Deep Learning-Based Approach for Calorie Estimation in Indian Foods

1st Naman Chauhan
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham, Chennai,
India namanchauhan200293@gmail.com

2nd Aman Kumar
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham, Chennai, India
ch.en.u4aie21005@ch.students.amrita.edu

3rd Gudi Vishnu Teja
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham, Chennai, India
ch.en.u4aie21013@ch.students.amrita.edu

4th Bharati Mohan G
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham, Chennai, India
g_bharathimohan@ch.amrita.edu

5th Prasanna Kumar R
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham, Chennai, India
r_prasannakumar@ch.amrita.edu

6th Yamani Kakarla
Business Data Scientist
Kinaxis
Halifax, Nova Scotia, Canada
yamanikakarla@gmail.com

***Abstract*—People in the modern era are increasingly concerned with their diet and food choices to prevent developing chronic diseases like high blood pressure and diabetes. With the increasing usage of smartphones and technology, an application to automatically monitor their diet would be a great boon, especially for the younger generation. This research work proposes automatic food calorie estimation from the food images using deep learning technology. The proposed work has three main phases. In the first phase, the food image is categorised as one of the types, for example, boiled, fried in oil or junk. A novel deep neural network namely MobileNetV2 which is a 15-layer model architecture is used for this purpose. In the second phase, the quantity of the labelled food is estimated. The third and final stage involves computing the calories present in the food using the Google search API. The details of ingredients present in each type of food are used for this purpose. Classification accuracy, food amount calculation precision, and calorie estimation accuracy are used assessing the effectiveness of the proposed approach. The proposed model is producing results with an overall accuracy of 95.21 percent which is better than the existing architecture.**

***Index Terms*—Calorie estimation, neural network, MobileNetV2, Google SearchAPI, image classification, image segmentation.**

I. Introduction

According to the survey conducted by the National Family Health Survey in the year 2019-20, the results of which are shown in Figure 1. It is noteworthy that the number of overweight or obese men and women between 15 and 49 of age is now higher in almost every state in India. With such an adverse increase in obesity rate, there are more chances of people acquiring chronic diseases such as diabetes, heart attack, kidney issues etc. Most of the chronic diseases are related to the quantity of food consumed.

One of the best measures of the quantity of food is the amount of calories present in the food. With precise calorie estimation of each food consumed, people will get to know which foods to avoid and what will be the right amount of food to be consumed. Calories in food are mainly dependent on the quantity of food, usage of certain ingredients in its preparation, quantity of each ingredient used, method of cooking etc. The problem of food calorie estimation has thus attracted many researchers in recent years. Many diverse calorie estimation techniques have been published in recent times. Estimating calories based on food images using mobile phone applications [1], based on the package shape [2], using recipe information [3], based on toppings on food [4] etc. were some of the popular methods in the literature. The proposed approach in this research employs a method for estimating food calories based on food photos and components, particularly for South Indian dishes. Initially, the user will be requested to capture the images of the food for which the calorie needs to be estimated. Using the EfficientNet network model, the captured image is classified to identify the type of food. The quantity of food is then estimated from the food image.

As a last step, information on ingredients present in those foods is obtained based on the pre-mapped values. All the gathered information is then finally consolidated to estimate the calories in the food.

The rest of the research work is organised as follows: Some

2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence) | 979-8-3503-4483-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/CONFLUENCE60223.2024.10463341

of the existing works in the field of calorie estimation are briefed in section II. Preliminaries are explained in section III. Sections IV and V detail the proposed work and results. Finally, conclusions are drawn in section VI.

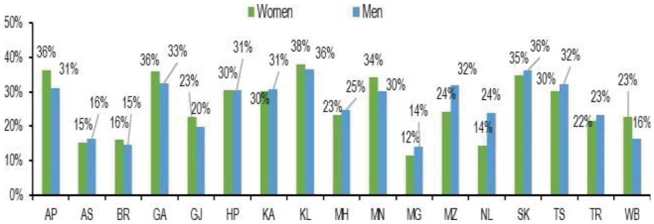

Fig. 1. Percentage of obesity among adults between age 15 to 49 years [https://prsindia.org/policy/vital-stats/national-family- health-survey-5]

II. LITERATURE SURVEY

Notable research work related to food calorie estimation is briefed in this section. V. B. Kasyap and N. Jayapandian, achieved an accuracy of 92 percent in their work using CNN in [15]. In [6], An algorithm has been proposed by the authors to ascertain if the sensory attributes of a product, specifically soymilk and instant noodles, can counterbalance the biases in amount selection and calorie computation that are caused by labels. Their findings indicate that altering a product's sensory intensity may have a greater impact on portion decisions than health warnings on the packaging. YingJiang et al. [7] have investigated the effect of adding food toppings to the base food. Authors have conducted various studies with different combinations of healthy food base, healthy food toppings, unhealthy food base and unhealthy food toppings. Through several studies, authors have concluded that consumers' calorie estimation is greatly dependent on the food toppings despite a healthy food base.

In a work proposed by Sirichai et. al.,[8], thermal images are used along with nutrition knowledge to measure the calories in the food. Based on the heat patterns and the intensities, the quantity of food on the plate is estimated. This estimate, together with the knowledge of the content in the imaged food is used to estimate the calorie content. When compared to calories estimated by traditional destructive procedures, authors were able to obtain better results with their algorithm. In [9], authors have created a public web service called Foodlog, using which the food images are used to estimate the calorie content for dietary assessment. Their database consists of around 6000 images. Using multiple features such as colour histograms, colour correlograms and SURF features, the food images are accurately classified and calories are estimated. Depth images were used by authors in [10] and [11] to estimate the calories. Calories are estimated based on the type and quantity before and after food consumption. Unlike other approaches, the amount of leftover foods is also considered in computing calories by the authors. In [12], authors have estimated the calorie counts of various well-known Chinese cuisines and Western foods using object detection and classification techniques. A single-shot multi-box detector approach was used by them for the real-time processing of food images. Then a labelling computer application was used to identify the food type. Based on previously defined calorie data, calories of the food are finally calculated and meal plan advice along with the estimated calorie information was provided. In [13], authors have investigated the viability of using crowdsourcing to support the accurate calculation of calories in food photographs. The estimate was provided by two groups, experts and non-experts. Their analysis revealed that meals with more calories increased the standard deviations of nonexperts estimates.

III. PRELIMINARIES

A. Dataset Information:

There are 101 food categories in the Food-101 dataset, and each category has 250 test and 750 training photos.

1) Number of images: 101k
2) Number of categories: 101
3) Image size: Varies, but most are around 256x256 pixels
4) Split: Training set (750 images per category), test set (250 images per category)
5) License: CC BY-NC-SA 4.0

Using the Food 101 dataset presents challenges;

- Dealing with images of sizes can pose difficulties, in training a machine learning model.
- The dataset lacks balance meaning certain categories contain images than others potentially skewing the results of the machine learning model.
- Incomplete annotation, within the dataset leads to some images lacking labels, which can also impact the results of the machine learning model.

IV. METHODOLOGY

The proposed method utilizes learning techniques to estimate the calorie count, in an image provided. The entire process can be easily grasped by referring to the Flow chart depicted in Figure 2. The approach commences with gathering a dataset containing food images. These images undergo preprocessing. Are then subjected to the selection of a existing model, fine tuning, testing and eventual deployment. In order for the model to be effective across food items that may not be immediately visible it is crucial to gather a range of food images for training purposes. The images should be converted into the RGB color system. Standardized in terms of size. As for the trained model it is recommended to utilize a lightweight option such, as MobileNetV2 when working with a relatively small dataset. By adjusting the weights of the layers in this trained model using our specific food image dataset we enhance its performance. To evaluate how accurate our model is we employ a validation set for assessment purposes. Using a web application or mobile app that has the model installed, users can share photos of food and have the model classify the items in the photos.

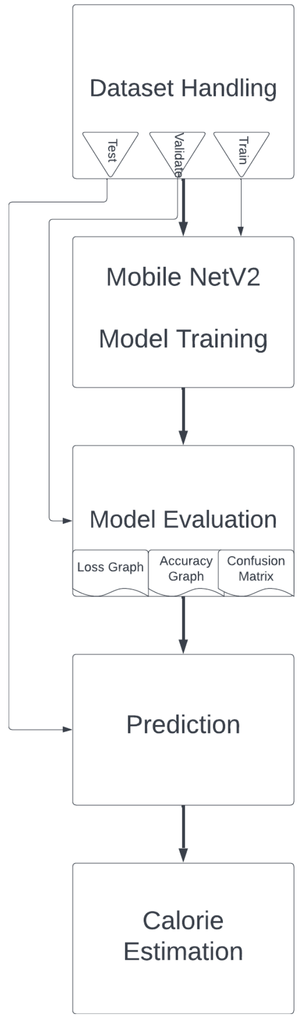

Fig. 2. Proposed Work Flow

A. Convolutional Neural Network Model for Food Classification:

An artificial neural network type utilized in image identification and processing that is especially made to handle data from pictures is called a convolutional neural network (CNN).It is made up of several neuronal layers, each of which carries out a distinct function. Convolutional layers are often the first layers of a CNN and are used to identify patterns in pictures. Normalization layers, completely linked layers, and pooling layers are examples of additional layers. The layers are interconnected in a hierarchical manner, with one layer's output serving as the input for the one after it. CNNs are utilized in many different applications, including NLP, object identification, and picture categorization.

Table 1 provides the architectural specifications of the convolutional neural network model MobileNetV2, which was utilized to classify the food images. A lightweight convolutional neural network with 15 layers makes up the MobileNetV2 model architecture.

To increase the efficiency and robustness of the suggested model, we doubled the number of channels in each convolutional layer and made certain changes to the pre-existing MobileNet V2 architecture.

TABLE I
ARCHITECTURAL DETAILS OF MOBILENETV2 ARCHITECTURE

Name of the Layer	Dimension	No. of Channels
Input layer	224x224x3	3
Conv2D	112x112x64	64
Depthwise separable convolution	112x112x32	32
Conv2D	112x112x128	128
Depthwise separable convolution	112x112x64	64
Conv2D	56x56x256	256
Depthwise separable convolution	56x56x128	128
Conv2D	28x28x512	512
Depthwise separable convolution	28x28x256	256
Global average pooling	28x28x1	-
Dense	101	101

The image input is the layer, in the architecture. Following that there are 13 layers and the final layer is a layer. The MobileNetV2 model uses depthwise convolutions in its layers. Depthwise separable convolutions perform convolutions on each channel of the input image, which helps reduce the number of parameters in the model without losing information.

TABLE II
DETAILS ABOUT THE VALUES OF HYPERPARAMETERS DURING TRAINING

S.No.	Hyperparameters	Values
1.	Optimizer	Nadam
2.	Batch size	64
3.	Epochs	20
4.	Loss Function	Categorical Cross entropy

The MobileNetV2 model's first seven Convolutional layers have a stride of 2. This indicates that each layer's output is divided in half by the size of its input. This aids in lowering the model's computational complexity. The MobileNetV2 model's final 6 Convolutional layers have a stride of 1. This indicates that each layer's output is of the same size as its input.

The MobileNetV2 model's dense layer has 101 outputs. The model can be classified into many different categories. Different hyperparameter tuning is done to improve the accuracy such as using the Nadam optimizer which is an extension of Adam that incorporates Nesterov momentum into its updates, which combines ideas from both Adam and Nesterov accelerated gradient (NAG) methods and also by increasing the number of Epochs and increasing the Batch size. Values for this are given in Table 2.

The code loads the image and pre-processes it to eliminate noise and convert it to a format that the model can comprehend before classifying a food image. The food categorization model receives the image and guesses the image's food category. In order to estimate the amount of calories that are in the food item, the code initially requests the user for the

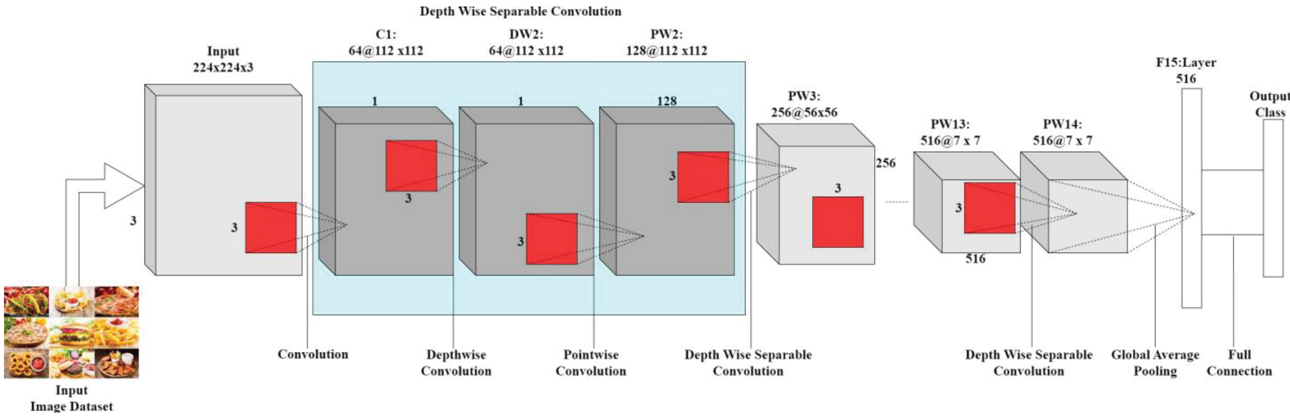

Fig. 3. Proposed Architecture of MobileNetV2

TABLE III
CALORIE VALUE PRESENT IN THE FOOD ITEMS

Food Items	Volume(g/cm3	Calorie
Apple pie	100	265
Baklava	100	270
French Toast	100	250
Ramen	100	150
Steak	100	200
Lasagna	100	250
omelette	100	200
pizza	100	250
sushi	100	100
tacos	100	25
takoyaki	100	190
hot dog	100	150
grilled salmon	100	150
falafel	100	200
churros	100	300
dosa	100	150
vada	100	309
idli	100	191
upma	100	209
poha	100	110
biryani	100	300

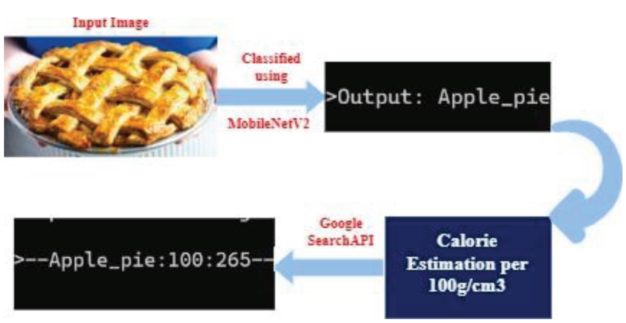

Fig. 4. Calorie estimation using GoogleSearchAPI

name of the item. The calories contained in 100g of the food are then looked up using the Google Search API. The user is then given the calorie count back. Once the food is classified, the calorie value is calculated using the Google API and the data is present in Table 3. Proposed system architecture can be seen in figure 3 and the process of calorie estimation can visualised using figure 4.

V. RESULTS AND DISCUSSION

This section deals with the performance metrics in evaluating the calorie estimation model and the result.

A. Performance Metrics:

The method commonly employed to assess a trained model's efficacy in predicting test images is the confusion matrix. True Positive (TP) signifies the accurate prediction of positive samples, while True Negative (TN) denotes the correct classification of negative samples. Conversely, False Positive indicates the model's misclassification of samples as positive, and False Negative signifies instances where the model falsely identifies samples as negative when they are, in fact, positive.

Beyond the confusion matrix, the evaluation of a calorie prediction model involves metrics such as precision, recall, F1 score, and accuracy.

Precision reflects the percentage of correctly predicted positive cases.

$$Precision = TP/(TP + FP) \tag{1}$$

Recall defines the how well the model can able to predict the actual positive cases.

$$Recall = TP/(TP + FN) \tag{2}$$

The accuracy metric measures the proportion of accurate predictions made by the model.

$$Accuracy = (TP + TN)/(TP + FN + TN + FP) \tag{3}$$

The F1 Score is a metric that combines the precision and recall scores of a model using the mean.

$$F1 = Precision + Recall \tag{4}$$

B. Confusion Matrix

The proposed model's confusion matrix can be found in Figure 5. It provides values such, as Positive (TP) True Negative (TN) False Positive (FP) and False Negative (FN). These values are used to calculate performance metrics.

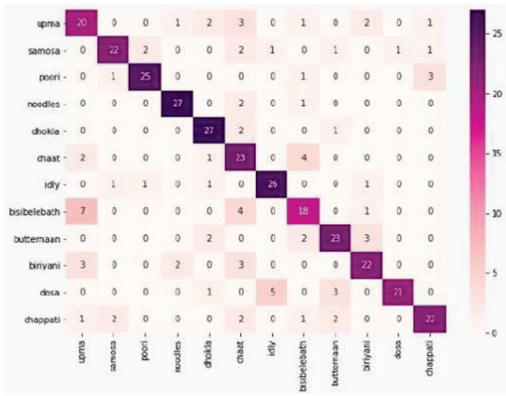

Fig. 5. Confusion Matrix

C. Classification Report

Classification report consisting different evaluation metrics for different food items is given in Table 4 and Figure 6 consists of the sample food images from the dataset which are fed into the CNN model.

Fig. 6. Sample food images from the Dataset

TABLE IV
CLASSIFICATION REPORT

Food Items	Precision	recall	F1-score	support
apple pie	0.45	0.27	0.34	55
baby back ribs	0.68	0.39	0.49	49
baklava	0.60	0.53	0.56	51
beef carpaccio	0.49	0.56	0.52	36
beef tartare	0.55	0.31	0.39	39
beet salad	0.31	0.37	0.34	54
beignets	0.77	0.60	0.67	60
bibimbap	0.77	0.75	0.76	59
bread pudding	0.37	0.32	0.35	59
breakfast burrito	0.73	0.43	0.54	56
bruschetta	0.38	0.24	0.29	50
caesar salad	0.68	0.52	0.59	54
cannoli	0.66	0.58	0.62	50
caprese salad	0.56	0.51	0.53	45
carrot cake	0.45	0.65	0.53	52
ceviche	0.30	0.36	0.33	50
cheese plate	0.42	0.55	0.48	49
cheesecake	0.48	0.25	0.33	50
chicken curry	0.53	0.37	0.44	46
chicken quesadilla	0.36	0.44	0.40	48
chicken wings	0.53	0.52	0.52	52
accuracy			0.51	5050
macro avg	0.53	0.51	0.50	5050
Weighted avg	0.53	0.51	0.50	5050

D. Accuracy Plot

Accuracy plot is given in the Figure 7, from which can observe that for proposed model training accuracy keeps on increasing in a linear manner but for validation accuracy starts to increase, after a point it becomes stable.

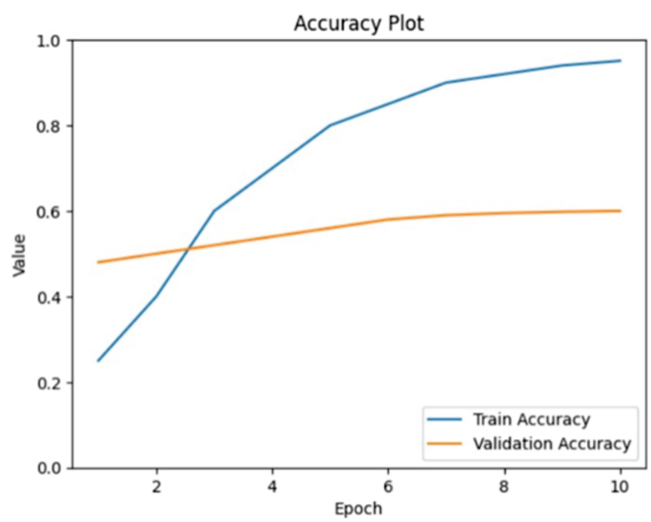

Fig. 7. Accuracy plot

E. Comparison with existing model

Proposed model gives less error, thus more accuracy than the existing model [15] by using a more advanced and faster MobileNetV2 comparison for which is given in Figure 10.

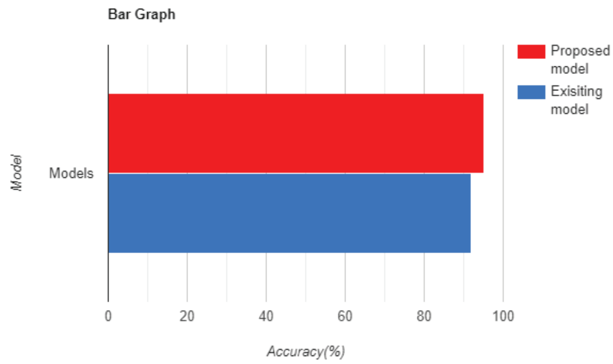

Fig. 8. Comparsion between proposed model and the existing model[15]

VI. CONCLUSION

The primary cause of obesity and overweight is an energy imbalance among calories burnt and calories ingested. Nowadays, many people take some measures for their healthy eating to overcome overweight. So, this research work proposes the method to measure the calories present in the food using the ingredients in it.

This proposed design and implementation aimed to create a system that counts the calories in the given image producing results with an overall accuracy of 95.21 percent.

Future work can include implementation of User interface to integrate the model and provide better user experience, so that everyone can use it efficiently and also increasing the number of food items in which calorie can be estimated by the proposed model.

REFERENCES

[1] Okamoto, K., Yanai, K. (2016, October). An automatic calorie estimation system of food images on a smartphone. In Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management (pp. 63-70)

[2] Koo, J., Suk, K. (2016). The effect of package shape on calorie estimation. International Journal of Research in Marketing, 33(4), 856-867

[3] Ege, T., Yanai, K. (2018). Image-based food calorie estimation using recipe information. IEICE TRANSACTIONS on Information and Systems, 101(5), 1333-1341.

[4] Sathish, S., Ashwin, S., Quadir, M. A., Pavithra, L. K. (2022). Analysis of Convolutional Neural Networks on Indian food detection and estimation of calories. Materials Today: Proceedings.

[5] McCrickerd, K., Tang, C. S., Forde, C. G. (2020). The independent and combined impact of front-of-pack labelling and sensory quality on calorie estimations and portion selection of commercial food products. Food Quality and Preference, 79, 103766.

[6] Jiang, Y., Lei, J. (2014). The effect of food toppings on calorie estimation and consumption. Journal of Consumer Psychology, 24(1), 63-69.

[7] Turmchokkasam, S., Chamnongthai, K. (2018). The design and implementation of an ingredient-based food calorie estimation system using nutrition knowledge and fusion of brightness and heat information. IEEE Access, 6, 46863-46876.

[8] Miyazaki, T., de Silva, G. C., Aizawa, K. (2011, December). Image-based calorie content estimation for dietary assessment. In 2011 IEEE International Symposium on Multimedia (pp. 363-368). IEEE.

[9] Kaneda, K., Ooba, T., Shimada, H., Shiku, O., Teshima, Y. (2021, October). Estimation method of calorie intake by deep learning using depth images obtained through a single camera smartphone. In 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE) (pp. 874-875). IEEE.

[10] Kaneda, K., Inoguchi, S., Shimada, H., Shiku, O., Teshima,Y. (2020, October). An estimation method of calorie intake using photograph and depth image of dishes in medical facilities. In 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE) (pp. 572-574). IEEE.

[11] Hu, H., Zhang, Z., Song, Y. (2020, December). Image Based Food Calories Estimation Using Various Models of Machine Learning. In 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE) (pp. 1874- 1878). IEEE.

[12] McAllister, P., Moorhead, A., Bond, R., Zheng, H. (2017, November). Automated adjustment of crowdsourced calorie estimations for accurate food image logging. In 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1059-1066). IEEE.

[13] Hu, H., Zhang, Q., Chen, Y. (2022). NIRSCAM: A Mobile Near-Infrared Sensing System for Food Calorie Estimation. IEEE Internet of Things Journal.

[14] R. Prasanna Kumar, Rithani M., ”A Review on Big data based Deep Neural Network Reinforcement Learning Algorithms”, IAES International Journal of Artificial Intelligence, 2023. (Accepted)

[15] V. B. Kasyap and N. Jayapandian, ”Food Calorie Estimation using Convolutional Neural Network,” 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 2021, pp. 666-670, doi: 10.1109/ICSPC51351.2021.9451812