Progress in Nutrition 2024; Vol. 26, N. 1: e2024014 DOI: 10.23751/pn.v26i1.15009

Original article

A cognitive-automated ingestive behavior and physical activity monitoring system using IoT

V P Jayachitra, M Mathangi
Department of Computer Technology, MIT Campus, Anna University, Chennai, India

Abstract. *Background and aim:* To have a healthy lifestyle and to reduce the risk of chronic diseases, a fully balanced diet with enough quantity of nutrients is required. Maintaining the required levels of consumption of food and regularity in eating habits is important for preserving a healthy life and live without any health-related diseases. In regards to the above statement, the proposed work has been defined to provide novel smart dietary solutions and to monitor and detect the nutrition intake each day. *Methods:* Comfortable sensors such as PPG sensor incorporating a BPW34FS photodiode, SFH4247 LED, and FG-23329D65 microphone are worn to record the observations in the form of signals during the consumption of foods. The data are pre-processed, and models are proposed to identify eating bouts and swallowing events, using the extracted features. The human body also needs physical movements which are recognized from an LIS3DH Triaxial accelerometer model to analyze the physical behavior. On the aggregation of both the eating and physical activity analyses, this paper proposes to monitor the ingestive habits and physical behavior related to health. *Results:* The detection of chewing occurrences was classified by RMWC, obtaining an accuracy of 97% while the determination of swallowing sounds to obtain accurate chewing rate, was performed by the proposed Temporally Efficient Bidirectional (TEB) algorithm attaining a higher accuracy of 97.51%. *Conclusions:* In this paper, misclassification of saliva from normal swallowing has been overcome. On the contrary, physical activities have been determined by constructing a hierarchy of clusters. Thus, the proposed paradigm determines the ingestive behavior and the energy balance by generating the daily report for the user in an effective way.

Key words: calorie detection, chewing analysis, swallowing detection, discrimination of saliva, neural network

Introduction

Carrying on with a good routine of health gives peace in our life. Variations from the norm in this parity can result in chronic inflammation, which weakens the immune system and heightens the susceptibility to infections, such as heftiness, anorexia (31), and other dietary problems leading to ceaseless maladies if not ideally treated. Since energy imbalance and unhealthy eating behavior are strongly related to chronic diseases, hence, observing the ingestive conduct of people and physical behavior helps evade such problems. Monitoring the Ingestive Behavior (MIB) of a person based on acoustical and visual approaches is salient to the research of eating habits in humans suffering from obesity and eating disorders. More recently, devices were opened to subdue erroneous readings from people's views and to lessen the self-reporting (30) burden by uniquely recognizing food intake episodes (11). Previous approaches have used different types of devices which are non-obtrusive and subtle (10) such as microphones to determine the crushing sounds (29), videofluorographic swallowing study (VFSS) to analyze bolus while swallowing (14), a piezoelectric sensor

to differentiate between liquid and solid food (15), wireless surface electromyogram (sEMG) to classify swallowing and chewing activities (2). Physical behavior analysis of an individual is essential for a healthy life to overcome energy imbalance (23). Presently, there are commercial activity trackers such as fit bit, Actigraph's Uniaxial GT1M, and Triaxial GT3X, which measures the intensities of activities to motivate people to exercise (16). Moreover, indirect calorimetry validates the Tritac-R3D to determine the sedentary behavior of adolescents (17). However, the ActiGraph accelerometers are frequently used for the accurate study of physical activity (12). At present, the Internet of Things (IoT) technology helps in effective activity monitoring system by interconnecting wearable devices with or without human intervention. At the same time, robust, automatic eating guidance based on individuals' ingestion behavior and dietary assessment still needs to be addressed (33).

To bring up a solution, the Ingestive and Physical Activity Monitoring (IPAM) system has been proposed that automatically determines eating behavior using various behavioral indicators through accurate detection of ingestion and estimation of energy intake and expenditure. The data from worn sensors are used to classify the chewing and swallowing activities. The novel Random Median Weighted Classifier (RMWC) built by merging the Random Forest Classifier (RFC) with the Median Weighted algorithm has been proposed to determine the normality of chewing behavior. Swallowing instances have been identified by finding relevant peaks and differentiated from chewing events as determined by the proposed TEB algorithm. Further, the misclassification of saliva from swallowing has been overcome to estimate the food intake precisely. In addition, the chewing rate is determined to estimate the energy intake from the foods consumed. Accordingly, the calories consumed by the subject are calculated depending on the food provided. Besides, the accelerometer device has been integrated to identify various body movements hierarchically, followed by classifying into distinct activities, which assists in estimating the energy expended through it. Finally, the robust IPAM system offers automatic dietary monitoring, assessing ingestive behavior, recognizing physical activities, computing calories, and generating comprehensive reports to provide individuals with a holistic view of their health, facilitating weight loss and promoting a healthy lifestyle, particularly beneficial for those focusing on diet control.

Related works

The ingestive behavior is monitored to understand the health of individuals with great attention to prevent them from spending for health and to avoid getting suffered from chronic diseases. A food intake monitoring system to estimate calories.

On the other hand, M. Farooq et al. (1), evaluated the capability of an accelerometer to monitor the intake in normal subjects providing an accuracy of 87.9%. This device also has the potential to record the physical activities using the accelerometer. Independent models of SVM were used to recognize the physical activities and to identify more postures and evaluate the algorithms in (7). Consequently, in (18), a wide range of different physical activities such as sedentary, household, the gym workouts, and indirect calorimetry was used to obtain Energy Expenditure (EE) reference data. Zui Rhang et al. (27), presented a bottom-up paradigm to detect and recognize the eating events and the individual cycles of chew, possessing a higher time complexity.

As summed-up above, various techniques have been established to identify the chewing and swallowing activities. The physical activities, recorded using a variety of equipment, have only been classified wherein the unsupervised learning approach has not yet been carried out. The misclassification of saliva from swallows and detecting the context of swallow, which plays a vital role to estimate the accurate chewing activity, was absent and not detected. Also, only the top-down and bottom-up methodologies of higher complexity have been followed to detect the chewing cycles.

This paper is organized as follows. Related works are described in Section II. The proposed work and the experimental evaluation are described in Section III and Section IV, respectively. Discussion of results is provided in Section V. The conclusion is given in Section VI.

Proposed work

Recording equipment

In this paper, the dataset used was collected from SPLENDID Chewing Detection Challenge (4). The prototype model combines an audio microphone and PPG sensor incorporated together, placed in the ear pinna (3). The BPW34FS photodiode, SFH4247 light-emitting diode (LED), and the FG-23329D65 microphone that detects the chewing sounds at 48 kHz were used. The PPG sensor captures the jaw movement depending on the blood flow variations when the photodiode generates the voltage, and the LED irradiates the skin. Triaxial accelerometer model LIS3DH by STMicroelectronics is worn around the waist to determine the physical activities. The position of sensors placed is shown in Figure 1. There were 22 subjects out of which 19 were female, and 3 were male. Due to hardware discrepancies, 26 sessions from 14 subjects have been collected and used for the analysis. The individuals with mean body-mass-index (BMI) of 28±2.3kg/m^2 and mean age of 22.9±1.9 years participated. Different varieties of food like meat, orange juice, etc., were provided to them and were allowed to consume the desired quantity. The subjects were informed to consider at least four physical activities and three eating events. Further, the dataset can be downloaded from https://dx.doi.org/10.17026/dans-zxw-v8gy, as given in (4).

Signal pre-processing

A high pass FIR filter with a cut-off frequency of 0.5 Hz, 20 Hz, and 1 Hz is applied to the PPG, microphone and accelerometer signals respectively as a pre-processing step. The PPG signal is low sampled at 21.3 Hz (3,4), and the accelerometer signal is sampled at the same frequency as the PPG signal. Then, adaptive amplification is performed on PPG to accurately detect the events. The microphone recordings are downsampled at 2 kHz, to match with the design parameters of the setup, undergo pre-processing as done in (3), and normalized by scaling the values to remove any unwanted external noise. Besides, the intensity of physical activity is obtained from the raw acceleration signals (13). Thus, the necessary pre-processing steps have been undertaken.

IPAM system

The workflow of the proposed novel IPAM system, as shown in Figure 2, depicts the behavioral analysis and the nutrient estimation for the individual.

Figure 1. Prototype sensors (4).

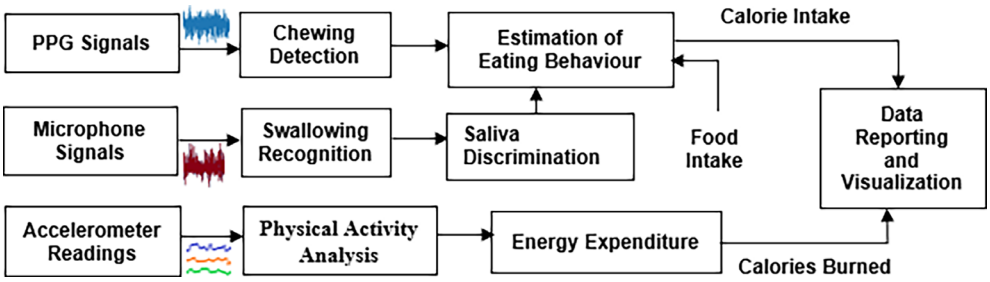

Figure 2. Workflow of proposed IPAM system.

The recording equipment consists of a PPG sensor, a Microphone, and a Triaxial Accelerometer. The PPG readings are used to recognize the chewing activity, and microphone recordings are used to detect the instances of swallow activity. To analyze the calorie intake of a person, there is a need to determine the chewing rate as the energy intake depends on it (32). To estimate the chewing rate accurately, the discrimination of saliva from swallow is necessary, and thus, the eating behavior can be determined. Parallelly, the physical activities are determined by the accelerometer, and the energy expended through it is calculated. A final report consisting of information about the eating behavior and calories consumed and expended is generated to stay hygienic and healthy.

C.1 Chewing detection

For the detection of eating episodes and absorbing the nutrients of the food well, monitoring of how the individual masticates and how their body behaves in response to ingestion is necessary. To extract the features, the time varying spectrum (TVS) of the signal is estimated using the Welch method (3) by calculating the Discrete Fourier transform (DFT) coefficients for each window. The sample size is considered as N=128, and the Hamming windowing technique is applied over each sliding window. The sample step is taken as 2, and the sliding window is moved for every 6 seconds. Since there are overlapping windows for each frame, the optimal average of the DFT coefficients is considered.

$$P_n[q] = \frac{1}{M}\sum_{n=0}^{M-1} X_n[q]e^{-j\pi qn/N} \qquad (1)$$

In equation (1), $P_n[q]$ denotes the DFT coefficients for q = 1,2…N for the n^{th} window, *X* represents the pre-processed data and *j* is a complex number. The $P_n[q]$ is computed only up to M=N/2 values since the signal is of real-valued. Thus, a ten-dimensional spectral log energy feature vector *F* is obtained consisting of energy values for each timestamp in frequency bands of 0.0–1.0, 1.0–1.8, 1.8–3.3, 3.3–5.9, and 5.9–10.7 (Hz). It is identified that only the secondary (1.0-1.8 Hz) and tertiary (1.8-3.3 Hz) bands show the highest chewing properties and are considered as the chewing bands.

Algorithm 1: RMWC Chew Monitoring

Input: Pre-processed Training data X, Extracted Features F

Parameters: Classifier C, Threshold Value V, Weights W, Median Array M, Estimated Value E, Matrix A, No. of chewing samples R, Loop variables i, j, Intermediate Value K

Output: Inference I

1: Select desired features from F and build an Ensemble learning model C

2: Assign W to F

3: **for** each F in X

4: Calculate M

for j=0 to R-1

$$\sum_{(i=0)}^{(R-1)} M_i = X_i + \frac{(X_{(j+R-1)} - X_i)}{2}; \text{j = j + R}$$

5: Combining W and M

for i=0, j=0 to R-1

$$A_{ij} = W_i * M_i; \text{j = j + R}$$

6: Determine V $V = \sum_{j=1}^{R} \frac{A_j}{F}$

7: Find E for testing data and calculate

$$I = \begin{cases} -1, E < V \\ 0, \quad E = V \\ +1, E > V \end{cases}$$

The next step after extracting the features is classification. The Random Median Weighted Classifier (RMWC) model is proposed where the random forest classifier (RFC) and the median weighted algorithm have been combined. RMWC model is employed as it could provide efficient estimations and has the power to handle a large amount of data. As the weighted median algorithm (19) is robust to outliers, the chewing behavior can be detected more accurately. The model uses 70% of data for training and the remaining 30% for testing. The features of the model are assigned with high weights and low weights, decided empirically, to the chewing bands and the second-order priority bands, respectively. The median value is computed on the training set for every feature and combined with weights to calculate the threshold (as step 6 in Algorithm 1) and to be compared with the estimated value in the testing data. The ensemble learning is applied by building sub-models that predict a class, and the class with maximum value is considered as the output. Thus, the output provides information about individual chewing behavior with respect to normality. The 10-fold cross-validation is used as it has lower variance and the classifier produced a higher accuracy of 97%, as discussed in Section IV. Besides, the data containing only the chewing signals is considered such that (-1) indicates below normal, (0) indicates normal and (+1) indicates above normal chewing activity. The non-chewing segments are not considered for detection purposes. Therefore, the three classes distinguish how the chewing behavior of an individual varies with respect to normal chewing. The normality refers to the normal chewing behavior.

C.2 Swallowing recognition

The necessity of a swallowing recognition algorithm is to avoid being prone to diseases such as Dysphagia and Parkinson's, which lead to difficulty in swallowing, especially in elderly people (22). Generally, the swallows can occur at unspecified times (28), which is to be identified. The swallowing events are most often clearly visible as peaks (6), and moreover, the liquid contents result in higher and sharper amplitude and magnitude in comparison with the solid items which are shallow. Therefore, there is a need to detect swallow events efficiently. In this paper, Temporally Efficient Bidirectional (TEB) Chew bout detection algorithm has been proposed to determine the swallow peaks, discrimination of salivary events, and the presence of consecutive chews based on the context of occurrence of swallow. At first, the input microphone signal X is univariate uniformly sampled into N samples, and the samples are processed by sliding the window W where $\{W_m = 2m | m = 1, 2, \ldots, L\}$, and $L = (N/2) - 1$. The data is detrended quadratically and polynomial function is applied by finding the least-squares fit in contrast to the existing approach to determine the obtained sequence of values. Then, a matrix of values is generated using a random number in uniform distribution around the local maxima within a specific time frame, as in (2). A minimum threshold value is evaluated by calculating the row wise summation of M. The global minimum value is picked up for calculating the column wise standard deviation of M and thus the peak vector is obtained.

$$\text{Value n in Matrix M} = \begin{cases} 0, \ X_{m-1} < X_m > X_{m+1} \\ \mu, \ otherwise \end{cases} \quad (2)$$

In equation (2), M represents L* N matrix, μ represents a random number in a uniform distribution in the range [0,1], and the set of values of n where μ+constant factor are assigned for values from n=1,…, m+1 and n=N-m+2,…, N. However, often there is misclassification of saliva and swallow instances leading to the problem of classifying saliva events as swallow occurrences. Hence, the discrimination of the saliva events from swallowing activity is necessary and is performed after identifying the peaks. Swallowing of saliva normally occurs only when there are not many eating instances. The probability of consuming saliva is high and is being detected only in the silent periods of the signal where much of eating events have not been taken place. Accordingly, the feature extraction is used as performed in (3) on microphone signals by

estimating the TVS, which captures frequencies till 1000 Hz, and the coefficients are collected in nine energy bands with the window size of 0.3 seconds. The values of Skewness and Kurtosis (21) plays a vital role in determining the characteristics of salivary consumption. The distribution of values, whether positively or negatively skewed and higher peaked value, determines the swallowing of saliva. The energy capacity of saliva swallow is of shorter amplitude and low production in comparison to that of any solid swallow (SS) or liquid swallow (LS) and is represented in equation (3).

Energy for Saliva (E_S) < Energy for Swallow (E_{SW}) (3)

Algorithm 2: TEB

Input: Pre-processed Input data X

Parameters: No. of Samples N, Ceil values array L, Matrix LSM, Time t, Polynomial function P, row-major calculation μ_k, global minimum β, Loop variable i, Sliding Window W_n, Peak vector P_Q, Previous Peak Index P_{pr}, Next Peak Index P_{nt}, Signal Values S, Window Index W, Duration D, Intermediate Value ct, Intermediate Value V, Condition Variable Left_Flag, Condition Variable Right_Flag, Skewness S_K, Kurtosis K, constant α, Energy of Saliva E_S, Saliva events vector S_Q

Output: Chewing Rate C_R, Solid Swallow SS, Liquid Swallow LS

1: Evaluate P(t) $\sum_{i=1}^{N} f(P(t)) \; for \; X_i$

2: Fit values of L=(N/2) – 1 for each W_n

3: Update values in LSM as given in (2) and find β = min(μ_k)

4: Peak vector P_Q is determined

5: **for** each i in P_Q

6: V = Calculate S_K, K and E_S

7: **if** V represents E_S as in (3), S_K, and K, append V to S_Q

8: Update P_Q removing S_Q

9: **for** each consecutive peak in P_Q

10: Calculate S $S_i = \sum_{i=P_{pr}}^{P_{nt}} X_i$

11: Initialize W = (P_{nt} – P_{pr})/2, Left_Flag=Right_Flag= False

12: **for** i=P_{pr} to W

$$W = \begin{cases} ct=ct+\alpha, & X_{i-1}<X_i<X_{i+1} \\ W - W/2, & otherwise \end{cases};$$

Left_Flag = True

13: **for** i=W+1 to P_{nt}

$$W = \begin{cases} ct=ct+\alpha, & X_{i-1}>X_i>X_{i+1} \\ W + W/2, & otherwise \end{cases};$$

Right_Flag = True

14: **if** Left_Flag=Right_Flag is **True**

15: $C_R = S_i/D$; print C_R and SS

16: **else**

17: print LS

These features largely contribute to the discrimination to avoid the misclassification, which is likely to obtain. After discriminating saliva from swallow events, for each phase between the detected peaks from the peak vector, a varying window is applied. This window is processed simultaneously from both sides back and forth; that is, the window is moved bidirectionally in order to find the presence of any chewing activity. This bidirectional movement is necessary to understand the contextualized occurrences of swallow. For e.g., the chewing of crispy foods into pieces will be classified as SS, as they must have been broken down before swallowing. Thus, the left and the right half are recursively processed to determine and track the continuous movement of the microphone signals. The signals may also produce continuous peaks representing the consumption of liquid intake without delay. Hence, the context, i.e., the swallow occurred because of liquid or solid food consumption, is determined to avoid any misclassification in swallow counts and chew signals. Thus, the swallow events are even more accurately determined. Besides, the chewing rate is estimated considering the duration and length of each signal from bottom to top and reaching bottom again, as stated in TEB. As the loop runs from both sides, TEB executes in lesser time than the usual peak detecting algorithm. This bidirectional way of determining the presence of chewing cycles and bouts overcomes the higher execution time of the bottom-up algorithm (27). Hence, the time complexity optimally reduces to O (log n), which is better than the linear processing of $O(n^2)$.

C.3 Physical activity analysis

Motion in a subject is usually characterized by the speed of movement or acceleration A. Static acceleration, owing to gravity and dynamic acceleration, can be detected by a device called triaxial accelerometer (7). The accelerometer captures the signals along x, y, and z axes, with the timestamp. Because of the gravity in the downward direction towards the earth's surface, an accelerometer at an idle or inert state may move upwards. Hence, a high pass FIR filter with 1 Hz is applied to eliminate the 1g acceleration for all the objects subjected to it (26).

$$\Delta A = abs(A(time_{i+1}) - A(time_i)) \quad (4)$$

The change in acceleration ΔA, calculated as in equation (4) is necessary to decide the activity being performed with respect to change in time ΔT. The calculated features such as acceleration, change in time and the tilting angles, as in equation (5), (6), and (7), namely pitch ρ, theta θ, and phi ϕ, between the dimensions, give us postures of the activities relative to the ground (25). For predicting different activities into groups, the rule-based classification based on the IF-THEN rules framed is used to satisfy two or more conditions simultaneously.

$$\rho = 180/\pi * arctan\left(x / \sqrt{y^2 + z^2} \right) \quad (5)$$

$$\varphi = 180/\pi * arctan\left(y / \sqrt{x^2 + z^2} \right) \quad (6)$$

$$\theta = 180/\pi * arctan\left(z / \sqrt{x^2 + y^2} \right) \quad (7)$$

The rules are manually identified based on the values present in the dataset. The optimally suitable rules are chosen to classify the activities and the rules generated are defined in Table 1.

The subset (ΔT, ΔA) provides the intensity and duration of the activity, and the tilt angle measurement subset (ρ, θ, ϕ) gives us the posture information and angle of twist. These multiple subsets (20) are formed and used as features to identify which subset gives optimal performance to distinguish one activity from another. The identified features are fed as input to the hierarchical clustering algorithm, which forms clusters starting from the single large cluster decomposed into smaller atomic clusters as the points were far apart. Euclidean distance is used to determine the closeness of two points, and proximity helps in recognizing similar actions. Then, using the similarity measure, the least similar clusters are grouped together to form five different activities. To distinguish the activities more accurately, the obtained cluster labels are considered a feature and fed as input to the RFC model for better classification and discussed in Section V.

Table 1. Generated rules for classification of activities.

Rule1	ΔA > 4.2 m/s^2 for ΔT of 3 seconds, Activity=Motion
Rule2	ΔA < 0.2 m/s^2 for ΔT of 6 seconds, Activity≠Motion
Rule3	Activity≠ Motion and ρ≈90°, θ≈0°, ϕ≈0°, Standing
Rule4	Activity≠ Motion and ρ≈ 90°, θ ≈ 90°, ϕ ≈ 0°, Sitting
Rule5	Activity=Motion and ρ≈90°, θ≈105°-115°, ϕ≈170°-180°, Walking
Rule6	Activity= Motion and ρ≈42°, θ≈9°-13°, ϕ≈13°-16°, Walking including stairs
Rule7	Activity= Motion and ρ ≈40°-50°, θ ≈ 140°-150°, ϕ ≈ 90°-100°, cycling

C.4 Eating behavior estimation

As saliva swallow instances play a vital role in monitoring the food intake and detecting the chewing activity, the discrimination of saliva from solid/liquid swallow can accurately estimate the eating behavior of an individual. Spectrograms are computed on chewing signals in time-frequency domain where the fast Fourier Transform (FFT) size=128 samples and sample rate=2000. The rate of chewing from TEB is used to

Algorithm 3: Estimation of Calories
Input: Change in Acceleration ΔA, Change in Time ΔT, Rho ρ, Phi ϕ, Theta θ, signal values (of an accelerometer) i, Clusters C

Parameter: MET Values V, Macronutrients M, Quantity Q, BMI B, Initial value I, Energy Equivalents EE, Intermediate Value A

Output: Calories Intake C_{in} and Calories Expended C_{out}

1: **for** each i
2: calculate average of ΔA with respect to ΔT
3: calculate the values of ρ, ϕ, θ from equations (5,6,7) for each day
4: C_{out} =0
5: **for** each activity in C
6: Determine V for activity
7: I = Combining B and V
8: C_{out} = C_{out} + I
9: print C_{out}
10: **for** each food
11: Determine Q
12: **for** each row in M
13: A = Multiply row and EE
14: C_{in} = C_{in} + A
15: return C_{in}

determine the calorie consumption as it is affected by the energy intake. Additionally, the human body also requires the intake of minerals and vitamins to maintain the dietary level of energy. The intake of calories depends on many determinants such as age, weight, height, gender, etc. The major macronutrients (M) consumed regularly are Carbohydrates, Fats, Proteins, Alcohol, and is considered for estimation. A standard 1 cup measurement is considered as 250ml, and 1 spoonful is considered to be 15 ml. The quantity (Q) of food is obtained from the dataset, containing the details about the foods consumed. The energy equivalents of the macronutrients are identified in kJ. The Q and the number of servings add on to estimate the intake of calories (36). The M is multiplied with its appropriate EE values in kcals and is summed up for all the Figure 3 macronutrients of the food. The higher and lower chewing rate per bite, ingested low and high energy into the body respectively. The base value, quantity of the food was considered from the dataset for ground truth and further estimated the required information. The estimation of calories is described in algorithm 3. The energy expenditure is calculated by combining the MET value of each activity multiplied with BMI value and the time taken to perform the activity. Similarly, it is calculated for each activity and summed up to obtain the total energy expended. A correlation graph illustrating the estimated calorie versus the actual calorie intake, demonstrating a high correlation coefficient (R^2) of 0.9057 and is shown in Figure 3.

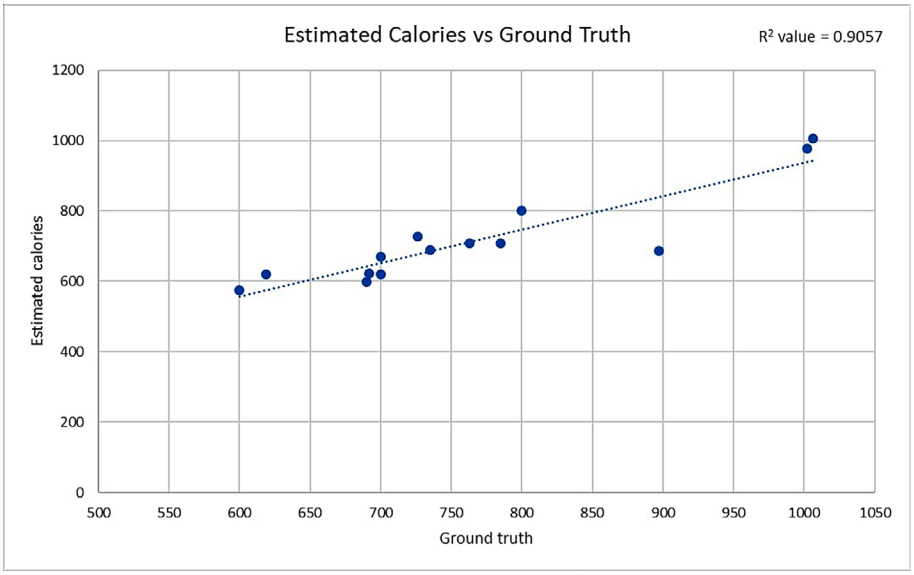

Figure 3. Correlation graph illustrating the estimated calorie versus ground truth.

Experimental evaluation

The experiments were conducted to evaluate and validate the performance of ingestive events and assess the physical activity. Primarily, the learning curve of the RMWC model indicating the cross-validation score in (%), is shown in Figure 4. Besides, the accuracy of the RMWC model turns out to be 97%, whereas 88% is obtained for SVM as given in (3). The parameters for SVM are regularization parameter $C\epsilon\{10^i, i = -2,-1,0,1,2\}$ and radial basis kernel $\gamma\epsilon\{\gamma_0^i, i = -2,-1,0,1\}$, where $\gamma_0 = D^{-1}$ is the standard value in libSVM library. Here, D represents the number of features (D = 15 for audio and D = 10 for PPG) (3).

The SVM classifier could not fit since the dataset is very large and not robust to outliers. On the contrary, the parameters considered for RMWC are max_features=sqrt (no. of features), n_estimators=110, min_samples_leaf=2 and min_samples_split=2. Thus, RMWC involved constructing smaller decision trees where the weights assigned were of energy features. Here, ensemble learning was performed achieving a significant performance. Furthermore, the precision vs. recall curve of RMWC model achieves an average precision of 0.98. In equation (8), True positive (TP) indicates the number of correctly classified chewing events.

$$\text{Precision} = TP/(TP+FP), \quad \text{Recall} = TP/(TP+FN) \tag{8}$$

False positive (FP) and False negative (FN) indicate the number of chewing events that were not considered above and below normal respectively. In order to continuously monitor and evaluate the RMWC model on different subjects, cumulative based evaluation is carried out. This evaluation of the precision and recall of all the subjects are shown in Figure 5 providing a clear depiction of the proposed RMWC model with higher accuracy.

Figure 6 shows the analysis between an individual and conventional (normal) chewing activity, indicating fast chewing behavior as the deviation is high. The y axis indicates the amount of variation between the two and x axis indicates the time in secs. Moreover, the completion of chewing activity falls in the range of [0,18000] seconds. This depicts that the chewing behavior is above normal conduct. The identification of fast chewing indicates that the person is not biting

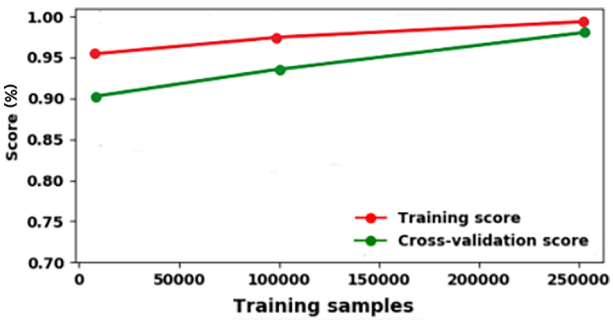

Figure 4. Learning Curve of RMWC.

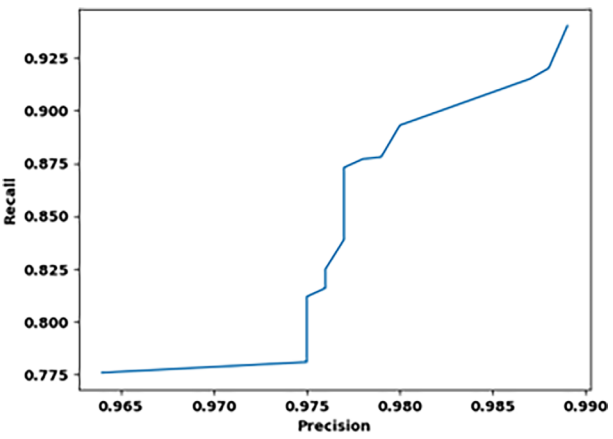

Figure 5. Evaluation of chewing activity on all subjects.

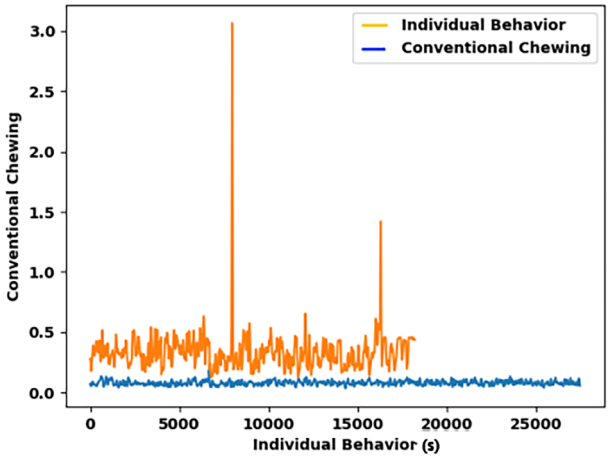

Figure 6. Chewing behavior analysis of subject 1.

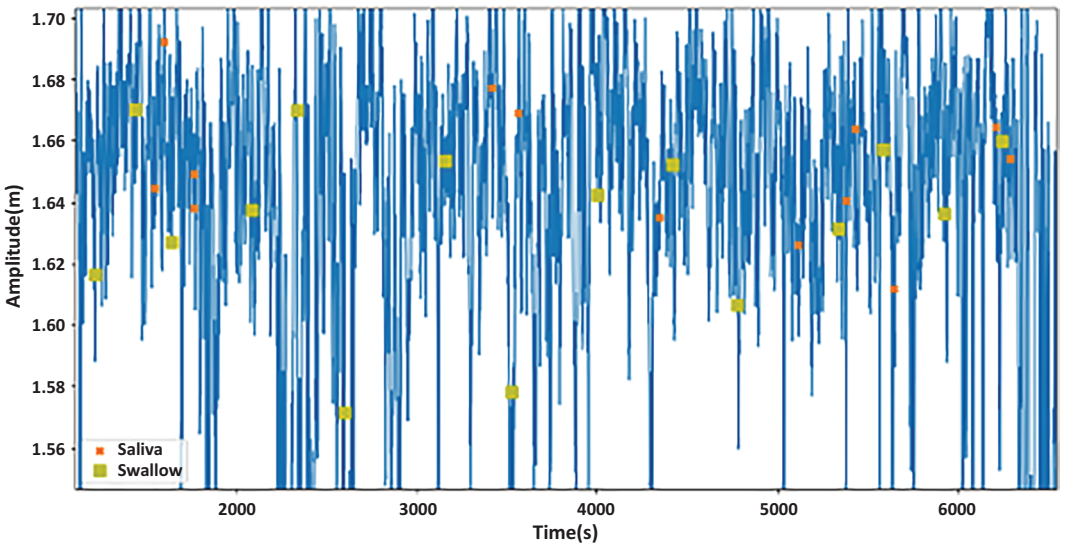

Figure 7. Saliva Vs. Swallow Discrimination (Zoomed).

the food into pieces and is consuming in a high-speed manner than normal, leading to indigestion and acid reflux problems. Moreover, the discrimination of saliva from the swallow instances is an unaddressed challenge and is distinguished based on the characteristics of saliva efficiently by the IPAM system. The features such as the energy during the silent periods in the signal, skewness, and kurtosis are used for discrimination, and the graph is shown in Figure 7. Determining the percentage of the presence and absence of saliva events in swallow occurrences by discriminating them correctly avoids any false assumption of ingestive activities.

The terms "Swallow with Saliva" is referred as "Saliva Swallow" and "Swallow without Saliva" is referred as "Swallow instances" or as "Solid Swallow/ Liquid Swallow". The comparison graph of the execution time between the existing Chewing cycle detection (27) algorithm with TEB has improved by 8% and is depicted in Figure 8. The swallow events are detected with higher accuracy as shown in Figure 9. TEB is validated by building a neural network model with 10-fold cross-validation consisting of 70% of training data and remaining for testing to reduce lot of variations that might affect performance. There are 32 nodes in the first hidden layer and 12 nodes in the second hidden layer with 'ReLU' activation function.

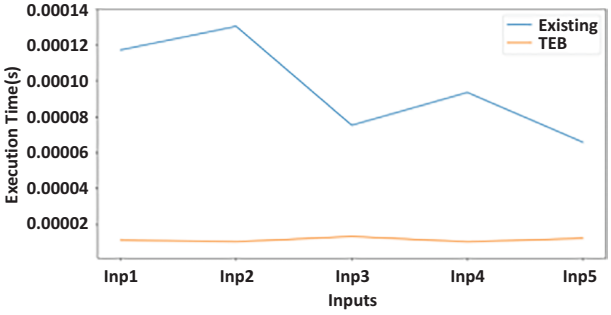

Figure 8. Comparison of Chewing cycle detection algorithm (27) vs. TEB algorithm.

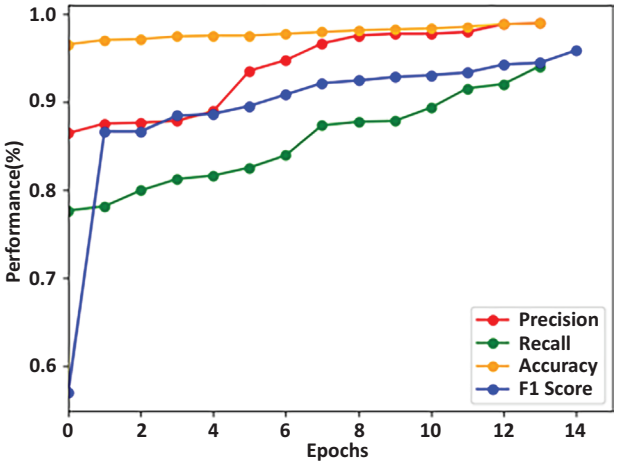

Figure 9. Summary of Metrics of swallowing using TEB algorithm.

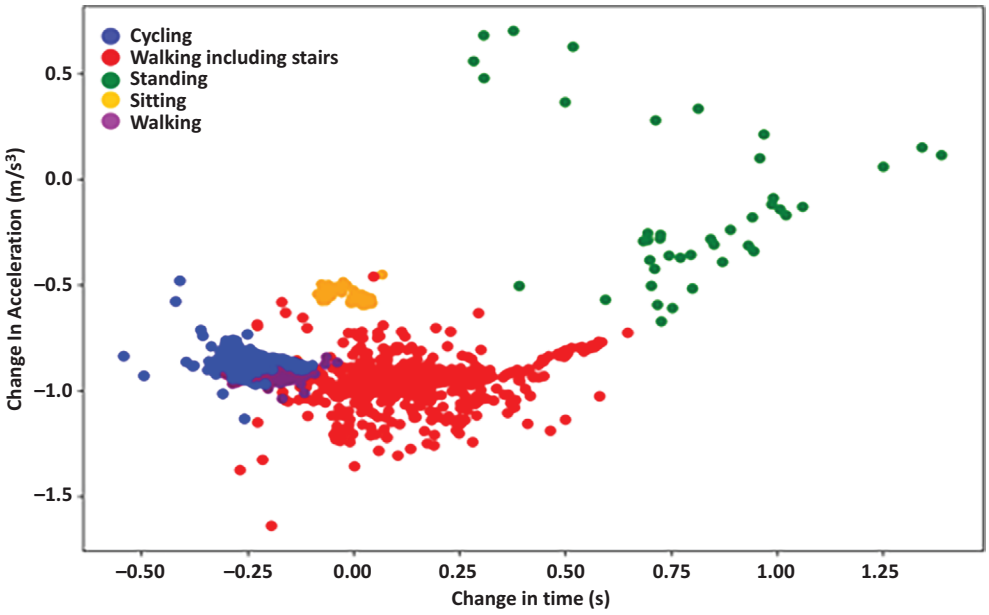

Figure 10. Physical activity analysis.

The classified output from the TEB model is validated against the ground truth labels in the neural network model wherein the label (1) indicates the presence of swallow, and label (0) indicates non-swallow.

An accuracy of 97.5% is obtained, which conveys that the new data can be identified more accurately. Furthermore, the coordination of all three activities chewing, swallowing and physical activities are verified, such that no two activities happen simultaneously.

The synchronization of all the recordings obtained from different devices helps us differentiate between distinct behavior such as chewing, swallowing, and physical activity to avoid any further misclassification between them. The occurrences were classified correctly, such that the events were getting coincided. Next is the recognition of the physical activity, which is a time series prediction, and the rate of change of acceleration provides an important measure in detecting the activity. The activities were discriminated considering the motion of the body, the intensity of movement and the angle of twist. A scatter plot is used to visualize the clusters of different physical activities, as in Figure 10. Moreover, the physical activities were further classified using RFC classifier with higher accuracy and are used to determine the energy expended by the body, as discussed in Section V. The average calories expended on each activity for the entire dataset is shown in Figure 11.

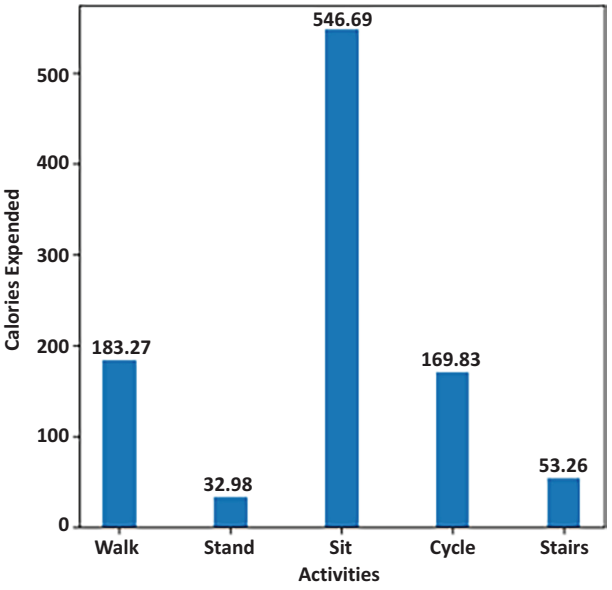

Figure 11. Physical activity evaluation.

Discussion

The objective of this paperwork is to propose an automatic dietary food intake monitoring system. The first step in determining the ingestive activities is to analyze the chewing behavior of an individual. Hence, the chewing behavior of a person is detected and distinguished using the RMWC method as above, below, or normal behavior. The proposed RMWC model produced 97% accuracy and an average precision of 0.98 with cumulative based evaluation. However, distinguishing swallowing events play a vital role in accurately determining the chewing activity. Therefore, to avoid incorrect predictions of chew and swallow counts, there comes a need to determine the context of swallow occurrences to know whether the swallow is a liquid or a solid swallow. Hence, a novel TEB algorithm has been proposed. The relevant swallowing peaks are identified, and the bidirectional method is used to determine the context of the occurrence of swallows. For e.g., the dominant sounds produced by chewing of crispy foods will be detected as the presence of chewing activity before swallowing and is classified as SS by TEB. Thus, the bidirectional execution is necessary. However, the energy intake of a person depends on the rate of chewing and is estimated by TEB. Thus, the detection of all these ingestive activities and chewing rate help determine the proper eating guidelines. The F1 measure of the swallowing evaluation of TEB on the entire dataset reaches about 94.8%. Furthermore, the swallow instances are validated against the ground truth labels by building a bidirectional neural network model achieving the highest accuracy of 97.5%. Next, the energy expended through physical activities is determined. The clustered labels of the physical activities determined are used as an input feature to the RFC model and validated achieving a recognition performance of 98.1% in distinguishing the activities more accurately. The physical activities are determined from the values present in the dataset of accelerometer. The Metabolic Equivalent of Task (MET) values are identified using (35) for each of the physical activity carried out by the participant and the energy expended based on mild or heavy activities were computed. The human weight impacts in estimating the expended energy. The body demands indispensable calories, but the excessive calories supplements to gaining weight. Thus, the physical activities performed by the participants help in calculating the calories expended through it. The IPAM system could now analyze ingestive behavior and nutrient estimation. The data has been gathered from all the detected events and examined to generate a report. The entire ingestive behavior of an individual, and the quantity of energy consumed and spent, have been reported to indicate the progress of the person to know their health. The eating behavior of all the subjects has been tabulated in Table 2 to study how the behavior of the

Table 2. Eating behavior of subjects.

S	SI	Th	AV	CA	TC	CI	CB
S 1	949	0.405	0.26	Below	689.5	621.97	778.05
S 2	258	0.339	0.16	Below	709.8	620.89	1055.6
S 3	770	0.074	0.28	Above	768.3	696.2	1111.7
S 4	100	0.166	0.33	Above	738.6	689.11	809.47
S 5	439	0.176	0.16	Normal	811.2	805.68	895.51
S 6	184	0.101	0.32	Above	728.2	675.24	977.2
S 7	181	0.181	0.34	Above	708.6	658.19	1002.4
S 8	161	0.279	0.07	Below	602.3	520.60	960.59
S 9	261	0.291	0.17	Below	682.4	592.31	962.84
S 10	103	0.286	0.27	Normal	735.9	730.14	694.71
S 11	900	0.281	0.14	Below	899.1	653.27	862.29
S 12	690	0.106	0.34	Above	1012.5	923.44	987.53
S 13	1137	0.300	0.31	Normal	648.2	639.21	924.38
S 14	189	0.278	0.27	Normal	1016.2	1004.3	981.54

events varies from individual to individual. In Table 2, the corresponding representations of S, SI, Th, AV, CA, TC, CI, CB are Subjects, Swallow Instances, Threshold, Amount of Variation, Chewing Analysis, Total Calories, Calorie Intake and Calories Burned respectively. The fact is that the increased chewing reduces the amount of energy intake (32). The above-mentioned statement has been proven to ensure that chewing at a normal rate prevents us from consuming additional foods. Thus, the rate of chewing has an impact on the energy obtained from meals. Consequently, Table 2 draws up the conclusion that the consumption of calories is strongly affected by their ingestive activities. As the IoT technology helps in connecting cost-effective wearable sensors, the IPAM system can be integrated as part of mobile applications.

Conclusion

Thus, the IPAM system helps in dietary food intake monitoring with greater accuracy. Therefore, the proposed work helps us to identify problems and provide a remedy in an efficient and effective way. This novel work helps to define the diet and maintenance of a healthy body by precisely discriminating the ingestive events based on the context of the events obtaining a lower time complexity of O (log n). Consequently, a person will know about their ingestive behavior, which could help them get better solutions to solve any disorders and health-related issues. However, further research is required to understand ingestive behavior considering many subjects of different age groups and to determine the calorie estimation. As a part of future work, detection of simultaneous events in specific time intervals, recommendations on what type of food to consume, and the diet to be followed can be provided for a healthy lifestyle of an individual.

Conflicts of Interest: Each author declares that he or she has no commercial associations (e.g. consultancies, stock ownership, equity interest, patent/licensing arrangement etc.) that might pose a conflict of interest in connection with the submitted article.

References

1. Farooq M, Sazonov E. Accelerometer-Based Detection of Food Intake in Free-Living Individuals. IEEE Sens J. 2018;18(9):3752-3758.
2. Hassan E A, Elbially M S, Morsy A A. Monitoring and evaluation of ingestive activities. IEEE EMBS Int Conf on BHI. 2018; 21-24.
3. Delopoulos. A Novel Chewing Detection System Based on PPG, Audio, and Accelerometry. IEEE J of BHI. 2017;21(3):607-618.
4. Papapanagiotou V, Diou C, Zhou L, et al. The SPLENDID chewing detection challenge. IEEE Int Conf of EMBC. 2017; 817-820.
5. Bi Y, Lv M, Song C, et al. AutoDietary: A Wearable Acoustic Sensor System for Food Intake Recognition in Daily Life. IEEE Sensors J. 2016;16(3):806-816.
6. Kalantarian, Haik, Nabil A, et al. Monitoring eating habits using a piezoelectric sensor-based necklace. Comput Biol Med. 2015;58:46-55.
7. Zhang S, Mccullagh P, Callaghan V. An Efficient Feature Selection Method for Activity Classification. Int Conf on Intelligent Environments. 2014;16-22.
8. Fontana JM, Sazonov ES. A robust classification scheme for detection of food intake through non-invasive monitoring of chewing. IEEE Int Conf of EMBS. 2012;4891-4894.
9. Sazonov ES, Makeyev O, Schuckers S, et al. Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior. IEEE TBME 2010;57(3):626-633.
10. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil. 2012 Apr 20;9:21.
11. Fontana JM, Sazonov ES. Evaluation of Chewing and Swallowing Sensors for Monitoring Ingestive Behavior. Sensor letters. 2013;11:560-565.
12. Migueles JH, Cadenas-Sanchez C, Ekelund U,et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports medicine 2017; 47:1821-1845.
13. Bujari A, Licar B, Palazzi CE. Road crossing recognition through smartphone's accelerometer. 2011 IFIP WD 2011;1-3.
14. Martin-Harris B, Jones B. The videofluorographic swallowing study. Phys Med Rehabil Clin N Am. 2008 Nov; 19(4):769-85, VIII.
15. Farooq M, Sazonov E. A Novel Wearable Device for Food Intake and Physical Activity Recognition. Sensors (Basel). 2016 Jul 11;16(7):1067.
16. Arnin J, Anopas D, Triponyuwasin P, Yamsa-ard T, Wongsawat Y. Development of a novel classification and calculation algorithm for physical activity monitoring and its application. APSIPA. 2014;1-4.

17. Chen KY, Sun M. Improving energy expenditure estimation by using a triaxial accelerometer. J Appl Physiol. 1998;83(6):2112-2122.
18. Altini M, Penders J, Vullers R, Amft O. Estimating Energy Expenditure Using Body-Worn Accelerometers: A Comparison of Methods, Sensors Number and Positioning. IEEE J of BHI. 2015;19(1):219-226.
19. Rauh A, Arce GR. A fast weighted median algorithm based on Quickselect. IEEE Int Conf on Image Processing. 2010;105-108.
20. Kobayashi H, Yoshifumi M. Distribution characteristics of salivary cortisol measurements in a healthy young male population. J Physiol Anthropol. 2015;34:30-33.
21. Cai J, Luo J, Wang S, Yang S. Feature selection in machine learning: A new perspective. Neurocomputing. 2018;300:70-79.
22. Sejdic E, Malandraki GA, Coyle JL. Computational Deglutition: Using Signal- and Image-Processing Methods to Understand Swallowing and Associated Disorders. IEEE Signal Process Mag. 2019;36(1): 138-146.
23. Voicu RA, Dobre C, Bajenaru L, Ciobanu RI. Human Physical Activity Recognition Using Smartphone Sensors. Sensors. 2019;19(3):458-476.
24. Jayatilake D, Ueno T, Teramoto Y, et al. Smartphone-Based Real-time Assessment of Swallowing Ability From the Swallowing Sound. IEEE JTEHM. 2015;3:1-10.
25. Rajesh R, Baranilingesan I. Tilt Angle Detector Using 3-Axis Accelerometer. IJSRST. 2018;4(2): 784-791.
26. Gjoreski H, Gams M. Accelerometer Data Preparation for Activity Recognition. International Multiconference Information Society. 2011;1014.
27. Zhang R, Oliver A. Retrieval and Timing Performance of Chewing-Based Eating Event Detection in Wearable Sensors. Sensors. 2020; 20(2):557-573.
28. Santoso LF, Baqai F, Gwozdz ML, et al. Applying Machine Learning Algorithms for Automatic Detection of Swallowing from Sound. Int Conf of the IEEE EMBC.2019;2584-2588.
29. Papapanagiotou V, Diou C, Delopoulos A. Chewing detection from an in-ear microphone using convolutional neural networks. Int Conf of the IEEE EMBC.2017;1258-1261.
30. Weber JL, Reid PM, Greaves KA, et al. Validity of self reported energy intake in lean and obese young women, using two nutrient databases, compared with total energy expenditure assessed by doubly labeled water. Eur J Clin Nutr. 2001;55:940–950.
31. Papapanagiotou V, Diou C, Lingchuan Z, et al. Fractal Nature of Chewing Sounds. ICIAP Workshops. 2015; 401-408.
32. Borvornparadorn M, Sapampai V, Champakerdsap C, Kurupakorn W, Sapwarobol S. Increased chewing reduces energy intake, but not postprandial glucose and insulin, in healthy weight and overweight young adults. Nutrition & dietetics. 2019; 76:89–94.
33. Selamat NA, Ali SHM. Automatic Food Intake Monitoring Based on Chewing Activity: A Survey. IEEE Access. 2020; 8:48846-48869.
34. Shengjie B, Caine K, Halter R, et al. Auracle: Detecting Eating Episodes with an Ear-mounted Sensor. Proceedings of the ACM on IMWUT. 2018;2:1-27.
35. MET values for 800+ Activities. ProCon.org. Access: https://golf.procon.org/met-values-for-800-activities/.
36. USDA Nutrient Database for Standard Reference (SR). FoodData Central (FDC). Access: https://fdc.nal.usda.gov/.

Correspondence:
Received: 13 July 2023
Accepted: 13 March 2024
V P Jayachitra
Department of Computer Technology, MIT Campus,
Anna University, Chennai, India
E-mail: jayachitravp@annauniv.edu