Hybrid Deep Learning Algorithm Based Food Recognition and Calorie Estimation

V.Shanmugapriya [1], S.keertiga [2], G.Jothika [3], A.vigneshwari [4], W.Newton David Raj[5]

[1,2,3,4]U.G Student, [5]Assistant Professor

Department of ECE, Kings College of Engineering, Punalkullam ,Thanjavur, Tamilnadu

vpriya13122002@gmail.com, skeertiga007@gmail.com

ABSTRACT

Every individual requires some sort of system that informs them about portions and calories of food, as well as providing them with directions on how to consume it. In our study, we propose a hybrid architecture that makes use of deep learning algorithms to forecast the number of calories in various food items on a bowl. This consists of three major components: segmentation, classification, and calculating the volume and calories of food items. When we use a Mask RCNN, the images are first segmented. Using the YOLO V8 framework, features are collected from the segmented images and the food item is categorized. In order to determine the dimensions of each food item, we identify the items first. In order to calculate the quantity of the food item, the estimated dimension must be used. The calories are then computed using the food item's volume. The aforementioned approaches, which were trained on the dataset's food images, that correctly identified and forecasted a food item's calories had an accuracy of 97.12%. To Provide directions on how to consume food is expected by individual and will be completed after knowing intake of volume of food.

INTRODUCTION

A lot of individuals enjoy consuming fast food and soft drinks, which are high in calories and sugar. Obesity levels are rising as a result of people getting less exercise, not knowing enough about nutrition, and having uncontrollable eating habits. Obesity-related conditions include hypertension, diabetes, heart problems, and breathing difficulties. Due to excess body weight, obesity leads to persons having knee or other joint ligament injuries. Using the recent technologies, a number of methods were created to track daily dietary consumption, where simply taking a photo of the meal was sufficient. To identify the calories present in a specific cuisine,these applications employ image processing techniques. Most existing technologies are inaccurate in recognizing food items, let alone estimating total calories. To address this issue, we describe in this paper a hybrid framework for food identification and calorie estimates that combines deep learning methods such as Mask RCNN and YOLO V5.A hybrid system for deep learning, in the context of managing dietary choices and calorie intake, combines the strengths of deep learning with other techniques or data sources to enhance its capabilities.

BLOCKDIAGRAM

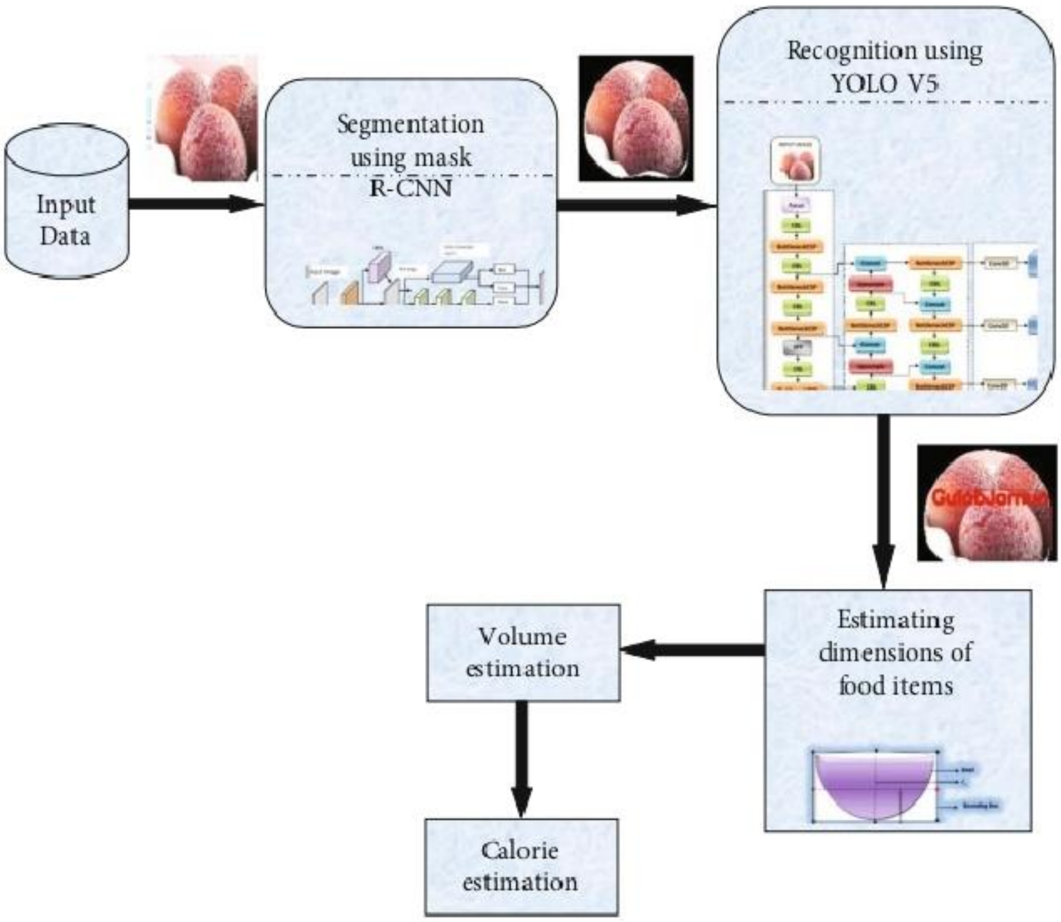

Figure 1: The overview of the proposed scheme.

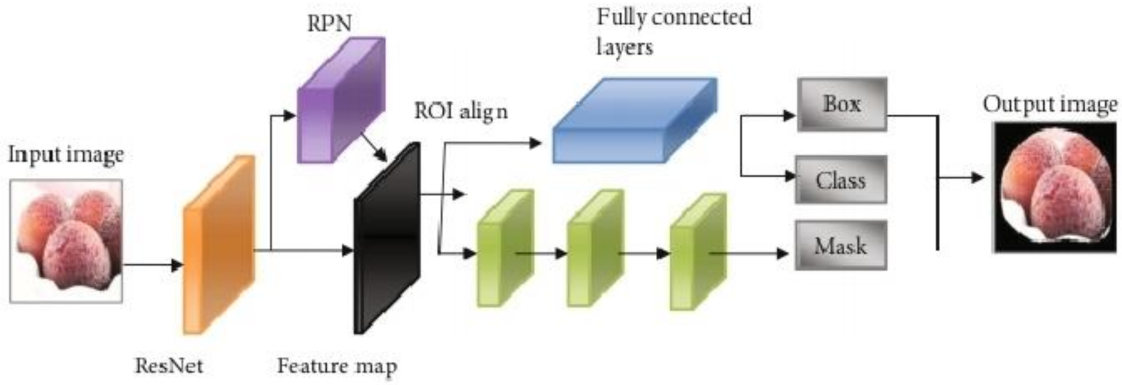

FIGURE 2: Architecture of Mask RCNN.

MASK RCNN

The data set is used to analyze the proposed scheme. After the preparation of the dataset, the images are sent into Mask RCNN for segmentation. The Mask RCNN is a deeplearning driven technique for detecting objects and segmenting instances. Instance segmentation gives different masks for different instances of the same class of objects. The components in Mask RCNN are explained as follows,

Resnet-FPN

ResNet-FPN is a backbone architecture used for feature extraction in Mask RCNN. A deep CNN called ResNet has proven to be quite successful in classifying images.

RPN

The RPN network receives the features that were extracted using ResNet101+FPN as input to produce ROIs. RPN may predict the front and back of an image if the fold length and breadth ratios are different.

ROI Align

ROI Align takes a set of rectangular region proposals as input and extracts feature from the feature map that correspond to each proposal.

YOLO V5

After segmentation, the segmented images are sent into the YOLO V5 classifier for further classification. Features are retrieved from those segmented images, and the food items are then identified using the YOLO V5 network. Redmon J. presented the YOLO object identification technique as the first singlestage object detection system. After feeding the full image into the network, the essential premise of YOLO is to immediately reverse the positions and types of the bounding boxes on the output layer, which is equivalent to Faster RCNN. The YOLO is made up of three major components:

- **Backbone:** Convolutional neural network that gathers and produces image characteristics based on image resolution.
- **Neck:** Group of network layers that aggregate and blend picture features before passing them on to the prediction layer.
- **Head:** In addition to producing bounding boxes and predicting categories, it can predict image attributes.

VOLUME ESTIMATION

The volume of the item will be measured following that the ROI has been clearly designated and the object has been classified. The volume of irregular food items is estimated by the pixel-per-meter approach. This approach determines the reference object's width in pixels. Using this ratio, the object's height along with breadth is further calculated. Theestimated volume (VolE) of the bowl is then determined.

CALORIE ESTIMATION

Calories can be computed when the volume has been determined. For both cooked and uncooked meals, there are standard calorie mapping charts available. Volume is transformed into grams and linked to the calorie table in order to determine the exact number of calories in food item that our system has identified. Here, CE is the calorific equivalent, and total calorie (Calorie total) is calculated by using the below

formula:

Calorie total = volume ∗ CE

Thus, the calorie is estimated using the volume of the food item.

PPERFORMANCE ANALYSIS

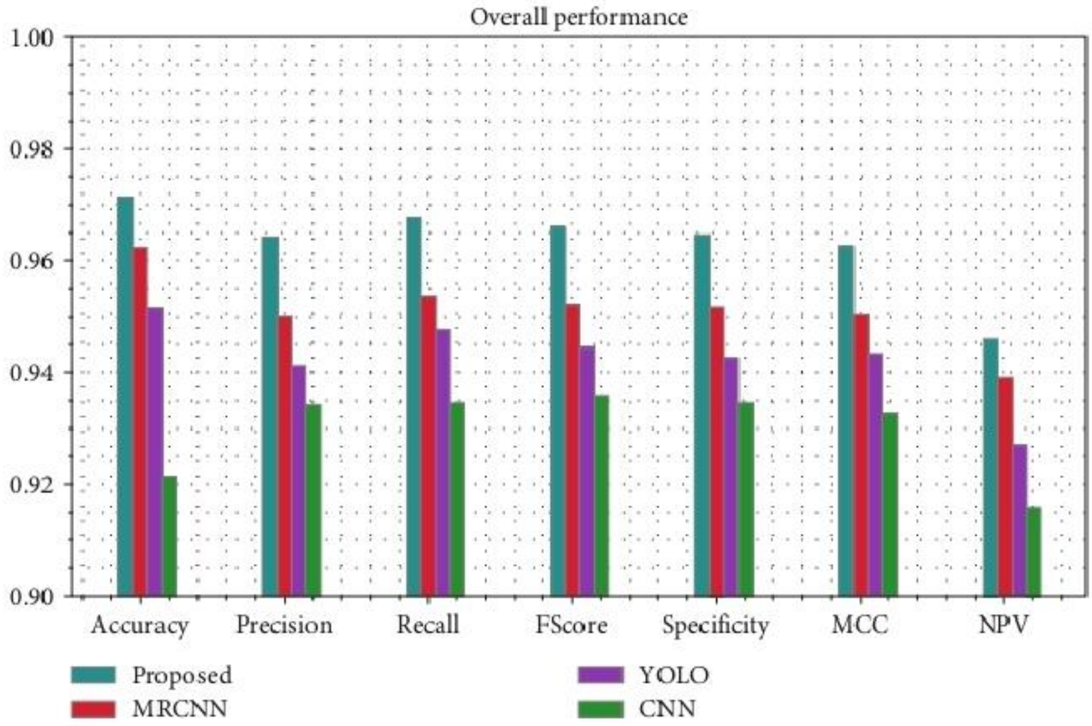

Fig 3: Analysis of already existing approaches with proposed methods

The performance of the proposed method is analysed in terms of accuracy, sensitivity, specificity, precision, recall, F-score, MCC, NPV, FPR, FDR, and FNR. Moreover, the efficiency of the proposed Mask RCNN+YOLOV5 is compared and analysed with other classification models such as Mask RCNN, YOLO, CNN.

CONCLUSION

In this study, a hybrid framework using deep learning algorithms to estimate the calorie content of food items has been proposed. Initially, the images were segmented using Mask

RCNN. From the segmented images, features were extracted,and the food item is classified using YOLO V5 framework. After recognizing the food items, dimension of each food item was determined. Based on the estimated dimension, food item's volume was calculated. Then, the calorie was estimated using the food item's volume. The efficiency of the proposed approach has been evaluated and compared with other classification models such as CNN, YOLO, and Mask RCNN using a variety of criteria, accuracy, NPV, sensitivity, specificity, FDR, precision, FPR, recall, F-score, MCC, and FNR. From those evaluation metrics, the proposed method was found to be very high with 97.12% accuracy compared to other models.

REFERENCES

1. P. Kadam, S. Pandya, S. Phansalkar et al., "FVEstimator: a novel food volume estimator wellness model for calorie measurement and healthy living," Measurement, vol. 198, article 111294, 2022.
2. R. D. Kumar, E. G. Julie, Y. H. Robinson, S. Vimal, and S. Seo, "Recognition of food type and calorie estimation using neural network," The Journal of Supercomputing, vol. 77, no. 8,pp. 8172–8193, 2021.
3. S. A. Ayon, C. Z. Mashrafi, A. B. Yousuf, F. Hossain, and M. I.Hossain, "FoodieCal: a convolutional neural network based food detection and calorie estimation system," in 2021 National Computing Colleges Conference (NCCC), pp. 1–6,Taif, Saudi Arabia, 2021.
4. Z. Shen, A. Shehzad, S. Chen, H. Sun, and J. Liu, "Machine learning based approach on food recognition and nutrition estimation," Procedia Computer Science, vol. 174, pp. 448–453, 2020.
5. M. A. Subhi, S. H. Ali, and M. A. Mohammed, "Vision-based approaches for automatic food recognition and dietary assessment: a survey," IEEE Access, vol. 7, pp. 35370–35381, 2019.