IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 787
Research Paper © 2012 IJFANS. All Rights Reserved UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

Food and nutrition recognition benchmark: using CAFE computer program to recognize nutrient values

K. Vanishree [1]

Assistant Professor,

Computer Science and Engineering,

J.N.N Institute of Engineering, Chennai, Tamil Nadu, India

A. Iyyanarappan [2], ***M. Murali*** [3], ***R. Subramani*** [4], ***M. Karthikeyan*** [5]

Assistant professor,

Computer Science and Engineering[2,4,5], Electronics & Communication[3]

J.N.N Institute of Engineering, Chennai, Tamil Nadu, India

Abstract

In nutritional epidemiology, food frequency questionnaires (FFQs) are commonly utilised; however, the creation of the auxiliary programmes and nutritional databases is not well documented in any of the articles. The structure and content of the The discussion also includes the development of cut-points for data removal and extreme nutrient levels. The CAFE computer application modifies dietary intake based on the amount of fat that is visible on meat as well as the specific fats that are utilised in food preparation. It combines many cereals for breakfast and differentiates text for brands and varieties. Due to the high reported frequency of specific items, several extreme values for carbohydrate, energy, fat, alcohol, and protein were still present when outliers in nutrient intake were excluded. A unique technique for matching text and adaptable, updatable databases are some of CAFE's capabilities. More research should be done to determine how extreme nutrient values affect the FFQ's accuracy in assessing diet in nutritional epidemiology.

Keywords: Nutrition recognition. CAFÉ computer application, foof frequency questionnaires, Food consumption.

Introduction

A list of foods or food kinds combined with a range of frequency estimate options makes up food frequency questionnaires (FFQs). They are employed to create regular or habitual food or nutrient consumption. Because semiquantitative FFQs are simple to administer and process, making them ideal for use in large-scale population studies, they are frequently employed to estimate nutritional intake, especially in epidemiology. However, creating programmes to calculate nutrients is labor-intensive and demands a high level of computational and nutritional knowledge. There are many publications on FFQ results, validity, and repeatability, but there aren't any methodological studies that explain the challenges in creating the nutritional databases and programmes that support them.

ISSN PRINT 2319 1775 Online 2320 787
Research Paper © 2012 IJFANS. All Rights Reserved UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

Numerous fully or semi-automated food analysis techniques have been developed to eliminate bias. A growing number of apps are being developed into mobile systems to facilitate food analysis, as a result of the quick adoption of mobile devices. Zhu et al. [12] presented, for instance, a segmentation assisted food classification method for dietary evaluation. Their objective is to use conventional computer vision techniques to automatically detect the regions in a picture that contain a certain food and to correctly identify those parts. Pouladzadeh et al. suggested a second mobile cloud-based food calorie measurement system. They trained their system to identify meals and determine the calories associated with each food item using cloud Support Vector Machine (SVM). Different algorithms and machine learning techniques, particularly in the field of deep learning, are becoming more and more advanced in everyday life as well as numerous industrial and professional processes, thanks to the rapid growth of hardware devices. These days, we see more self-driving cars on the road, AI gamers who can play video games, and even friendly digital robots who possess empathy. Since software is more precise, efficient, and error-free these days, we can almost entirely rely on it to complete our tasks.

Modern neural network-based food recognition and nutrition analysis technologies manifest similarly to other industries. Patients and physicians were given access to an intelligent and convenient assistive calorie measurement system through the work, which allowed them to track and quantify their daily caloric intake. These days, the majority of food analysis programmes classify foods from images using a classification algorithm. Despite having a very high accuracy rate, the categorization approach is unrealistic due to one flaw. It is limited to classifying a single object at a time. The initial step in analysing the nutrition contents and calories in the food photographs is to identify different food items from a single image. Rather than using a classification method, we require an item detection system that can identify several foods from a same image simultaneously in order to identify them all. Thus, the first difficulty is to apply a food detection system for numerous food photos, rather than a food classification system. Initially, the Wellnavi approach was applied with camera-enabled PDAs to authenticate an image-based food record. Users were obliged to take 45-degree angle photos of food both before and after eating episodes. The PDA's stylus was positioned next to the food items on the table to serve as a visual aid for estimating portion sizes. Users were needed to supply written information on the screen after the photographs were taken, particularly for dishes that were thought to be difficult to judge solely by looking at the images. Wireless transmission of the food photographs and descriptions to a server allowed registered dietitians and nutritionists to manually analyse the images. A quick survey was utilised to get further data on dietary practises, such as adding sugar to drinks and often utilised condiments, in order to facilitate analysis.

Methods

Collection of data

The EPIC-FFQ was distributed to 26478 people in the 40–79 age group EPIC-Norfolk cohort research, which is detailed in greater detail elsewhere. The FFQ was sent to participants, who then brought it back for a health evaluation. Nursing personnel then reviewed and filled it out as needed.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 787

Research Paper **© 2012 IJFANS. All Rights Reserved** UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

PLEASE PUT A TICK (✓) ON EVERY LINE

FOODS AND AMOUNTS	AVERAGE USE LAST YEAR								
BREAD AND SAVOURY BISCUITS (one slice or biscuit)	once a ay	Once a week week	2-4 per week	2-3 per week	5-6 per week	1-3 per month	6+ per day	4-5 per day	Never or less than once month
Brown bread and rolls									
Crispbread e.g. Ryvita		✓				✓			
Cream crackers, cheese biscuits				✓					
Porridge, Readybrek								✓	
Breakfast cereal such as cornflakes, muesli etc.		✓			✓				
CEREALS (one bowl)									
Whole meal bread and rolls						✓			
White bread and rolls		✓							

Figure 1 morning breads and cereals are discussed in the food frequency questionnaire illustration.

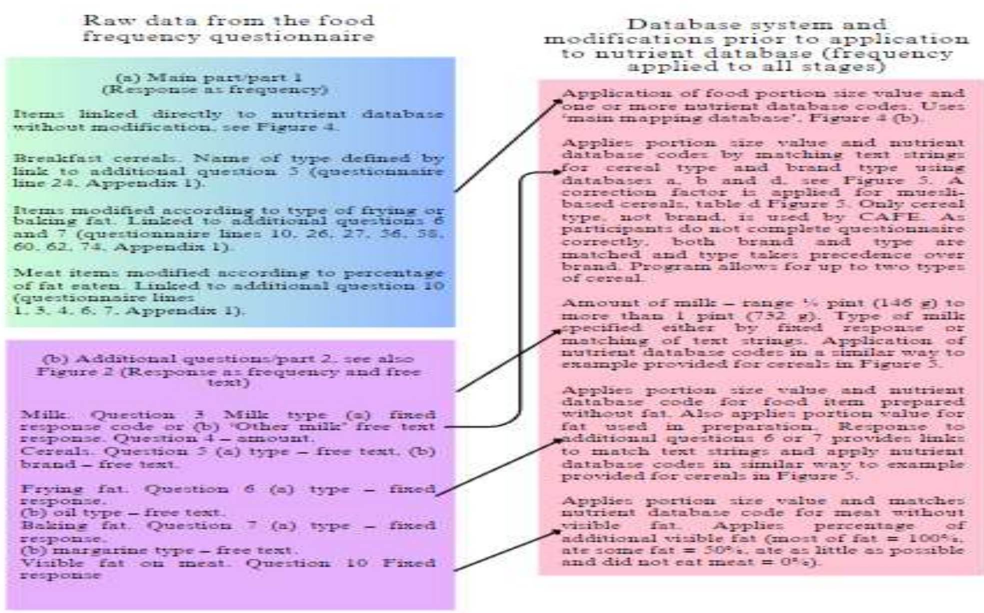

Figure 2 CAFE programme

Data input

We looked for codes outside of the permitted ranges when entering numbers. In order to look for

ISSN PRINT 2319 1775 Online 2320 787
Research Paper **© 2012 IJFANS. All Rights Reserved** UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

mistakes within the anticipated numerical ranges, a number of questionnaires were also visually examined and confirmed to be accurate. The text-matching system created specifically for the programme was used to fix any typographical problems. The projected entry time does not account for this because it was completed at a later period. Even with the initial verifications for accuracy, certain participants failed to include certain food items or provided more than one answer. The CAFE programme regarded these answers as missing data and assigned special codes during entry. Afterwards, those with more than ten missing data lines were removed from the dataset. Many respondents indicated how often they consumed cereal, how much milk they drank, or how often they baked or fried with fat, but they did not specify which kind in part 2 of the questionnaire. Questions were created to find the mismatch between parts 1 and 2 of the questionnaire because leaving out this information would have resulted in underestimates of nutrient consumption.

The CAFE program

Part 1 food items

As will be explained later, the "main mapping database" was designed to offer a portion size value and one or more nutrient database codes. Figs. 3 (the "a" main part) and 4 (the "b" main mapping database) show how most food items in part 1 (89%) were linked directly via this database. Eleven percent of the food products had a connection to the second section of the survey.

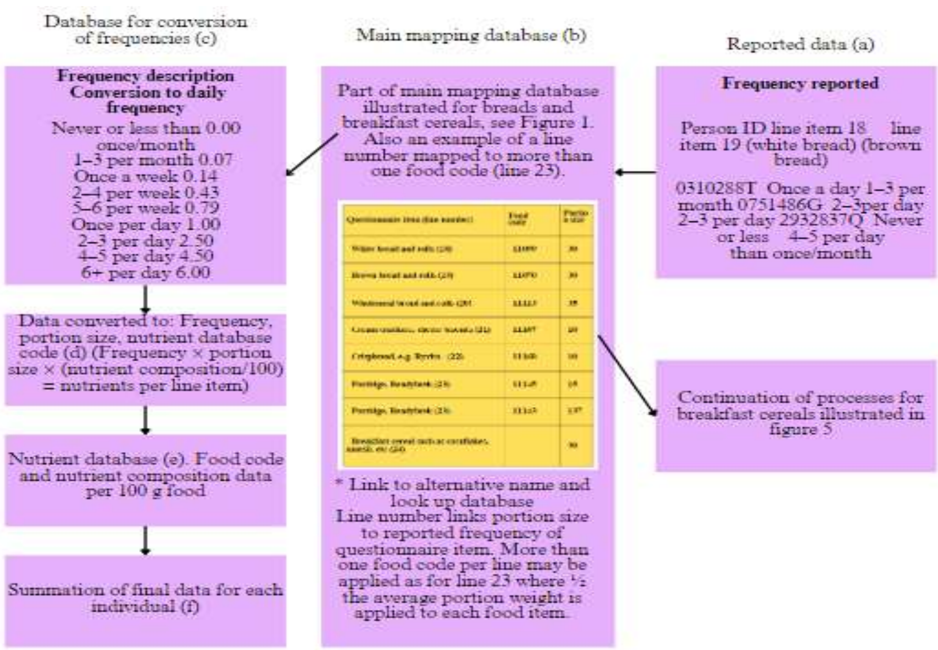

Figure 3 Diagrammatic illustration of the many forms of raw data that the CAFE programme collected and the Food frequency questionnaires (FFQ) used to modify the data.

Participants were asked to identify the brand and variety of cereals they had eaten, such as cornflakes

ISSN PRINT 2319 1775 Online 2320 787
Research Paper **© 2012 IJFANS. All Rights Reserved** UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

and Kellogg's, respectively. A maximum of four answers were permitted. However, because they did not think the distinction between brand and type was significant, a large number of respondents filled out the questionnaire inaccurately. In order to solve this issue, cereal brands were matched by another database (not shown), which prevented the need for additional processing. 2 different kinds of cereal were permitted per person under the programme, which gave preference to type over brand. Portion weights were also modified to take the number of types into consideration as seen in Fig. 2.

Visible fat

In order to determine the portion values in the mapping file for visible fat (question 10), data pertaining to the lean and separable fat components of meat (beef, lamb, or hog fat) were computed. The weight of separate fat was applied as follows: 100% if all fat was consumed, 50% if some fat was consumed, and 0% if no fat was consumed. Lean meat with no discernible fat and separable fat (beef, lamb, and hog) each have their own unique nutrient database codes within the mapping database.

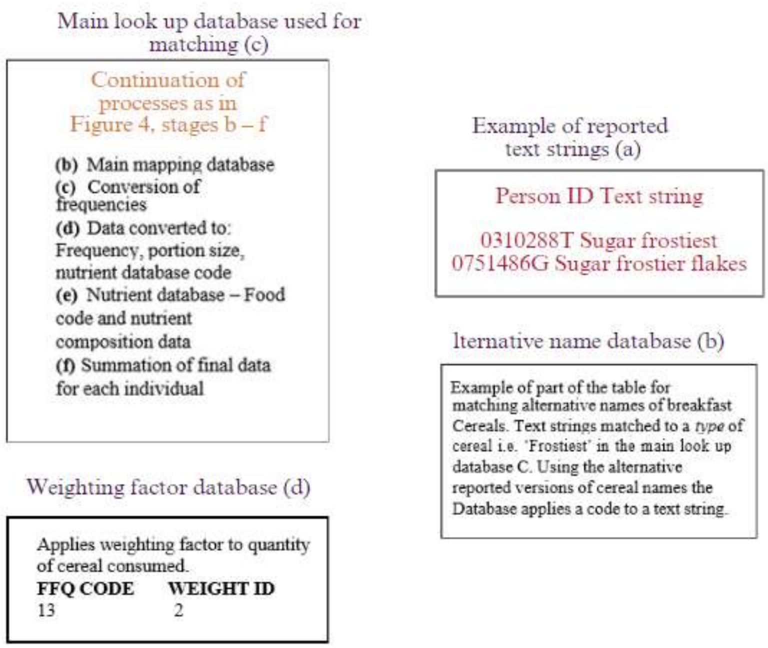

Figure 4 An example of the procedures used by the CAFE programme on the first section of the questionnaire. Each food's quantity is determined by applying a quantity from the "main mapping database" (b); databases (c) to the nutritional database, the final data (f) are obtained by applying data in the format (d)

Database:

Unless their nutrient makeup was equal, lines representing two or more foods were mapped to different nutrient database items; in that case, the more often consumed item (or items) was utilised. The data used to make decisions was for those between the ages of 40 and 74. Line items could be

ISSN PRINT 2319 1775 Online 2320 787
Research Paper © 2012 IJFANS. All Rights Reserved UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

represented by up to six different food codes.

The Indian population statistics and weighted records of study participants aged 40–74 years were used to calculate the masses of standard measures and average sizes of portions . The mean value was utilised if there were gender variations in the typical portion size. part sizes were distributed equally based on the number of line items that included two or more nutrient data codes; In order to consider summertime fruit consumption, certain portion sizes have been divided by 3.

Database of nutrient using programs

For eight of the eighteen nutrients listed here, the nutrient data is complete. Less than 3% of the values for carbohydrate, sugars, vitamin C, beta-carotene, iron, and potassium were missing for the remaining 10 nutrients; 3.6% of the data were missing for dietary fibre and magnesium, 6.9% for the antioxidant vitamin D, and 8.6% for folate. For questions 3–7, the nutritional information for generic answers to different kinds of cereals, fats, or milks was computed. This included equal amounts of "cornflakes," "muesli," and "weetabix" for cereals. Equal amounts of vegetable oil, butter, lard, compound cooking fat, and mixed margarines were utilised for non-specific fats. 40% semi-skimmed milk, 50% whole milk, and 10% skimmed milk were used to get the data for nonspecific milk.

Identification of outliers

Table 2 shows that the FFQ produced some implausibly high and low nutrient intakes, some of which fell outside the ranges discovered by a second dietary assessment approach, the EPIC 7-day dairy, which is an illustration of a record method . For comparison, diary data for 4271 women and 3891 males were provided . As a result, appropriate cut-points had to be created for after the bottom and top 0.5% of EI: BMR were excluded (see Table 1).

Consumption frequency of dietary groups by status report

The food list items were separated into 11 groups, and the average daily portion count for each group was calculated by adding the total reported frequency for each group. Nonalcoholic drinks and various snacks were not included. Due to the unpredictable nature of portion sizes, milk was also eliminated. After excluding those who had missing data for any questionnaire item, 17 083 people remained for study. BMR was estimated for these people using sex-specific formulae that took their age and body weight into account.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 787
Research Paper © 2012 IJFANS. All Rights Reserved UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

Table 1 Cutoff points utilising the ratio of consumed energy for the exclusion criterion for excessive energy values

Sex	MJ# Between	Bottom 0.5% EI : BMR*	Top 0.5% EI : BMR* Between	
EI : BMR known				
Male	#0.47	$2.45		
Female	#0.55	$3.43		
Where body weight not known				
Male	-	-	$2.56	#3.12
Female	-	-	$4.454	#2.65

Results

Table 2's data for people before outliers were eliminated indicate that, as would be predicted, men consumed more mean intake of all nutrients than women. Protein had the lowest CV whereas alcohol had the highest. Carotene, vitamin D, total fat, vitamin C, polyunsaturated, saturated, and monounsaturated fats were among the other nutrients having a high CV. For women, energy intake ranged from 0.69 to 38.23 MJ, whereas for men it was between 1.8 and 106.02 MJ. In our cohort, all maximum values for nutrient intakes were higher than those discovered by the 7-day diary approach.

Table 2 The impact of eliminating the nutritional values that fall outside of the upper and lower 0.5% range of the energy intake to anticipated body mass ratio (EI : BMR) on the food frequency questionnaire (FFQ)

	Data prior to exclusion of outliers* (n ¼ 11 360)						***Data after exclusion of outliers (n ¼ 11 248)***			
	Median	SD	Mean	Maximum	Minimum	CV %	Median	SD	Mean	*Minimum*
Vitamin D (mcg)	3.13	3.64	2.23	0.03	73.22	61.3	3.13	3.62	2.06	*0.03*
Alcohol (g)	6.7	12.3	16.2	0	177.9	131.7	6.7	12.3	16.1	*0*
MJ	8.	9.22	2.99	1.8	106.02	32.	8.	9.17	2.63	*3.12*

ISSN PRINT 2319 1775 Online 2320 787
Research Paper © 2012 IJFANS. All Rights Reserved UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

	9					4	9			
Iron (mg)	12	12.5	4.1	2.2	92.2	32.7	12	12.49	3.88	*2.55*
Vitamin E (mg)	12.8	14.5	7.2	2.1	187.3	49.7	12.8	14.4	6.8	*2.1*
Polyunsaturated fat (g)	13.7	15.3	7.5	0.9	180.7	49.0	13.7	15.2	7	*1.4*
Fibre (Englyst, g)	17.3	18.1	6.6	1.3	115.6	36.5	17.3	18	6.4	*1.3*
Monounsaturated fat (g)	27.9	29.8	13.3	3.4	479.3	44.6	27.9	29.5	11.8	*4.7*
Saturated fat (g)	30	32.4	15	3.2	404.1	46.3	30	32.2	13.8	*3.3*
Fat (g)	78.7	83.5	34.9	10.5	1114.8	41.8	78.7	82.9	31.2	*13.4*
Protein (g)	83.2	85.1	23.6	13.4	477.7	27.7	83.2	84.9	22	*23.3*
Vitamin C (mg)	105	113	53	2	669	46.9	105	113	52	*10*
Starch (g)	122	127	49	10	1648	38.6	122	127	45	*10*
Carbohydrate (g)	259	271	98	48	3589	36.2	259	269	87	*48*
Folate (mcg)	316	328	100	23	1547	30.5	316	327	96	*77*
Calcium (mg)	1018	1039	316	187	6565	30.4	1018	1037	300	*189*
Carotene (mcg)	2714	2803	1490	66	27 870	53.2	2715	2798	1447	*107*
Potassium (mg)	**3802**	**3875**	**956**	**1186**	**15 640**	**24.7**	**3802**	**3868**	**908**	***1284***

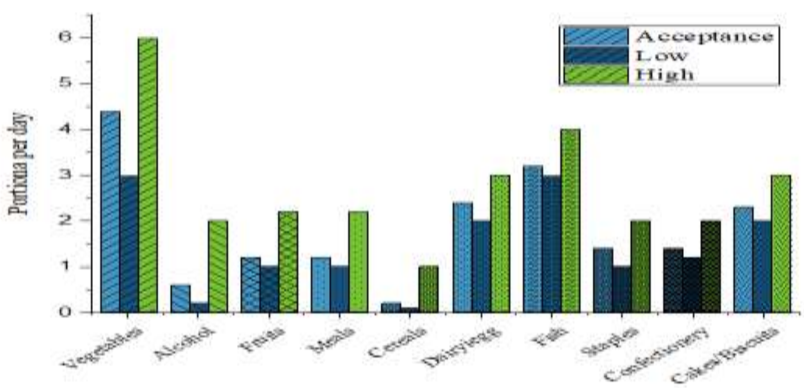

Figure 6 energy intake basal metabolic rate

Research Paper © 2012 IJFANS. All Rights Reserved

Discussion

It is also easily updateable and versatile. Additionally, the programme includes a variety of morning cereals and differentiates text for brands and varieties. There aren't any additional programmes that are specifically written for nutritional analysis that we are aware of.

Only a portion of the questionnaire made use of the text-matching capability to connect precise reported data with the nutritional database. This idea, though, might be expanded upon and used more widely as a method of managing data for scientific research. The ability to record data precisely as reported and handle it systematically at a later time is one advantage of free text entry. The number of cereal varieties was restricted in the previous FFQ edition, but because cereals are so important for nutrition, it was decided to include more information in the UK. The majority of the portion weights utilised in the initial validation researchSome people may have meant to answer "never," which could account for missing data in this questionnaire. Other possible explanations include individual negligence or inattention.

Conclusion

To sum up, we have created a versatile programme that generates nutritional information from the FFQ and it has been applied to a sizable population research involving both men and women in their middle years. The nutrient intakes were comparable to what was reported for the SHHS and Willett FFQs. Without additional research on portion sizes, the EPIC-FFQ is unlikely to be suitable for usage in other age groups. We determined that this cohort was following a standard UK diet by using frequencies. Although an approach to eliminate extreme energy values was devised, certain extreme nutrient values persisted in the data due to high and low food frequency reports, which may be a sign of inaccurate reporting.

Reference

1. Shen, Z., Shehzad, A., Chen, S., Sun, H., & Liu, J. (2020). Machine learning based approach on food recognition and nutrition estimation. Procedia Computer Science, 174, 448-453.
2. Kirk, D., Kok, E., Tufano, M., Tekinerdogan, B., Feskens, E. J., & Camps, G. (2022). Machine learning in nutrition research. Advances in Nutrition, 13(6), 2573-2589.
3. Oliveira Chaves, L., Gomes Domingos, A. L., Louzada Fernandes, D., Ribeiro Cerqueira, F., Siqueira-Batista, R., & Bressan, J. (2023). Applicability of machine learning techniques in food intake assessment: A systematic review. Critical Reviews in Food Science and Nutrition, 63(7), 902-919.
4. Morgenstern, J. D., Rosella, L. C., Costa, A. P., de Souza, R. J., & Anderson, L. N. (2021). Perspective: Big data and machine learning could help advance nutritional epidemiology. Advances in Nutrition, 12(3), 621-631.
5. Ma, P., Li, A., Yu, N., Li, Y., Bahadur, R., Wang, Q., & Ahuja, J. K. (2021). Application of machine learning for estimating label nutrients using USDA Global Branded Food Products

ISSN PRINT 2319 1775 Online 2320 787
Research Paper **© 2012 IJFANS. All Rights Reserved** UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

Database,(BFPD). Journal of Food Composition and analysis, 100, 103857.

6. Naik, P. A. (2020). Intelligent food recommendation system using machine learning. Nutrition, 5(8).
7. Hu, G., Ahmed, M., & L'Abbé, M. R. (2023). Natural language processing and machine learning approaches for food categorization and nutrition quality prediction compared with traditional methods. The American Journal of Clinical Nutrition, 117(3), 553-563.
8. Yunus, R., Arif, O., Afzal, H., Amjad, M. F., Abbas, H., Bokhari, H. N., ... & Nawaz, R. (2018). A framework to estimate the nutritional value of food in real time using deep learning techniques. IEEE Access, 7, 2643-2652.
9. Sundaravadivel, P., Kesavan, K., Kesavan, L., Mohanty, S. P., & Kougianos, E. (2018). Smart-log: A deep-learning based automated nutrition monitoring system in the iot. IEEE Transactions on Consumer Electronics, 64(3), 390-398.
10. Ma, P., Lau, C. P., Yu, N., Li, A., & Sheng, J. (2022). Application of deep learning for image-based Chinese market food nutrients estimation. Food Chemistry, 373, 130994.
11. Picone, G., Mengucci, C., & Capozzi, F. (2022). The NMR added value to the green foodomics perspective: Advances by machine learning to the holistic view on food and nutrition. Magnetic Resonance in Chemistry, 60(7), 590-596.
12. Li, T., Wei, W., Xing, S., Min, W., Zhang, C., & Jiang, S. (2023). Deep Learning-Based Near-Infrared Hyperspectral Imaging for Food Nutrition Estimation. Foods, 12(17), 3145.
13. Al-Saffar, M., & Baiee, W. R. (2022). Nutrition information estimation from food photos using machine learning based on multiple datasets. Bulletin of Electrical Engineering and Informatics, 11(5), 2922-2929.
14. Davies, Tazman, Jimmy Chun Yu Louie, Rhoda Ndanuko, Sebastiano Barbieri, Oscar Perez-Concha, and Jason HY Wu. (2022). "A machine learning approach to predict the added-sugar content of packaged foods." The Journal of Nutrition 152, no. 1 : 343-349.
15. Davies, Tazman, Jimmy Chun Yu Louie, Tailane Scapin, Simone Pettigrew, Jason HY Wu, Matti Marklund, and Daisy H. Coyle. (2021) "An innovative machine learning approach to predict the dietary fiber content of packaged foods." Nutrients 13, no. 9 : 3195.
16. Pouladzadeh, Parisa, Pallavi Kuhad, Sri Vijay Bharat Peddi, Abdulsalam Yassine, and Shervin Shirmohammadi. (2016) "Food calorie measurement using deep learning neural network." In 2016 IEEE international instrumentation and measurement technology conference proceedings, pp. 1-6. IEEE.