IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 30, 2018, accepted January 22, 2019, date of publication February 1, 2019, date of current version February 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2895874

Harnessing Commodity Wearable Devices to Capture Learner Engagement

YU LU, JINGJING ZHANG, BAOPING LI, PENGHE CHEN, AND ZIJUN ZHUANG
Advanced Innovation Center for Future Education, School of Educational Technology, Faculty of Education, Beijing Normal University, Beijing 100875, China

Corresponding authors: Jingjing Zhang (jingjing.zhang@bnu.edu.cn) and Penghe Chen (chenpenghe@bnu.edu.cn)

This work was supported in part by the Humanities and Social Sciences Foundation of the Ministry of Education of China under Grant 17YJCZH116, in part by the National Natural Science Foundation of China under Grant 61702039 and Grant 61807003, and in part by the Fundamental Research Funds for the Central Universities.

ABSTRACT With the aim of capturing learner engagement, we propose and advocate the use of commodity wearable devices and their built-in sensors (e.g., accelerometer, gyroscope, and magnetic sensor) to detect the fine-grained learning activities (e.g., writing notes or raising the hand in class). Next, by leveraging the established theory that links learner activities to learner engagement, the detected learner activities can be used to further infer the learner engagement levels, durations, and other key information. We thus designed a hassle-free and non-intrusive system running on the latest wrist-worn commodity wearable devices, which adopts the latest activity recognition and sensor data fusion techniques. We conducted the system-level evaluation, survey, and interviews involving both students and teachers. The evaluation results show that our system can accomplish the accurate learner activity recognition task, and meanwhile effectively capture the learner engagement. We also provide the engagement-based intervention service during class to illustrate the unique usefulness of the proposed system.

INDEX TERMS Educational technology, wearable sensors, learner engagement, activity recognition.

I. INTRODUCTION

Learner engagement (or student engagement) often refers to ''the quality of effort students themselves devote to educationally purposeful activities that contribute directly to desired outcomes'' [1]. This concept is closely related to *active learning* in the way in which students *engage* cognitively with learning material [2] rather than passively receiving the knowledge [3]. As a result, highly engaged students are more likely to improve their academic performance and learning outcomes [4]. Thus, understanding student engagement is key to the evaluation of the learning outcomes [5], yet accurately and effectively assessing student engagement is still one of the challenging and difficult tasks that requires further investigation [6], [7].

Recent studies [8] have shown that the extent to which a learner engages in a learning task can be reflected in the overt activities that the learner voluntarily exhibits and undertakes. That is, the level of student engagement can be explored through observation of the fine-grained student learning behaviors, such as taking notes or asking questions.

In practice, experienced teachers understand the engagement of their students from daily observations of their behaviors. Nevertheless, understanding the engagement resulting from teachers' daily observation is not completely satisfactory. Important behaviors often go unnoticed under circumstances where a large number of students behave differently at the same time, and even the observed behaviors might not be accurately interpreted. Efforts are sought to solve these problems given the fact that the latest sensor technology with machine learning techniques may overcome the inability of traditional observation that is adopted in educational studies. The schools and industry players, espcially in K-12 education, also attempt to provide proper solution to tackle the challenges:

- *In what ways and to what extent student learning behaviors can be properly recorded and recognized simultaneously?*
- *How to utilize the captured learning behaviors to determine the engagement level of individual learners?*

The rapid advances in sensor-based activity recognition techniques and the availability of the commodity wearable devices offer a possible solution to the questions raised above. Commodity wearable device refers to smart electronic

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

equipment worn by users, including wrist-worn devices (e.g., Apple's smartwatch [9] or Fitbit's wristband [10]) and eyewear devices (e.g., Google glass [11]). Different from the dedicated or customized equipment (e.g., smart textiles [12]), commodity devices usually can be easily purchased from the market at a relatively low price. On the other hand, many built-in sensors, such as accelerometer, magnetic and gyroscope sensors, together with a powerful multicore processor have become the de facto components of most commodity wearable devices, which can be used to conduct human activity recognition and complicated data analytics tasks. In addition, commodity wearable devices are usually equipped with Wi-Fi or 3G/4G modules to support wireless communication with the backend servers and cloud. Compared with the vision-based solution that employs single or multiple cameras to capture human activities, wearable devices do not require any fixed infrastructure and thus have almost no maintenance cost. Moreover, the vision-based systems often suffer from low performance even in natural lighting conditions [13].

Driven by the above research questions and motivated by the latest wearable technology, we propose and advocate the use of commodity wearable devices and sensor-based activity recognition techniques to identify overt learning activities, and subsequently infer the learner engagement level. The captured learner engagement level information can be directly used for proper intervention and service design. We accordingly implement such a system on the commodity wearable device, which provides a hassle-free and non-intrusive solution to achieve the above objectives. The main innovations in the system include (i) the use of real time data from multiple built-in sensors to accurately detect multiple types of learner physical activities, (ii) the use of the detected physical activities to infer learner's engagement states, and (iii) the aggregation of the activity and engagement information to provide novel interventions to learners. Finally, we carried out the system evaluation in the classroom environment with both qualitative and quantitative approaches.

The rest of the paper is organized as follows: Section II introduces the related work. Section III depicts the system design. In section IV, we describe the evaluation design, along with the empirical evaluation results in section V. The discussion is given in section VI, and we finally conclude in section VII.

II. RELATED WORK

A. STUDENT ENGAGEMENT AND ICAP FRAMEWORK

Drawing on Bloom's taxonomy of educational objectives [14], student engagement has enjoyed considerable attentions from academia and educators [5], [15], [16]. In general, it can be defined as "the quality of effort students themselves devote to educationally purposeful activities that contribute directly to desired outcomes" [1], or as "the extent to which students are engaging in activities that higher education research has shown to be linked with high-quality learning outcomes" [17]. Following these definitions, a sound body of literature has established the positive correlations between student engagement and learning outcomes [4], [18]. Meanwhile, researchers identify several dimensions of the student engagement, typically including behavioural engagement, emotional engagement and cognitive engagement, where cognitive engagement is the key dimension influencing learning outcomes [19]. Simply speaking, cognitive engagement refers to the student cognitively investing in their learning, seeking to go beyond requirements and relishing challenges [20].

Researchers [8] recently proposed the so-called ICAP framework (denoting *Interactive*, *Constructive*, *Active*, and *Passive*), and it reveals that the amount of cognitive engagement can be assessed by observation of the fine-grained learning behaviors. In other words, the ICAP framework shows that different overt student activities can reveal their different engagement levels. As explained by the authors [8], "although far from perfect, overt behaviors are a good proxy to reflect different modes of engagement," and accordingly the ICAP framework proposes four modes to categorized student overt activities and the corresponding engagement levels.

1) **Passive Mode**: *learners are oriented toward and receive information from the instructional materials without doing anything else*. For example, a student passively listening to the lecture without any other actions.
2) **Active Mode**: *learners undertake some form of overt motoric action or physical manipulation*. For example, a student quickly taking down lecture notes.
3) **Constructive Mode**: *learners generate or produce additional externalized outputs or products beyond what was provided in the learning materials*. For example, a student in class raising hand to ask or answer questions.
4) **Interactive Mode**: *learners interact with partners and meet two criteria: a) both partners' utterances must be primarily constructive; b) a sufficient degree of turn taking must occur*. For example, a student conducting a stimulating group discussion.

The above four modes, from *passive* to *active* to *constructive* to *interactive*, describe the student engagement levels from low to high. Moreover, the ICAP framework also shows that the corresponding cognitive outcomes increase as the engagement level increases, which elicits different knowledge-change and learning process. The ICAP framework has been widely adopted for the student engagement assessment [21], [22]. In this work, we also adopt this established framework to link student activities to student engagement status.

B. WEARABLE TECHNOLOGY FOR EDUCATION

Wearable device usually refers to electronic equipment that can be directly worn on user's body, typically including wrist-worn smart watches. They have been introduced and applied for education context and learning analytics [23]–[25]. Gua *et al.* [26] utilized wearable devices data to assist in language learning for young children. Ngai *et al.* [27] deployed a wearable computing platform for computing education. In addition, Muller *et al.* [28] employed wearable sensors to enable reflective learning for employees. Compared with traditional vision-based systems and technology, wearable devices do not need any fixed infrastructure, which incurs a high maintenance cost over time and suffers from low performance in natural lighting conditions [13].

On the other hand, high-tech companies such as Apple and Google have launched a variety of commodity smart wearable devices with desired functionalities [9], [11], typically including sleep monitoring and fitness tracking. For example, Adidas has produced a wrist-worn device for school children to track their fitness [29]. Such commodity wearable devices usually equip with a high sensing and computational capability, and thus they can be an ideal platform to collect the information from the learner side and conduct analytics for the educational purposes.

C. SENSOR-BASED ACTIVITY RECOGNITION

Rapid advancements have been made in sensor-based activity recognition [30], [31]: recent studies have utilized versatile built-in sensors on wearable devices, such as the three-axis accelerometer and gyroscope, to understand shopping gestures [32], eating activities [33] and social context [34]. From the perspective of data science, the latest data processing and machine learning techniques [35] have been employed to identify distinct human activities. In this work, we adopt the similar techniques, but innovate in recognizing learning activities for the purpose of engagement capture.

Among the previous studies, the most relevant to our work in terms of data type and activity recognition is [36]: it identifies teacher's physical actions, such as explanation and questioning, using multiple wearable sensors (eye-tracking, electroencephalogram, accelerometer, audio, and video) during classroom enactment. However, our work mainly focuses on the student side and utilizes commodity wearable devices in tandem with the established framework to capture and understand learner engagement.

III. SYSTEM DESIGN

Continuously and concurrently collecting student engagement information in class at the individual level is significantly useful but also critical: the engagement information would not only help to understand the amount of effort that each student invests, but also help teachers to provide the necessary interventions and feedback. It is difficult to be conducted by using a teacher's personal observation or using a vision-based system with cameras, which suffers from low detection accuracy and high maintenance cost. Hence, in this paper, the system design mainly targets on providing an effective solution to tackle this problem.

The block diagram of the proposed system is illustrated in Figure 1, and its hierarchical architecture consists of three modules, from the bottom to the top, namely *sensor data collection module*, *activity recognition module* and *engagement inference module*.

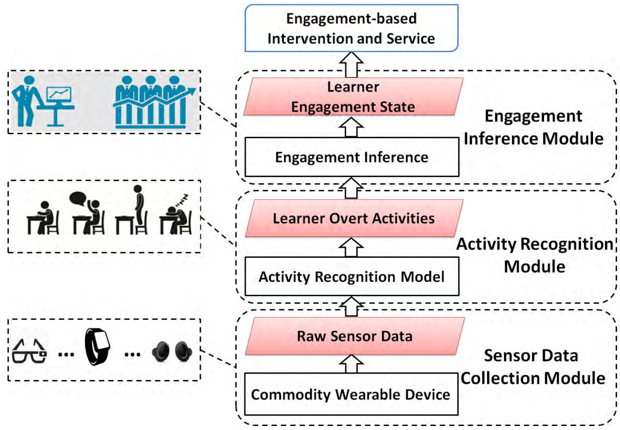


FIGURE 1. Block Diagram of the System.

A. SENSOR DATA COLLECTION MODULE

Running at the bottom of the system, the sensor data collection module mainly conducts the raw sensor data collection from the commodity wearable device hardware, such as wrist-worn devices. These commodity wearable devices are usually equipped with multiple built-in sensors. Taking wrist-worn devices as an example, on-device sensors typically include *accelerometer sensor* (measuring the linear acceleration along three-axis), *gyroscope sensor* (measuring the rotational accelerations along the three-axis), and *magnetic sensor* (measuring the direction and strength of the magnetic field along the three-axis). In addition to the above-described sensors, other sensors such as heart rate sensors and temperature sensors are also often embedded in the latest commodity wearable devices.

We currently choose the wrist-worn device among different commodity wearable devices available on the market, as this work mainly focuses on the student behavior and engagement in classroom, where the wrist-related physical actions are crucial (e.g., writing notes, hand-up and hand-down). In practice, we adopt one type of smartwatch called Ticwatch [37], which is equipped with the desired sensors, including accelerometer, gyroscope, and magnetic sensors. Moreover, it is also equipped with a heart rate sensor, location sensor (GPS), and powerful communication interfaces (e.g., cellular 3G/4G modules). Ticwatch is currently running android operating system, and provides the open SDK [38] for developers. Figure 2 shows the hardware of the selected Ticwatch smartwatch, whose price is less than US$200 each.

In short, this module is mainly responsible for collecting the rich but raw sensor data from multiple on-device sensors, and then sends them to the upper activity recognition module. Note that collecting the raw sensor data needs permission from the learner side, even the data would not be used for any commercial purpose.

FIGURE 2. Hardware of the System Device.

B. ACTIVITY RECOGNITION MODULE

This module gathers the raw sensor data from the sensor data collection module, and then performs the learner activity recognition tasks. Briefly speaking, it first extracts the key features from the pre-processed sensor data on both the time domain and frequency domain. Using the extracted features, it adopts the specifically designed machine learning models to identify student overt learning activities (e.g., student raising a hand or writing notes in class). The identified student overt activities are sent to the upper engagement inference module in real time.

1) FEATURE EXTRACTION

Once the raw sensor data collection is triggered, multiple types of sensors, such as accelerometer, gyroscope, and magnetic sensors, are activated. The system would first apply a low-pass filter to the raw sensor measurements to remove the jitters and noises from the raw sensor data. After such a sensor data preprocessing step, the system segments the data into non-overlapping fixed-size frames, and computes the features. They typically include both time-domain features (e.g., mean and variance), and frequency-domain features (e.g., spectral energy and entropy). Once the desired features are extracted successfully, the system enables the upper classification task for activity recognition.

2) ACTIVITY RECOGNITION

To conduct the fine-grained activity recognition, one or multiple specific classification models need to be implemented. The model can be selected from the existing machine learning models, typically including decision tree, naive bayes or artificial neural network. The classifier usually has different output options, such as *writing* or *hand-up-down*. These output options are used to label each non-overlapping short-duration time window (e.g., 2s). For different use cases and learning environment, the system may adopt different models and options. Moreover, the model building usually requires a so-called training process to properly select the model type and model parameters. The training process needs to collect the training data, which are also the sensor data but labeled with the activity information.

For the classroom learning environment, the system currently detects five typical student overt activities as follows.

1) *Stationary*: hand is placed in one place and stays relatively stable, which indicates a student in a motionless state.
2) *Writing*: hand is continuously moving on a horizontal plane, which indicates a student taking notes in class.
3) *Hand-Up-Down*: hand is quickly raised or lowered, which indicates a student notifying the teacher that they wish to pose or answer questions.
4) *Head-Scratching*: hand is raised and keeps touching head, which indicates a student pondering or puzzling over something.
5) *Other-Moving*: hand keeps changing its positions in other ways, which reflects a student taking other physical actions (excluding the above ones).

The system mainly employs the accelerometer data to conduct the activity recognition for the above five types of activities, where the magnetic sensor and gyroscope sensor data are used to transform the accelerometer data from device coordinates to earth coordinates [39]. The system adopts both time-domain features (i.e., mean, variance, correlation and interquartile range), and frequency-domain features (i.e, spectral energy, entropy and wavelet magnitude). The feature space consists of 19 features in total. We then collected the training data from multiple participants to build the activity classification model. Among multiple supervised learning models, including decision tree, naive bayes, and artificial neural networks, we adopted the decision tree C4.5 for the classification model. This is mainly because the tree-type model can be easily interpreted and meanwhile achieve a satisfactory classification accuracy.

In the practical operation, the inputs of the activity classification model are the latest three-axis transformed accelerometer data, and outputs of the model are one physical activity to label the current time window. For example, the built model outputs *writing* to label the current time window, meaning it detects that the student is taking down notes during this time period. The time window size for this activity recognition model needs to be relatively small, as a large window may cover a long period that a student may perform multiple activities, which would decrease the model accuracy. Finally, the sequence of the detected student activities would be sent to the engagement inference module to further process.

C. ENGAGEMENT INFERENCE MODULE

By leveraging on the recognized overt activities together with the established theories for engagement, this module mainly ascertains the student engagement state, which can be used to provide the engagement-based services. As mentioned earlier, we introduce the ICAP framework to link the student overt activities to different engagement level. Accordingly, we adopt the heuristic approach to design a practical algorithm, called the *engagement state analysis* (ESA) algorithm, to infer distinct student engagement levels.

The basic idea of the ESA algorithm is simple: according to the ICAP framework, different activities reflects different modes of engagement. For example, *stationary*, *writing* and *hand-up-down* can be categorized into the *passive* mode, *active* mode and *constructive* mode of engagement, respectively. Furthermore, we simply regard both *active* mode and *constructive* mode as *high* engagement mode, and the *passive* mode as *low* engagement mode. Given any time period, the ESA algorithm would categorize a student into the so-called *high engagement state*, when the student always exhibits overt activities in the *high* engagement mode; similarly, the ESA algorithm would categorize a student into the so-called *low engagement state*, when the student always exhibits overt activities in the *low* engagement mode. For other cases, the ESA algorithm would categorize a student into the *normal engagement state*. For all the above three engagement states inference, the *other-moving* duration should not occupy too long, otherwise the algorithm would simply categorize a student into the *unclear state*. As the inputs of the ESA algorithm, assuming N_s, N_w, N_h, N_p and N_o are the number of the detected *stationary*, *writing*, *hand-up-down*, *head-scratching* and *other-moving* activities, respectively. The complete algorithm is shown in Algorithm 1.

Algorithm 1 Engagement State Analysis (ESA) Algorithm

Input: N_s, N_w, N_h, N_p, N_o, classification window length l, thresholds θ_h, θ_l and θ_u.

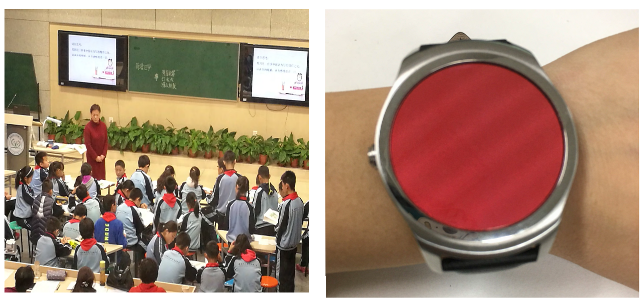
Output: Student engagement state for the current time period T.

$T_{high} = N_w * l + N_h * l + N_p * l$;
$T_{low} = N_s * l$;
$T_{unclear} = N_o * l$;
$r_e = \frac{T_{high}}{T_{low}}$;
$r_u = \frac{T_{unclear}}{T_{high}+T_{low}+T_{unclear}}$;
if $r_e > \theta_h$ and $r_u < \theta_u$ **then**
  Infer *high engagement state* for the time period T;
else if $r_e < \theta_l$ and $r_u < \theta_u$ **then**
  Infer *low engagement state* for the time period T;
else if $\theta_h \geq r_e \geq \theta_l$ and $r_u < \theta_u$ **then**
  Infer *normal engagement state* for the time period T.
else
  Infer *unclear state* for the time period T.

The given ESA algorithm mainly consists of three steps: firstly, it calculates the time duration of the three student engagement states, using the number of the detected activities multiplying the window length l, which is the window size used by the activity classification model and normally a small value (e.g., 1 s). Hence, the *high engagement state* duration T_{high} is the total lasting time of *hand-up-down*, *writing* and *head-scratching* activities; the *low engagement state* duration T_{low} is the lasting time of *stationary* activity. The *unclear state* duration $T_{unclear}$ is the lasting time of *other-moving* activity. After that, the ESA algorithm calculates the so-called engagement ratio, denoted as r_e, namely the duration of *high engagement mode* divided by the duration of *low engagement mode*. Similar logic applies to calculating the so-called unclear ratio, denoted as r_u. Finally, the ESA algorithm infers the student in a *high engagement state* using the conditions that the engagement ratio r_e is greater than the threshold values θ_h and meanwhile the unclear ratio r_u is lower than the threshold θ_u. Similar logic applies to inferring the student is in a *low engagement state* and *normal engagement state*. For any other situations, the ESA algorithm infers the student is in an *unclear state*.

To aggregate the information for the ESA algorithm, the system needs to buffer all the recognized student activities during the current time period T, which is set as a fixed-size moving time window in the practical implementation. Compared with the activity classification, engagement inference may require a relatively long-term time window (e.g., 5 min to 60 min) to collect enough learner activities and information. The inferred engagement state information can be transferred in a timely manner to the external services via the wireless communication interfaces, or used directly by the services on local wearable devices.

The inferred student engagement states can be used to support and build upper interventions and services. For example, we implement the near-real-time intervention on the smartwatch devices: during the class time, when the consecutive *low* engagement states are detected and their total duration T_{low} is significantly long, this indicates that it is highly probable that the student is not in a normal learning state. In that case, the device would automatically vibrate to send the student an explicit alert. Compared with the teacher directly warning the student in class, the intervention in such a private way can help to avoid embarrassing the students in public. Figure 3a shows a class where the students are wearing the smartwatch, and Figure 3b shows the smartwatch vibrating with the colored screen colored to remind a student in a private way. Similarly, a variety of interventions and services can be designed and implemented using the derived student engagement states and the recognized student activities.



(a) (b)

FIGURE 3. System Deployment and Implemented Service. (a) Students with the Smartwatch. (b) Smartwatch Vibrating.

In short, the three modules described above work cooperatively and automatically, and the entire process does not require any configuration or input from the learner side. Therefore, the designed system is a hassle-free and non-intrusive solution for learner activity recognition and engagement inference.

IV. SYSTEM EVALUATION

We conducted the evaluations on the proposed system, which mainly focused on the validation of the captured students' learning activities and their engagement levels. As defined earlier, the student *active* and *constructive* learning activities, including *writing*, *hand-up-down* and *head-scratching*, belong to ***high engagement mode (HEM)***, while the student *passive* learning activities, such as *stationary*, belong to ***low engagement mode (LEM)***. The system firstly recognizes both *HEM* and *LEM* modes of activities and then use the proposed ESA algorithm to infer the engagement levels.

A. EVALUATION PROCEDURE AND PARTICIPANTS

The evaluation process mainly consisted of three parts: the first part mainly built the activity recognition model and evaluated the model's accuracy. The second part applied the proposed ESA algorithm to divide the participating students into three groups according to their engagement levels, namely low engagement group, normal engagement group and high engagement group. The third part of the evaluation validated the three groups's engagement state by using the engagement assessment questionnaire. Moreover, we also conducted a post-evaluation semi-structured interview with the participating students and teachers, where they could openly express their perceptions towards the use of the system. Their feedback was audio-taped and transcribed by the interviewers.

The participating students ranged in age of from 11 years to 13 years (MEAN = 11.98, SD = 0.46) studying in year 7 and 8 of two local schools. A total of 71 students (35 Male, 36 Female) were invited to participate the evaluation, which reflected the current composition of the general student body in their schools.

In the first part of the evaluation, we collected the students' activity data using our system during a two-week period as the model training data. The students mainly use their dominant hand to wear the smartwatch during the data collection process. After the data collection, we adopt multiple models, including decision tree, neural networks and naive bayes, to construct the classification model for the activity recognition.

In the second part of the evaluation, all the participating students were required to wear our smartwatch for their data collection and engagement analytics during multiple lessons, which mainly consisted of teacher's instruction and the question-answer interactions during the 40-minute class time. The proposed ESA algorithm was performed to classify the students into GROUP-1 (low engagement state), GROUP-2 (normal engagement state) and GROUP-3 (high engagement state). Note that some students are possibly classified into the *unclear state* rather than none of the above three groups by the ESA algorithm, as their *unclear ratio* r_u is too high (i.e., bigger than the threshold θ_u) in the ESA algorithm.

In the third part of the evaluation, we adopt a well-recognized student course engagement questionnaire (SCEQ) [40], which consists of 23 items and converges with the national survey of student engagement (NSSE) [41]. The items used a six-point Likert scale (i.e., from 6 = strongly agree to 1 = not at all agree). The questionnaire results are used to validate the inferred student engagement levels.

B. EVALUATION METHODOLOGY

For the activity recognition task, F1 score is a common metric to evaluate the accuracy of the classification model, which usually ranges from 0 to 1. Normally, a classification model with a F1 score above 0.8 can be regarded as a good model, as the F1 score considers both precision (namely the number of correct positive results divided by the number of all positive results) and recall (namely the number of correct positive results divided by the number of positive results that should have been returned).

For the engagement level validation, we mainly adopt one-way ANOVA test [42], as more than 2 groups are involved in the evaluation. Moreover, we also compute the effective size (ES) using Cohen's f, namely $ES = \sqrt{F/n}$, where F is the F-statistic in ANOVA test and n is the number of subjects in each group.

V. EVALUATION RESULTS

A. PART I: STUDENT ACTIVITY RECOGNITION

To build the activity recognition model, we firstly collected the training data from the participating students during a two-week period. The training data means the sensor data labeled with the learning activity name, and they are mainly marked by an independent observer. Among multiple models, the decision tree C4.5 is finally adopted to construct the model, as it achieves the highest performance and can be easily implemented. Table 1 summarizes the built model's accuracy: the F1 scores of all the activities are above 0.8, indicating that the built classification model can well distinguish the five types of student activities, although the F1 score of

TABLE 1. Activity recognition evaluation results.

	Precision	Recall	F1 Score
Hand-Up-Down	0.890	0.814	0.850
Stationary	0.905	0.952	0.928
Writing	0.878	0.859	0.869
Head-Scratching	0.888	0.829	0.858
Other-Moving	0.866	0.918	0.891

hand-up-down and *head-scratching* are slightly lower than others. The overall classification accuracy is 88.1% with a 10-fold cross-validation, which ensures the feasibility to use the detected student activities for student engagement analytics. The sampling frequency for accelerometer, gyroscope and magnetic sensors are set to 50 Hz, and the time window length was set to 1 second.

B. PART II: ENGAGEMENT INFERENCE AND GROUPING

We ran the designed ESA algorithm on the recognized student activity data collected from all the participating students. The algorithm automatically classifies the students into three groups, namely GROUP-1 (low engagement state), GROUP-2 (normal engagement state) and GROUP-3 (high engagement state). The threshold ratios θ_h, θ_l and θ_u are set to 2.0, 0.5 and 0.3 respectively in the ESA algorithm, which can be further fine-tuned and adjusted. The time period T is set to the duration of a lesson, namely 40 minutes (i.e., 2400 seconds).

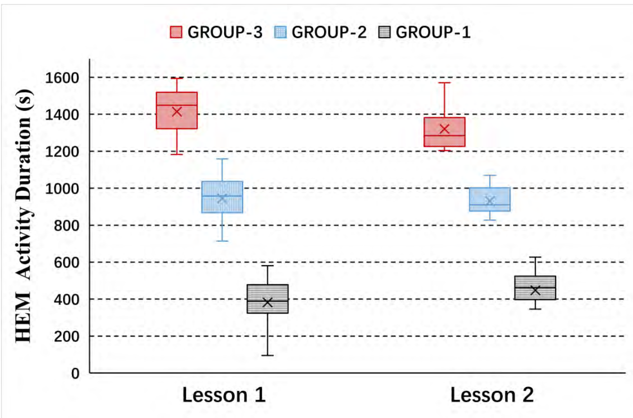


FIGURE 4. Boxplot of Action Durations in HEM.

Figure 4 gives the boxplot of the HEM activity durations for all the three groups at 2 different lessons. We see that for both lessons, the HEM activity durations of GROUP-3 are longer than GROUP-2, and the corresponding durations of GROUP-2 are longer than GROUP-1. Specifically, the HEM activities in GROUP-3 took around 1400 seconds and 1320 seconds on average in Lesson 1 and Lesson 2 respectively, which means GROUP-3 students spent more than 50% time on the HEM activities during the class (e.g., writing notes and raising hand).

Figure 5 gives the boxplot of the LEM activity durations for all the three groups at the same 2 lessons. We see that for both lessons, the LEM activity durations of GROUP-1 are longer than GROUP-2, and the corresponding durations of GROUP-2 are longer than GROUP-3. Specifically, the LEM activities in GROUP-1 took around 1600 seconds and 1430 seconds on average in Lesson 1 and Lesson 2 respectively, which means GROUP-1 students spent more than 60% time on the LEM activities during the class.

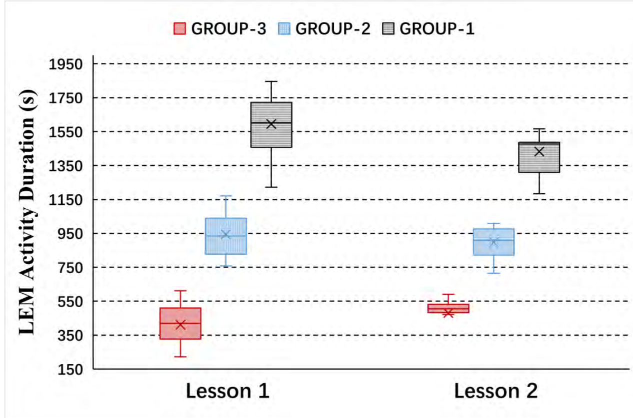


FIGURE 5. Boxplot of Action Durations in LEM.

C. PART III: ENGAGEMENT LEVEL VALIDATION

We randomly choose 21 students from each group to complete the SCEQ questionnaire, and thus a total of 63 students' data are collected for the engagement level validation. The typical questions in SCEQ questionnaire include "*You are often listening carefully in class*", and "*You are often finding ways to make the course interesting to yourself*". We then compute the average scores of the students in each of the three groups.

TABLE 2. ANOVA test result for engagement level.

Source of Variation	Sum of Squares	DOF	Mean Squares	F	P-value
Between Groups	34.59	2	17.29	109.32	9.84E-21
Within Groups	9.49	60	0.16		
Total	44.08	62			

The aggregated results show that the average engagement level of GROUP-3 students is higher than GROUP-2 (MEAN $= 5.27 > 4.41$), and GROUP-2 is higher than GROUP-1 (MEAN $= 4.41 > 3.46$). Table 2 shows the ANOVA test results, which indicate the significant difference among the three groups (F = 109.32, P = 9.84E-21 < 0.05, ES = 2.28). We further conduct the post-hoc test, where the two-tailed t-test shows the significant difference between GROUP-3 and GROUP-2 (P = 1.37E-10 < 0.05), and the significant difference between GROUP-2 and GROUP-1 (P = 3.88E-8 < 0.05).

In short, all the above analytics results validate the feasibility and effectiveness of using the proposed system to capture the student engagement.

D. INTERVIEW

On the student side, all the participants were interviewed and overall satisfied with the use of the system (MN = 4.83, SD = 0.81), and wearing the smartwatch did not interfere their learning process in class (MN = 4.54, SD = 0.97). Regarding the implemented intervention service on the device, the positive comments include "It is cool to remind me using the smartwatch when I almost fall asleep in class", "This watch seems know my head in the clouds like my teacher", and "I like the watch rather than the teacher to warn my distraction during class". However, some students felt that "It is not comfortable to always wear a heavy watch on the wrist" and "no interesting stuff and only a few functions on the watch". These encourage us to further improve both hardware and software of the system, and meanwhile introduce more functionalities by leveraging on the captured activity and engagement information.

On the teacher side, 11 teachers were involved in our interview: they overall agreed that the system is helpful for teachers (MN = 4.82, SD = 0.87) and the importance to automatically recognizing student activity and engagement information at the individual level (MN = 4.91, SD = 0.83). Some important comments include "develop a specific APP for teachers to properly visualize the collected student activity and engagement information" and "simplify the teacher configuration process and provide a user-friendly interface". We are currently working on resolving the issues raised by the teachers.

VI. CURRENT LIMITATIONS AND FUTURE SCOPE

While the ICAP framework reveals the strong connections between student learning activities and the engagement level, how to more effectively and continuously capture the student engagement is still an open and interesting research problem. For example, the evaluation results show that the students in both high engagement group and low engagement group exhibit a significant duration of *stationary* in class, while the same *stationary* activity may indicate different cognitive status of the learners. We believe that a deeper understanding of human cognitive mechanism and learner activities would help to further explain the engagement concept, and eventually in turn to establish a way of simplifying and optimizing the current engagement capture process.

On the other side, the current system mainly focuses on recognizing the overt student activities and engagement information in the classroom context, which can be further extended to the school and home environment. Accordingly, more types of student activities, including both curricular and extracurricular ones, need to be captured and translated using the ICAP framework or other theories. Meanwhile, more types of sensors and wearable devices may need to be adopted to capture the implicit student activities (e.g., using wearable devices to detect learner's eye blinking frequency and heart rate variance).

VII. CONCLUSION

To effectively capture the student engagement and activity information, we have proposed and implemented a novel and practical system by harnessing commodity wearable devices. The system firstly utilizes the built-in sensors on the commodity wearable devices, together with the latest activity recognition techniques, to detect the fine-grained student activities. After that, the established ICAP framework is introduced to link the overt student activities to their engagement levels. The evaluation results show that the designed system can accurately recognize the desired typical student activities in class, and meanwhile validate the feasibility and effectiveness of using the proposed system to capture the student engagement level in class.

While our system design is novel and the design objectives are achieved, we believe that the main impact of this work is to illustrate the broader possibility of creating a practical solution to capture student engagement and learning status, based on a combination of the commodity wearable products, pervasive computing techniques, and the theories from learning science. We are currently working with the industry partners to deploy the system in more than 60 local schools.

REFERENCES

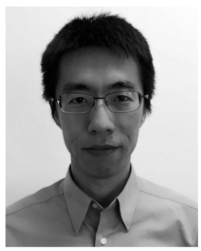
[1] S. Hu and G. D. Kuh, "Diversity experiences and college student learning and personal development," *J. College Student Develop.*, vol. 44, no. 3, pp. 320–334, 2003.
[2] C. C. Bonwell and J. A. Eison, *Active Learning: Creating Excitement in the Classroom*. San Francisco, CA, USA: Jossey-Bass, 1991.
[3] A. King, "From sage on the stage to guide on the side," *College Teach.*, vol. 41, no. 1, pp. 30–35, 1993.
[4] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 23, pp. 8410–8415, 2014.
[5] V. Trowler, "Student engagement literature review," *Higher Edu. Acad.*, vol. 11, no. 1, pp. 1–15, 2010.
[6] G. D. Kuh, "Assessing what really matters to student learning *inside the national survey of student engagement*," *Change, Mag. Higher Learn.*, vol. 33, no. 3, pp. 10–17, 2001.
[7] Y.-L. Hu, G. S. Ching, and P.-C. Chao, "Taiwan student engagement model: Conceptual framework and overview of psychometric properties," *Int. J. Res. Stud. Educ.*, vol. 1, no. 1, pp. 69–90, 2011.
[8] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," *Educ. Psychol.*, vol. 49, no. 4, pp. 219–243, 2014.
[9] Apple Inc. *Apple Smartwatch*. Accessed: Jan. 2019. [Online]. Available: http://www.apple.com/watch/
[10] Fitbit Inc. *Fitbit Wristband*. Accessed: Jan. 2019. [Online]. Available: https://www.tbit.com/uk/ex2/
[11] Google Inc. *Google Glass*. Accessed: Jan. 2019. [Online]. Available: https://www.google.com/glass/
[12] M. Stoppa and A. Chiolerio, "Wearable electronics and smart textiles: A critical review," *Sensors*, vol. 14, no. 7, pp. 11957–11992, 2014.
[13] T. Gill, J. M. Keller, D. T. Anderson, and R. H. Luke, "A system for change detection and human recognition in voxel space using the Microsoft Kinect sensor," in *Proc. IEEE Appl. Imag. Pattern Recognit. Workshop (AIPR)*, Oct. 2011, pp. 1–8.
[14] B. S. Bloom, *Taxonomy of Educational Objectives: The Classification of Educational Goals*. New York, NY, USA: McKay, 1956.
[15] T. Phan, S. G. McNeil, and B. R. Robin, "Students' patterns of engagement and course performance in a Massive Open Online Course," *Comput. Educ.*, vol. 95, pp. 36–44, Apr. 2016.
[16] C. R. Henrie, L. R. Halverson, and C. R. Graham, "Measuring student engagement in technology-mediated learning: A review," *Comput. Educ.*, vol. 90, pp. 36–53, Dec. 2015.
[17] K.-L. Krause and H. Coates, "Students' engagement in first-year university," *Assessment Eval. Higher Educ.*, vol. 33, no. 5, pp. 493–505, 2008.
[18] E. T. Pascarella and P. T. Terenzini, *How College Affects Students: A Third Decade of Research*. San Francisco, CA, USA: Jossey-Bass, 2005.
[19] L. Corno and E. B. Mandinach, "The role of cognitive engagement in classroom learning and motivation," *Educ. Psychol.*, vol. 18, no. 2, pp. 88–108, 1983.
[20] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," *Rev. Educ. Res.*, vol. 74, no. 1, pp. 59–109, 2004.

[21] T. Nokes-Malach, E. J. Richey, and S. Gadgil, "When is it better to learn together? Insights from research on collaborative learning," *Educ. Psychol. Rev.*, vol. 27, no. 4, pp. 645–656, 2015.
[22] X. Wang, W. Wen, and C. P. Rose, "Towards triggering higher-order thinking behaviors in MOOCs," in *Proc. ACM Int. Conf. Learn. Anal. Knowl. (LAK)*, 2017, pp. 398–407.
[23] M. Bower and D. Sturman, "What are the educational affordances of wearable technologies? *Comput. Educ.*, vol. 88, pp. 343–353, Oct. 2015.
[24] Y. Lu, S. Zhang, Z. Zhang, W. Xiao, and S. Yu, "A framework for learning analytics using commodity wearable devices," *Sensors*, vol. 17, no. 6, p. 1382, 2017.
[25] *New Media Consortium and EDUCAUSE Learning Initiative. Wearable Technology*, EDUCAUSE, Louisville, CO, USA, 2016, pp. 46–47.
[26] E. de la Guía, V. L. Camacho, L. Orozco-Barbosa, V. M. B. Luján, V. M. R. Penichet, and M. L. Pérez, "Introducing IoT and wearable technologies into task-based language learning for young children," *IEEE Trans. Learn. Technol.*, vol. 9, no. 4, pp. 366–378, Oct./Dec. 2016.
[27] G. Ngai, S. C. F. Chan, J. C. Y. Cheung, and W. W. Y. Lau, "Deploying a wearable computing platform for computing education," *IEEE Trans. Learn. Technol.*, vol. 3, no. 1, pp. 45–55, Jan. 2010.
[28] L. Müller, M. Divitini, S. Mora, V. Rivera-Pelayo, and W. Stork, "Context becomes content: Sensor data for computer-supported reflective learning," *IEEE Trans. Learn. Technol.*, vol. 8, no. 1, pp. 111–123, Jan./Mar. 2015.
[29] *Adidas Brings Fitness Tracking to Schoolchildren*. Accessed: Jan. 2019. [Online]. Available: http://go.nmc.org/adizone
[30] F. Attal *et al.*, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.
[31] T. Okoshi, Y. Lu, C. Vig, Y. Lee, R. K. Balan, and A. Misra, "QueueVadis: Queuing analytics using smartphones," in *Proc. ACM Conf. Inf. Process. Sensor Netw. (IPSN)*, 2015, pp. 214–225.
[32] M. Radhakrishnan, S. Sen, S. Vigneshwaran, A. Misra, and R. Balan, "IoT+Small Data: Transforming in-store shopping analytics & services," in *Proc. 8th Int. Conf. Commun. Syst. Netw.*, 2016, pp. 1–6.
[33] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, "The case for smartwatch-based diet monitoring," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2015, pp. 585–590.
[34] J. Knighten, S. McMillan, T. Chambers, and J. Payton, "Recognizing social gestures with a wrist-worn smartband," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops*, Mar. 2015, pp. 544–549.
[35] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1432–1462, 2014.
[36] L. P. Prieto, K. Sharma, P. Dillenbourg, and M. Jesús, "Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors," in *Proc. ACM 6th Int. Conf. Learn. Anal. Knowl. (LAK)*, 2016, pp. 148–157.
[37] *TicWatch Version 2*. Accessed: Jan. 2019. [Online]. Available: https://en.ticwear.com/
[38] *TicWatch SDK for Developer*. Accessed: Jan. 2019. [Online]. Available: http://developer.chumenwenwen.com/en/doc/ticwear.html
[39] Y. E. Ustev, O. D. Incel, and C. Ersoy, "User, device and orientation independent human activity recognition on mobile phones: challenges and a proposal," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publication*, 2013, pp. 1427–1436.
[40] M. M. Handelsman, W. L. Briggs, N. Sullivan, and A. Towler, "A measure of college student course engagement," *J. Educ. Res.*, vol. 98, no. 3, pp. 184–192, 2005.
[41] *National Survey of Student Engagement (NSSE)*. Accessed: Jan. 2019. [Online]. Available: http://nsse.indiana.edu/
[42] G. González-Rodríguez, A. Colubi, and Á. Gilb, "Fuzzy data treated as functional data: A one-way ANOVA test approach," *Comput. Statist. Data Anal.*, vol. 56, no. 4, pp. 943–955, 2012.

YU LU received the Ph.D. degree from the National University of Singapore, in 2012. He was a Research Scientist with the Institute of Infocomm Research (I2R), A*STAR, Singapore. He is currently an Associate Professor with the School of Educational Technology, and the Director of the Artificial Intelligence (AI) Laboratory, Advanced Innovation Center for Future Education, Beijing Normal University. His current research interests include data mining, pervasive computing, and AI in education.

JINGJING ZHANG received the Ph.D. degree from the University of Oxford. She was a Clarendon Scholar and a member of the Brasenose College (funded by ORS Scholarship). She is currently an Associate Professor with the Faculty of Education, Beijing Normal University, specializing in learning and technology. Her current research interest includes how learning occurs between people engaging in communication that is situated in daily work environments online and offline.

BAOPING LI received the Ph.D. degree in educational economy and management from Beijing Normal University (BNU). She was a Senior Visiting Scholar with the Education University of Hong Kong. She is currently a Lecturer with the Faculty of Education, BNU. Her research interests include smart learning spaces and design and development of digital learning activities.



PENGHE CHEN received the B.S. and Ph.D. degrees in computer science from the National University of Singapore. He has been with the Advanced Digital Sciences Center, Singapore. He is currently a Principle Researcher with the Advanced Innovation Center for Future Education, Beijing Normal University, China. His research interests include knowledge graph construction, data mining, and learning analytics.

ZIJUN ZHUANG received the B.S. degree in software engineering from Beijing Union University. She was a Research Engineer with the Artificial Intelligence Laboratory, Advanced Innovation Center for Future Education, Beijing Normal University. Her research interest includes data mining and human–computer interaction.