IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Food Image Detection and Calorie Calculation Using Deep Learning

[1]Anoop Krishnan P , [2]Mr. Pramod K

[1]MCA Scholar, [2] Associate Professor

Department of MCA

Nehru College of Engineering and Research Centre, Pambady,

ABSTRACT

The increasing demand for health and fitness applications has driven the development of AI- powered systems for food recognition and calorie management. This project utilizes Convolutional Neural Networks (CNN) and Artificial Intelligence (AI) to create a robust food image detection and calorie calculation platform. The system allows users to upload images of food, which are processed using a CNN model trained to recognize a wide variety of food items. Once the food is identified, users can input additional details such as portion size, ingredients, drinks, or preparation method to enhance the accuracy of the calorie estimation.

The core functionality of the system revolves around accurate food recognition through CNNs, which are well-suited for image classification tasks. The network is trained on a large dataset of labeled food images to classify the food items and estimate their nutritional content. Based on the identified food, the system calculates the total calorie count using a deep learning model integrated with a nutritional database. In addition to calorie calculation, the system provides personalized exercise recommendations based on the user's calorie intake. By leveraging AI algorithms, the platform suggests various types of exercises (e.g., running, cycling, yoga) and the duration needed to burn the calories consumed, ensuring that users can make informed decisions about their physical activity to maintain a healthy lifestyle. The system can access this information via a Google Gemini API or the pre-trained LLAMA model from Hugging Face for more complex, real-time conversational interactions.

This platform not only provides accurate food detection and calorie tracking but also empowers users to make healthier choices by offering actionable insights for maintaining balance between calorie intake and expenditure. The integration of AI technologies ensures a seamless, interactive, and personalized experience for the end-users.

Key words: Mask R-CNN, ROI (Region of Interest), IOU (Intersection Over Union), and Resnet 50

I. INTRODUCTION

In the modern world, it is crucial that individuals understand what they are eating and how it will affect their bodies. Thus, a system that might assist people in maintaining the number of calories they consume is crucial. Most people on the planet reside in nations where obesity and overweight cause more deaths than any other illness. Having enough food is not the issue here; rather, it is the people's ignorance about their nutrition. People could choose how much calories they wish to eat with ease if they could calculate how many calories they would take in a day. But controlling caloric intake is a difficult undertaking that requires people to manually record the foods they eat during the day, and they must calculate how much of calories they have taken in. Because the calorie estimation depends on both what you are eating and how much you are having, this method is not only manual but also imprecise.

The need for image recognition models has increased due to advancements in image processing techniques. Researchers are using picture recognition technology extensively as a model for a number of applications, including video frame analysis, cancer diagnosis, and self- driving cars. Researchers are also interested in using the image to determine how many calories the food item contains. Using the provided photos, researchers have estimated the number of calories using a variety of machine learning and deep learning algorithms.

II. LITERATURE SURVEY

Although the subject of automated food image classification is still in its infancy, it has the potential to revolutionise how people live and eat. CNNs are a type of deep learning system that is very good at classifying images. CNNs operate by removing characteristics like edges, shapes, and colours from pictures. The image is then categorised into a particular group using these attributes [1]. CNNs are a type of deep learning system that is very good at classifying images.

CNNs work by recognising features such as edges, colours, and shapes in pictures. These features are then used to categorise the image into a certain category. Because food looks vary so much and are so complicated, automatic food recognition is a difficult undertaking. For babies in particular maintaining the proper dietary balance is crucial. Serious disease and organ damage can result at any time the body is lacking essential nutrients and this can cause significant health problems as an adult [2].

Numerous improvements in image processing have been made possible via deep learning. In particular,there have been significant developments in the classification of food images using deep learning techniques [3]. More than 65% of Sub-Saharan countries are predicted to be malnourished as a result of universal poverty, and some agricultural regions are experiencing drought [4]. With the highest employability among the many industries in the world, the food preparation and processing industry is the most significant [5].

Using natural language translation, the method combines deep learning for the identification of false food images with food matching and standardisation [6]. By automating reactive composting, the aforementioned compost maturity forecasting approach reduces labour costs [7]. Since food is the cornerstone of societal advancement, stability, and human well-being, food quality and safety are vital concerns for the entire community [8].

III. METHODOLOGIES

1) Overview

Our model receives an image of the food item as input and outputs the calorie content of the item. Several intermediate steps are taken in order to accomplish this. First of all, the collected image identifies the food item whose calorie content needs to be predicted. The food item's size, volume, and final calorie estimate are calculated to 128*128 pixels after it has been identified and used in the model. Images are recognised using the Mask R-CNN algorithm, and calories are predicted using the approximate percentage method format.

2) Dataset

There are six distinct food items in our sample. Images of food items are chosen from the internet and from pre-existing datasets, such as Food 101, to create a bespoke dataset. Food items include hot dogs, oranges, pizza toast, bananas, and oatmeal. Since it's crucial to scale the photos that are collected from various sources, they are first resized to 128*128 pixels before being utilized in the model. Additionally, images are manually annotated using the Pixel Annotation tool, and masks are made for model training.

Food Item	Image Count
Banana	113
Pizza Toast	104
Orange	101
Idli	113
Hot Dog	102
Omelet	105

Table 1. Dataset details

3) Food-Item Identification

To construct a pixel-wise mask of the image's objects, we employ instance segmentation. This method allows us to comprehend the food data in the image at a finer level. In this case, we utilize For picture segmentation, Mask R-CNN employed ROI and IOU to create bounding boxes, supply labels, and create a mask. as shown in picture 2. The food item is given a label that reads "omelette," and a bounding box is made across it.

Fig.2 Mask output of omelet

A deep neural network called Mask R-CNN was created to address instance segmentation in machine learning. It distinguishes between various things in a picture or a video and displays the bounding boxes, classes, and masks for objects. The mask R-CNN has two stages. The first step involves using the input image to suggest regarding the possible existence of an item. The second step involves predicting the object's class, fine-tuning the border boxes, and creating a mask at the pixel level. The backbone structure, which aids in feature extraction, is connected to both phases. Here, the backbone is Resnet 101. Preloaded weights are used to initialise the Mask R-CNN Model.

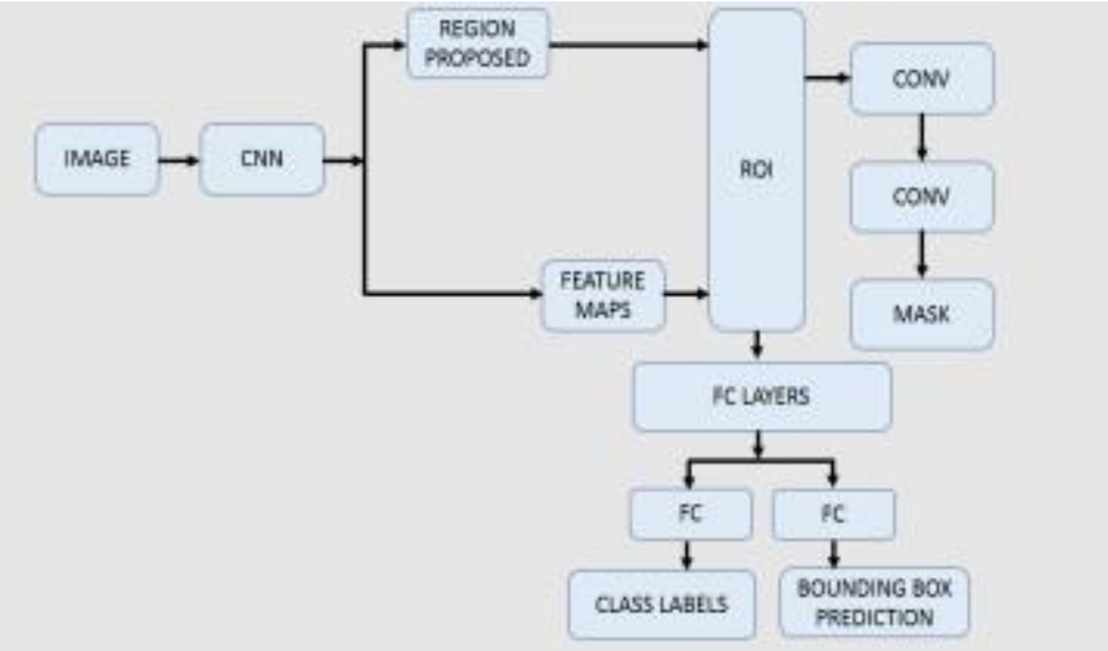

Fig.3 Mask R-CNN Model

Deep learning techniques require a large number of photos in order to train a Mask R-CNN- based image recognition model. Given that the dataset we gathered is insufficiently substantial to train the model,We applied the Matterport repository's transfer learning methodology [7]. Using transfer learning, we began with a weights file that had already been trained on the COCO dataset rather than starting from scratch when creating a model. The trained model weights have already learnt many of the characteristics found in realistic photos thanks to the large number of images (~120K) in the COCO dataset, which is very beneficial.

4) Food Calorie Prediction

We need a way to assess the size of the meal in a real-world situation or compute the calories because the same food can be photographed at different depths to produce varying picture sizes. Once the targeted food items and their masks have been identified, we require the actual object sizes, which cannot be obtained from a pinhole camera image alone. In order to determine the precise size of the food in that particular image, we employ a referencing technique that compares the food objects to the size of the known object.

The approximation of proportions is the approach used to predict the number of calories in meals. This method predicts the calorie of the input by using the calorie per mask of the food class as a reference. picture. A spreadsheet with the following fields has been created: "Class," "Calorie Per Mask," "Minimum Calories," and "Maximum Calories."

Food	Calorie_per_mask	Minimum_Calorie	Maximum_Calorie
Banana	0.02739726	72	135
Hot Dog	0.0375	100	290
Omelette	0.01725	50	80
Orange	0.0385	130	320
Pizza Toast	0.051555556	200	242
Idli	0.009428571	25	40

Table 2. Calorie details used in predicting food calorie

Each image has been resized to 128*128 pixels using the proportion approximation approach, and the segmented area of the food class is computed and then multiplied by calories according to Mask.

IV. RESULTS

1) Dataset Inspection

Six meal items—banana, pizza toast, omelette, orange, and idli are included in our dataset

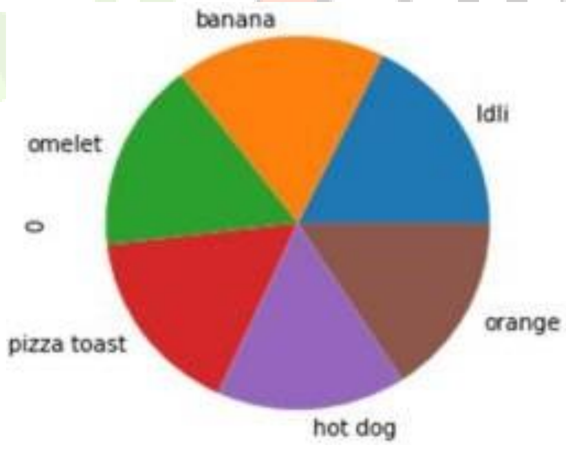

Fig.5 Food classes in dataset

Since the lengths of the images vary, we converted them to 128*128 before loading them. In a similar vein, mask pictures also transformed them into 128*128 pixels. We performed data augmentation and added extra photos because the original dataset was class imbalanced. It was discovered during the visual dataset assessment that the classes were initially out of balance; this was corrected by including additional photos. The image size has also been adjusted to 128*128. Class images have a minimum of

101 and a maximum of 113.

Fig.6 Initial Visual representation of dataset

2) Food Item Identification

Once the elements of the image have been extracted, our objective is to determine the type of food item. Here, the image is labelled, a mask is applied, and a bounding box is generated using Mask R-CNN. After preprocessing, pictures and masks are analysed and converted to 128 * 128 pixels.

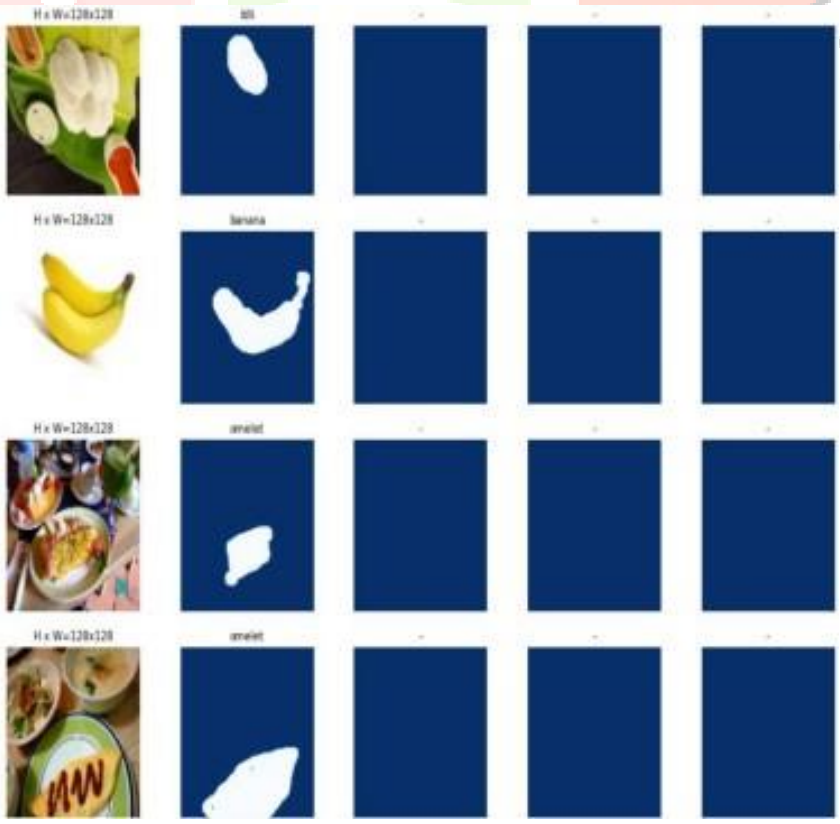

Fig.7 Analyzing images and mask after preprocessing

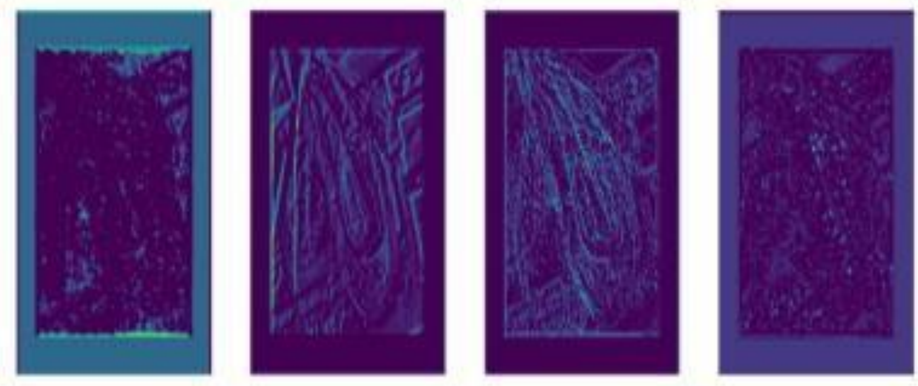

Fig.8: Analyzing activation of hidden layers

3) Calorie Prediction

A referencing technique that determines the true size of the food in that particular image by comparing the food objects to the size of the known thing. See figure 4.

Fig.9 Actual Image- Image before Mask R-CNN and calorie prediction

Fig.10 shows 242 calories for pizza toast - Image after Mask R-CNN and calorie prediction

4) Visualization Of Model Accuracy & Loss

Fig.11 Overall loss while training the Train dataset

Fig.12 Overall loss while Validating the Test dataset

Fig.13 Classification loss while training the Train dataset

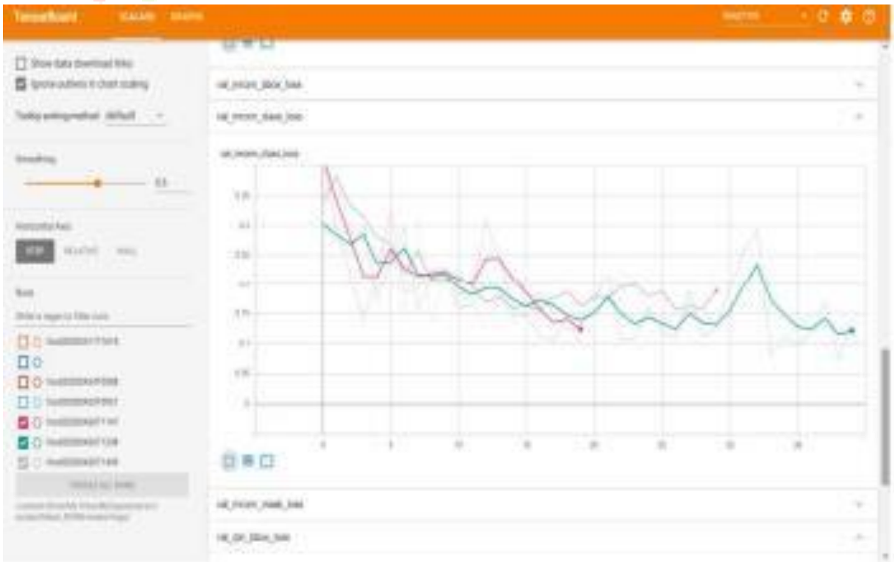

Fig.14 Classification loss while Validating the Test dataset

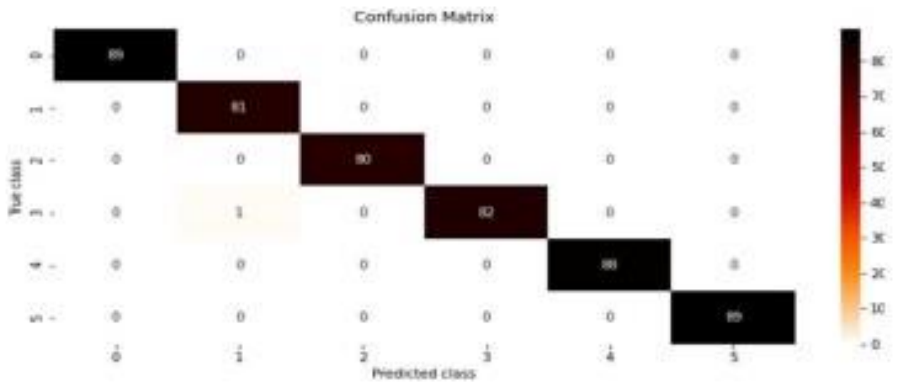

Fig.15: Confusion matrix while Training the Train data (0=BG ,1=Orange, 2=hot dog, 3= omelet ,4=banana, 5=pizza toast,
6=idli)

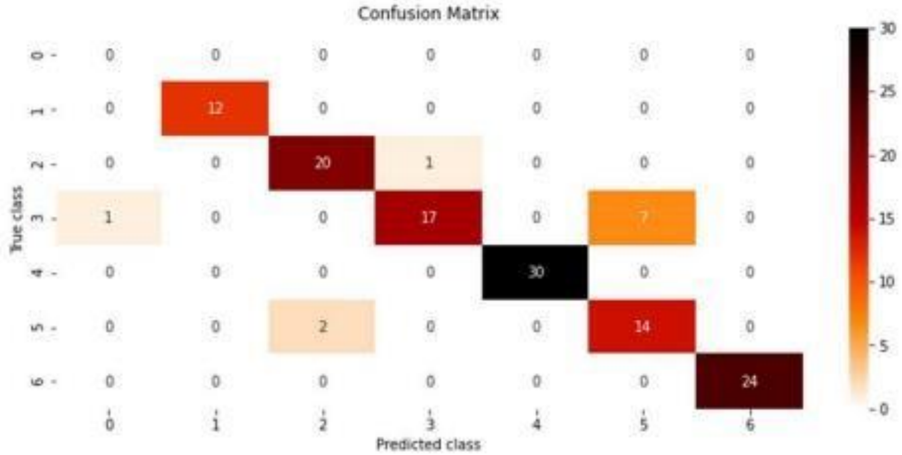

Fig.16: Confusion matrix while Validating/testing the Test data (0=Background ,1=Orange, 2=hot dog, 3= omelet,4=banana, 5=pizza toast , 6=idli)

V. CONCLUSION

In this study, we predicted the total number of calories in the food item shown in the picture using a deep learning-based model. This solution was developed using the Mask R-CNN method for making bounding boxes and masks. This further enabled the model to anticipate the calories associated with each food item in a satisfactory manner by assisting it in determining the surface area filled by the various food items in the image. Mathematical algorithms that compare the percentage of the image that each food item occupies and calculate the calories associated with it is used to estimate calories.

VI. CHALLENGES AND FUTURE WORK

Food image recognition and calorie prediction systems have the potential to revolutionize the way people manage their diets and monitor their health. However, these systems face several challenges. Variability in food appearance due to differences in lighting, angles, and preparation methods makes it difficult for models to consistently identify food items. Additionally, estimating portion sizes accurately from images remains a significant challenge, as slight discrepancies can lead to incorrect calorie estimations. The diversity of food types across different cultures and regions also complicates the task of developing models that can generalize well across various cuisines. Furthermore, obtaining large, well-labeled datasets for training these models is resource-intensive and often incomplete. Current approaches often rely on manual user input, which can be inaccurate and cumbersome. Future research should focus on enhancing the accuracy of food recognition models by incorporating more diverse datasets, using advanced techniques like 3D reconstruction for better portion size estimation, and integrating wearable devices for personalized calorie predictions. Additionally, improving real-time processing and user interfaces will be crucial for making these systems accessible and practical for everyday use. Addressing these challenges will enable food image recognition and calorie prediction systems to become more effective tools for managing health and promoting healthy eating habits.

REFERENCES

1. Bishop, T. R., von Hinke, S., Hollingsworth, B., Lake, A. A., Brown, H., & Burgoine, T. (2021). Automatic classification of takeaway food outlet cuisine type using machine (deep) learning. *Machine learning with applications*, *6*, 100106.

2. Sundaravadivel, P., Kesavan, K., Kesavan, L., Mohanty, S. P., C Kougianos, E. (2018). Smartlog: A deep- learning based automated nutrition monitoring system in the iot. *IEEE Transactions on Consumer Electronics*, *c4*(3), 390-398.

3. Pan, L., Pouyanfar, S., Chen, H., Qin, J., C Chen, S. C. (2017, October). Deepfood: Automatic multi-class classification of food ingredients using deep learning. In *2017 IEEE 3rd international conference on collaboration and internet computing (CIC)* (pp. 181-189). IEEE.

4. Farooqui, N. A., Mishra, A. K., C Mehra, R. (2022). IOT based automated greenhouse using machine learning approach. *International Journal of Intelligent Systems and Applications in Engineering*, *10*(2), 226-231.

5. Kumar, I., Rawat, J., Mohd, N., C Husain, S. (2021). Opportunities of artificial intelligence and machine learning in the food industry. *Journal of Food Quality*, *2021*, 1-10.

6. Mezgec, S., Eftimov, T., Bucher, T., C Seljak, B. K. (2019). Mixed deep learning and natural language processing method for fake-food image recognition and standardization to help automated dietary assessment. *Public health nutrition*, *22*(7), 1193-1202.

7. Wan, X., Li, J., Xie, L., Wei, Z., Wu, J., Tong, Y. W., ... C Zhang, J. (2022). Machine learning framework for intelligent prediction of compost maturity towards automation of food waste composting system. *Bioresource Technology*, *3c5*, 128107.

8. Zhu, L., Spachos, P., Pensini, E., C Plataniotis, K. N. (2021). Deep learning and machine vision for food processing: A survey. *Current Research in Food Science*, *4*, 233-249.

9. Kumar, R., C Singhal, V. (2022). IoT enabled crop prediction and irrigation automation system using machine learning. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, *15*(1), 88-97.

10. Khan, M. I., Acharya, B., C Chaurasiya, R. K. (2022). Automatic Prediction of Glycemic Index Category from Food Images Using Machine Learning Approaches. *Arabian Journal for Science and Engineering*, *47*(8), 10823-10846

11. Sood, S., C Singh, H. (2021). Computer vision and machine learning based approaches for food security: A review. *Multimedia Tools and Applications*, *80*(18), 27973-27999.

12. van den Bulk, L. M., Bouzembrak, Y., Gavai, A., Liu, N., van den Heuvel, L. J., C Marvin, H. J. (2022). Automatic classification of literature in systematic reviews on food safety using machine learning. *Current Research in Food Science*, *5*, 84-95.

13. Salim, N. O., Zeebaree, S. R., Sadeeq, M. A., Radie, A. H., Shukur, H. M., C Rashid, Z. N. (2021, July). Study for food recognition system using deep learning. In *Journal of Physics: Conference Series* (Vol. 1963, No. 1, p. 012014). IOP Publishing.