled scoring metric. In this case, we assume that a day is considered unusual if the deviation of the extracted transition times from the reference pattern is higher than a predefined threshold.

To explain our dissimilarity scoring function, we begin with the following definitions of *Reference Routine* and *Transition Times*.

[1] file:LSC2018$_{m}etadata.xml$



(a) Average Transition Times on a weekday



(b) Average Transition Times on the weekend

Figure 3: Average Transition Times extracted based on different number of segments over all 27 days.

Definition 4.1. Transition Times$_k$ is a vector of integers: $TT_k = [t_1, t_2, ..., t_k - 1]$, where $t_i, i \in [1, K)$ is the *i-th* transition time for the given set of time series and so on. **K** indicates the number of segments which is one more than the number of transitions.

Definition 4.2. Reference Routine, RR, is a set of TT for the referenced routine pattern:

$$RR = \{RR_i \mid 1 < i < K\},$$

where K is the number of segments and RR$_i$ is *Transition Times* of the reference model.

As mentioned before, we have extracted *Transition Times* for **K** between 4 to 11. According to previous sections, we calculated the Reference Routine for both Weekday and Weekends based on all sets of *Transition Times* for whole set of the days (Figure 3a and 3b).

For each day of the lifelogging dataset, we use the following concept of dissimilarity to identify the behavior deviations automatically:

Definition 4.3. A *Dissimilarity Matrix D* of a given set of time series is a vector of **0**s and **1**s which each value indicates that the standard deviation corresponding TT's is higher than a predefined threshold. we consider a day as inconsistent if all the values of the D is set to **1**.

Here we defined the threshold as the average of deviation through all the dataset. The *Dissimilarity Matrix D* is calculated based on the following algorithm:

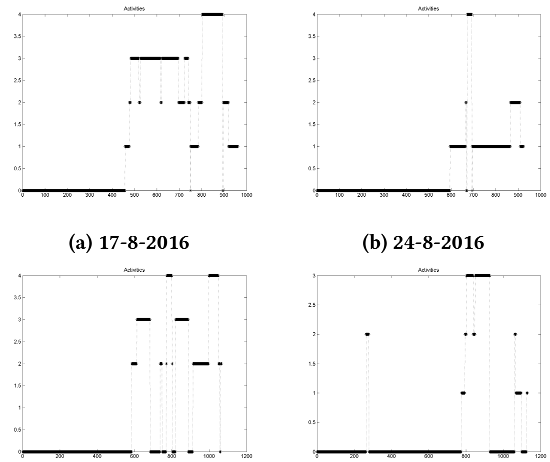
```
//Input = RR, TT, K
//output = D
 initialization\;
 D = zeros(1,K);   //Initialize Dissimilarity Matrix
 for  k in K   //Iterate over all nember of segments
    for  i = 1 to k   //Iterate over all ...
          suggested transition times
    //calculate Standard Deviation of ith ...
          transition time in all k−segment TTs.
     TTdev(i) = ...
           calculateDeviationFromRR(TT(k,i),RR(k));
     end
avg_dev(k) = average(TTdev);    //Average over ...
      calculated deviations
end
D(avg_dev(:) > threshold) = 1;
```

Comparing daily logs with the extracted routine pattern in Figure 3 lead us to identify some of the days that has an unusually higher deviation from the reference transition times. Table 1 lists the days that include unusually high deviation from the reference transition times. The reasons behind this deviation are derived from investigating the extracted ground truth based on image dataset. Figure 4 shows the related activity annotations for all inconsistent day. Because there was not any overlap between the images that were judged by annotators, there is no agreement between the three judges for the manual image labelling task. So we did not extend our comparison because of the reliability issue.

Table 1: List of all inconsistent days

Date	Reason of deviation
17-8-2016	Quit his work earlier and had shopping and lunch
24-8-2016	He did not go to work.
29-8-2016	caught a bus instead of driving
30-8-2016 and 8-9-2016	He had a flight to somewhere and get back to work again.

Although the proposed metric that we use for this paper is based on the pure deviation from dissimilarity matrix D, other scoring function can be leveraged, such as Kullback-Leibler (KL) divergence score [5] to address more complex problems for interleaving and overlapping of multiple human activities in a task. In [6], the divergence score is used to maximize the class separability for temporal segmentation of multiple human activities.



(a) 17-8-2016 **(b) 24-8-2016**

(c) 29-8-2016 **(d) 30-8-2016**

Figure 4: Activity annotation for the inconsistent days.

5 CONCLUSION AND FUTURE WORK

In this paper, we present a pipeline to extract daily work routine patterns and identify behavior deviation from these patterns. In our experiment, we used an Information Gain-based approach and consider the distribution of different heterogeneous input channels to identify the best task boundaries. Then, we extracted a routine pattern for working and non-working days with the objective to find behavior deviations. From the threshold-based technique that we propose in this paper, all of the inconsistent days can be identified automatically. Moreover, behavior deviations on these inconsistent days are then aligned and validated manually with the ground-truth image data. Conclusively, the outcome of our experiment on lifelogging data can be beneficial to enhance image-based approaches for task identification, given that the deviations can be identified seamlessly from non-vision (i.e. biometric) sensor data.

Although this paper is limited to the identification of behavior deviations in work routines, there are many immediate challenges that should be addressed for future work. Given that a task is composed of low-level activities, identifying the new task and recognize these activities are non-trivial problems. To tackle these problems, reliable methods need to be designed to fuse the features from images and wearable data to extract representative patterns for detecting intents, and possibly information needs that can be concluded from such activities. Our experiment results provide early indications that models trained using visual and temporal data can lead to more accurate results than those using visual data alone. Integrating these aspects of task recognition will be considered in future works.

There are plenty of jobs where people's activity changes on a daily basis. Although it may be thought that this method is suitable for only situations with predictable routines, we believe that the idea of merging those auxiliary data along with images can be significantly informative for the situations that do not conform to a regular pattern.

ACKNOWLEDGMENTS

This work was supported by Data61, an arm of the Commonwealth Scientific and Industrial Research Organisation (CSIRO), under Grant No.: 500588.

The authors would like to thank Dr. Amin Sadri for providing the MATLAB code of the *IGTS* method.

The authors would also like to thank Lissa, Siro, and Erica for volunteering in annotating the task labels on the image data.

Finally, we want to thank our reviewers for their valuable comments.

REFERENCES

[1] 2018. LSC 2018 @ ICMR 2018. http://lsc.dcu.ie/index.html

[2] Jason Brownlee. 2018. A Gentle Introduction to a Standard Human Activity Recognition Problem. https://machinelearningmastery.com/how-to-load-and-explore-a-standard-human-activity-recognition-problem/

[3] Jesús Martínez Del Rincón, Maria J Santofimia, and Jean-Christophe Nebel. 2013. Common-sense reasoning for human action recognition. *Pattern Recognition Letters* 34, 15 (2013), 1849–1860.

[4] Cathal Gurrin, Klaus Schoeffmann, Hideo Joho, Bernd MÃijnzer, Rami Albatal, Frank Hopfgartner, Liting Zhou, and Duc Tien Dang Nguyen. 2019. *A Test Collection for Interactive Lifelog Retrieval: 25th International Conference, MMM 2019, Thessaloniki, Greece, January 8âĂŞ11, 2019, Proceedings, Part I.* 312–324. https://doi.org/10.1007/978-3-030-05710-7_26

[5] Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics* 22, 1 (1951), 79–86.

[6] Jonathan Liono, A Kai Qin, and Flora D Salim. 2016. Optimal time window for temporal segmentation of sensor streams in multi-activity recognition. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services.* ACM, 10–19.

[7] Amin Sadri, Yongli Ren, and Flora D Salim. 2017. Information gain-based metric for recognizing transitions in human activities. *Pervasive and Mobile Computing* 38 (2017), 92–109.

[8] Allah Bux Sargano, Plamen Angelov, and Zulfiqar Habib. 2017. A comprehensive review on handcrafted and learning-based action representation approaches for human activity recognition. *applied sciences* 7, 1 (2017), 110.

[9] Le Wang, Xuhuan Duan, Qilin Zhang, Zhenxing Niu, Gang Hua, and Nanning Zheng. 2018. Segment-Tube: Spatio-Temporal Action Localization in Untrimmed Videos with Per-Frame Segmentation. *Sensors* 18, 5 (2018), 1657.