# The Potential of a Classification-based Algorithm to Calculate Calories in Real-Time via Pattern Recognition

Marhaini M.S[1], Mohamed Ariff Ameedeen[1, 2]
*[1]Faculty of Computer System and Software Engineering*
*Universiti Malaysia Pahang, Malaysia.*
*[2]IBM Centre of Excellence, Universiti Malaysia Pahang, Malaysia.*
*pcc14005@stdmail.ump.edu.my*

*Abstract*—Calories refer to a unit of energy that people should consume based on total energy needed. Thus, a system for health monitoring applications that can measure calories and nutrition can be very useful. This research is mainly focused on creating a new algorithm based on classification technique to calculate food calorie intake in real-time. Enhancement on Extreme Learning Machine (ELM) algorithm will be done to get better results in terms of accuracy and speed of calculating the food calorie. The ELM algorithm will be applied to an ultra-mobile Near Infrared (NIR) spectrometer. While the algorithm helped to classify different types of wavelengths produced from the sensor, a classification-based algorithm via Pattern Recognition Method will be used to classify and match the food components. The results will display the total amount of calories consumed per day, per week and per month with total amount of calories left in a mobile application.

*Index Terms*—Classification; Extreme Learning Machine (ELM); Food Calorie; Pattern Recognition; Ultra-Mobile Near Infrared (NIR) Spectrometer.

## I. INTRODUCTION

A calorie is a unit of energy. It supplies energy to the human body. It has strong relationships with human health conditions. Nowadays, many people do not only lack health awareness, but there are a few whom are unable to practice vital health tips due to their busy schedules. People are, generally, getting more and more unconcerned about what kinds of food they eat every day. They keep on eating variety of food either heavy meal or some kinds of snack meal without checking the nutrition labels on the food packaging. They are also not practicing a healthy lifestyle which is important to maintain and improve human health and body fitness. Thus, this kind of unhealthy lifestyle and unhealthy food intake are the root causes of many kinds of diseases such as obesity

The increase in obesity worldwide has significant impact on health impairment and reduced quality of life [1]. In fact, [2] has found out that obesity has contribute to a few serious problems such as cardiovascular disease, diabetes mellitus, cancer, osteoarthritis, sleep apnea and work disability. Although obesity is a multifactorial disease, an expenditure and imbalance between energy intake is commonly cited as its only cause [3]. This is because most of the people cannot balance the amount of calories taken with total energy they need for everyday used. Thus, this situation may cause the number of calories taken accessing the number of energy they need which is, it will lead into energy imbalanced. The energy imbalanced will be one of the causes that contribute into obesity disease which promote lipogenesis [4]. Proper management of the amount of calorie intake, therefore, would be one of the vital steps to overcome the obesity problem.

The rest of this paper is organized thus: section two presents an overview of previous related researches; section three explains the proposed works that are going to be implemented on how to calculate the food calorie intake in real-time based on the pattern recognition method and classification technique; section four presents the discussion and conclusion.

## II. FOUNDATION

This section in this paper is an overview of previous conducted research to provide the foundation for this current research. This includes an overview about calories, how to calculate calories and how it is related to human health. Related studies on methods that had been used for food classification also being reviewed in order to classify the food component for calorie calculation.This also includes reviews on System on Chip that will be used as an intermediate tool to apply an algorithm into the sensor.

### A. Calories Calculator

Within the past three decades, the number of obesity cases had been continuously increased until the present year. This obesity problem will contribute into other disease such as diabetes and cardiovascular disease resulting from unhealthy food intake and unhealthy lifestyle.

In view of this problem of unhealthy dietary habits, developers have designed calories' calculator to assist in determining the exact amount of daily calorie intake. In order to calculate the total calorie intake per person, they should know how much calories a particular kind of food has and then add up the number of calories consumed in a particular point in time. Then, the total amount of calories will be deducted from the estimated amount of calories that a person should have per day.

The previous calories calculator has its own limitations. They need a few information of the food in order to calculate the food calories. The previous method simply obtains and calculates the food's component weight data and the data is placed in a simple mathematical equation for calorie calculation as shown in Figure 1.
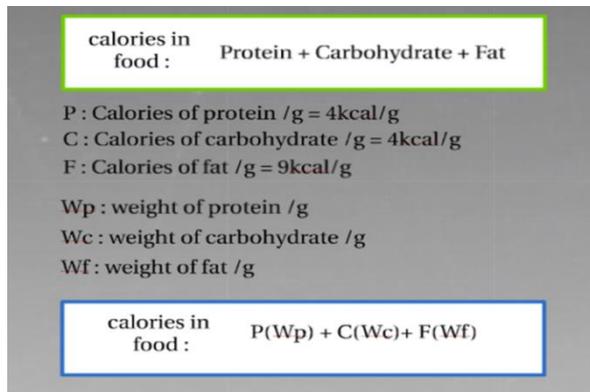


Figure 1: Formula for food calories calculation

Basal Metabolic Rate (BMR) can be indicate as a daily calorie count [5]. BMR is the amount of calories that a human burn at rest. Study show that, a human body uses about 60% of calories in order to maintain with natural processes at rest. The number of BMR increases as a human body gains more lean muscle mass. Thus, another remain energy are used for doing physical activities and digestion process. 30% of calories that has been converted into energy are used for physical activities while another 10% is used for the digestion process. Thus, by eating smaller meals often, people can burn more calories too because of the digestion processes.

BMR has a big significant to regulate the right number of calories needed daily before continuing with the process of calculating the calories need. To determine the BMR, a few information of a person need to be provide such as weight, height and age. Once the BMR number has been retrieved, the rate of physical capability need to be identify as shown in Table 1. By considering the table below, the amount of calories burned during the exercise will be found. So this is how the calorie intake formula will be used to calculate the daily's calorie intake.

Table 1
Rate to be multiply with bmr based on human activity level [6]

| BMR Calculation Rate | Level of Activeness |
|---|---|
| BMR x 1.2 | Low intensity : activities and leisure activities |
| BMR x 1.375 | Light exercise : doing leisure walking for 30-50 minutes 3-4 days/week like golfing, house chores |
| BMR x 1.55 | Moderate exercise for 3-5 days per week for 30-60 minutes/session |
| BMR x 1.725 | Moderate to high intensity exercise: Active individuals doing exercise 6-7 days/week at about 45-60 minutes/session |
| BMR x 1.9 | Heavy/intense exercise : Extremely active individuals doing exercise 6-7 days/week for 90 minutes/session and more) like heavy manual labor, heavy lifting, endurance athletes and competitive team sports athletes |

### B. Atwater System

Calculating the calories initially started from the food testing laboratories. Several researches on food calories have been done by a number of researchers till date. The Atwater conversion factors are used to calculate the amount of energy that can be acquired from eating certain types of food. It is used to isolate each of the food components in order to get their net weight. This has been approved by the USA Department of Agriculture (USDA) and the International Association of Analytical Communities (AOAC). This method provides a crude calorie count per gram of a list of numerous food types [5]. Basically food contains a few components such as protein, carbohydrate, fat and alcohol. Each of the components has the fix number of calorie per gram. In order to calculate the calories from a portion of a food sample, the net weight of each component needs to be retrieved first. Later the weight will be multiplied by the number of calorie per gram and all the numbers of the calories will be added up to get the actual number of calories from a food sample.

According to the National Data Lab (NDL), calories values of food that are kept in industry food tables (IFT) are mostly based on indirect calories estimation. The estimation numbers are obtained using Atwater System [5]. Based on this system, calorie is not determined directly by burning the food. The calories of food types are calculated by adding up the calories from each of the food components. Each of them can get the amount of calories based on the calories provided by energy-containing nutrients. But, this method still has its own limitations in its derivation.

Table 2
Average Value of Calories per Gram (g)

| Food Component | Calorie value per gram |
|---|---|
| Protein | 4kcal/g |
| Carbohydrate | 4kcal/g |
| Fat | 9kcal/g |
| Alcohol | 7kcal/g |

Table 2 shows the average value of calorie per gram for each of the basic food components. The amount of calorie of a food can be calculated by identifying the weight of the protein, carbohydrate, fat and alcohol from the energy bar on the food labeling that is usually printed on the food packaging material. Then, the weight of each of the food components is multiplied by the value of calorie per gram (g).

### C. Food Classification

Nowadays, the classification of foods is very important and has many practical applications such as brand classification [7], aging time detection [8] geographic origin discrimination and so on. Previously, the old methods of using the wet chemical for classification process are expensive and time-

consuming. They involve a relatively large amount of manual work.

As the time keeps on, some researcher has introduced Near-infrared (NIR) or Mid-infrared (MIR) spectroscopy as a substitute method since they have a few benefits as a good analytical method such as non-destructive, simple, and fast. It needs less work to do and can be executed easily [9],[10].There are a few algorithm that had been developed for the spectroscopy-based classification such as k-nearest neighbor (KNN) [7], partial least-squares discriminant analysis (PLS-DA) [11],[12], artificial neural networks (ANN) [13] and support vector machine (SVM) [14][15][16].

KNN is a simple unsupervised classification algorithm that is also used to assign the category of a new sample. It can be implemented easily. However it also has a few problems such as its classification accuracy depends closely on the dataset and has slow running speed.

PLS-DA is known as a variant of partial least-squares regression (PLS) for the optimal separation of classes. Based on [17][18] they have successfully applied a number of PLS-based classification applications. Nevertheless, PLS has sometimes a difficulty in yielding satisfactory performance because of nonlinearity and over-fitting.

ANN has been employed to figure out the classification problems. It is a non-parametric and non-linear regression technique. The most commonly used ANN is the Back propagation (BP) neural network. Broadly, gradient descent algorithms are used to determine the parameters of the BP neural networks. Based on [19], it was summaries that this technique also has a few problems that need to be faced. They have many overlapping issues such as stopping criteria, local minima and learning rate. They are also relatively slow.

Based on previous studies on the classification algorithm, SVM has shown good generalization performance. It also performs well on different data sets. However, SVM still needs to face a problem related to the parameters which are hard to be tuned. The least-squares support vector machines (LS-SVM) is a version of reformulation of SVM. A previous study has discovered that LS-SVM can prevent a complex calculation. This can be done using a set of standard linear equations. They also conclude that the prediction performance of LS-SVM is better than that of PLS or ANN. Even so, LS-SVM does not really suitable for real-time scenario since it is originally created for binary classification which need to be extend for multi-class classification that needs to consider the running time.

It was stated from [19] that Extreme Learning Machine (ELM) could be an algorithm that could learn fast with high generalization performance. Furthermore, because of its network structure, it could implement the multi-class classification quickly. However, [19] claimed that ELM still not being used food classification based on near or mid-infrared spectroscopy.

### D. System on a Chip (SoC)

A system-on-a-chip (SoC) is a microchip that is being used as a parts of a few existing system such as smartphones and computers. It has all the necessary electronic circuits that is being designed on a single integrated circuit (IC). A SoC might include might include a microprocessor, an audio receiver, a memory, an analog-to-digital converter (ADC), and the input/output logic control of a user for a sound-detecting device. It's all was designed on a single small chip. It is used in tiny chip with increasingly complex consumer electronic devices. It can be used as the intermediate device to integrate a sensor with a software.

In order to ensure the ability to map different algorithms, a wise combination of programmable and hard-wired processing elements is needed. This was done in the system architecture level while using lower power than fully programmable architectures. Therefore in [20] previous work, SoC is being used to demonstrated the first sub-microrowatt electroencephalograph (EEG) seizure detection and integrate an ARM Cortex-M3 processor. While based on [21] study, they have identified that configuration functionality with low-power consumption for portable ECG monitoring applications can be reached by using a mixed-signal ECG.

In the future, SoC that already equip with nanorobots might act as programmable antibodies to fend off previously incurable diseases. SoC video devices might be embedded in the brains of blind people, allowing them to see and SoC audio devices might allow deaf people to hear. Handheld computers with small whip antennas might someday be capable of browsing the Internet at megabit-per-second speeds from any point on the surface of the earth. SoC is evolving along with other technologies such as silicon-on-insulator (SOI), which can provide increased clock speeds while reducing the power consumed by a microchip.

### III. THE PROPOSED CALCULATION AND IMPLEMENTATION

In this study, the main target is to create a new algorithm that can calculate the calorie intake in a real-time. This research will be focused on how to calculate the food calories without key in any data to the system just by directly detecting how much calories exist in a small portion of food taken by a person. Thus, this system should be able to calculate calories per intake.

Moreover, this process is planned to be done based on Pattern Recognition method using classification techniques to classify and match the food component based calories such as protein, carbohydrate and fat. Thus, the system will be evaluated based on a pattern recognition standard such as accuracy, precision, recall, confusion matrix. Enhancement on Extreme Learning Machine (ELM) algorithm as a classifier will be done to get a better result in term of accuracy and speed of calculating the food calories. Extreme Learning Machine (ELM) will be combined with Support Vector Machine (SVM) classification algorithm to become a hybrid algorithm.

Other than that, this research also looking forward on applying the enhanced ELM algorithm on an ultra-mobile Near Infrared (NIR) spectrometer embedded with Digital Light Processing (DLP) chip. This selected sensor will be used for detecting the components of the food inside. A newly derived hybrid algorithm will be injected into the sensor using System on Chip (SoC) that will be functioning as the intermediate tool between sensor and the software. The hybrid algorithm that will be embedded into the software in the selected sensor will help to classify the food component. In

other word, the system will work as an automatic food calorie measurement system via mobile phones that can help patients and dietitians to measure and manage their daily food intake. The results are expected to display the total amount of calories consumed per day, per week and per month with total amount of calories left in a mobile application. So this hybrid algorithm will be invented as a new enhanced classification algorithm for pattern recognition in the food analysis process.

In the first place, calculating the food calories needs a little information about the object. In this case, it will be the sample of the food. The mathematical formula that is going to be used for calculating the food calories consist of the weight of each of the food components such as protein (Wp), fat (Wf) and carbohydrate (Wc):

$$\text{Calories in the food} = \text{calories of (protein + carbohydrate + fat)} \quad (1)$$
$$\text{CalFood} = \text{Cal\_P(Wp)} + \text{Cal\_C(Wc)} + \text{Cal\_F(Wf)} . \quad (2)$$

With reference to Equation (2), each of the food component weight is multiplied by the amount of the estimated calories per gram that are already listed in the National Data Lab (NDL). The values of calories per 1gram that are kept in the industry food table are mostly based on indirect calorie estimation. Cal P will be the calories of protein for 1gram, Cal C will be the calories of the carbohydrate for 1gram and Cal F will be the calories of fat for 1gram. The amount of the food calories will be retrieved by adding up all the amount of calories for protein, carbohydrate and fat.

In order to get the information about the weight of the food component, the component of the main elements in the food should be captured first. Protein, carbohydrate and fat should be recognized and identified before measuring their quantity. Then their weight will be measured using the weight scale. Thus, this is the main focus of the research study: to identify the way on how to capture the chemical elements in the food. By identifying those chemical elements, this will help to differentiate the type of the food components that exist in the tested food's sample.

After all, Pattern Recognition method has been chosen as a method to recognize the chemical elements of the food by using the classification technique. This classification technique helps to classify the existing component in the food to make it easier to differentiate between protein, carbohydrate and fat according to their classes. So, the benchmark data or dataset that is used in this study include any kind of food that consists of protein, carbohydrate and fat.

A sensor functions as an instrumental tool. An ultra-mobile Near Infrared (NIR) spectrometer embedded with Digital Light Processing (DLP) chip is used as analytical instrument to detect and identify the target object. It is also used to retrieve the signal and send it to the data receiver. NIR spectrometer has been chosen as the instrumental tool since NIR or mid-infrared (MIR) spectroscopy has become a good analytical technique. This is due to the fast, simple and nondestructive features they have. This selected sensor is embedded together with a software with a hybrid classification algorithm in order to process the complex information that has been retrieved by the sensors. The signal obtained from the sensor is then processed using the specified software to get the final result. Practically, this sensor detects the food components and sends the signal or information to the data receiver.

Data receiver could be a smartphone that displays all the information received. Signals from the sensor are interpreted into meaningful data by the software and sent into the smartphone via a Bluetooth connection. In this research, the newly derived algorithm also helps to improve the connectivity of the Bluetooth connection so that it will produce a seamless pairing connectivity between two devices.

Information retrieved from the sensor is calculated and programmed directly on the smartphone. So, no server is needed to save the data. This procedure prevents any delay during network congestion in the process of transferring data. This process also increases the speed rate of receiving the information.

This study also creates a temporary storage for data archived that will be integrated with the sensor. This temporary storage is created in order to avoid missing data if the sensor is disconnected from the data receiver or the smartphone. The smartphone might probably be out of battery during the process of transferring data. Thus, by having this kind of temporary storage, it secures all the retrieved data

A mobile application is developed for displaying the information of the food taken such as the amount of calories and weight of the food. The application will record the calorie intake of a person every time they are using the sensor that already integrated with the hybrid classification algorithm. The calorie consumption per day will be calculated at the application so that the user can monitor and keep track of the amount of calorie intake. So, people can manage their amount of calorie intake that are suitable with the total energy they use daily.

## IV. DISCUSSION

Our bodies need energy to perform. This energy is gained by eating food. Human bodies will digested and process the food and it will convert into energy. The energy then, will be burnt and used during physical activities and also to perform the basic bodily operations. If one takes more energy than is needed, the extra energy will be converted into fat that will be stored in the body. Thus, it is extremely important to control the amount of calories consumed in order to stay healthy. In the meantime, a system that can measure and manage calorie intake in everyday meals can be very useful. As people across the world are becoming more concerned in observing their weight, avoiding obesity and eating more healthily.

Recently, due to the ubiquity of mobile devices such as smart gadgets, the health monitoring applications are approachable by the people practically all the time. Thus, this paper is literally describing the ideas on the proposed work. The work with the implementation of the classification algorithm based on pattern recognition on food calories will be carried out later with further details in the experimental process. Based on the proposed work, the idea will be mainly focused on creating a new algorithm based on classification technique to calculate food calorie intake in real-time. Enhancement on Extreme Learning Machine (ELM) algorithm will be done to get a better result in term of accuracy and speed of calculating the food calories based on Pattern

Recognition Method. In the meanwhile, the system will be also evaluated based on pattern recognition standard. Based on [19], it was presented that ELM has gained a better accuracy in most cases than the others existing competitors. Furthermore, previous research has found out that ELM has execute much faster than BP-AN, KNN and LS-SVM in the classification stage. Thus, these advantages of ELM can be a promising method for a real-time food classification with a comparable accuracy based on near or mid-infrared spectroscopy.

Furthermore, this will also look forward on applying the enhanced ELM algorithm on an ultra-mobile NIR spectrometer embedded with DLP chip. The sensor will be used for detecting the components of the food inside and the algorithm will help to classify the food component. In other word, the system will work as an automatic food calorie and nutrition measurement system via mobile that can help patients and dietitians to measure and manage daily food intake The results is expected to display the total amount of calories consume per day, per week and per month with total amount of calories left in a mobile application.

Calculation of the calories will be done using a sensor as the instrumental tools that would help the process to measure the weight of each of the food components by placing the sensor on a surface that can touch the food each time people eat. This idea has been used to ensure that the calories of food can be calculated based on how much food had been taken. The algorithm that embedded in the software will make the process of calculating the food calories. Thus, the main focus would be the efficiency of how a classification algorithm can calculate the calories in real-time. This experiment on applying the algorithm with the sensor for detecting the food calorie intake will be conducted to show the validity of data and captured the main objective of this research. Calculating food calorie intake in real-time would be the main objective in this study by classifying the main component of food using proposed classification algorithm. The proposed algorithm also will be used to evaluate the efficiency or the feasibility of the current research.

However, there are a few potential difficulties that might be faced later. The classification algorithm should be able to discriminate the sample that might be very tricky to be differentiated. It is also should be able to extract the testing sample which would be relatively small in size. With some accuracy problems will also be faced later.

So, there are a few potential ways that can be done to solve the problems. For example, more training sample should be done for better data accuracy. Combining two classifiers as one new hybrid classifier also can be one of the solution so that it can solve the problems of any sample that cannot be separate by the single linear classifier. However, these ideas will might not be the best solution or can be done until the experiment is completely finished.

### REFERENCES

[1] Taylor, V.H., Forhan, M., Vigod, S.N., McIntyre, R.S., Morrison, K.M.: The impact of obesity on quality of life. Best Pract. Res. Clin. Endocrinol. Metab. 27, 139–146 (2013).

[2] Seidell, J.C., Halberstadt, J.: The global burden of obesity and the challenges of prevention. Ann. Nutr. Metab. 66 Suppl 2, 7–12 (2015).

[3] Hall, K.D., Heymsfield, S.B., Kemnitz, J.W., Klein, S., Schoeller, D.A., Speakman, J.R.: Energy balance and its components: implications for body weight regulation. Am. J. Clin. Nutr. 95, 989–994 (2012).

[4] Simopoulos, A.P., Bourne, P.G., Faergeman, O.: Bellagio report on healthy agriculture, healthy nutrition, healthy people. Nutrients. 5, 411–423 (2013).

[5] Coating, P., Complete, T., Handbook, F., Data, P.P., Parker, P.E.J.: System and method for weighing food and calculating calorie content. 2, (2010).

[6] Black, a E.: Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. Int. J. Obes. 24, 1119–1130 (2000).

[7] Tan, C., Qin, X., Li, M.: Comparison of chemometric methods for brand classification of cigarettes by near-infrared spectroscopy. Vib. Spectrosc. 51, 276–282 (2009).

[8] Shen, F., Ying, Y., Li, B., Zheng, Y., Zhuge, Q.: Multivariate classification of rice wines according to ageing time and brand based on amino acid profiles. Food Chem. 129, 565–569 (2011).

[9] Chen, J., Arnold, M. a., Small, G.W.: Comparison of combination and first overtone spectral regions for near-infrared calibration models for glucose and other biomolecules in aqueous solutions. Anal. Chem. 76, 5405–5413 (2004).

[10] Cen, H., He, Y.: Theory and application of near infrared reflectance spectroscopy in determination of food quality. Trends Food Sci. Technol. 18, 72–83 (2007).

[11] Bevilacqua, M., Bucci, R., Magrì, A.D., Magrì, A.L., Marini, F.: Tracing the origin of extra virgin olive oils by infrared spectroscopy and chemometrics: A case study. Anal. Chim. Acta. 717, 39–51 (2012).

[12] Vitale, R., Bevilacqua, M., Bucci, R., Magri, A.D., Magri, A.L., Marini, F.: A rapid and non-invasive method for authenticating the by NIR spectroscopy and chemometrics. Chemom. Intell. Lab. Syst. 121, 90–99 (2013).

[13] Palacios-Morillo, A., Alcázar, Á., De Pablos, F., Jurado, J.M.: Differentiation of tea varieties using UV-Vis spectra and pattern recognition techniques. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 103, 79–83 (2013).

[14] Zhao, J., Chen, Q., Huang, X., Fang, C.H.: Qualitative identification of tea categories by near infrared spectroscopy and support vector machine. J. Pharm. Biomed. Anal. 41, 1198–1204 (2006).

[15] Acevedo, F.J., Jiménez, J., Maldonado, S., Domínguez, E., Narváez, A.: Classification of wines produced in specific regions by UV-visible spectroscopy combined with support vector machines. J. Agric. Food Chem. 55, 6842–6849 (2007).

[16] Xie, L., Ying, Y., Ying, T.: Classification of tomatoes with different genotypes by visible and short-wave near-infrared spectroscopy with least-squares support vector machines and other chemometrics. J. Food Eng. 94, 34–39 (2009).

[17] Al-Jowder, O., Kemsley, E.K., Wilson, R.H.: Mid-infrared spectroscopy and authenticity problems in selected meats: A feasibility study. Food Chem. 59, 195–201 (1997).

[18] Holland, J.K., Kemsley, E.K., Wilson, R.H.: Use of Fourier transform infrared spectroscopy and partial least squares regression for the detection of adulteration of strawberry purées. J. Sci. Food Agric. 76, 263–269.

[19] Zheng, W., Fu, X., Ying, Y.: Spectroscopy-based food classification with extreme learning machine. Chemom. Intell. Lab. Syst. 139, 42–47 (2014).

[20] Sridhara, S.R., DiRenzo, M., Lingam, S., Lee, S.-J.J., Blazquez, R., Maxey, J., Ghanem, S., Lee, Y.-H.H., Abdallah, R., Singh, P., Goel, M., Blázquez, R., Maxey, J., Ghanem, S., Lee, Y.-H.H., Abdallah, R., Singh, P., Goel, M.: Microwatt Embedded Processor Platform for Medical System-on-Chip Applications. IEEE J. Solid-State Circuits. 46, 721–730 (2011).

[21] Kim, H., Kim, S., Van Helleputte, N., Artes, A., Konijnenburg, M., Huisken, J., Van Hoof, C., Yazicioglu, R.F.: A configurable and low-power mixed signal SoC for portable ECG monitoring applications. IEEE Trans. Biomed. Circuits Syst. 8, 257–267 (2014).