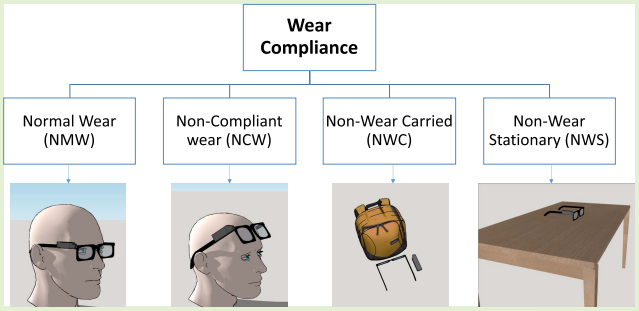
Detection of Food Intake Sensor's Wear Compliance in Free-Living

Tonmoy Ghosh, *Student Member, IEEE*, Delwar Hossain, and Edward Sazonov, *Senior Member, IEEE*

***Abstract*—Objective detection of periods of wear and non-wear is critical for human studies that rely on information from wearable sensors, such as food intake sensors. In this paper, we present a novel method of compliance detection on the example of the Automatic Ingestion Monitor v2 (AIM-2) sensor, containing a tri-axial accelerometer, a still camera, and a chewing sensor. The method was developed and validated using data from a study of 30 participants aged 18-39, each wearing the AIM-2 for two days (a day in pseudo-free-living and a day in free-living). Four types of wear compliance were analyzed: 'normal-wear', 'non-compliant-wear', 'non-wear-carried', and 'non-wear-stationary'. The ground truth of those four types of compliance was obtained by reviewing the images of the egocentric camera. The features for compliance detection were the standard deviation of acceleration, average pitch, and roll angles, and mean square error of two consecutive images. These were used to train three random forest classifiers 1) accelerometer-based, 2) image-based, and 3) combined accelerometer and image-based. Satisfactory wear compliance measurement accuracy was obtained using the combined classifier (89.24%) on leave one subject out cross-validation. The average duration of compliant wear in the study was 9h with a standard deviation of 2h or 70.96% of total on-time. This method can be used to calculate the wear and non-wear time of AIM-2, and potentially be extended to other devices. The study also included assessments of sensor burden and privacy concerns. The survey results suggest recommendations that may be used to increase wear compliance.**



***Index Terms*— Accelerometer, adherence, compliance, egocentric camera, food intake, wear compliance.**

I. Introduction

MONITORING of food intake is crucial to understand and investigate the factors that contribute toward the development of obesity [1], eating disorders [2], and diabetes [3]. Traditional methods of dietary assessments are food diaries, food frequency questionnaires [4], and 24-hour dietary recalls [5], [6] are not accurate and impose a heavy burden on the user. For those reasons, wearable sensors received increased attention as a potential solution for objective monitoring of food intake and ingestive behavior.

Manuscript received September 22, 2021; accepted October 19, 2021. Date of publication October 29, 2021; date of current version December 14, 2021. This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award R01DK100796. The associate editor coordinating the review of this article and approving it for publication was Prof. Kea-Tiong Tang. *(Corresponding author: Tonmoy Ghosh.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by The University of Alabama Institutional Review Board under Approval No. 17-005-ME, August 22, 2017.

The authors are with the Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL 35401 USA (e-mail: tghosh@crimson.ua.edu; dhossain@crimson.ua.edu; esazonov@eng.ua.edu).

Digital Object Identifier 10.1109/JSEN.2021.3124203

Numerous wearable sensors have been proposed to detect food intake events, their duration, environment, and to estimate consumed energy. Wrist-worn sensors were used to detect food-related hand-to-mouth gestures in [7] and [8]. A wearable ear pad sensor was introduced by [9], where an acoustic sensor was used to detect chewing sounds. A sensor system in [10] used electroglottography (EGG) signals to detect food intake. A piezoelectric sensor was proposed in [11] to detect and characterize chewing. Monitoring systems combining a sensor for food intake detection and a wearable camera were introduced in [12] and [13]. The electrical proximity sensor-based prototype was proposed in [14] with a textile capacitive neckband. Estimation of mass and energy intake was proposed in [15] using a chewing sensor and video observation. Moreover, a just-in-time feedback sensor system was proposed in [16] to reduce energy intake. In order to monitor the pattern of food intake, eyeglass-based wearable devices were introduced in [13], [17], and [18]. An accelerometer sensor was used to detect food intake by [19] in free-living individuals. All of the above-mentioned wearable sensors did not analyze wear compliance (the amount of time the study participants wear the device as prescribed by instructions). While the measurement accuracy and battery lifetime are essential to the deployment

of a wearable sensor, they come second to compliance because a sensor is meaningless if it remains unused. Thus, there is a research gap in understanding the wear compliance of wearable food intake sensors, and factors that impact wear compliance.

Few studies have been proposed that analyze wear compliance. The only study investigating compliance in a food intake sensor was presented in [20], where three placements of the sensor in the human body were considered factors affecting compliance. The wear compliance of electronic scoliosis brace was investigated in [21] and [22], and found a 78% compliance. A threshold in temperature sensor readings determined brace on and brace-off states. Wear and non-wear compliance were detected in [23] by using an accelerometer sensor, but the objective of that system is to monitor free-living physical activity. In a study of 30 participants, the accelerometer sensor was worn on the wrist or ankle. The wear compliance was detected by combining the accelerometer and skin temperature data. A threshold-based classifier was developed on average temperature over a time window and the standard deviation of accelerometer data, reporting a sensitivity of 94%. Patient compliance with wearing contact lenses was investigated in [24] by interviewing patients, demonstrating an 86% compliance rate. Wear compliance of waist-worn accelerometer in children (76.7%) was reported in [25]. Spectacle-wear compliance in school children was studied in [26], reporting compliance of 58%. To find compliance, researchers interviewed students and their teachers. Compliance analysis was also performed by examining adherence in activity monitoring wearable devices [27]–[29]. In the literature, there was no clear definition of good or bad compliance. However, in general, greater than 80% compliance was regarded as adequate [30]. To the best of our knowledge, no research has been performed on detection of compliance in food intake sensors. In this paper, a wear compliance measurement method is proposed on the example of Automatic Ingestion Monitor (AIM-2), which is a wearable food intake sensor. The compliance measurement relies either on accelerometer data or on images from an egocentric camera or both. These three methods are compared in terms of the accuracy of compliance measurement. Wear time analysis can help to understand compliance with wear regimens in studies relying on wearable sensors, in this particular case, studies of diet and ingestive behavior. Previously we reported on a method to detect eating episodes from the data of AIM-2 [13]. However, such a method can only detect eating, but does not differentiate periods of wear when there is not eating from periods when device is not worn. Both result in non-detection of eating. Thus, assessment of compliance with device wear is as important as detection of eating. This paper addresses this important gap.

This paper is organized as follows: section (II) presents the wearable sensor system, the definition of different types of wear compliance, data preparation, feature extraction method, classifier, and method for privacy concern and sensor burden assessment, and in section (III) the simulation results are presented and discussed, lastly, the paper is concluded in section (IV).

Fig. 1. Representation of AIM-2 wearable sensor device system: (a) AIM2 device; (b) AIM2 attached with the eyeglasses.

II. Method

A. Wearable Sensor System

The food intake sensor used in the study was the Automatic Ingestion Monitor, version 2 (AIM-2) [13]. A second-generation food intake monitoring device consists of a low-power 3D accelerometer (ADXL362 from analog devices, Norwood, MA, USA), a chewing sensor (SpectraSymbol 2.2" flex sensor), and a 5-megapixel camera with a 170 - degree wide-angle gaze-aligned lens (Fig. 1). The sensor system was attached to the right side of the frame of eyeglasses. The camera captures images periodically at a rate of one image per 15-second interval. The accelerometer signal was sampled at 128 Hz and all the collected images and signals were stored in the SD card.

A human study was conducted among 30 volunteer participants (65% males and 35% females, age range from 18 to 39 years). The participants represented five races, African American, Asian/Pacific Islander, Hispanic/Latin, White/Non-Hispanic, and others. The University of Alabama institutional review board approved the study (IRB Protocol #17-005-ME, dated Aug 22, 2017), and participants were compensated for their participation. The participants spent one day in pseudo-free-living conditions (all meals consumed at the lab, no other restriction) and free-living (no restrictions). In general, the participant was asked to use the device for the whole day (from waking up to going to sleep). The participants were instructed not to wear the device in any situations where privacy had to be protected (for example, using a restroom/bathroom), during water activities, and sleeping. A total of 60 days of sensor data was collected. In pseudo-free-living, the on-time logged by the AIM-2 (unplugged from the charger, cumulative of wear and non-wear time) was 10 to 15 hours, with an average on-time of 12.33 +/− 1.25 h (mean +/− standard deviation). In free-living, the on-time time was 8.5 hours to 15.5 hours (12.1 +/− 1.5h).

B. Wear-Compliance Definition

Ideally, for a wearable food intake detection sensor, a person should wear the device all time except for sleeping. But, in practice, people are not always compliant with the prescribed wear regimen. Besides, participants were asked to take the device off for any confidential activities that may incur privacy concerns. From the study data, it was empirically found that there were four types of wear compliance: 1) normal-wear, the device being worn as prescribed (Fig. 2(a)); 2) non-compliant-wear, the device is worn, but not as it is supposed to be worn, for example, eye-glasses lifted to the forehead or hanging from the neck; 3) non-wear-carried, where the device is carried on the body, for example, inside a bag or in a pocket;

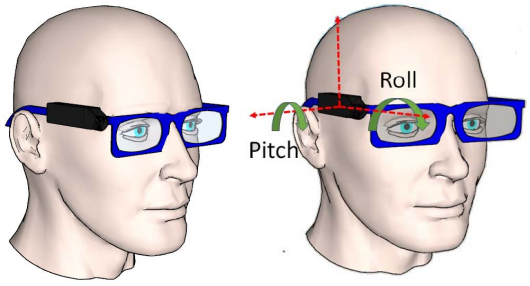


Fig. 2. Demonstration of wear and rotation of head movements, (a) normal-wear of AIM device, (b) pitch and roll angle.

4) non-wear-stationary, where the device is completely off the body, such as when it is placed on a desk.

C. Ground Truth Annotation

A total of 757 hours of data and 180,000 images of 30 participants (2 days for each, a total of 60 days) were captured in this study. Ground truth was obtained by manually reviewing the images captured by the camera. Each image was checked to verify it contains the egocentric point of view of a person performing daily activities, which signifies normal wear. Such images were marked as 'normal-wear' (Fig. 3(a)). If an image contained visual information contradicting a period of 'normal-wear', then immediately preceding and following images in the captured sequence were reviewed. If there were no changes among the images, then they were labeled as 'non-wear-stationary' (Fig. 3(d)). The rationale for such labeling is that head movements are very frequent and images even from a sedentary person vary significantly over time. If the immediately preceding and following images indicated a change in the scene, but not from the egocentric view of the participant, then it was labeled as 'non-compliant-wear' (Fig. 3(b)). Further, the 'non-wear-stationary' images were reviewed. If the image revealed the inside of a bag or pocket, the image was labeled 'non-wear-carried' (Fig. 3(c)). If the image indicated the device being placed on a stationary object (e.g. a desk), then the image was labeled 'non-wear-stationary'. A sample of the images from these four compliance classes is presented in Fig. 3. The accelerometer signal was marked into these four compliant classes by matching the time stamp of captured images.

A total of 180, 637 images were labeled during the ground truth annotation. From those, 91,202 were collected from the pseudo-free-living day, with 69,180 'normal wear', 545 'non-compliant-wear', 2,883 'non-wear-carried', and 18,559 'non-wear-stationary'. Similarly, 89,435 data samples were obtained from the free-living. Among them 63,461 were 'normal wear', 2,006 were 'non-compliant wear', 2,148 were 'non-wear-carried', and 21,820 were 'non-wear-stationary'. The time-matched labels from the ground truth annotation were used as labels to feature vectors in compliance classification. A subset of this dataset (a total of 5 participants data) was annotated by another person to evaluate the level of agreement in labeling. Because the compliance classes are categorical, the inter-rater reliability [31] was measured using a Cohen's kappa coefficient (κ). We obtained a kappa value of 0.96, indicating

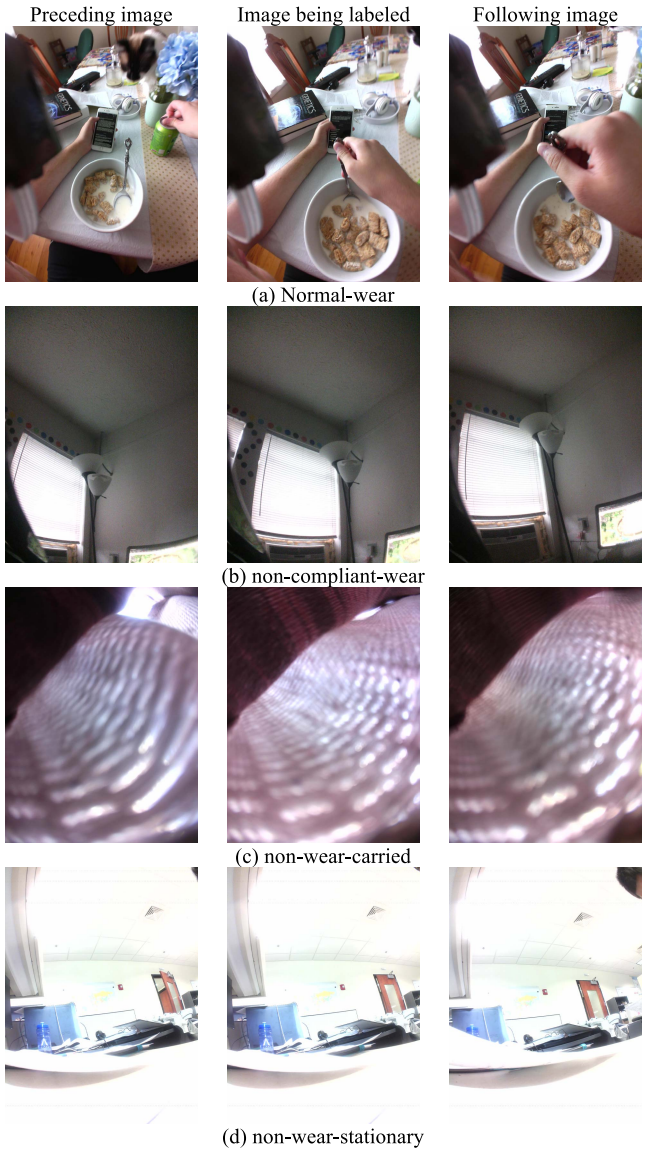


Fig. 3. Examples of images for different compliance classes.

that the labeling is well agreed upon by both annotators. It should be noted that kappa value 1 is regarded as ideal.

D. Feature Extraction

In this paper, we attempt to measure compliance based on the accelerometer signals and in combination with the features extracted from the images. Computed accelerometer features describe the movement and orientation of the sensor on the body. If a participant wears a device on the eyeglasses, the accelerometer data should indicate changes in the movement and certain orientations of the device in the field of gravity. For the 'non-compliant-wear' and 'non-wear-carried' classes, the movement may be similar to 'normal-wear', but the orientation may be infeasible for the normal body postures. The 'non-wear-stationary' class has comparatively low movement and, quite often, infeasible orientations.

The movement of the sensor was measured by the amplitude of acceleration. Similarly, the orientation of the sensor in the gravity field was measured by pitch and roll angles (Fig. 2). Features were computed from the accelerometer signals

(X, Y, and Z-axis), net acceleration, pitch angle, roll angle, and image data. Net acceleration, pitch, and roll angle [32] at the time t were computed as follows

$$netA(t) = \sqrt{(Ax^2 + Ay^2 + Az^2)} \quad (1)$$

$$Pitch, \theta(t) = \tan^{-1}\left(\frac{-Ax}{sign(Az)\sqrt{Ay^2 + Az^2}}\right)\frac{180}{\pi} \quad (2)$$

$$roll, \varphi(r) = \tan^{-1}\left(\frac{Ay}{sign(Az)\sqrt{Ax^2 + Az^2}}\right)\frac{180}{\pi} \quad (3)$$

where, Ax = X-axis value of the accelerometer at time t

Ay = Y-axis value of the accelerometer at time t

Az = Z-axis value of the accelerometer at time t

From net acceleration, a standard deviation was computed that represents changes in acceleration. Similarly, from the pitch and roll values, the average angle was considered a feature. Features were computed in non-overlapping 15 seconds of data (it is the same as image interval, which made sensor features, image features, and ground truth were in the same time window). To keep it computationally light, raw accelerometer sensor data were used to extract features without any preprocessing. A computed image feature was chosen on the observation that the camera captures the same image repeatedly during non-wear. Thus, the mean square error between two consecutive images is extremely low or close to zero. The mean square error is calculated as follows

$$I^k_{mse} = \frac{1}{P * Q}\sum_{i=1}^{P}\sum_{j=1}^{Q}(I^k_{i,j} - I^{k-1}_{i,j})^2 \quad (4)$$

where P and Q are the dimensions of the image, I^k_{mse} is the pixel intensity difference and I^k, I^{k-1} are the kth and (k-1)th images, respectively. The feature vector was formed with four features: Standard deviation (σ) of net acceleration, average pitch angle, average roll angle, and mean square error (I^k_{mse}) extracted from each 15-second epoch. In this study, we focused on exploring simple features and developing algorithms with a low level of computational complexity. Less complex algorithms in wearable sensor systems consume less battery power (resulting in longer battery life) and are easier to implement. The proposed method, similar to real-time food intake detection [33], may aid in detecting compliance in real-time.

E. Classifier

Extracted features were computed from two types of sensors: an accelerometer and camera, thus three distinct classifiers were trained using the feature extracted from 1) accelerometer, 2) image, and 3) both the accelerometer and image.

There were 30 participants, and each participant had two days of data (Pseudo-free-living and free-living day). A leave-one-subject-out cross-validation technique was used to validate the classifier. This means that the data from one participant were held for testing while the data from the remaining participants were used to train the classifier, and the process was repeated for all participants. During this process, the trained classifier never saw the data from the testing participant.

TABLE I
CLASSIFICATION PERFORMANCE OF DIFFERENT CLASSIFIERS

Classifier	Accuracy
Decision Tree	66.93%
Linear discriminant analysis (LDA)	61.34%
Quadratic discriminant analysis (QDA)	63.77%
Gaussian naïve bayes	66.61%
k-nearest neighbors (KNN)	79.37%
Support-vector machine (SVM)	67.17%
Random Forest	89.24%
Artificial Neural Network (ANN)	79.45%

A large number of state-of-the-art classifiers (decision tree, linear discriminant analysis, quadratic discriminant analysis, support-vector machine, k-nearest neighbor, random forest, and Artificial Neural Network (ANN)) were investigated. The best classifier was selected by analyzing the performance.

We used an under-sampling and over-sampling technique to balance the training dataset because the database was unbalanced in terms of compliance classes. We oversampled 'non-compliant-wear' and 'non-wear-carried' to match with 'non-wear-stationary' using the Safe-level-synthetic minority over-sampling technique (SMOTE) [34]. On the other hand, we under-sample 'normal wear' to match 'non-wear-stationary.' It should be noted that we did not modify the test dataset.

F. Privacy Concern and Sensor Burden

The privacy and sensor burden may impact the wear compliance of the AIM device. The privacy concern and sensor burden questionnaire were developed from [35] and all participants responded to the questionnaire quantifying emotion, attachment, harm, perceived change, movement I, movement II, anxiety, and sensor size. The questions were 1) emotion: It is acceptable how I look when I wear the sensor; 2) attachment: The sensor is well attached to my body. I do not feel it is moving; 3) harm: The sensor does not cause some pain and/or tickling; 4) perceived change: I do not feel awkward wearing the sensor; 5) movement I: The sensor does not affect the way I move; 6) movement II: The sensor does not restrict the way I eat; 7) anxiety: I feel secure wearing the device; 8) sensor size: A sensor of smaller size would increase my level of acceptability. There were 11 levels of agreement with the statement, starting from 10 to 0, where 10 means strongly agreed and 0 means did not agree. Moreover, in the case of continuous image capturing, participants did not wear the device where privacy needs to be protected, resulting in lower compliance. Thus, the participant was asked how concerned they are about the continuous image capture over an image captured during food intake only on a scale of 1 to 7 (1 is not concerned and 7 is extremely concerned).

III. RESULTS AND DISCUSSION

Fig. 4 shows the box plots of the extracted features. It is to be noted that NMW, NCW, NWC, and NWS are acronyms of 'normal-wear', 'non-compliant-wear', 'non-wear-carried', and 'non-wear-stationary' classes. The standard deviation of net acceleration is shown in Fig. 4(a), showing a moderately tight range for compliant 'normal-wear'. On the other hand, for the

class of the 'non-compliant-wear', the distribution is similar to the 'normal wear' with a higher average and range. Due to the movement during 'non-wear-carried', the standard deviation is extremely high. Also, in the case of 'non-wear-stationary', the value is mostly zero except for some of the outliers. From the figure, it is observed that there are significant differences in the values of standard deviation among those four compliance classes.

In terms of pitch angle (Fig. 4(b)), the average value is 10 in the 'normal-wear' class, which represents the head movement of up and down directions. The average value of the pitch angle is zero in the 'non-wear-stationary' class due to most of the time device is placed horizontally on a surface (e.g. table). However, in the case of 'non-wear carried', there are no particular directions to keep the device, resulting in a pitch range from 180 to −180 degrees, but its average value of 180 degrees represents, most of the time the device is kept horizontally flipped. Whereas for 'non-compliant-wear' the pitch angle should be higher than 80 degrees, if the device is worn on the forehead, or smaller than −50 degrees if the device is hung from the neck, which is reflected in the figure. From the figure, it is also observed that the median value of pitch angles is well separated to use in a classifier.

The roll angle values are represented in Fig. 4(c). For the 'normal wear' compliance class, side to side head movement is limited, thus a small roll angle range with a slightly negative average value is observed. A similar characteristic is also observed in 'non-wear-carried', where the average value is −130 degrees. A wide range of roll angles is witnessed in 'non-wear-stationary' and 'non-compliant-wear', but their respective average values are different. The roll angles are well separated among those four classes except the median values of 'normal-wear' and 'non-wear-stationary'. Still, it can be used as a potential feature along with standard deviation and pitch angle.

The fourth feature, the mean square error is presented in Fig. 4(d). In the case of 'non-wear-stationary', the average intensity difference is close to zero due to no movement of the device. However, the other three classes, show a higher intensity difference value with different median values.

The classification performance on the 30-fold cross-validation (leave-one-subject-out) is reported in Table I using different types of classifiers. The performance results are presented in terms of epochs (15 sec time window). Among different classifiers k-nearest neighbors (KNN), random forest, and ANN offered above 75% accuracy. The best classification accuracy (89.24%) was achieved using a random forest classifier. Thus, in this study, we selected random forest classifier. Random forest is an ensemble decision tree classifier, which combines the results of many decision trees, which reduces the effects of overfitting and improves generalization. It selects a random subset of predictors to use at each decision split as in the random forest algorithm [36]. The main reasons for using random decision trees classifiers are that it is simple, easy to implement in wearable devices, less computationally expensive, and sufficiently accurate.

Classification accuracy for accelerometer-based, image-based, and combined classifiers is reported in Table II.

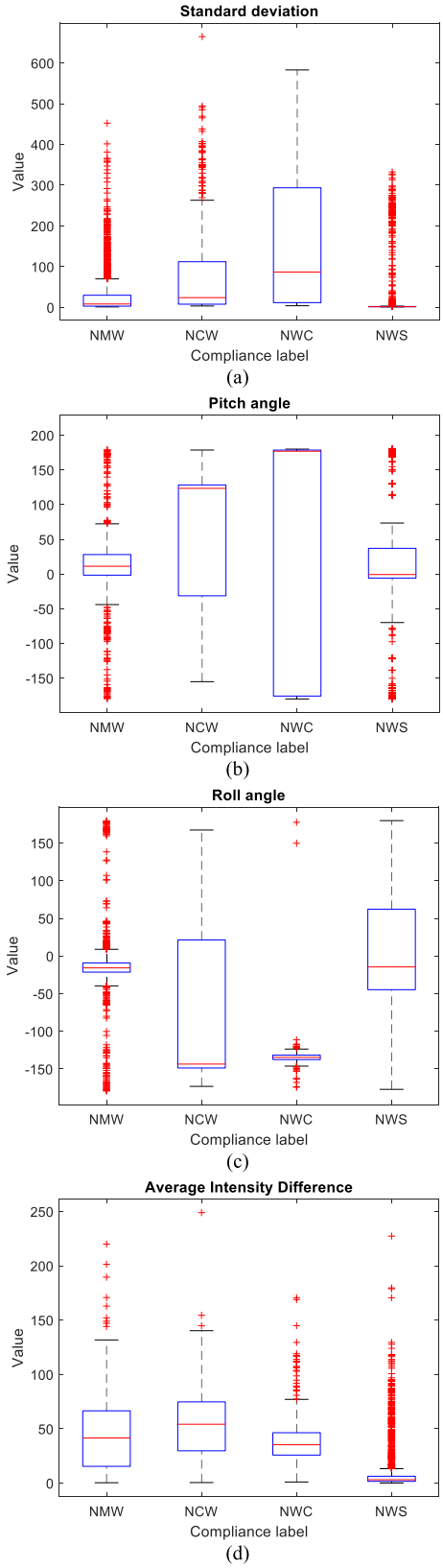


Fig. 4. Boxplot of extracted features, (a) standard deviation of net accelerometer signal, (b) pitch angle, (c) roll angle, (d) average intensity difference of two consecutive images. Here, NMW = 'normal-wear', NCW = 'non-compliant-wear', NWC = 'non-wear-carried', and NWS = 'non-wear-stationary' compliance classes.

The accelerometer-based classifier produced higher accuracy (76.46%) than the image-based classifier (63.57%), while

TABLE II
CLASSIFICATION PERFORMANCE OF DIFFERENT FEATURE SOURCES

Signal source	Features	Accuracy
Accelerometer	Std, pitch and roll	76.46%
Image	Mean square error	63.57%
Combined accelerometer and image	Std, pitch, roll and average intensity	89.24%

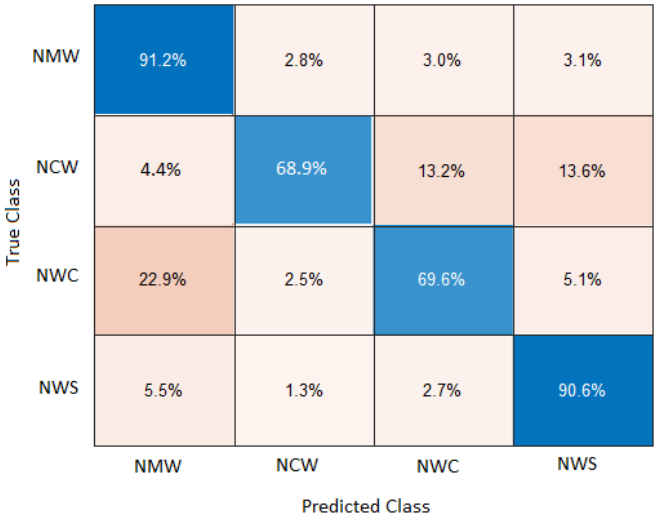


Fig. 5. Confusion matrix for the accelerometer-based classifier, NMW = 'normal-wear', NCW = 'non-compliance-wear', NWC = 'non-wear-carried', and NWS = 'non-wear-stationary.'

the combined classifier produced an improvement (89.24%) of 13%. These results suggest that the performance of the combined classifier is superior, but the accelerometer classifier can be used as the primary means of compliance detection where only an accelerometer signal is available. This is an important finding from two perspectives. First, if compliance is detected by an accelerometer, the camera does not have to take continuous images and could only be triggered during food intake. As the questionnaire results demonstrate, this greatly alleviates privacy concerns. Second, the camera demands significantly more power than an accelerometer. Delegating compliance detection to the accelerometer has the potential to dramatically extend battery life, which is extremely important for a wearable device. Thus, in terms of accuracy and privacy protection accelerometer-based classifier is a good option.

The confusion matrix for the accelerometer-based classifier is shown in Fig. 5. The trained classifier can detect the 'normal-wear' and 'non-wear-stationary' classes with a very high true positive rate. Also, the 'non-compliant-wear', and 'non-wear-carried' detection rates were satisfactory. However, 13.2% of 'non-compliant-wear' was misclassified as 'non-wear-carried' due to its difficulty to separate 'non-compliant-wear' from 'non-wear-carried'. On the other hand, 22.5% 'non-wear-carried' was wrongly predicted as 'normal-wear', likely because, sometimes 'non-wear-carried' shows similar characteristics as 'normal-wear'. Moreover, 13.6% 'non-compliant-wear' also detected as 'non-wear-stationary'.

Wear compliance analysis of 30 participants is reported in Table III using manual annotation. The wear compliance was calculated by taking the ratio of 'normal wear' duration

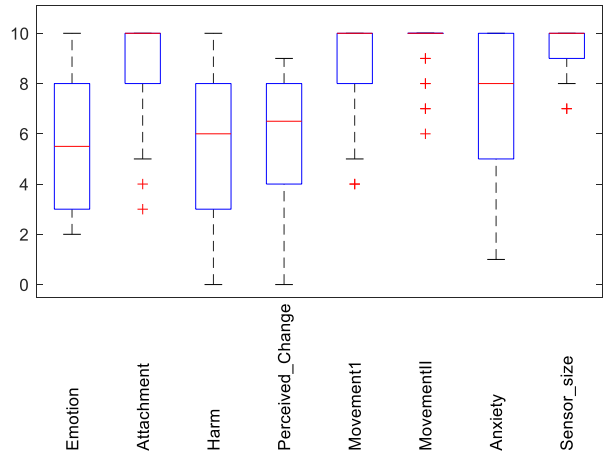


Fig. 6. Box plot of score of sensor burden questionnaires.

and device 'on' duration. In the pseudo-free-living day, the average duration of 'normal wear' was 9 hours 36 minutes with a standard deviation of one hour 45 minutes and the wear compliance was 75.85%. However, on free-living days, the average duration of 'normal wear' was 8 hours 49 minutes with a standard deviation of 2 hours and the wear compliance was 70.96%. In free-living, the wear compliance was lower than the pseudo-free-living days. The accuracy of detecting compliance did not differ significantly between semi-controlled (pseudo free-living) and uncontrolled (free-living) environments, suggesting that the accuracy of the method is generalizable across experimental conditions. The potential changes in the level of compliance across time in longer experiments need to be investigated in future studies. Participants did not wear the device during specific events due to privacy issues, such as inside the restroom, attaining a lecture, or exam.

Moreover, sensor burden assessments are presented using a box plot in Fig. 6. The average value and standard deviation of attachment, movement-1, movement-2, anxiety, and sensor-size are 8.70 ± 1.97, 8.67 ± 1.92, 9.47 ± 1.11, 7.5 ± 2.37, and 9.40 ± 0.97 respectively. This means that the participants like the sensor attachment with their bodies. The sensor does not affect participant movement as well as the way of eating, a good number of participants are not anxious wearing the device, and the sensor size is acceptable to most of the participants. Since the average values of these factors are close to 10 and the standard deviation are low, it hardly impacts the wear compliance of the sensor, On the other hand, the average value and standard deviation of emotion, harm, and perceived change are 5.56 ± 2.75, 5.67 ± 2.68, and 5.27 ± 2.78 respectively. The average values were approximately 5, which means that most of the participants are not strongly agreeing with how they look when they wear the device. The sensor does cause some pain or tickling, and they may feel somewhat awkward wearing the sensor. Since the average values of these factors are in the middle of the 0 to 10 scale, thus, there is a chance that they may affect wear compliance. These considerations may be used to improve sensor design and further improve wear compliance.

In the case of privacy concerns, participants are more concerned about continuous image capture as opposed to image

TABLE III
WEAR COMPLIANCE ANALYSIS

Day	Measurement	Device 'on' (hh:mm:ss)	Normal-wear (hh:mm:ss)	Non-compliant-wear (hh:mm:ss)	Non-wear-carried (hh:mm:ss)	Non-wear-stationary (hh:mm:ss)	Wear compliance
Pseudo free-living	Total	380:12:49	288:15:27	2:16:12	12:06:00	77:24:22	75.85%
	Mean	12:43:26	09:36:25	0:04:54	0:24:20	2:32:31	
	Standard Deviation	1.24:48	1.45:21	0:01:55	0:06:12	1.59:34	
Free-living	Total	372:32:33	264:25:12	8:21:48	8:57:00	90:55:21	70.96%
	Mean	12:24:45	8:48:58	0:16:42	0:17:57	3:02:07	
	Standard Deviation	1:30:42	2:03:00	0:05:12	0:04:48	2:31:05	

capture only during eating [13]. Results indicate that if the image capture is performed only during food intake that would reduce privacy concerns from 5.0 ± 1.6 (somewhat concerned to concerned) to 1.9 ± 1.7 (not concerned). A reduction in privacy concerns may lead to better compliance with device wear.

The primary goal of this sensor system is to detect food intake and estimate daily energy consumption. If the user does not wear the device as instructed, the sensor system will not be able to fulfill its function. Lower compliance suggests that the device was left unattended for a longer period. The wear compliance of 71% in free-living indicates that the AIM was worn for the majority of the time.

This study has a few limitations. Because the participants were compensated, there could be a bias in compliance. However, the compensation was not contingent on compliance and there was no requirement that participants wear the device for a specific amount of time. The participants were free to wear the device for as long as they want and to remove it if they felt uncomfortable or needed privacy. Furthermore, all the participants in this study were students.

A minor limitation of the current study is the use of egocentric camera data for interpretation of the device user's compliance. There are possible scenarios where human annotation of compliance may be incorrect, for example, a participant sleeping with the device may result in camera views infeasible during normal wear. Such cases could potentially be resolved by direct observation methods (such as a research assistant following the participant during daily activities), however, that would make the study prohibitively expensive. We believe that cases of incorrect annotation are extremely rare, and the findings closely reflect real-world compliance with device use. In the future, we will look into more complex models, such as deep neural networks (DNN) or convolutional neural networks (CNN) that perform image recognition, to improve the accuracy of wear compliance detection. However, the advantage of the proposed method that it is relying on computationally lightweight features and classifier, and thus may potentially be deployed on the edge, providing real-time assessment of compliance and feedback (e.g., in a case where participant forgets to wear the device).

IV. CONCLUSION

In a study of 30 participants, we developed a method that can detect wear compliance of a wearable device (AIM-2). The accelerometer-based random forest classifier provided sufficient accuracy to measure wear compliance while preserving the privacy of the users. The proposed classifier was tested on 30 participants, 2 days each. In pseudo-free-living days, the wear time of the device was 09:36:25 +/− 01:45:21 (75.85% compliant wear). In the case of free-living, the wear time was 08:48:58 +/− 02:03:00 (70.96% of compliant wear). The accuracy of the proposed method was 89.24% in the free-living environment. After analyzing the sensor burden and privacy questionnaires it can be concluded that compliance may potentially be increased if certain modifications are made, such as image capture only during food intake, and implementation of a contactless chewing sensor.

ACKNOWLEDGMENT

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] J. Mayer and D. W. Thomas, "Regulation of food intake and obesity," *Science*, vol. 156, no. 3773, pp. 328–337, Apr. 1967.
[2] J. Polivy and C. P. Herman, "Causes of eating disorders," *Annu. Rev. Psychol.*, vol. 53, no. 1, pp. 187–213, 2002.
[3] N. D. Volkow, G.-J. Wang, and R. D. Baler, "Reward, dopamine and the control of food intake: Implications for obesity," *Trends Cogn. Sci.*, vol. 15, no. 1, pp. 37–46, Jan. 2011.
[4] E. A. Krall and J. T. Dwyer, "Validity of a food frequency questionnaire and a food diary in a short-term recall situation," *J. Amer. Dietetic Assoc.*, vol. 87, no. 10, pp. 1374–1377, Oct. 1987.
[5] R. L. Karvetti and L. R. Knuts, "Validity of the 24-hour dietary recall," *J. Amer. Dietetic Assoc.*, vol. 85, no. 11, pp. 1437–1442, 1985.
[6] R. L. Carter, C. O. Sharbaugh, and C. A. Stapell, "Reliability and validity of the 24-hour recall," *J. Amer. Dietetic Assoc.*, vol. 79, no. 5, pp. 542–547, 1981.
[7] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, "Detecting periods of eating during free-living by tracking wrist motion," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 4, pp. 1253–1260, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.
[8] H. Kalantarian and M. Sarrafzadeh, "Audio-based detection and evaluation of eating behavior using the smartwatch platform," *Comput. Biol. Med.*, vol. 65, no. 1, pp. 1–9, Oct. 2015, doi: 10.1016/j.compbiomed.2015.07.013.
[9] O. Amft, "A wearable earpad sensor for chewing monitoring," in *Proc. IEEE Sensors*, Kona, HI, USA, Nov. 2010, pp. 222–227, doi: 10.1109/ICSENS.2010.5690449.
[10] M. Farooq, J. M. Fontana, and E. Sazonov, "A novel approach for food intake detection using electroglottography," *Physiol. Meas.*, vol. 35, no. 5, p. 739, 2014.
[11] M. Farooq and E. Sazonov, "Comparative testing of piezoelectric and printed strain sensors in characterization of chewing," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7538–7541.
[12] J. Liu *et al.*, "An intelligent food-intake monitoring system using wearable sensors," in *Proc. 9th Int. Conf. Wearable Implant. Body Sensor Netw.*, May 2012, pp. 154–160.

[13] A. Doulah, T. Ghosh, D. Hossain, M. H. Imtiaz, and E. Sazonov, "'Automatic ingestion monitor version 2'—A novel wearable device for automatic food intake detection and passive capture of food images," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 2, pp. 568–576, May 2020, doi: 10.1109/JBHI.2020.2995473.
[14] J. Cheng *et al.*, "Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publication*, Sep. 2013, pp. 155–158.
[15] X. Yang *et al.*, "Statistical models for meal-level estimation of mass and energy intake using features derived from video observation and a chewing sensor," *Sci. Rep.*, vol. 9, no. 1, p. 45, Dec. 2019.
[16] M. Farooq, M. A. McCrory, and E. Sazonov, "Reduction of energy intake using just-in-time feedback from a wearable sensor system," *Obesity*, vol. 25, no. 4, pp. 676–681, 2017.
[17] J. Chung, J. Chung, W. Oh, Y. Yoo, W. G. Lee, and H. Bang, "A glasses-type wearable device for monitoring the patterns of food intake and facial activity," *Sci. Rep.*, vol. 7, no. 1, p. 41690, Mar. 2017.
[18] M. Farooq and E. Sazonov, "A novel wearable device for food intake and physical activity recognition," *Sensors*, vol. 16, no. 7, p. 1067, Jul. 2016, doi: 10.3390/s16071067.
[19] M. Farooq and E. Sazonov, "Accelerometer-based detection of food intake in free-living individuals," *IEEE Sensors J.*, vol. 18, no. 9, pp. 3752–3758, May 2018.
[20] R. Alharbi *et al.*, "Investigating barriers and facilitators to wearable adherence in fine-grained eating detection," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops (PerCom Workshops)*, Kona, HI, USA, Mar. 2017, pp. 407–412, doi: 10.1109/PERCOMW.2017.7917597.
[21] T. Rahman *et al.*, "Electronic monitoring of scoliosis brace wear compliance," *J. Children's Orthopaedics*, vol. 4, no. 4, pp. 343–347, Aug. 2010, doi: 10.1007/s11832-010-0266-6.
[22] T. Rahman, J. R. Bowen, M. Takemitsu, and C. Scott, "The association between brace compliance and outcome for patients with idiopathic scoliosis," *J. Pediatric Orthopaedics*, vol. 25, no. 4, pp. 420–422, Jul. 2005, doi: 10.1097/01.bpo.0000161097.61586.bb.
[23] S.-M. Zhou *et al.*, "Classification of accelerometer wear and non-wear events in seconds for monitoring free-living physical activity," *BMJ Open*, vol. 5, no. 5, May 2015, Art. no. e007447, doi: 10.1136/bmjopen-2014-007447.
[24] T. H. Bui, H. D. Cavanagh, and D. M. Robertson, "Patient compliance during contact lens wear: Perceptions, awareness, and behavior," *Eye Contact Lens, Sci. Clin. Pract.*, vol. 36, no. 6, pp. 334–339, Nov. 2010, doi: 10.1097/ICL.0b013e3181f579f7.
[25] C. Tudor-Locke *et al.*, "Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the international study of childhood obesity, lifestyle and the environment (ISCOLE)," *Int. J. Behav. Nutrition Phys. Activity*, vol. 12, no. 1, p. 11, Dec. 2015, doi: 10.1186/s12966-015-0172-x.
[26] F. B. von-Bischhoffshausen, B. Muñoz, A. Riquelme, M. J. Ormeño, and J. C. Silva, "Spectacle-wear compliance in school children in Concepción Chile," *Ophthalmic Epidemiol.*, vol. 21, no. 6, pp. 362–369, Dec. 2014, doi: 10.3109/09286586.2014.975823.
[27] L. R. Yingling *et al.*, "Adherence with physical activity monitoring wearable devices in a community-based population: Observations from the Washington, DC, cardiovascular health and needs assessment," *Transl. Behav. Med.*, vol. 7, no. 4, pp. 719–730, 2017.
[28] J. Müller, A.-M. Hoch, V. Zoller, and R. Oberhoffer, "Feasibility of physical activity assessment with wearable devices in children aged 4–10 years—A pilot study," *Frontiers Pediatrics*, vol. 6, p. 5, Jan. 2018.
[29] T. S. Marin *et al.*, "Examining adherence to activity monitoring devices to improve physical activity in adults with cardiovascular disease: A systematic review," *Eur. J. Preventive Cardiol.*, vol. 26, no. 4, pp. 382–397, Mar. 2019.
[30] N. Elmagboul *et al.*, "Physical activity measured using wearable activity tracking devices associated with gout flares," *Arthritis Res. Therapy*, vol. 22, no. 1, pp. 1–9, Dec. 2020.
[31] M. L. McHugh, "Interrater reliability: The Kappa statistic," *Biochem. Med.*, vol. 22, no. 3, pp. 276–282, 2012.
[32] M. Pedley, "Tilt sensing using a three-axis accelerometer," *Freescale Semicond. Appl. Note*, vol. 1, pp. 2012–2013, Mar. 2013.
[33] T. Ghosh, D. Hossain, M. Imtiaz, M. A. McCrory, and E. Sazonov, "Implementing real-time food intake detection in a wearable system using accelerometer," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Mar. 2021, pp. 439–443.
[34] C. Bunkhumpornpat, K. Sinapiromsaran, and C. Lursinsap, "Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem," in *Proc. Pacific–Asia Conf. Knowl. Discovery Data Mining*, 2009, pp. 475–482.
[35] J. F. Knight and C. Baber, "A tool to assess the comfort of wearable computers," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 47, no. 1, pp. 77–91, Mar. 2005, doi: 10.1518/0018720053653875.
[36] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

Tonmoy Ghosh (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical electronic engineering (EEE) from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, USA. Before this, he was a Lecturer with the Electrical and Electronic Engineering Department, Pabna University of Science and Technology (PUST), Pabna, Bangladesh. His research interests include applying signal processing and machine/deep learning-based method to address engineering problems in images, wearable sensors, and interdisciplinary research with a particular focus on computer-aided detection and health monitoring applications. He also serves as a Reviewer for several journals, including IEEE, Elsevier, and other publications.

Delwar Hossain received the bachelor's degree in electrical engineering from the Khulna University of Engineering and Technology, Khulna, Bangladesh. He is currently pursuing the Ph.D. degree in electrical engineering with The University of Alabama, Tuscaloosa, AL, USA. His research interests include the development of wearable systems, sensor networks, and machine learning algorithms for preventive, diagnostic, and assistive health technology with a special focus on physical activity and dietary intake monitoring.

Edward Sazonov (Senior Member, IEEE) received the Diploma of Systems Engineer degree from the Khabarovsk State University of Technology, Russia, in 1993, and the Ph.D. degree in computer engineering from West Virginia University, Morgantown, WV, USA, in 2002. In 2020, he served as the Fulbright Distinguished Chair of the University of Newcastle, Australia. Currently, he is a Professor with the Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL, USA, and the Head of the Computer Laboratory of Ambient and Wearable Systems (http://claws.eng.ua.edu). His research interests span wearable devices, sensor-based behavioral informatics, and methods of biomedical signal processing and pattern recognition. Devices developed in his laboratory include a wearable sensor for objective detection and characterization of food intake; a highly accurate physical activity and gait monitor integrated into a shoe insole; a wearable sensor system for monitoring of cigarette smoking; and others. His research was recognized by several awards, including best paper awards, the President's Research Award at The University of Alabama, and others. His research has been supported by the National Institutes of Health, the National Science Foundation, the National Academies of Science, as well as by state agencies, private industry, and foundations. Dr. Sazonov serves as an associate editor for several IEEE publications.