Development of a Method for Compliance Detection in Wearable Sensors

Siavash Esfandiari Fard
Department of Electrical and Computer Engineering
University of Alabama
Tuscaloosa, AL 35487, USA
sesfandiarifard@crimson.ua.edu

Tonmoy Ghosh
Department of Electrical and Computer Engineering
University of Alabama
Tuscaloosa, AL 35487, USA
tghosh@crimson.ua.edu

Delwar Hossain
Department of Electrical and Computer Engineering
University of Alabama
Tuscaloosa, AL 35487, USA
dhossain@crimson.ua.edu

Megan A McCrory
Department of Health Sciences
Boston University
Boston, MA 02215, USA
mamccr@bu.edu

Graham Thomas
Psychiatry and Human Behavior
Warren Alpert Medical School of Brown University 196 Richmond St, Providence, RI, 02903, USA
john_g_thomas@brown.edu

Janine Higgins
Division of Endocrinology, Metabolism, and Diabetes
University of Colorado Anschutz Medical Campus
Aurora, Colorado, USA
Janine.Higgins@childrenscolorado.org

Wenyan Jia
Department of Neurological Surgery
University of Pittsburgh
Pittsburgh, PA 15261, USA
jiawenyan@gmail.com

Tom Baranowski
USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine
Houston, TX 77030, USA
tom.baranowski@bcm.edu

Matilda Steiner-Asiedu
Department of Nutrition and Food Science
University of Ghana
Legon, Ghana
tillysteiner@gmail.com

Alex K. Anderson
Department of Nutritional Sciences
University of Georgia
Athens, GA 30602, USA
fianko@uga.edu

Mingui Sun
Department of Neurological Surgery
University of Pittsburgh
Pittsburgh, PA 15261, USA
drsun@pitt.edu

Gary Frost
Section for Nutrition Research, Department of Metabolism, Digestion and Reproduction, University of Imperial College London
London SW7 2BX, UK
g.frost@imperial.ac.uk

Benny Lo
Hamlyn Centre, Department of Surgery and Cancer
Imperial College London London SW7 2AZ, UK
benny.lo@imperial.ac.uk

Edward Sazonov
Department of Electrical and Computer Engineering
University of Alabama
Tuscaloosa, AL 35487, USA
esazonov@eng.ua.edu

2023 International Conference on Electrical, Computer and Energy Technologies (ICECET) | 979-8-3503-2781-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICECET58911.2023.10389483

***Abstract*— One of the crucial elements in studies relying on wearable sensors for quantification of human activities (like physical activity or food intake) is the assessment of wear time (compliance). In this paper, we propose a novel method based on the Automatic Ingestion Monitor v2 (AIM-2), deployed for measuring nutrient and energy intake. The proposed method was developed using data from a study of 30 participants for two days each (US dataset) and tested with an independent dataset (Ghana dataset) on 10 households (30 Participants, 3 days for each, a total of 90 days). The signals from the accelerometer sensor of the AIM-2 were used to extract features and train the gradient-boosting tree classifier. To reduce the error in the classification of non-compliance in situations where the sensor changes its position with respect to gravity, a two-stage classifier followed by post-processing was introduced. Previously, we developed an offline compliance classifier, and this work aimed to develop a classifier for a cloud-based feedback system. The accuracy and F1-score of the developed two-phase classifier based on K-fold validation for the training and validation dataset were 95.37% and 96.93%, and for the Ghana dataset, were 95.86% and 92.56%, respectively, showing satisfactory performance results. The trained classifier can be deployed to monitor compliance with device wear in real-time applications.**

***Keywords*— Compliance Detection, Accelerometer Sensor, Embedded Systems, Wearable Sensors**

***Clinical Relevance*— Food Intake and physical activity studies can contribute to detecting, controlling, and even improving eating or physical activity-related problems, like obesity, diabetes, and eating planning. To ensure effective monitoring, compliance with the wearing of the device is crucial.**

I. Introduction

The use of wearable sensors to track human activities gives us more useful information than traditional self-report methods like questionnaires or recalls. Wearable sensors may provide objective information on human activity only if the user complies with the prescribed wear regimen. Thus, monitoring wear compliance is an important consideration in human studies utilizing wearables.

A few studies investigated compliance detection. In [1], video cameras, neck-worn, and wrist-worn sensors were used to detect compliance in food-intake activities. In the study, participants were free to select the sensors based on their preferences. The data was collected and used to capture eating episodes with the selected sensors. The results of the study show that participants mostly prefer wrist-worn camera sensors. In [2], a sensor-based wearable events classification was proposed to monitor free-living activities. Combined skin temperature and acceleration data were used to train an automated model to detect wear time versus non-wear time based on a predefined threshold. The method was evaluated on 50 participants (30 females and 20 males), including 23 children and 27 adults. The proposed model for combined accelerometer and temperature sensors demonstrated sensitivity of 94%. Studies on wear time compliance for children were reported in [3-5]. In [3], the participants' wear

979-8-3503-2781-6/23/$31.00 ©2023 IEEE

time characteristics were investigated using the waist-worn accelerometer sensors in 9–11-year-old children in 24 h monitoring protocol. The valid cases were defined as having ≥4 days with ≥10 hours of waking wear time per day. In [4], a study to evaluate compliance with wearing eyeglasses was conducted on 270 schoolchildren. The study's overall compliance rate was about 58%, in which older children were less compliant than younger children. The effect of the location of the accelerometer sensor was investigated in [5]. The results showed that in 9-12 year-old children a wrist sensor can lead to greater compliance than the hip sensor. Another study used a dual accelerometer system to capture 24-hour movement patterns in children and adults [6]. The sensors were placed in the participants' thighs and lower back, containing an inbuilt skin temperature sensor, showing higher wear time in adults than in children (71.6% against 41.7%). In [7], a non-wear period detection algorithm based on the wrist-worn accelerometer and temperature sensors was developed to detect periods of non-wear start, demonstrating an F1-score and accuracy of 94.2% and 99.6%, respectively. In [8], a study was performed to differentiate sleep from non-wear based on wrist-worn accelerometer data and is evaluated in 47 children during school. This study showed that non-wear accuracy ranged between 86% and 95%, with the best performance for the standard deviation of the vector magnitude.

To detect and monitor food intake, we developed the Automatic Ingestion Monitor (AIM-2), worn as an attachment to eyeglasses, and developed a method to detect compliance with the device [9]. This paper further develops a novel, lightweight two-stage classifier, which uses only three features calculated from accelerometer data, which can effectively classify compliant/non-compliant events. Also, evaluating the developed model on an independent dataset shows satisfactory results in the case of accuracy and F1-score.

This paper is organized as follows: Section (II) presents the AIM-2 device, overall information about the participants in the study, the definition of compliant wear, feature extraction based on the accelerometer measurements, and different classifiers. In section (III), the results of the different classifiers are shown. In addition, the problem of classifiers' limitations for some stationary scenarios is illustrated, and a two-stage classifier to address this issue is proposed. Lastly, a post-processing method to reduce false negatives and positives is presented. In section IV, the paper is concluded.

II. Methodology

A. Wearable Sensor

The Automatic Ingestion Monitor, version 2 (AIM-2) [10-13], shown in Fig. 1., was used to develop a compliance detection algorithm. The device consists of a low-power 3D accelerometer (ADXL362 from analog devices, Norwood, MA, USA), a chewing sensor, and a 5-megapixel camera with a 170–degree wide-angle gaze-aligned lens. AIM-2 was installed on the right side of the eyeglass, sampled the accelerometer signal at 128Hz, and stored data on an SD card.

B. Data Collection

Two different datasets were used to develop the compliance detection algorithm. In the first dataset, two days of data were collected from 30 volunteer student participants (65% male and 35% female with an age range from 18 to 39 years old) to train and validate the algorithm. On the first day, the participants consumed all the meals in the lab (which was their only restriction), and for the second day, there were no restrictions for meal consumption. There were no specific instructions to wear the device daily, just not wearing it while sleeping, doing water activities, or for any privacy-related reasons. This dataset is collected at the University of Alabama and approved by IRB Protocol #17-005-ME, dated Aug 22, 2017.

An independent dataset was used to evaluate the proposed algorithm's performance. The data collection was done in rural and urban areas in Ghana with 30 households. To conduct this study, two adults and one adolescent for each household were recruited. (Study's approval was granted by The Imperial College Research Ethics Committee under Application No. 18IC4780 and 18IC4795). In the study, the AIM-2 (or another wearable, eButton [14]) was assigned for fathers and adolescents' children, while mothers were given both devices. The collected data were stored on an internal SD card, and after completion of the data collection, they were uploaded to a cloud-based server for processing [14]. From this dataset, 10 households (30 Participants, 3 days for each, a total of 90 days) have been chosen.

D. Compliance Criteria

This study aimed to develop an algorithm to detect whether a person is compliant with wearing the AIM-2. Ideally, users should wear AIM-2 except during sleeping and water activities. However, in real cases, users do not use eyeglasses for multiple reasons, like privacy-related situations, during heavy exercise, or just for no specific reason.

In our definition, we have two classes (compliance and non-compliance), and the non-compliance class has three different scenarios. These scenarios consist of putting the device on a surface (stationary non-wear), wearing the device in an improper way (such as putting it on the forehead), and carrying the device (putting it in a backpack or pocket). Every non-compliance scenario has distinctive characteristics and patterns. For example, for the stationary non-wear, because the position of the eyeglasses is fixed, we expect low change in net acceleration. For carrying the device, despite the high changes in the net acceleration, we expect a limited range of

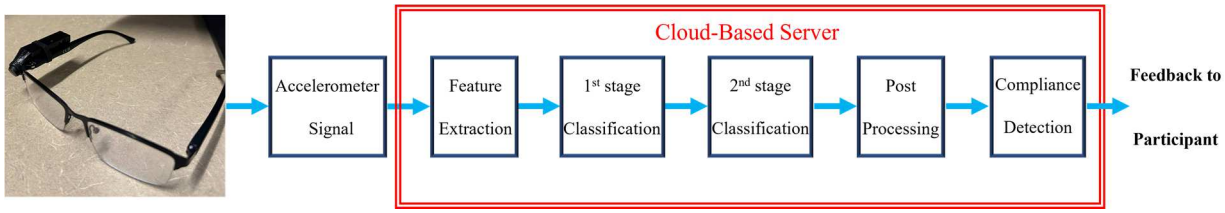

Fig. 1. The workflow of the proposed real-time compliance detection algorithm using the AIM-2 device.

changes in pitch and roll angles compared to the compliant wearing the eyeglasses. Lastly, improperly wearing eyeglasses, like putting them on the forehead, causes different pitch and roll angle ranges compared to the compliant wear. Hence, calculating net acceleration, pitch angle, and roll angle can be beneficial in classifying compliance/non-compliance classes.

E. Feature Extraction

The accelerometer sensors measure the change in the velocity of the device in three directions, which can be used to detect the movement and orientation of the device. The amplitude of acceleration (net acceleration) measures the movement, and pitch and roll angles are the orientations of the sensor in the gravity field. Therefore, for the stationary position of the device, we anticipate that the value of net acceleration was low, while in other cases, it should be relatively high. Moreover, based on the different patterns for the measurement, we can define some features which can distinguish between compliance/non-compliance. For this purpose, we calculated the net acceleration, pitch, and roll angles as follows:

$$Net\ Acceleration\ = \sqrt{Ax^2 + Ay^2 + Az^2} \tag{1}$$

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

$$Roll\ Angle\ = \tan^{-1}\left(\frac{Ay}{Az}\right)\frac{180}{\pi} \tag{3}$$

where Ax, Ay, and Az are the value of the accelerometer at the X-axis, Y-axis, and Z-axis at time t, respectively. This compliance detection algorithm works alongside the food-intake detection algorithm, which classifies 8-second signal intervals. Therefore, to be consistent with the food intake algorithm, we used the same interval decision-making process to detect compliance in 8-second intervals and used the mean value for the net acceleration, pitch, and roll angles in 8-second intervals [15].

F. Classifier

The computed features were used to train the compliance detector. For the classifier training, the collected data of the US dataset was used. The K-Fold technique (with K=6) was implemented to validate the classifier. Therefore, the data from 24 participants were used as the training data, and the remaining data were used as the validation data. This process repeated for all K-Folds, and the classifier never saw the test data during the training.

We trained AdaBoost, Gradient-Boosting, Hist Gradient Boosting, Random Forest, Logistic Regression, Linear Discriminant Analysis (LDA), K Neighbors, Decision Tree, Gaussian NB, and Support Vector Machine (SVM) classifiers to compare the accuracy from different machine learning methods, which is shown in TABLE I.

The classifiers failed to detect some scenarios as non-compliance. These scenarios were mostly related to non-compliance carried and the stationary position of the device. These issues were addressed by introducing the second stage of classification.

TABLE I. Performance of different classifiers for compliance detection

Classifier	Classification Performance				
	Accuracy	Precision	Recall	F1-score	Time (sec)
AdaBoost	94.45 %	95.27 %	97.50 %	96.37 %	0.38
Gradient Boosting	94.46 %	95.19 %	97.59 %	96.37 %	0.0187
Hist Gradient Boosting	93.68 %	93.96 %	97.86 %	95.86 %	0.0469
Random Forest	93.06 %	92.98 %	98.13 %	95.47 %	0.787
Logistic Regression	80.99 %	86.50 %	88.73 %	87.47 %	0.003
LDA	79.88 %	83.20 %	92.56 %	87.32 %	0.003
K Neighbors	92.50 %	92.58 %	97.85 %	95.13 %	1.33
Decision Tree	90.65 %	91.61 %	96.32 %	93.89 %	0.0125
Gaussian NB	86.74 %	87.17 %	96.57 %	91.58 %	0.003
SVM	92.92 %	95.28 %	96.67 %	95.92 %	173.69

G. Developing a Two-Stage Classifier

To improve the classifier's performance in classifying stationary scenarios, we added another stage just before the machine learning classifier. In the first stage, if the standard deviation of the pitch and roll angles over 8 second interval were lower than a threshold, we classified data as non-compliance. If the position changed, we took the classifier's output as the final output.

The first stage classifier established a threshold on the standard deviation of the pitch and roll angles. The value of the threshold was established by splitting the data from the US dataset into training and validation sets, and finding the threshold that produced the lowest misclassification for stationary data (TABLE II).

TABLE II. Different thresholds to identify stationary data in the first stage.

Threshold	**Classification Performance**			
	Accuracy	***Precision***	***Recall***	***F1-score***
0	93.69 %	94.21 %	97.28 %	95.70 %
0.03	93.92 %	94.96 %	97.01 %	95.97 %
0.04	94.27 %	95.63 %	96.87 %	96.11 %
0.05	94.95 %	96.48 %	96.66 %	96.56 %
0.06	94.91 %	96.54 %	96.55 %	96.53 %
0.07	94.88 %	96.59 %	96.45 %	96.51 %
0.08	94.85 %	96.64 %	96.37 %	96.49 %

H. Post-Processing

The goal of post-processing is to reduce false positives and false negatives originating from changes in orientation of the device in the field of gravity. We used majority voting in a fixed-length window (with the length of W), composing the current and L-1 previous predictions.

TABLE III. Effect of choosing different parameters for post-processing.

W_W		Results after Applying Post-Processing				
		L=1	*L*=2	***L=3***	L=4	*L*=5
5	Accuracy	95.01%	95.37%	95.30%	94.87%	93.71%
	F1-Score	96.75%	96.93%	96.90%	96.59%	95.76%
6	Accuracy	94.86%	95.23%	95.30%	95.11%	94.60%
	F1-Score	96.65%	96.88%	96.91%	96.77%	96.40%
7	Accuracy	94.71%	95.12%	95.25%	95.15%	94.91%
	F1-Score	96.56%	96.81%	96.88%	96.81%	96.62%
8	Accuracy	94.56%	94.99%	95.16%	95.14%	94.99%
	F1-Score	96.47%	96.74%	96.83%	96.81%	96.69%

TABLE *III* shows the effect of changing the values of W and L.

During the post-processing phase, we endeavored to reduce classifier by using a trend-based filter. This filter operates with parameters where W denotes the total length of the filter, and L signifies the maximum acceptable error of the classifier. By observing the last W-1 and current estimation, if the count of detected compliance (or noncompliance) instances is equal to or less than L, we classify the current estimation as an error of the detector, and we designate noncompliance (or compliance) as the resulting output. Although the use of this filter may result in misclassifications during the initial L measurements following a transition from a compliant period to a noncompliant one, it enhances overall performance by compensating for the classifier's errors. However, if a substantial number of consecutive misclassifications occur, this post-processing algorithm can only correct the initial L outputs. In this filtering scheme, setting L to 0 implies no post-processing, and as L increases

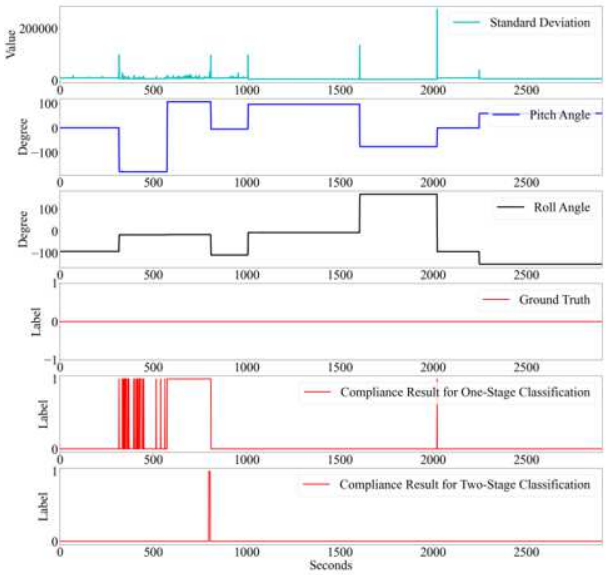

Fig. 2. Compliance Results for One-Stage and Two-Stage Classifier for the stationary data in different scenarios.

in value, more measurements are categorized independently of the classifier's output. In the extreme scenario where W equals L, the current compliance/noncompliance status is estimated solely based on past values.

I. Test Procedure

To assess the performance of the two-stage classifier on detecting non-compliance with a stationary device, we collected a dataset where the AIM-2 device remained stationary in 8 possible orientations for 500-1000s (Fig. 2.). The non-compliant wear was assigned as the ground truth label for the entire stationary dataset.

To evaluate the real-life performance of the trained two-stage classifier, we used the Ghana dataset. From every day of every participant in the dataset, 50 timestamps were randomly generated with a uniform distribution. The compliance label (compliant/non-compliant wear) for every timestamp was verified by manually reviewing 3 consecutive images captured by the AIM-2 camera. The assigned labels were used as the ground truth for classifier validation.

III. RESULTS

A. Classifier Selection

The main factors in evaluating the classifiers were the F1-score and the response time. As seen in TABLE I, the Gradient Boosting classifier has the best accuracy and fastest response over other classifiers. Therefore, based on these considerations, we chose the Gradient-Boosting classifier.

B. Stationary Data Classification

As shown in TABLE II, by increasing the threshold for the standard deviation of the pitch and roll angles, accuracy and precision improve because more stationary data can be detected and relabeled as non-stationary. Conversely, increasing the threshold led to lower recall, mainly because of the situations in which participants wore the eyeglasses while they were lying down. Therefore, we calculated the standard deviation of the pitch angle and roll angle for the 8 seconds intervals (1024 measurements) and assigned the value of 0.05 degrees in the standard deviation of the pitch and roll angles as the threshold.

Fig. 2. shows the stationary data for different eyeglass positions after application of two-stage classification with the selected threshold. As shown in Fig. 2., introducing the first stage classification leads to a significant reduction in false positives and improvement in accuracy (95.37% and 96.93% for accuracy and F1-score, respectively).

E. Result on Ghana Dataset

The Ghana dataset was used as an independent dataset on our cloud-based server to evaluate the results of the proposed algorithm. Regarding accuracy and F1-score, the proposed algorithm achieved 95.86% and 92.56%, respectively. The recall and precision were 97.37% and 88.20%, respectively.

IV. DISCUSSION AND CONCLUSION

We developed a two-phase compliance detector based only on the accelerometer measurements. This paper addresses the misclassification problem of some stationary data by using a

two-phase classifier and performing a postprocessing method to reduce false positives and false negatives. We trained and validated the proposed algorithm on 30 participants. According to the proposed method, accuracy and F1-score for the US dataset were 95.37% and 96.93%, respectively. To evaluate the performance of the developed algorithm, we used the Ghana dataset as an independent dataset, which achieved 95.86% and 92.56% for the accuracy and F1-score, respectively. The difference in the accuracy may be explained by non-compliance carrying of the device, found more frequently in Ghana dataset. The classifier may misclassify the output as compliant wear depending on the AIM's position and movement. Considering the results for these two datasets, the algorithm's performance was satisfactory for compliance detection in human studies using AIM-2 device.

REFERENCES

[1] R. Alharbi *et al.*, "Investigating barriers and facilitators to wearable adherence in fine-grained eating detection," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Mar. 2017, pp. 407–412. doi: 10.1109/PERCOMW.2017.7917597.

[2] 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, p. e007447, May 2015, doi: 10.1136/bmjopen-2014-007447.

[3] 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. Nutr. Phys. Act.*, vol. 12, no. 1, p. 11, Feb. 2015, doi: 10.1186/s12966-015-0172-x.

[4] F. Barria 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.

[5] G. McLellan, R. Arthur, and D. S. Buchan, "Wear compliance, sedentary behaviour and activity in free-living children from hip-and wrist-mounted ActiGraph GT3X+ accelerometers," *J. Sports Sci.*, vol. 36, no. 21, pp. 2424–2430, Nov. 2018, doi: 10.1080/02640414.2018.1461322.

[6] S. Duncan *et al.*, "Wear-Time Compliance with a Dual-Accelerometer System for Capturing 24-h Behavioural Profiles in Children and Adults," *Int. J. Environ. Res. Public. Health*, vol. 15, no. 7, 2018, doi: 10.3390/ijerph15071296.

[7] A. Vert *et al.*, "Detecting accelerometer non-wear periods using change in acceleration combined with rate-of-change in temperature," *BMC Med. Res. Methodol.*, vol. 22, no. 1, p. 147, May 2022, doi: 10.1186/s12874-022-01633-6.

[8] M. N. Ahmadi, N. Nathan, R. Sutherland, L. Wolfenden, and S. G. Trost, "Non-wear or sleep? Evaluation of five non-wear detection algorithms for raw accelerometer data," *J. Sports Sci.*, vol. 38, no. 4, pp. 399–404, Feb. 2020, doi: 10.1080/02640414.2019.1703301.

[9] 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, Feb. 2021, doi: 10.1109/JBHI.2020.2995473.

[10] T. Ghosh, D. Hossain, and E. Sazonov, "Detection of Food Intake Sensor's Wear Compliance in Free-Living," *IEEE Sens. J.*, vol. 21, no. 24, pp. 27728–27735, Dec. 2021, doi: 10.1109/JSEN.2021.3124203.

[11] E. S. Sazonov and S. Schuckers, "The Energetics of Obesity: A Review: Monitoring Energy Intake and Energy Expenditure in Humans," *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 1, pp. 31–35, Feb. 2010, doi: 10.1109/MEMB.2009.935470.

[12] M. Farooq, A. Doulah, J. Parton, M. A. McCrory, J. A. Higgins, and E. Sazonov, "Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment," *Nutrients*, vol. 11, no. 3, 2019, doi: 10.3390/nu11030609.

[13] A. Doulah *et al.*, "Meal Microstructure Characterization from Sensor-Based Food Intake Detection," *Front. Nutr.*, vol. 4, 2017, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnut.2017.00031

[14] J. Qiu *et al.*, "Egocentric image captioning for privacy-preserved passive dietary intake monitoring," *IEEE Trans. Cybern.*, 2023.

[15] T. Ghosh and E. Sazonov, "A Comparative Study of Deep Learning Algorithms for Detecting Food Intake," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2022, pp. 2993–2996. doi: 10.1109/EMBC48229.2022.9871278.