ADVANCED SENSOR RESEARCH Open Access
www.advsensorres.com

Exploring the Elusive Mind: A Multimodal Wearable Sensor Solution for Measuring Mind Wandering in University Students

*Sara Khosravi, Haobo Li, Ahsan Raza Khan, Ahmad Zoha, and Rami Ghannam**

Mind-wandering is a typical daily phenomenon during which attention shifts from external stimuli to internal trains of thought. It affect students' learning by impairing comprehension, diminishing academic achievement, impeding critical thinking, and encouraging a lack of attention and engagement in classroom activities. This study aims to introduce a new method of detecting and tracking mind wandering in university students. This approach involves using wearable sensors, including galvanic skin response (GSR), photoplethysmography (PPG), and eye-trackers, along with machine learning techniques. The study provides a proof of concept for this multisensory approach. The association between longer fixation duration and mind wandering, and the influence of an instructor's presence on fixation allocation, and, consequently, the frequency and occurrence of mind wandering is investigated. Furthermore, the feasibility of using eye-trackers in conjunction with GSR and PPG sensors for detecting mind wandering through a wearable multisensory data collection system is assessed. The wearable multisensory device is evaluated by ten participants (university students, males/females aged between 21-30). Two distinct machine learning methods, support vector machine (SVM) and gated recurrent unit (GRU), are used as classification models. With sensor fusion, the SVM and GRU models yielded maximum accuracies of 86.53% and 89.86%, respectively. Moreover, participants are observed to fixate on instructors more often, just before instances of mind wandering.

1. Introduction

We live in a world where there are many distractions that can quickly capture our attention, making it difficult to stay focused. Even though educators try to provide engaging and interactive material to students, many still struggle with disinterest and lack of attention. This phenomenon, where the mind drifts away from the task at hand and becomes preoccupied with other thoughts, is known as mind-wandering.

Mind wandering is closely connected to learning, and many research studies have investigated its impact during lectures.[1,2] The two most widely recognized techniques for gathering data on mind wandering in the literature are the "self-report" and "probe-caught" methods.[3,4] The "probe-caught" approach involves interrupting participants and inquiring about their mental state at that moment, while the "self-report" technique requires participants to self-report their mind-wandering experiences. This study utilizes the probe-caught method and incorporates a multisensory approach, including biological, physiological, and gaze-tracking sensors, to collect the data for experiments.

Over the last decade, scientists have shown increasing interest in developing and deploying innovative technologies to track and understand mind wandering in educational settings, intending to improve learning outcomes. This has involved using mobile devices such as phones and tablets to gather self-reported data from students, as well as sensors that can detect biological markers, such as temperature and pressure. However, these methods can be intrusive for participants and require extensive data analysis. Moreover, the data collection process through a multisensory approach is more reliable as it involves multiple data sources that can be validated through experiments. Therefore, a multisensory solution involving wearable sensors such as wearable galvanic skin response (GSR), photoplethysmography (PPG), and eye-trackers combined with machine learning offers a more efficient solution for monitoring mind wandering in students,[5] as they are comfortable to wear and allow for real-time, rapid data analysis, as illustrated in **Figure 1**. Due to their lightweight design and ease of wear, students find these technologies comfortable to use. Furthermore, they can be integrated with data analysis algorithms, allowing for swift and real-time analysis of data.

S. Khosravi, H. Li, A. R. Khan, A. Zoha, R. Ghannam
James Watt School of Engineering
University of Glasgow
Glasgow G12 8QQ, UK
E-mail: rami.ghannam@glasgow.ac.uk

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adsr.202300067

DOI: 10.1002/adsr.202300067

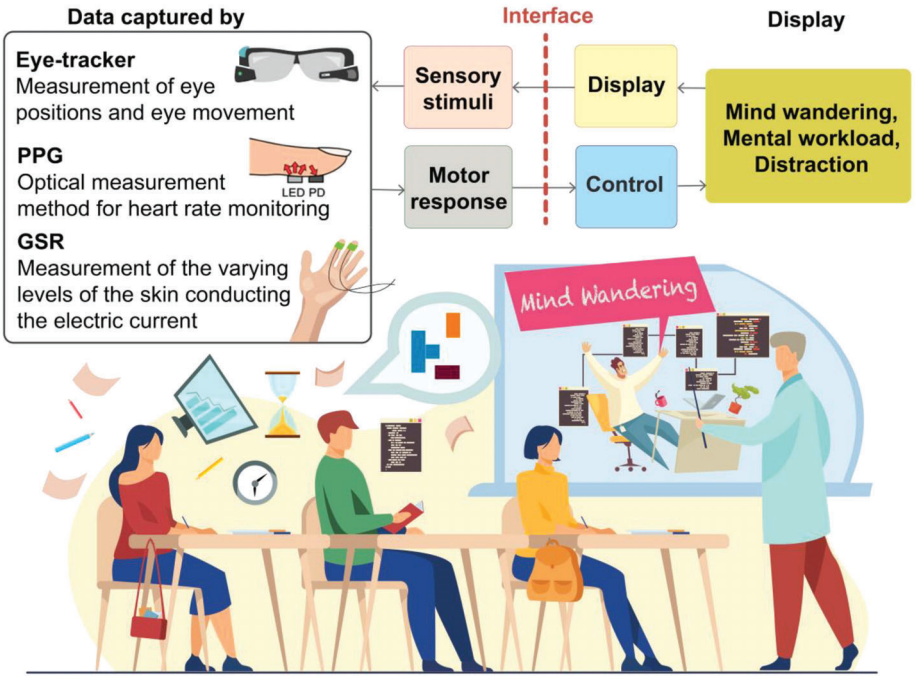

Figure 1. Conceptual schematic of the multisensory methodology for mind-wandering detection in the learning environment and an envisioned block diagram for its measurements in students using PPG, GSR, and eye-trackers.

We have previously shown how students react to educational materials, such as lecture slides, by assessing their visual attention through the use of eye-trackers.[6–8] Visual attention is only one aspect of paying attention, and even if a person appears to be focused on a task and their gaze is directed toward the intended subject, they may be thinking about something entirely different. The connection between eye movements and cognitive processing is undeniable.

To estimate cognitive load, attention, focus, and mind wandering, gaze behavior can be captured and analyzed, which enables the monitoring of visual attention.[9] Cutting-edge wearable devices enable the measurement of gaze movements on specific parts of learning materials, allowing for the identification of the source of distraction. Our previous work involved using a wearable eye-tracker to monitor students' visual attention while reading educational materials and distinguish between different behaviors in text- and graphic-based visuals.[10] However, as visual attention alone does not indicate a full focus on a task, we need an additional indicator to validate the results obtained from the wearable eye-tracker.

In this study, we sought to gain deeper insights into mind-wandering by using dynamic stimuli, such as video, rather than static stimuli. Following the approach taken by,[11] we aimed to capture mind-wandering during a video lecture. To enhance the accuracy of the experiment, we employed the use of a pupil core wearable eye-tracker, which records data at a frequency of 200 Hz.

The following two hypotheses will be tested to study the impact of mind-wandering:

1) **Hypothesis 1**: There is a positive correlation between participants' fixation duration and their mind-wandering while engaging in video-based learning sessions.
2) **Hypothesis 2**: In video-based teaching and learning, the presence of an instructor will lead to a higher level of mind-wandering during lectures.

In our earlier work, we hypothesized that the length of visual attention or fixation duration would be longer while a person is mind wandering. We also hypothesized that the existence of an instructor in the video could increase student mind wandering. In this study, we aim to demonstrate the impact of both of these hypotheses using a combination of multimodal sensors. The main motivation of this work is the use of eye tracker, PPG, and GSR sensors and the processing of their collected data by machine learning methods, including support vector machine (SVM) and gated recurrent unit (GRU), that can be easily adopted in classroom settings compared to electroencephalography (EEG), which is very difficult and inconvenient for students. The eye-tracker, PPG, and GSR signals are well-suited for realization in a wearable device that can gather data from students without causing any discomfort or privacy concerns.

The structure of this paper is as follows: Section 2 offers an overview of the current developments in multisensory design for educational purposes. Section 3 outlines our methodology. Section 4 describes our process for developing a multisensory design. The results of our experiment and data analysis, showcasing the proposed system, are presented in Section 5. Finally, Section 6 provides concluding remarks.

2. State-of-the-Art Wearable Devices for Mind-Wandering Detection

Advances in technology for education have made it technically feasible to put together electronic devices to enhance learning

Table 1. Pros and cons of various sensors for measuring mind wandering.

Sensors	Pros	Cons
Respiration/pressure	Unique to its specific purpose	Lack of movement, discomfort
EEG	Provides data on cognitive state	Sensitive to environment
Heart-rate sensors	Ease of access	Need for validation
GSR	Measuring the emotional state	Sensitive to movement
Eye-tracker	Can be used with moving target	Only collect visual data
PPG	Provides a more indirect measurement	Inaccuracy during daily activities

and teaching. By utilizing a multisensory system that includes various sensors for collecting and processing diverse physiological and neurological data, wearable devices can enhance the accuracy of attention and mind-wandering monitoring. In state-of-the-art studies, the multisensory approach has been used to assess mental workload, identify emotional states, and analyze the learning and information processing process. For instance, Campisi et al.[12] investigated the mental workload of students while web browsing using augmented reality, Mutlu-Bayraktar et al.[13] examined the effects of split attention in multimedia learning environments using eye-tracking and EEG sensors, Saffaryazdi et al.[14] analyzed people's emotional state during stimuli exposure using GSR and PPG sensors, and Srivastava et al.[15] investigated the process of information learning and visual attention using eye-trackers.

In this article, we aim to demonstrate how a fusion of sensors, including GSR, PPG and eye-trackers sensors, can be used for assessing mind wandering. GSR is a widely used physiological measure of emotional arousal, as it provides an indirect indication of emotional arousal through the measurement of sweat gland activity. Increased sweat gland activity is often associated with increased emotional arousal. However, it is essential to note that GSR is not a direct measure of emotions but rather a measure of physiological arousal that may or may not be related to emotions. Furthermore, GSR readings can be affected by various factors, such as skin hydration levels, temperature, and skin pressure.[16]

2.1. Sensors for Mind-Wandering Measurements

Mind wandering has been detected using various physiological biomarkers such as heart rate, skin conductance,[17] and respiration. Among them, pressure sensors,[18] EEG, GSR, PPG and eye-trackers have been deemed as the most accurate options.[16] However, each of these technologies comes with its own benefits and limitations, as outlined in **Table 1**. The selection of the technology to be used depends on the specific context and environment.

In practical applications, it is not feasible to measure all the signals, so in our study, we opted to focus on analyzing the GSR, PPG, and eye-tracker sensors. These signals are well-suited for implementation in a basic wearable device that can gather data from students without causing any discomfort or privacy concerns.

2.2. Wearable Eye-Trackers for Gaze Measurements

Most desktop eye-trackers only rely on a webcam for measuring participant eye data. Affordability and ease of use are the main advantages of such technologies that caused their widespread use by researchers. However, desktop eye-rackers use a camera at a distance beyond 50 cm and can interfere with background lights such as sunlight. On the other hand, wearable eye-trackers can precisely monitor pupil movement because of their close proximity to the eye. Furthermore, they provide the ability to monitor and collect the gaze point that happened outside of the monitor's frame and follow the participants' gaze map to identify the source of distraction. In this experiment, we recorded data using Pupil Core eye tracking glasses at 200 HZ in a 400 × 400 pixel resolution.[6] Pupil Core eye-trackers can measure gaze behavior in 2D and 3D formats. The 3D gaze collection uses pye3d for 3D pupil detection.[7] In addition, a confidence level is provided to set a threshold for proper pupil detection. The threshold is between 0, indicating no pupil detection, and 1, the highest possibility of pupil detection. To calculate fixations, the Pupil Core employs a dispersion-based algorithm[8] with the ability to be implemented both online and offline. In this study, we used an online method for both pupils and gaze detection and calculated the fixation based on a dispersion level of 1.5° in terms of degrees of visual angle with a minimum duration of 100 ms.[9] Data with a confidence value lower than ≈0.6 were eliminated from the experiment.

The presentation of information to students can be a determining factor in their success or failure. The format in which the information is presented, whether in text or graphics, can have a significant impact. Our recent study focused on examining the difference in gaze behavior between text and graphic representations in lecture slides. Our experiment was based on the study in,[11] using an 18-min video lecture on international comparisons in education. Participants were asked five questions prior to watching the video and 18 questions afterward. We defined areas of interest (AOI) to distinguish between the slide and the teacher's window. In gaze tracking technology, an AOI refers to a defined region or specific part of a visual stimulus that is being observed by a person's gaze. AOIs can be used to analyze and identify which parts of an image are being attended to as well as how long they are being attended to

2.3. Wearable GSR and PPG for Physiological Multisensory Measurements

We developed emotion classification models using lightweight, small, and compact PPG and GSR Shimmer3 sensors.[19] We developed a complete application that utilized a database for storing data and used pictures from the Geneva affective picture database (GAPED) to elicit emotions from participants. The post-processing process involved using statistical parameters and

power spectral density (PSD) as features and support vector machine (SVM) and k-nearest neighbors (KNN) as classifiers.

In the literature, researchers previously used other physiological multisensors, such as EMG and GSR signals, as sources of information for real-time emotion detection in gaming scenarios.[20] Our work categorized emotions using arousal and valence scales, with arousal being divided into normal, high, and very high categories and valence being split into positive and negative categories. Another study, outlined in,[21] employed EEG, GSR, EMG, BVP, EOG, and skin temperature signals to create affective databases for emotion recognition. These databases contained speech, visual, or audio-visual data, with emotional labels including neutral, anxiety, amusement, sadness, joy, disgust, anger, surprise, and fear. In another relevant study measured EEG, ECG, GSR, and facial activity to understand the relationship between emotional attributes and personality traits.[22] Emotions were elicited using various stimuli (audio, video), and expressed in different ways (facial expressions, speech, and physiological responses), requiring the creation of a multimodal database for implicit personality and affect recognition.

3. Methodology and Multisensory Design

3.1. Wearable Multisensory Design

Electrodermal activity, which refers to the skin's resistance to the flow of electricity, can be measured using a GSR sensor. This is achieved by placing two electrodes at different points on the skin and applying a small electric charge, usually 1.0 V, and recording the current that passes between the electrodes. This current is then converted into $\mu\Omega$ (or μA), which provides an indication of the skin's resistance to the flow of electricity. Emotional stimuli can influence this resistance, making it a valuable measure for detecting changes in emotional arousal. Sweat gland function is regulated by the sympathetic nervous system, which is responsible for the body's "fight-or-flight" response. GSR sensors can detect even minor changes in electrical conductivity, making it possible to detect emotional stimuli without visible sweating. However, it is important to note that GSR alone cannot determine the cause of emotional arousal and that other physiological and contextual factors must be considered when interpreting results. Therefore, controlling for external factors in the experimental design is essential to interpret GSR results accurately.

An increase in sweat gland activity leads to more perspiration and lower skin resistance. To measure this, the standard unit for GSR is conductance, which is the inverse of resistance and is measured as Conductance = 1 / Resistance. This makes it easier to interpret the signal, as higher sweat gland activity results in higher skin conductance.

To measure a GSR signal for emotional research purposes, the most common method is to use a constant voltage system, also known as the exosomatic method. In this method, the GSR sensor applies a constant voltage of usually 0.5 V to the two electrodes that are in contact with the skin. A small resistance is in series with the voltage supplier and the electrodes to measure the skin conductance and its variation using Ohm's law (Voltage = Intensity × Resistance = Intensity/Conductance). Since the voltage is kept constant, any fluctuation in the current flow through the electrodes is due to a change in the electrical properties of the skin, specifically the sweat gland activity.

The connection between sympathetic nervous activity, as measured by skin conductance (SC) and skin temperature (ST), and attentional states has led to the use of physiology to monitor mind wandering.[23] Previous research has shown that higher levels of mind wandering are associated with overall lower levels of SC.[24] Despite this finding, there have been no efforts to develop automated mind wandering detectors based on SC or ST signals, nor has there been any research on the relationship between ST and mind-wandering. To address this gap, we collected a large dataset where students were periodically asked to report instances of mind wandering during computerized learning from instructional texts. Machine learning models were then trained on these signals to predict mind wandering.

3.2. Participant Group Selection and Test Environment

The wearable multisensory device was assessed by 10 participants who volunteered to take part in our experiments, participants were university students studying postgraduate engineering (males and females aged between 21 and 30). The experiments were conducted with the approval of Glasgow Univeristy's ethics committee and participants provided their consent through a consent form.

3.3. Data Collection Process

The stimulus was an 18-min lecture on international education.[11] There was a pre-lecture questionnaire with five questions regarding the knowledge of the participants about the concept of the lecture. Afterward, the lecture was presented to participants in the form of a video. After the lecture, the post-lecture questionnaire will be presented to the participants, containing 18 questions regarding the presented lecture and two questions asking participants to rate their overall experience.

Figure 2 shows the schematic diagram of the experimental setup. Participants were asked to sit in front of the computer in a comfortable position that allowed them to read the screen. The eye-tracking glasses, the GSR sensor and the PPG sensor, were placed on the participant's head and hand. The calibration and validation process were done for the eye tracker. Two GSR sensors and one PPG sensor were attached to the participant's fingers on the non-dominant hand.

Participants were also instructed on the process of the test and how to answer the questions, and what to consider as mind wandering or not. **Figure 3** shows the raw collected data from each individual sensor (eye-tracker, GSR and PPG) with and without mind-wandering. The acquisition frequency for eye-tracker is 200 Hz, and for GSR and PPG, sensors are 50 Hz.

The extracted characteristic information from the data in Figure 3 encompasses signal attributes such as amplitude, shape, frequency, as well as the precise positions and values of signal peaks. The SVM classifier derives this characteristic information from the original sensor data through a feature-extraction process, where statistical parameters such as mean, standard deviation, median, and others detailed in **Table 2** are computed

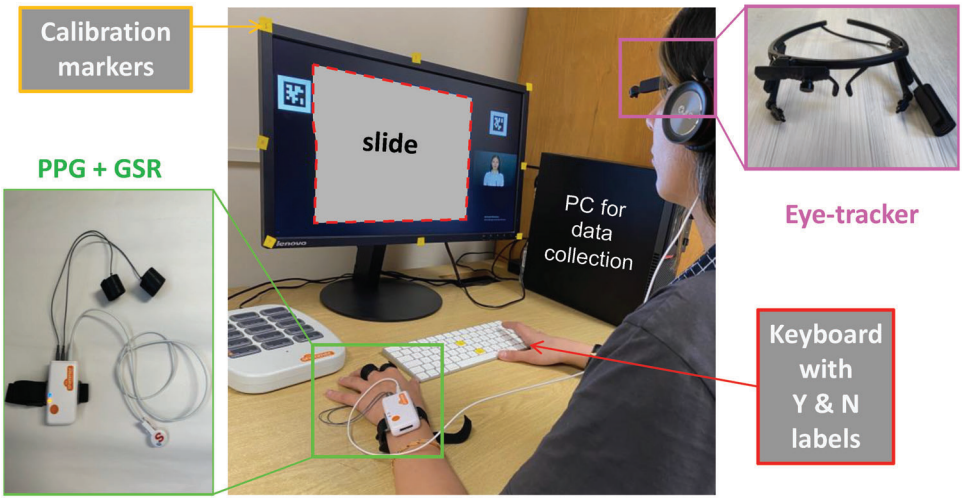

Figure 2. The experimental setup for the wearable device with a combination of multisensory eye-tracking, GSR and PPG sensors. The Figure shows the calibration procedure and the Pupil Core headset, Shimmer GSR and PPG sensors, and a participant taking the test.

and subsequently provided to the classifier for training and testing.

In contrast, the GRU network obtains the characteristic information by directly analyzing the time dependency within the original signal. This is accomplished through the update and reset gates of each GRU unit. In other words, GRU network finds the temporal relationship between data points at different time locations and uses it to represent the characteristic information hidden in raw data.

The processing of these data and their fusion using machine learning are described in Section 4.

4. Multisensory Fusion and Machine Learning

Prior to the classification, numerical features are extracted from the raw sensor data measurements. The features we used in this work are listed in Table 2. The whole feature pool contains regular statistical parameters such as mean, median,

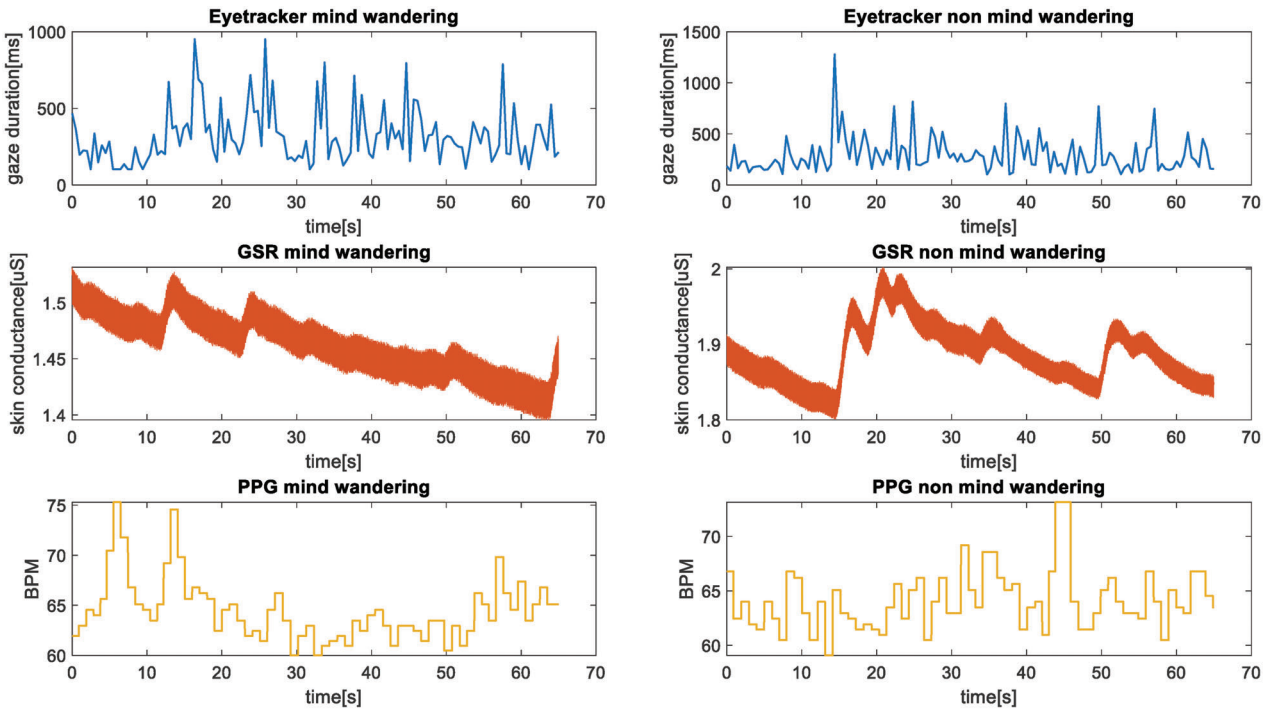

Figure 3. The raw data were collected from each sensor of the Pupil Core eye-tracker, Shimmer GSR, and PPG sensors, respectively. Graphs on the left show the collected data with mind-wandering, whereas the graphs on the right are for data with non-mind-wandering.

Table 2. The list of numerical features extracted from the raw sensor data measurements.

Features	
Mean	Skewness
Standard deviation	Kurtossis
Median	Minimum
Mad	Range
25th quantile	Mean of the autocorrelation
75th quantile	Standard deviation of the autocorrelation
Inter quartile range	Signal energy

minimum and standard deviation, and higher-order statistical parameters such as skewness and kurtosis. Besides that, correlation-based features and signal energy features are also involved.

Two different machine learning methods, notably, a conventional SVM and GRU, are utilized as the classification models. The SVM[25,26] is known as a robust classifier in the field of indoor human activity recognition, gesture recognition and biometric identification, it intends to build a hyperplane to separate the feature points of different classes based on the distribution of the features. The support vectors are the feature points close to the decision boundary, and they are able to control the position and orientation of the hyperplane, whereas the margin between the positive and negative hyperplane needs to be maximized through those support vectors.

The mathematical representation of a linear SVM hyperplane and its objective function is given as follows[27]:

$$h : x' W + b = 0 \tag{1}$$

$$\min_{W,b,\xi_i} \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{n}\xi_i (C > 0, \xi_i \geq 0) \tag{2}$$

where W denotes the normal vector to the hyperplane and b is the bias value. C refers to the regularization parameter, also known as the penalty factor, which is highly correlated with the tolerance of misclassification. The penalty factor is always greater than zero and the larger factor will create a hard margin, and vice versa (soft margin), its value needs to be determined carefully since hard margin may result in overfitting of the classifier. In our case, the penalty factor is set as one in the training of the classification model. ξ_i represents the slack variable related to the classification error, the SVM algorithm automatically allocates a slack variable for the feature points between the hyperplane and its margin, whereas the value of the slack variable ($0 \leq \xi_i \leq 1$) is proportional to the distance of feature points to the hyperplane. In the circumstance that the feature points beyond the hyperplane (misclassification), the slack variable is larger than one.

If a linear hyperplane is not able to separate the feature points, the features can be mapped to a higher-dimensional space through a kernel function, where a linear boundary is available. The conventional kernel function includes higher order polynomial (quadratic, cubic) and Gaussian function, whereas the choice of the kernel function depends on the data distribution and the optimal hyperplane to separate them. In this paper, we chose the quadratic kernel function. SVM algorithm is suitable to implement on a multiclass problem by utilizing multiple binary classifiers via "one versus one" approach, for instance, if there are N classes to distinguish, $N(N-1)/2$ times binary SVM will be computed to construct hyperplanes between each individual class.

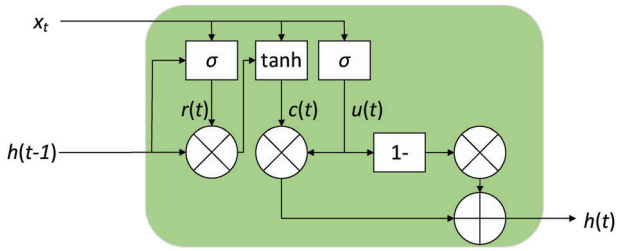

Figure 4. The block diagram of a simple GRU to process the recorded raw data with faster processing speed and less memory cost.

GRU is the improved version of the regular recurrent neural network (RNN).[28–30] A standard GRU consists of two different gates, namely, the update gate and reset gate; compared to other variants of RNN, such as long short-term memory (LSTM), GRU has comparable performance with a simpler architecture, faster processing speed and less memory cost. The block diagram of a simple GRU is illustrated in **Figure 4**, and the mathematical expression of GRU is given below[28,29]:

$$r(t) = \sigma(W_{\mathrm{rh}} h_{t-1} + W_{\mathrm{rx}} x_t + b_r) \tag{3}$$

$$u(t) = \sigma(W_{\mathrm{uh}} h_{t-1} + W_{\mathrm{ux}} x_t + b_u) \tag{4}$$

$$c(t) = \tanh(W_{\mathrm{ch}} r(t) \odot h(t-1)) + W_{\mathrm{cx}} x_t + b_c) \tag{5}$$

$$h(t) = z(t) \odot c(t) + (1 - z(t)) \odot h(t-1) \tag{6}$$

where $r(t)$ and $u(t)$ represent the output of the reset and update gate, respectively. σ denotes the sigmoid activation function. W refers to the weight index of gated units, whereas b is the bias value. $\odot\odot$ refers to the Hadamard product of two vectors. $c(t)$ represents the output of the tanh operator, which receives a linear combination of the current input $x(t)$ and the result of the Hadamard product between $r(t)$ and $h(t-1)$. $h(t)$ represents the current output of the GRU. The update gate $u(t)$ controls the ratio of current input information and historical information. When $u(t)$ is close to 1, most of the historical information is forgotten, and a larger volume of input information is taken from the current moment, and vice versa. In this paper, as GRU has the ability to extract useful time-dependent information from raw data, the input of GRU is a time-series signal rather than features.

The confidence level of the classifier is a probability matrix with its size equal to nXm, n is the number of samples and m is the number of classes. It is used to measure the certainty of classifier decision-making, whereas, for each sample, the class yielding the highest confidence level will be chosen as the output class. The value of the confidence level is converted from the

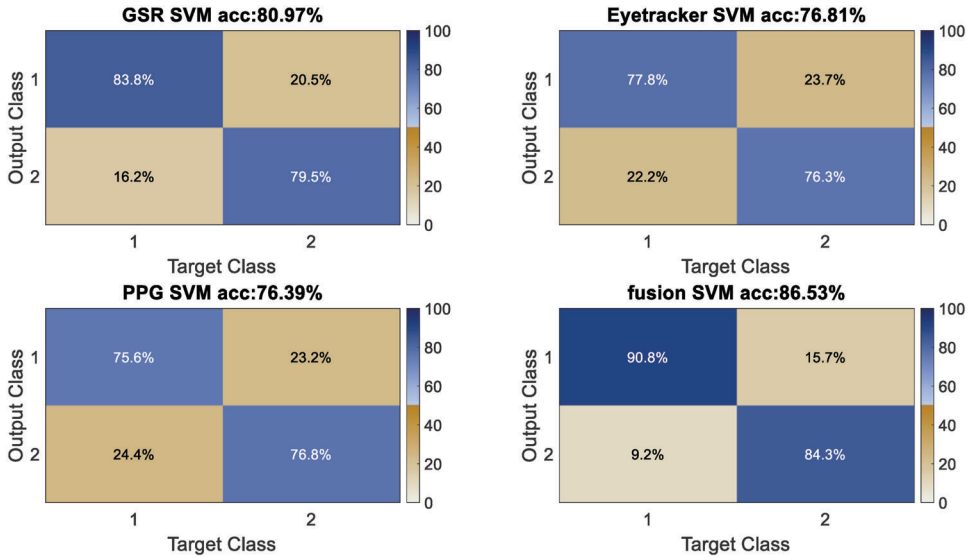

Figure 5. The SVM classification for GSR, eye-tracker, PPG, and fusion. Class 1 notes mind-wandering, and class 2 indicates non-mind-wandering.

unnormalized classifier output through a softmax function, as shown below:

$$P_c = \frac{e_c}{\sum_{k=1}^{K} e_k} \tag{7}$$

where class c is the class of interest, P_c the confidence level of class c, e_c, and e_k denote the unnormalized classifier output of class c and class k ($k \leq K$), respectively, and K is the number of classes.

In this paper, we developed a decision-level fusion method to combine the confidence level of all sensing approaches. The fusion process[26,31,32] is depicted in Equation (8):

$$P_{Fusion}(s, c) = \sum_{n=1}^{N} P_n(s, c) \tag{8}$$

where P_n denotes the confidence level matrix for classifier n, P_{Fusion} is the confidence level matrix after fusion. s refers to the index of samples, and c indicates the class number. The confidence level matrix of different sensors shares the same dimension. The fusion of data from multiple sensors is the straightforward accumulative of their respective output confidence level matrices. The new prediction label is the class with the highest fusion confidence level. In our paper, n equates to 3, representing g classification results from eye-tracker, GSR, and PPG, respectively. It is important to notice that the fusion of data is unrelated to the acquisition frequency of sensors because data fusion occurs at the decision level, following classification.

5. Results and Discussion

The schematic diagram of the experimental setup is presented in Figure 2. The training and testing procedure has been repeated for ten times and the results listed in **Figures** 5 and 6 are the average of ten iterations. In addition, the data selected for training and testing is different for every iteration.

Figure 5 illustrates the results of SVM classification, where class 1 and class 2 indicate mind-wandering and non-mind-wandering, respectively. The row elements of the confusion matrix denote the output class, whereas the column elements denote the target class. The sum of the target class should be equal to 100%. The right diagonal elements represent the correctly classified rate, and the left diagonal elements represent the misclassification rate. The results show that the processed for individual sensors with accuracies of 80.97%, 76.81%, and 76.39% for GSR, eye-tracker and PPG, respectively. Our results deploying SVM show that the accuracy of sensor fusion is 86.53%.

Figure 6 demonstrates the results of GRU classification, which processed for individual sensors with accuracies of 85.69%, 81.67%, and 80.42% for GSR, eye-tracker and PPG, respectively. Using GRU, the processing results confirm that the accuracy of sensor fusion reaches 89.86%. The fusion using GRU provides a subsequent improvement of around 3.3% on top of fusion through SVM algorithm.

Figure 7 illustrates the boxplot of the ten iterations of "training and testing". The blue circle represents the mean value of ten different classification results, whereas the red line in the middle of the "box" represents the median value. The upper and lower boundaries denote the maximum and minimum accuracy values of ten iterations of "training and testing" separately. The edges of the blue "box" denote the 25th and 75th percentile of the classification results. It is observed that fusion with SVM and GRU not only increases the mean classification accuracy, but also the variance of classification accuracy. In other words, the stability of the classification system improves thanks to the fusion process. Additionally, compared to the SVM classifier, GRU-based recurrent neural networks provide a higher gain in mean and variance of ten iterations "training and testing". The fluctuations in performance displayed by the wearable sensors in Figure 7 result from the classifier's selection of varying datasets for training and testing in each of the ten iterations. The classifier achieves a high classification accuracy when exposed to favorable data segments

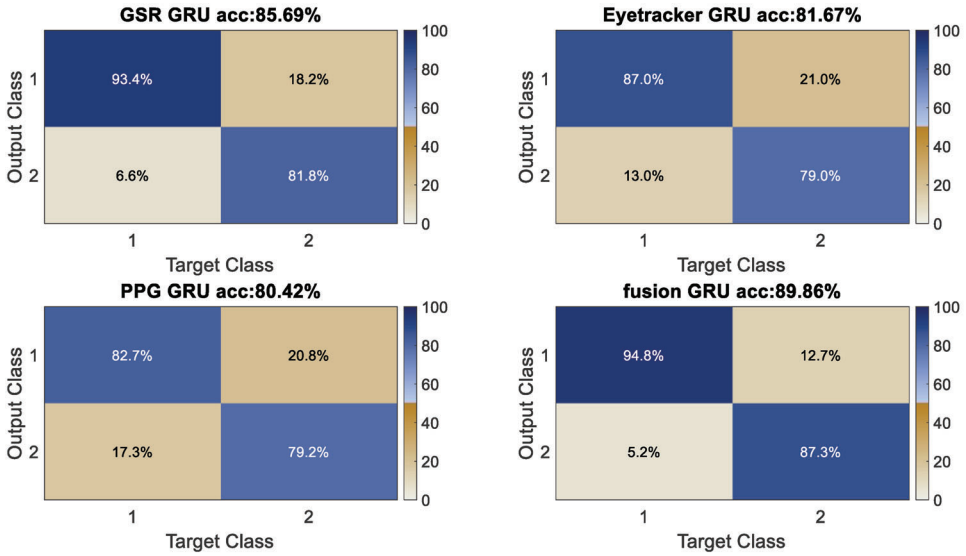

Figure 6. The GRU classification for GSR, eye-tracker, PPG, and fusion. Class 1 notes mind-wandering, and class 2 indicates non-mind-wandering.

and conversely, encounters challenges when presented with less favorable data segments.

6. Conclusion

In this paper, a proof-of-concept wearable multisensory were designed and implemented to work with machine learning technique in order to enhance the students learning and concentration detection. A decision-level fusion method was developed to combine the confidence level of all sensing techniques. The wearable multisensory device was tested by 10 participants (university students, males and females aged between 21 and 30). We used two different machine learning methods: SVM and GRU, for the classification models. Our results showed that using SVM and GRU, process the accuracy of sensor fusion with accuracies of 86.53% and 89.86%, respectively. In the future study, this work will be enhanced further by incorporating more participants with an integrated multisensory device.

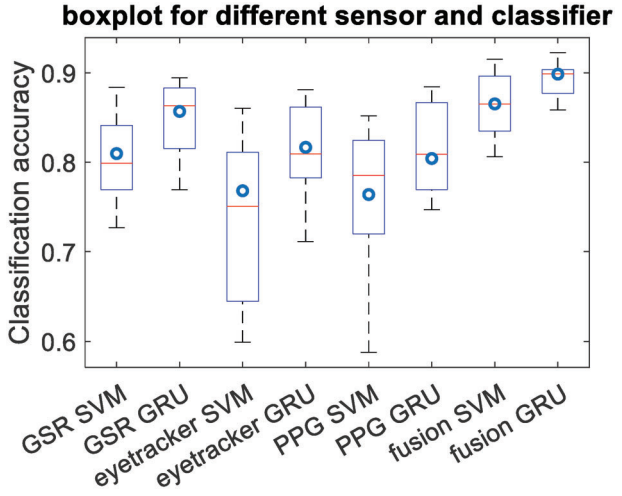

Figure 7. The boxplot of 10 iterations of training and testing using each individual sensor and their fusion using SVM and GRU.

Acknowledgement

S.K. was supported by the College of Science and Engineering Scholarship at the University of Glasgow.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Keywords

education, machine learning, mind-wandering, multisensory, sensor-fusion, wearable devices

Received: April 20, 2023
Revised: September 30, 2023
Published online: November 15, 2023

[1] S. Hutt, A. Wong, A. Papoutsaki, R. S. Baker, J. I. Gold, C. Mills, *Behav Res Methods* **2023**, https://doi.org/10.3758/s13428-022-02040-x.
[2] M. T. Chu, E. Marks, C. L. Smith, P. Chadwick, *Conscious Cogn.* **2023**, *108*, 103463.
[3] M. J. Kane, B. A. Smeekens, C. C. Von Bastian, J. H. Lurquin, N. P. Carruth, A. Miyake, *J. Exp. Psychol.* **2017**, *146*, 1649.
[4] E. F. Risko, N. Anderson, A. Sarwal, M. Engelhardt, A. Kingstone, *Appl. Cogn. Psychol.* **2012**, *26*, 234.

[5] S. Khosravi, S. G. Bailey, H. Parvizi, R. Ghannam, *Sensors* **2022**, *22*, 7633.
[6] S. Khosravi, A. R. Khan, A. Zoha, R. Ghannam, presented at *The IEEE Global Engineering Education Conf.* Tunis, Tunisia, March **2022**.
[7] S. Khosravi, A. R. Khan, A. Zoha, R. Ghannam, presented at *29th IEEE Int. Conf. on Electronics, Circuits and Systems (ICECS)*, Glasgow, UK October **2022**
[8] I. Brishtel, A. A. Khan, T. Schmidt, T. Dingler, S. Ishimaru, A. Dengel, *Sensors* **2020**, *20*, 2546.
[9] Y. Weinstein, *Behav. Res. Methods* **2018**, *50*, 642.
[10] A. R. Khan, S. Khosravi, S. Hussain, R. Ghannam, A. Zoha, M. A. Imran, presented at *The IEEE Global Engineering Education Conf. (EDUCON)*, Tunis, Tunisia March **2022**.
[11] H. Zhang, K. F. Miller, X. Sun, K. S. Cortina, *Appl Cogn Psychol* **2020**, *34*, 449.
[12] C. A. Campisi, E. H. Li, D. E. Jimenez, R. L. Milanaik, *Augmented Reality in Education*, Springer International Publishing, Berlin, Germany **2020**, 111.
[13] D. Mutlu-Bayraktar, P. Ozel, F. Altindis, B. Yilmaz, *Multim. Tools Appl.* **2022**, *81*, 8259.
[14] N. Saffaryazdi, Y. Goonesekera, N. Saffaryazdi, N. Daniel Hailemariam, E. Girma Temesgen, S. Nanayakkara, E. Broadbent, M. Billinghurst, presented at *27th Int. Conf. on. Intelligent User Interfaces*, Helsinki Finland March **2022**.
[15] N. Srivastava, S. Nawaz, J. Newn, J. Lodge, E. Velloso, S. M. Erfani, D. Gasevic, J. Bailey, presented at *LAK21: 11th Int. Learning Analytics and Knowledge Conf.*, Irvine, CA, USA April **2021**.
[16] K. S. McNeal, M. Zhong, N. A. Soltis, L. Doukopoulos, E. T. Johnson, S. Courtney, A. Alwan, M. Porch, *CBE Life Sci. Educ.* **2020**, *19*, ar50.
[17] J. Smallwood, J. W. Schooler, *Annu. Rev. Psychol.* **2015**, *66*, 487.
[18] Y. Zheng, D. Wang, Y. Zhang, W. Xu, *Front. Psychol.* **2019**, *10*, https://doi.org/10.3389/fpsyg.2019.00216.
[19] G. Udovičić, J. Đerek, M. Russo, M. Sikora, in *Proceedings of the. 2nd Int. Workshop on Multimedia for. Personal Health and Health Care*, ACM New York, October **2017**, 53.
[20] A. Nakasone, H. Prendinger, M. Ishizuka, in *Proceedings. of the 5th International Workshop on Biosignal Interpretation*, Citeseer, Tokyo, Japan **2005** 219.
[21] M. Soleymani, J. Lichtenauer, T. Pun, M. Pantic, *IEEE Trans. Affect. Comput.* **2012**, *3*, 42.
[22] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, N. Sebe, *IEEE Trans. Affect. Comput.* **2018**, *9*, 147.
[23] J. Wagner, J. Kim, E. André, presented at *2005 IEEE Int. Conf. on Multimedia and. Expo*, Amsterdam, Netherlands, July **2005**.
[24] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, U. Ehlert, *IEEE Trans. Inf. Technol. Biomed.* **2009**, *14*, 410.
[25] T. D. Bufler, R. M. Narayanan, *IET Radar, Sonar Navig.* **2016**, *10*, 1468.
[26] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, F. Fioranelli, *IEEE J. Electromagn. RF Microw. Med. Biol.* **2018**, *2*, 102.
[27] C. Cortes, V. Vapnik, *Mach. Learn.* **1995**, *20*, 273.
[28] M. Wang, Y. D. Zhang, G. Cui, *Digit. Signal Process.* **2019**, *87*, 125.
[29] M. Wang, G. Cui, X. Yang, L. Kong, *IET Radar, Sonar Navigat.* **2018**, *12*, 1046.
[30] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, arXiv preprint arXiv:1406.1078, **2014**.
[31] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, F. Fioranelli, *IEEE Sens. J.* **2019**, *20*, 1191.
[32] C. Chen, R. Jafari, N. Kehtarnavaz, *IEEE Sens. J.* **2015**, *16*, 773.