

## A non-invasive and non-wearable food intake monitoring system based on depth sensor

Muhammad Fuad Kassim, Mohd Norzali Haji Mohd, Mohd Razali Md Tomari,  
Nor Surayahani Suriani, Wan Nurshazwani Wan Zakaria, Suhaila Sari  
Department of Electronic Engineering, Faculty of Electrical and Electronic Engineering,  
Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia

### Article Info

#### Article history:

Received Dec 17, 2020  
Revised Feb 28, 2020  
Accepted Apr 7, 2020

#### Keywords:

Depth sensor  
Food intake  
Gesture recognition  
Kinect V2  
Non-wearable sensor

### ABSTRACT

The food intake counting method showed a good significance that can lead to a successful weight loss by simply monitoring the food intake taken during eating. The device used in this project was Kinect Xbox One which used a depth camera to detect the motion of a person's gesture and posture during food intake. Previous studies have shown that most of the methods used to count food intake device is worn device type. The recent trend is now going towards non-wearable devices due to the difficulty when wearing devices and it has high false alarm ratio. The proposed system gets data from the Kinect camera and monitors the gesture of the user while eating. Then, the gesture data is collected to be recognized and it will start counting the food intake taken by the user. The system recognizes the patterns of the food intake from the user by following the algorithm to analyze the gesture of the basic eating type and the system get an average accuracy of 96.2%. This system can help people who are trying to follow a proper way to avoid being overweight or having eating disorders by monitoring their meal intake and controlling their eating rate.

*This is an open access article under the [CC BY-SA](#) license.*



### Corresponding Author:

Mohd Norzali Haji Mohd,  
Departement of Electronic Engineering,  
Faculty of Electrical and Electronic Engineering,  
Universiti Tun Hussein Onn Malaysia (UTHM),  
86400 Parit Raja, Johor, Malaysia.  
Email: norzali@uthm.edu.my

## 1. INTRODUCTION

The study of the food intake counting is carried out due to the escalating interest in this unique material and its immense potential in reducing weight. Obesity is the 5th major cause of death worldwide and about 2.8 million people died each year from a disease related to obesity [1]. Mostly all the diet plan approaches rely on calories from food label and this is not effective because not all food at the grocery has calories label on packaged food. There is an ongoing debate of obesity recognition as a condition or a disease, motivated at least in part by the desire of researchers to increase options for its treatment and reduce the stigma and discrimination experienced by the obese. It has been shown that monitoring and counting food intake count reduce overweight people drastically [2]. Technology is all about helping people, which created a new opportunity to take serious action in managing their health care. Moreover, most of the dietary approach is not tracking and detecting the right calorie intake for weight loss, but currently used tools such as food diaries require users to manually record and track the food calories, making them difficult to be utilized for daily use.

This paper presents a food intake non-wearable system of counting motion consisting of a simple algorithm that is capable of detecting in real time information with regards to food intake during a meal. The algorithm focuses on detecting the user hand gesture motion movement using Kinect sensor to determine their food intake count. With the algorithm created and GUI design, it can tell that the target intake has been reached and it is time to stop eating, thereby helping people to create long-term healthy eating patterns and can prevent obesity. Despite the many efforts to encourage healthier diets, obesity continues to be a serious public health concern in Malaysia and across the world. Weight control can be assisted by self-monitoring of intake consumption, which has been consistently related to successful weight loss. Self-report tools for measuring energy intake in free-living include diet records, 24-hour recalls, food frequency questionnaires, and food photography methods. These methods require time-consuming data entry, recording food types and portion sizes, and linking data with extensive dietary databases.

Gesture recognition is the process to describe movement by which gestures are made by the user such as upper limb joints, facial face and lower limb joints which give information and control electronic device. Previously there was a device that we made just to track and detect a gesture from users to give information that can be used in 2-D and 3-D cameras [3]. Gesture Recognition has solved many problems such as sign language [4], activity action recognition [5] and human-computer interaction [6]. Most of the literatures related to gesture focused on the field where emotions are from face and also hand joint human gesture. The bite counter is a device to measure how much people eat, was created by Eric Muth, a psychologist, and Adam Hoover, an electrical engineer, both at Clemson University. Muth and Hoover launched their own company, Bite Technologies, and licensed the technology from Clemson University Research Foundation (CURF) in 2010 [7].

In this project, a new method was used to measure food intake. The food counting is measured using wrist joint motion during eating using depth-sensing cameras which is Kinect Xbox One [8-12]. By detecting a characteristic pattern, it can identify when a food intake has been taken. The GUI can monitor food intake in real-time and provide feedback to the user. The feedback gives information to tell the user to stop eating after a target intake had been reached a specific threshold which being set at the start-up on the GUI and can help the user track long-term eating patterns [13-19]. Target intake is being calculated using formula of kilocalories per food intake (KPFI). Generally, hand is aimed downwards to pick something up and sideways to place it into the mouth. This pattern holds regardless of the type of food or utensil. The first objective of this study was developing a system that monitors the motion of food intake. This objective was supported by showing that GUI system based on food intake was tracking accurate gesture of food intake. The Kinect sensor was placed in front of the user eating area and the system tracked user gesture thus they do not have to wear sensors on body to track the gesture of food intake [20]. The training of food intake is tested with several threshold values before we get a suitable value that can track the gesture.

The second objective was that food intake counting algorithm based on gesture of food intake derived from the algorithm of the food intake motion based on a threshold of joint and orientation has been successfully developed to classify food intake with encouraging results up to 96.2% of accuracy. As the nature of this approach is solely focusing on detection of gesture from the user when taking food, hence it required a continuous monitoring system. Lastly, the third objective was that developing a graphical user interface (GUI) that displays the user food count daily limit. The objectives have been achieved with the GUI of the system showed the "limit" on the screen when the user food intake has reached a specific threshold which have been set depending on the user height, weight and age.

## 2. RESEARCH METHOD

This section describes the plans and methods to achieve the objectives of this paper. The paper focuses on the methodology of this system to track and monitor the food intake of a person by detecting and setting threshold of their gesture combination [21]. Apart from that, the environment setup of food intake is also described where the approximate accuracy of the system with distance will be analysed. Figure 1 shows the flow of our proposed project.

At the beginning, a combination of data threshold is being detected. However, not all the data are suitable for eating classification since there is overlapping joint and it is not detected accurately. To overcome these problems, we set a specific threshold to track and detect the posture when eating gesture is detected which includes the methods to address the issues in overlapping joint tracking [22-23]. To perform well in gesture recognition compared to the previous method, we overcome this problem by setting a threshold.

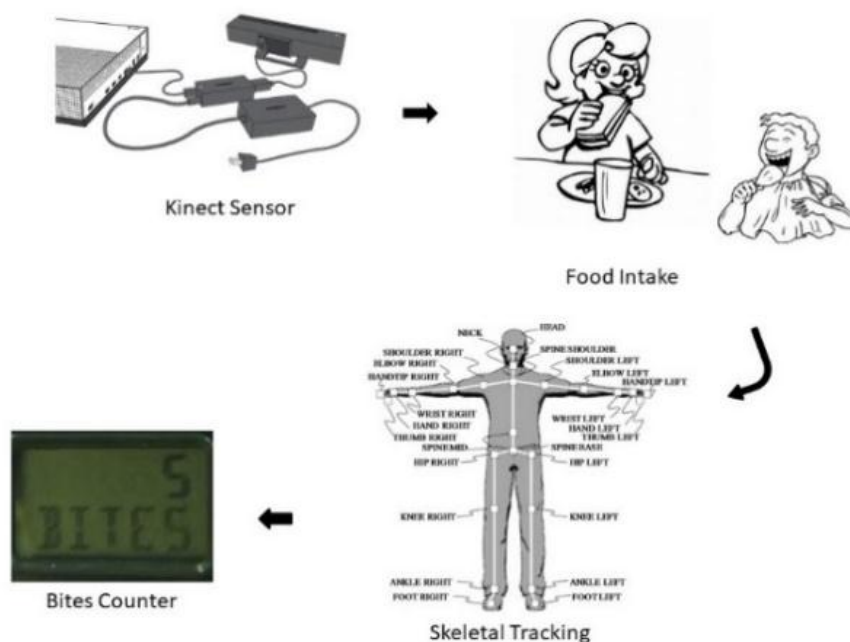


Figure 1. Image flow of proposed project

For this project, we have developed an algorithm for food intake counting monitoring which implemented as follows:

```

Food count = 0
Loop:
  Let F be the measured distance at joint point hand wrist.
  Let K be the measured distance at joint point of mouth.
  When F = K
    K=10cm
  if elbow joint pitch > X and hand wrist angle > Y
    then,
    Food count = 1
  
```

The variable food count means the first event of the cycle of roll motion food intake. The thresholds K means the roll distance that must be exceeded to trigger the detection of the events when the motion gesture of food intake happens. The threshold K is set to 10cm because the distance of hand when at mouth gives the value of approximately around 10cm and less from the raw value Kinect v2 that we captured.

Next, which data that is useful needs to be decided. For each meal, the sensors can record the movement of the wrist in three orientations: pitch, yaw, and roll. From these thresholds, we can come out with a hypothesis stating that; although we cannot detect all eating activity by only depending on the orientation, we can use the roll and pitch data while discarding the yaw data to detect food intake gesture during a meal intake since the value can be used for gesture training [24-25]. As a result, we will only use the pitch orientations and angle of hand wrist to develop the food intake detection gesture in our algorithm as mentioned above. The value of X and Y given in the algorithm will be discussed further in result and analysis of joint orientation.

### 3. RESULTS AND DISCUSSION

The basic performance measures evaluated for the proposed system are the accuracy, sensitivity, and specificity. Accuracy is the ratio of the number of correct predictions (in both the class) to the total number of observations in two classes. Sensitivity also called recall or true positive rate (TPR) is a ratio of correctly predicted positive observations to the total observations in the actual class (fall events). Meanwhile, specificity is a measure of correctly predicted negative class values over the total negative class observations. Accuracy and sensitivity are calculated. For example, if we want to calculate the sensitivity eating detection gesture, first, we get the number of true tracked in the system which is 48 and it is then divided by the true and false detection value thus we get the sensitivity of 96.2% as shown in Table 1.

Table 1. Food intake evaluation detection system

N=50	Eating	Sit/Rest	Drink	Wrist Roll	HAE	HAH	HAC	HAN	HAP	LHM
Eating	48	0	2	0	0	0	0	0	0	0
Sit/Rest	0	50	0	0	0	0	0	0	0	0
Drink	10	0	40	0	0	0	0	0	0	0
Wrist Roll	0	0	0	50	0	0	0	0	0	0
HAE	0	0	0	0	50	0	0	0	0	0
HAH	0	0	0	0	0	50	0	0	0	0
HAC	0	0	0	0	0	0	50	0	0	0
HAN	7	0	0	0	0	0	0	43	0	0
HAP	0	0	0	0	0	0	0	0	50	0
LHM	0	0	0	0	0	0	0	0	0	50
Sensitivity	96%	100%	80%	100%	100%	100%	86%	100%	100%	100%

### 3.1. Graphical user interface (GUI) output system

User interface comprise of everything the user can use to interface with the GUI system of computer. Basically, the human need to interact by input the data and the computer system can give output. Figure 2 below shows the GUI slider input of weight, height and age. The user first must enter the data of their information so that the GUI can run the program. When the data is collected the system calculated the user BMI value and displays it. The user needs to select the gender since our system required height, age, weight and gender. Figure 3 shows the GUI output of the system when the hand is at the table. Our system consists of counting algorithm thus it needs to make a gesture of eating activity for the system to start counting in the GUI. When the system not recognize the gestures from the user the GUI will do nothing which means there is no counting and the system will monitor the user continuously until there is a food intake gesture from the user. The GUI output of the system when the user exceeds their normal food intake count and their calories daily intake. The brown tab will be displayed “limit” when the threshold of the user data is reached. Different person will have different threshold depending on the user information.

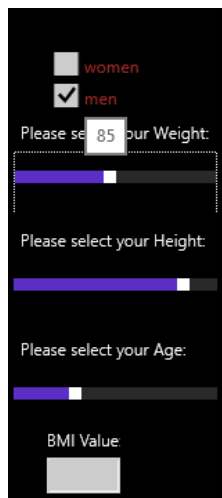


Figure 2. GUI slider input setting

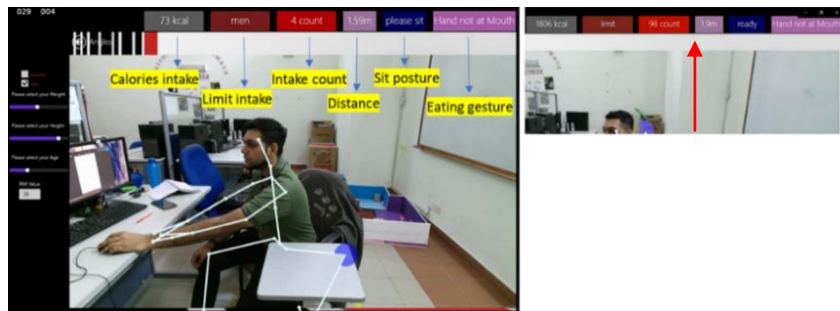


Figure 3. System GUI of food intake monitoring and display limit intake of over food intake

### 3.2. Analysis of eating and non-eating activity

Monitoring non-eating and eating is essential in monitoring the food intake detection so that the accuracy and the parameter used can be classified while setting the threshold of the subject. The activity of non-eating is shown below. The code we created is straightforward. Obviously, we cannot expect the user to do everything right since there is a non-eating activity that will perform during the gesture detection. One might eat but not perform the entire movement correctly and another might just perform the gesture, but it is too quick for the joint to detect. When developing this gesture using Kinect platform, we must be aware of all the issues occurred and creates a condition loop so that this problem can be solved and minimized.

From both Figures 4 and 5, it has been shown that different type of gesture of non-eating activity has different threshold value that we set for eating gesture. By monitoring these non-eating values, we can differentiate when the gesture of eating and non-eating happened. The evaluation is being categorized as

eating, sit/rest, drink, wrist roll, hand at ear (HAE), hand at head (HAH), hand at chin (HAC), hand at nose (HAN), hand at phone (HAP) and left-hand at mouth (LHM).

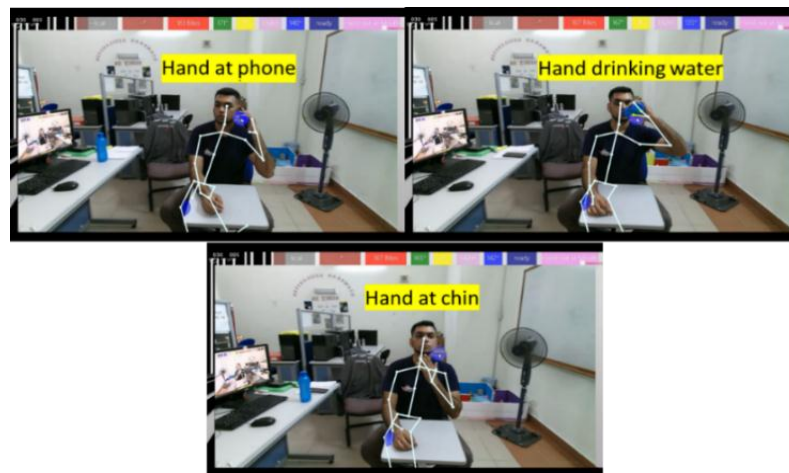


Figure 4. Gesture of non-eating activity of HAP, HAC and drink

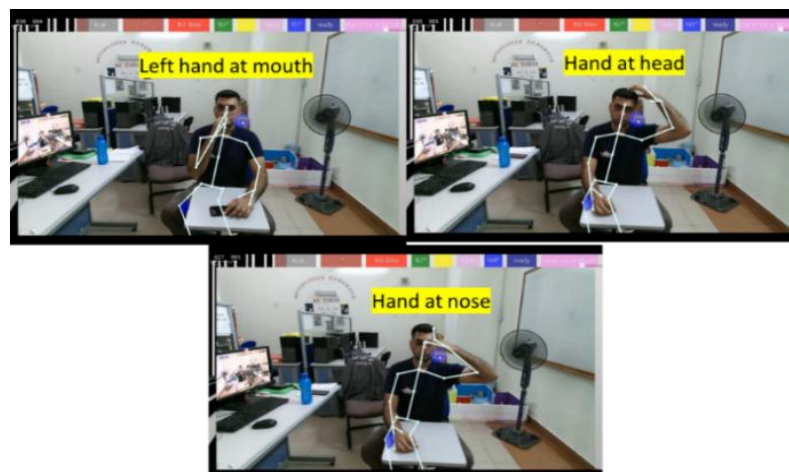


Figure 5. Gesture of non-eating activity of HAN, HAH and LHM

### 3.3. Joint orientation of roll pitch yaw

The Kinect Xbox has the ability to read joint orientation value as a quaternion. Quaternion is also known as Versors. In 3-dimensional space, Euler's theorem stated that any cycle of rotations of a solid body about a fixed point is equivalent to a single quaternion by a given angle that runs through a fixed point. Euler angles are the easiest way to describe an orientation which involves the rotation of X, Y and Z axes. Each joint around each axis which is X, Y and Z can be measured using Kinect SDK. The quaternion can then be changed into 3 sets of numerical values.

To create a rotation in 3-dimension, the axis and position of rotation needs to be determined. The X, Y, and Z of user's point are reported based on a coordinate system where the origin of the sensor and users are specified with Kinect's skeleton coordinate frame. Translations are in meters. The user's joint point is captured by three angles which are pitch, roll, and yaw. This can be used for tracking and monitoring the rotation of each joint around the X, Y and Z axis. Euler angle is used to represent the rotation in 3-dimension space.

The following Figure 6 and 7 show the setup of data used for food eating recognition of joint tracking position in terms of the action performed while hand is at the table and at the mouth. The following figures graphically illustrate the roll, pitch, and angle to describe and analyze with the help of the quaternion to get the approximate value to do the coding. Next, it can help the system track the user's eating gesture

more accurately by combining this data. During food intake gesture, the pitch will change as the arm is raised to the mouth and back to the table. Thus, we hypothesized that a simple gesture using the information threshold on pitch could be used to trigger food intake detection as well as the value of angle joint elbow [9].

The value below is captured by using Kinect by recording the user joint orientation of elbow and their angle of elbow using the formula of angle rotation phase the data below is gathered by combining both men and women to have a variety of different sex. From both graph illustrations, we can conclude that there is a different value of gesture. The GUI was programmed to capture the data of the selected axes so we could monitor the potential value that might be useful for food intake detection gesture recognition. The value for X and Y for the algorithm in above was selected and set in the code by using below data value of food intake monitoring

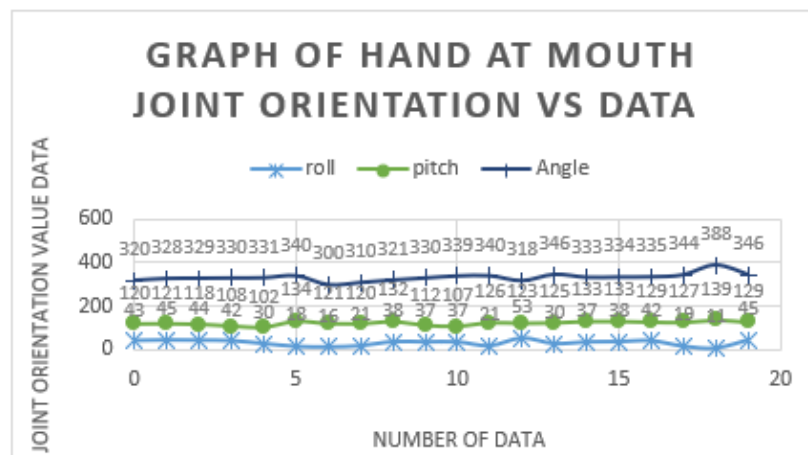


Figure 6. Illustration data rotation of elbow when hand is at mouth

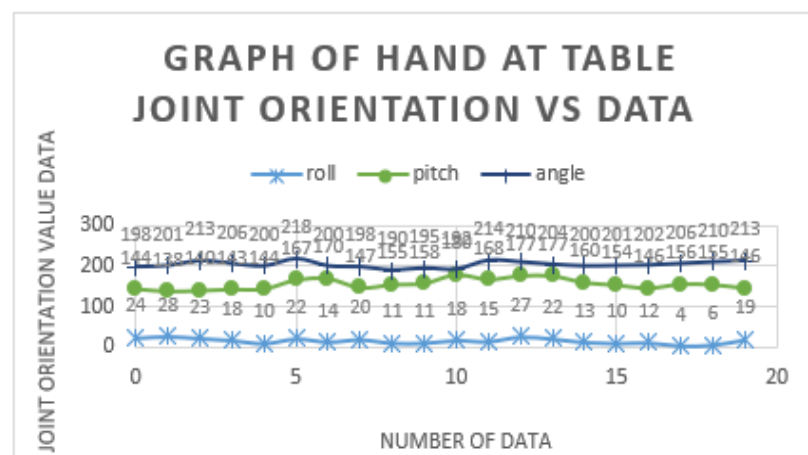


Figure 7. Illustration data rotation of elbow when hand is at table

From both graph illustrations, we can conclude that there is a different value of gesture. The GUI was programmed to capture the data of the selected axes so we could monitor the potential value that might be useful for food intake detection gesture recognition. The value for X and Y for the algorithm in above was selected and set in the code by using above data value of food intake monitoring.

The system mainly purposed to track food intake count thus we need to know how many calories per day that we need to maintain a good weight as suggested in the previous paper [10] if you want to lose one kilogram of weight each week you must decrease your daily calorie needs by 500 calories each day. If you want to lose two kilogram per week, then the decrease of our calorie needs to be of 1000 calories per day. Most women suggested a number around 1,200 calories per day to lose weight while men need to reach calories intake with a number close to 1,800 calories per day to lose weight. It doesn't matter how you count



calories, whether you use high tech apps, or a simple pen or paper method that you need to be as consistent as possible. The equation below shows how we track the calories intake using food intake detection [10].

Male:

$$\text{KPFIm} = (0.24455 \times \text{Height } x) + (0.0415 \times \text{Weight}) - (0.2597 \times \text{Age}) \quad (1)$$

Female:

$$\text{KPFIf} = (0.1342 \times \text{Height}) + (0.0290 \times \text{Weight}) - (0.0534 \times \text{Age}) \quad (2)$$

The formula above were calculated to determine the limit of kilocalories with food intake KPF (kilocalories per food intake) where height is set to cm and weight is set to kg and age is set to a maximum of 90 years old. Eating disorders are serious illness that affects a person's lifestyle. The formula used above is included in GUI coding for limitation when the user's food intake has reached its limit per day.

#### 4. CONCLUSION

The conclusion of this project was derived from the result based on the gesture of joint distance threshold derivation and joint orientation. From the research that has been carried out, it can be concluded that the objective is fully achieved. The significance of this project and the recommendations will be clarified in this chapter. Although the system we developed has limitations, our study has successfully revealed the ways in which the identified shortcomings can be addressed in the future.

This research involves detecting the food intake count of eating motion to control the over intake of food which led towards obesity. A food intake-based measure of kilocalorie intake shows that for individual use, self-monitoring is ought to be used for monitoring free-living kilocalorie intake. It is easily collected and based on the motion; it is believed to have the ability to be refined to a more accurate estimation of kilocalorie intake. However, the performance of the food intake event detection task still needs optimization and it could be used to analyze the food intake of people under laboratory settings. The recent development of new sensors that allow tracking important parts of the human body has resulted in a proliferation of different approaches to gesture recognition and their practical applications. Therefore, any new improvement of previous solutions towards better recognition accuracy or shortening recognition time is better for this project to be developed.

#### ACKNOWLEDGEMENTS

This research is supported by a Fundamental Research Grant Scheme (FRGS) FRGS/1/2019/ICT04/UTHM/02/2 sponsored by the Ministry of Education (MOE). The research also received funding from the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC), UTHM.

#### REFERENCES

- [1] M. Wilson, "Assessing the bite counter as a weight loss tool," Thesis, Clemson University, 2014.
- [2] K. Kearns, A. Dee, A. P. Fitzgerald, E. Doherty, and I. J. Perry, "Chronic disease burden associated with overweight and obesity in Ireland: the effects of a small BMI reduction at population level," *BMC Public Health*, vol. 14, no. 143, pp. 1-10, 2014.
- [3] S. Berman and H. Stern, "Sensors for gesture recognition systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 3, pp. 277-290, May 2012.
- [4] S. Goldin-Meadow and D. Brentari, "Gesture, sign and language: The coming of age of sign language and gesture studies," *Behavioral and Brain Sciences*, vol. 40, pp. 46-1-60, 2017.
- [5] M. Asadi-Aghbolaghi, A. Clapes, M. Bellantonio, H. Escalante, V. Ponce-López, X. Baró, I. Guyon, S. Kasaei, and S. Escalera, "Deep learning for action and gesture recognition in image sequences: a survey," *Gesture Recognition, Springer Verlag*, pp.539-578, 2017.
- [6] A. Haria, A. Subramanian, N. Asokkumar, S. Poddar, and J. S. Nayak, "Hand gesture recognition for human computer interaction," *Procedia Computer Science*, vol. 115, pp. 367-374, 2017.
- [7] J. L. Scisco, E. R. Muth, and A. W. Hoover, "Examining the utility of a bite-count-based measure of eating activity in free-living human beings," *Journal of the Academy of Nutrition and Dietetics*, vol. 114, no. 3, pp. 464-469, 2014.
- [8] M. F. Bin Kassim and M. N. Haji Mohd, "Tracking and counting motion for monitoring food intake based-on depth sensor and UDOO board: A comprehensive review," *IOP Conference Series: Materials Science and Engineering*, vol. 226, pp. 012089-1-9, 2017.

- [9] M. Fuad bin Kassim and M. Norzali Haji Mohd, "Food intake gesture monitoring system based-on depth sensor," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 2, pp. 470-476, 2019.
- [10] Y. Nizam, M. N. Haji Mohd, and M. M. Abdul Jamil, "Development of a user-adaptable human fall detection based on fall risk levels using depth sensor," *Sensors (Basel)*, vol. 18, no. 7, pp. 2260-1-14, 2018.
- [11] Y. Nizam, M. N. Haji Mohd, M. M. Abdul Jamil, "Human fall detection from depth images using position and velocity of subject," *Procedia Computer Science*, vol. 105, pp. 131-137, 2017.
- [12] Y. Nizam, M. Norzali M., and M. M. Abdul Jamil, "Biomechanical application: Exploitation of kinect sensor for gait analysis," *ARN Journal of Engineering and Applied Sciences*, vol. 12, no. 10, pp. 3183-3188, 2017.
- [13] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, "Comparison between human and bite-based methods of estimating caloric intake," *Journal of the American Academy of Nutrition and Dietetics*, vol. 116, no. 10, pp. 1568-1577, Oct. 2016.
- [14] Y. Nizam, M. N. Haji Mohd, and M. M. Abdul Jamil, "Classification of human fall from activities of daily life using joint measurements", *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 8, no.4, pp. 145-149, 2016.
- [15] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, "Measuring calorie and nutrition from food image," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 8, pp. 1947-1956, Aug. 2014.
- [16] H. Kalantarian and M. Sarrafzadeh, "Audio-based detection and evaluation of eating behavior using the smartwatch platform," *Computers in Biology and Medicine*, vol. 65, pp. 1-9, 2015.
- [17] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner, "A wearable system for detecting eating activities with proximity sensors in the outer ear," *Proceedings of the 2015 ACM International Symposium on Wearable Computers - ISWC '15*, pp. 91-92, 2015.
- [18] H. L. McClunga, C. M. Champagneb, H. R. Allenb, S. M. McGrawc, A. J. Youngc, S. J. Montainc, and A. P. Crombie, "Digital food photography technology improves efficiency and feasibility of dietary intake assessments in large populations eating ad libitum in collective dining facilities," *Appetite*, vol. 116, pp. 389-394, 2017.
- [19] F. J. Pendergast, N. D. Ridgers, A. Worsley, and S. A. McNaughton, "Evaluation of a smartphone food diary application using objectively measured energy expenditure," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 14, no. 1, pp. 1-10, 2017.
- [20] P. J. Stumbo, "New technology in dietary assessment: a review of digital methods in improving food record accuracy," *Proceedings of the Nutrition Society*, vol. 72, no. 1, pp. 70-76, 2013.
- [21] Y. He, N. Khanna, C. J. Boushey, and E. J. Delp, "Image segmentation for image-based dietary assessment: A comparative study," *International Symposium on Signals, Circuits and Systems ISSCS2013*, pp. 1-4, 2013.
- [22] G. Luzhnica, J. Simon, E. Lex, and V. Pammer, "A sliding window approach to natural hand gesture recognition using a custom data glove," In *2006 IEEE Symposium 3D User Interfaces (3DUI)*, IEEE, pp. 81-90, 2016.
- [23] Y. Nizam, M. N. H. Mohd, R. Tomari, and M. M. A. Jamil, "Development of human fall detection system using joint height, joint velocity and joint position from depth maps," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 8, no. 6, pp. 125-131, 2016.
- [24] H. Ishiyama and S. Kurabayashi, "Monochrome glove: A robust real-time hand gesture recognition method by using a fabric glove with design of structured markers," In *2006 IEEE Virtual Reality (VR)*, IEEE, pp. 187-188, 2016.
- [25] J. Dulayatrakul, P. Prasertsakul, T. Kondo, and I. Nilkhamhang, "Robust implementation of hand gesture recognition for remote human-machine interaction," In *2005 7th International Conference on: Information Technology and Electrical Engineering (ICITEE)*, IEEE, pp. 247-252, 2015.