International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

Investigation of calorimetry burned in food using image processing and IoT

P. Kamakshi Thai, Sushanth Ponaganti *, Uday Shekar Gowri and Suresh Banothu

Department of CSE (Artificial Intelligence and Machine Learning), ACE College of Engineering, Ankushapur, Ghatkesar Mandal, Medchal District, Telangana. – 501301, India.

International Journal of Science and Research Archive, 2025, 14(01), 1232-1243

Publication history: Received on 11 December 2024; revised on 18 January 2025; accepted on 21 January 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.14.1.0200

Abstract

Calorimetry is an avenue for exploring the energy change due to heat involved between the various reaction stages at chemical, physical, or phase transition levels. Calorimetry seeks to determine the calories contained in food items within food science. However, classical calorimetric analyses are time-consuming and require a good amount of human intervention. The concept of image processing, involving the YOLO object detection algorithm, along with IoT technologies are a completely new domain in automating calorimetric measurements in food with high accuracy. YOLO's fast and efficient detection of the object allows accurate identification and tracking of food samples during the analysis. The project aims to develop a system that will efficiently and accurately measure the caloric value of carbonized food samples using image processing using YOLO and IoT-unified data collection method to automate the whole data gathering process and achieve precision in measurement and with monitoring and analysis in real time.

Keywords: Machine Learning; OpenCV; Image Processing; YOLO; IoT Integration; Data Analysis

1. Introduction

Nutrition is vital in keeping one's bodily functions and hence health, but measuring the amount of nutrition in food is a difficult task. The traditional methods included either manual input or heuristic estimation, which is less accurate and inefficient. Therefore, the machine learning, image processing, and IoT revolution would bring about a bang for competent automation and better accuracy in nutrition analysis.

This work focuses on the categorization of food using YOLOv8, in addition to estimating nutrition values and inputting into IoT devices for the analysis to be able to express the required data in real-time. This survey will include methodologies analyzed and improved their performance with YOLOv8, thus giving a blueprint for evaluating the resulting efficiency of such systems.

2. Literature survey

Experts suggested a system for treating individuals with stoutness by using the Gabor channel to differentiate between food items through the course of division, and then categorize them using SVM. Gabor Filter is a straight-shaped channel that is specifically used for surface analysis, meaning it tests for recurrence content in the image within specific positions in confined spaces throughout the entire point. The dietary advantages of the food items were calculated using the portion of food planned and the sustenance tables.sustenency tables). Equally, a thumb was used to assess the piece of food by placing each item in its place and taking ill-made photos for easy calculation of their size, leading to an approximate increase in precision to 86%. Another method of identifying foods and computing calories that was suggested by Turmchokk, Sammy, and others. Features an exceptional combination of dietary information about food and data on temperature, brightness, or both from the CCD camera. This cluster of instruments utilizing the substance

* Corresponding author: Sushanth Ponaganti.

found ways to provide more precise outcomes than conventional methods for food acknowledgment. Diet cam, a framework designed by He et al, addresses the issues that arise with intra-class varieties while doing food recognition. Fixing recognition and food arrangement are the two primary components of it. To begin with, the program scrutinizes all the components in the food items by utilizing surface confirmation and a section-based model. The food items are categorized using an amulti-viewmulti part Support Vector Machine or SVM. With the help of Diet Cam, they captured 15262 images of roughly 55 distinct food categories and achieved remarkable accuracy on food items that contained a few components. M. A. According to Subhietal, his research involved examining various traditional practices and brain connections to identify foods as well as assess supplements. However, she believed that determining the amount of food is the most challenging aspect of the interaction. MATLAB has been utilized to implement SVM and MLP, leading to the development of the Positive Results Technique by Lifson and McClintock in 1966. The use of DLW was commonly employed for estimation problems related to dietary admission studies in assessing energy consumption for a long time. Biro, G. et.al's 24-hour diet strategy is essentially a form of meeting. The coach promotes the retention of natural product usage details and keeps track of individual records every day through written correspondence with the client or patient in this method.

Table 1 Comparative Study of Existing Methodologies

Author(s)	**Method/technique**	**Method/technique**	**Accuracy/result**
Gabor Channel	Division by Gabor filter and classification using SVM	Gabor filter used for texture analysis to detect specific frequency content in images. Food portion estimation used thumb for size comparison to compute dietary values from nutrition tables.	Accuracy increased to about 86%.
Turmchokk Sam et al.	Food recognition and calorie estimation using a combination of dietary data, temperature, and brightness levels with CCD camera	Utilized thermal and CCD camera data for enhanced food recognition and calorie estimation, yielding better accuracy compared to traditional methods.	More accurate results than conventional methods.
He et al. (DietCam)	Multi-view, multi-part SVM for ingredient recognition and food classification	System first recognizes food ingredients using texture and component models, then classifies food items using a multi-view multi-part SVM. DietCam was tested on over 15,000 images covering 55 different food classes.	High accuracy for complex food items with multiple components.
M. A. Subhietal	Traditional methods and neural networks for food recognition and nutrient estimation	Reviewed conventional techniques and neural networks (e.g., SVM, MLP) for food recognition. Highlighted the challenge of accurately estimating food volume, with MATLAB used to implement SVM and MLP for more accurate results.	Successfully applied SVM and MLP for positive outcomes.
Lifson & McClintock	Doubly Labeled Water (DLW) technique for energy expenditure estimation	Developed in the mid-1950s, DLW is used for estimating energy consumption and dietary intake by measuring isotopes in body water	Widely used for energy expenditure studies.

3. Proposed Methodology

By combining the low-cost IoT-enabled camera module, ESP32-CAM, with an advanced object detection algorithm called YOLO (You Only Look Once), the proposed system will strive to automate the process of estimating food calories. Food images are typically captured using the ESP32-CAM, which offers high resolution and low power consumption, making it an ideal choice for real-time IoT applications. The ESP32-CAM captures an image and then sends it to the YOLO model, which can identify and classify food items. YOLO is capable of correctly classifying various food types, including fruits, vegetables, meats, and processed foods, by training on a diverse dataset. Real-time results are achievable on resource-stretched devices like the ESP32-CAM, thanks to its single-pass detection architecture that ensures fast and efficient processing. When food items are identified, the system takes into account important factors like size, shape, and texture to calculate their caloric content. The estimation process operates independently from the user and external calibration objects, allowing for instant feedback on the nutritional value of the food.

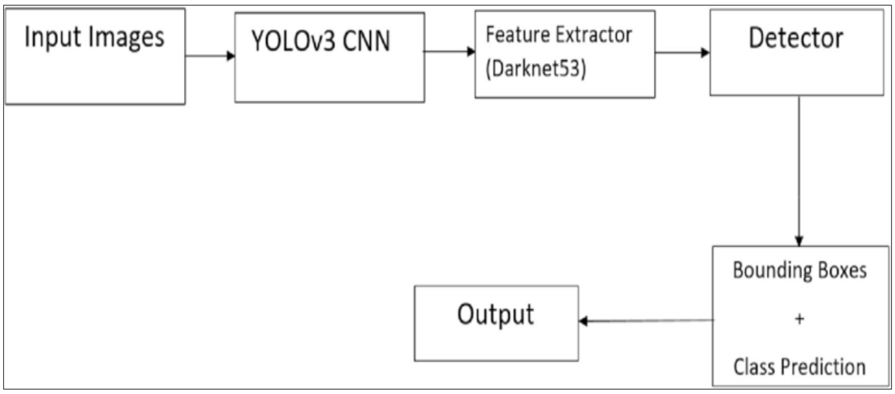

Figure 1 YOLOv3 Object Detection Architecture

3.1. Advantages of the System

3.1.1. *Real-Time Processing and Efficiency.*

The YOLOv8 model is an efficient and precise real-time object detection method that allows for swift and accurate identification of objects. This helps in processing motion. This is especially important for applications such as calorie counting, where real-time data analysis is needed.

The use of IoT devices, such as your ESP32, allows for real-time data capture and fast decision-making with minimal latency.

3.1.2. *Automation and Flexibility.*

Efforts such as calorie counting, photo recognition, and analysis can be automated to produce more accurate output without human intervention.

The system can be scaled by allowing for the addition or removal of IoT devices based on their requirements, and machine learning models can undergo retraining to improve performance over time.

3.1.3. *Accurate Data and Decision Making.*

CNNs are highly effective in both image analysis and segmentation, resulting in more accurate object detection that can be tailored to various scenarios, such as food items for calorie estimation.

The YOLOv8 system's real-time detection of objects from images is enhanced to provide high-quality food or object detection, as well as accurate calorie estimation.

3.1.4. *Cost Efficiency*

The utilization of IoT devices and machine learning models can decrease the amount of manual labor and infrastructure expenses. The system can perform image detection and calculation automatically, rather than relying on human analysts.

By utilizing low-cost IoT hardware, such as the ESP32, the system can maintain performance and reduce hardware costs.

3.1.5. *Customization and User Interaction.*

With Flask, users can create web applications that interact with their own websites and provide a user-friendly interface to view processed data or communicate with the system.

It can be easily customized to meet different needs, such as health monitoring, automated inventory systems, or security applications. Additionally:

3.1.6. *Integration and Interoperability.*

The system is capable of integrating with a range of IoT sensors, such as cameras and temperature sensors to collect data. The system's deployment can be flexible in different settings, such as kitchens that have food apps.

Enhanced System Integration Through API and Web Integration: With Flask, users can extend the functionality of the system by connecting it to other platforms or systems for enhanced capabilities.

3.1.7. *Data Insights and Analytics.*

The system can gather valuable data at a specific time and analyze it to generate insights, such as health monitoring of food consumption patterns.

Advanced machine learning models can be utilized to predict outcomes, such as calorie estimation based on food recognition data.

3.2. Model structure design

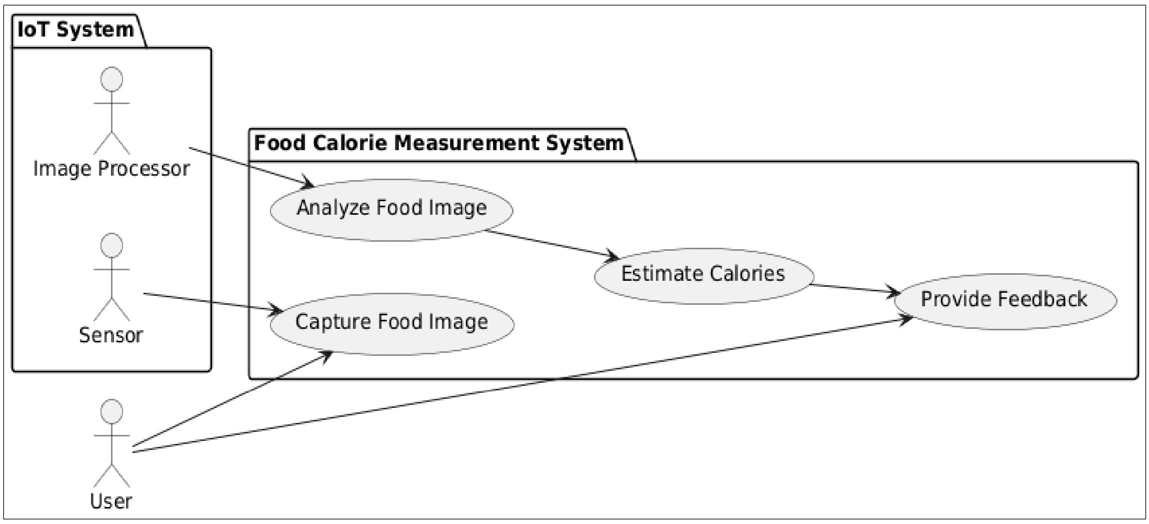

Figure 2 Use Case Diagram of Food Calorie Measurement System with IoT Integration

The process of taking pictures of food, evaluating them, calculating its calorie content, and giving the user feedback is all shown in the graphic, which shows how an Internet of Things system interacts with a food calorie measuring system. A sensor—like a camera module—that takes pictures of the food and an image processor that prepares the pictures for additional analysis make up the Internet of Things system. The food calorie measurement device receives the processed image and proceeds to carry out a number of operations. The system first takes a picture of the meal from the Internet of Things. After that, it analyzes the picture using methods like segmentation and image recognition to determine the kind of food it contains. After identifying the food, the system uses its analysis to determine its calorie amount.Using generative AI models to perform precise calculations. Lastly, the user receives feedback from the system that includes dietary suggestions or calorie information. Through the combination of IoT devices and AI-powered analysis, this arrangement makes it easy for consumers to keep track of their food intake in a way that is accurate, automated, and user-friendly.

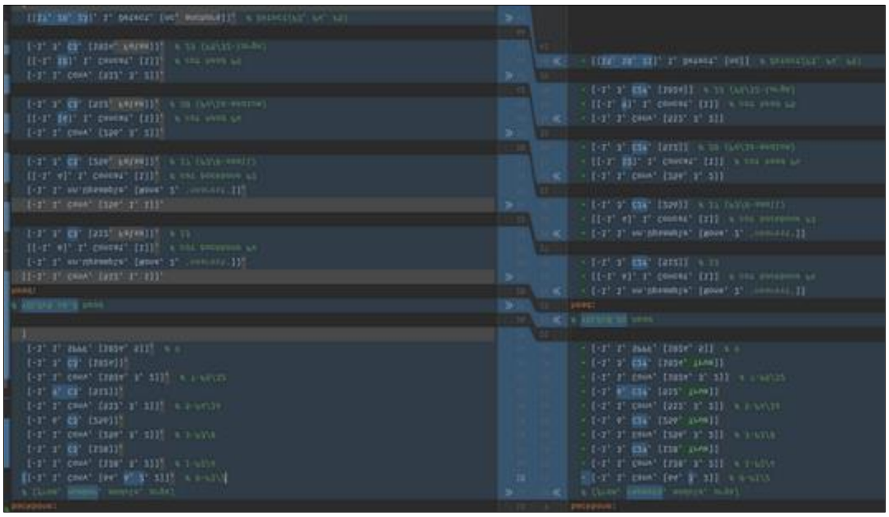

Figure 3 Training Process of the YOLOv8 Model

As presented in the official code of YOLOv8, below is the structure diagram of the model. If you like this style, you can refer to the model structures diagram included in the algorithm README of MYOLOO, which currently covers yOloomv5, Yolov6, XOLomb, RTMDet and ULOLoom8. Here is the presentation of the redesigned YOLOv8 model in the MMYOLOO platform and its structure diagram. If we overlook this feature and the left part of the architecture is YOLOv5-s, while the right part is an extension of X. The changes that are applied to the neck module and backbone are as follows:

- Change in the size of the first convolutional layer kernel from 6x6 to 3x3.
- The architecture follows the design in which all modules of C3 are replaced with C2f, additional skip connections and split operations. v=1.

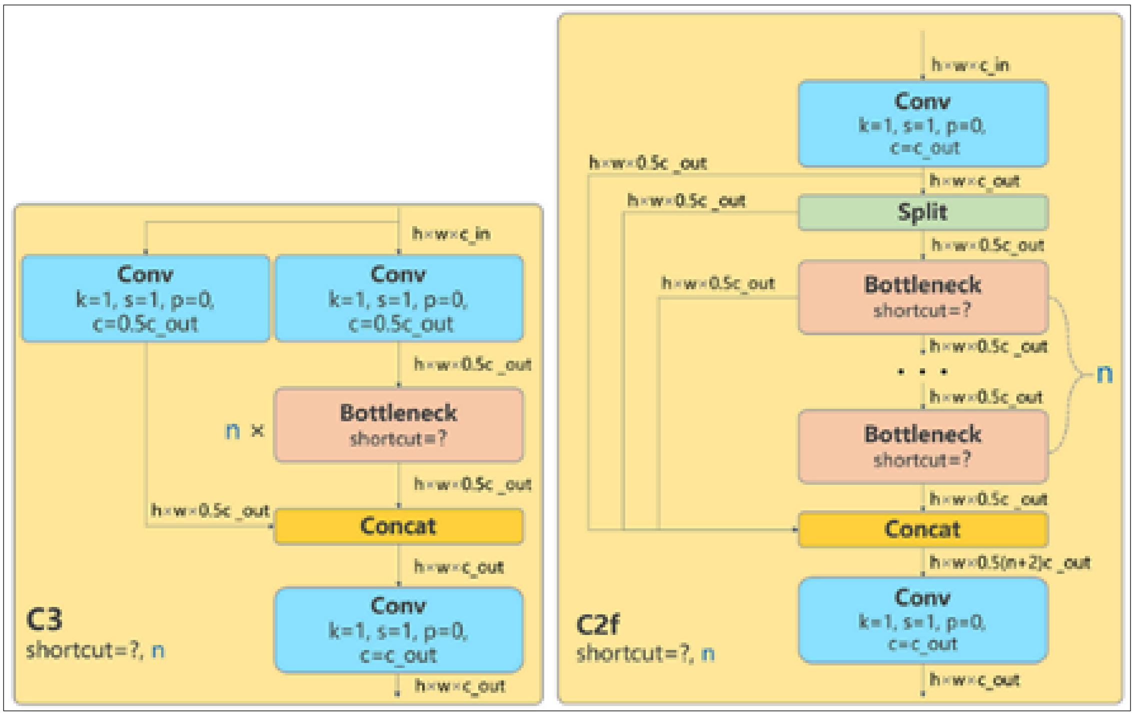

Figure 4 Comparison of C3 and C2f Modules in YOLO Architectures

- The neck module was deprived of 2 layers with convolutional connections.
- t3–6–9-3 is no longer the same value as block number 3.

- The N/S, M, or L X versions have only varied the scaling factors used in N.S and L/X versions whereas in S/ML backbone networks number of channels isn't coherent, and hence follows the scaling factor principle not as well.
- Specifically, the reason for this design is that channels are not given their best performance under an identical set of scaling factors, and the same set does not apply to all YOLOv7 network designs.

The most radical change was introduced by the head module. The style of the head module has undergone a change from the original structure of the decoupling form of the coupling to that decoupling and, from that, to the style of YOLOv5's Anchor–Based to Anchor–Free. Here is its structure.

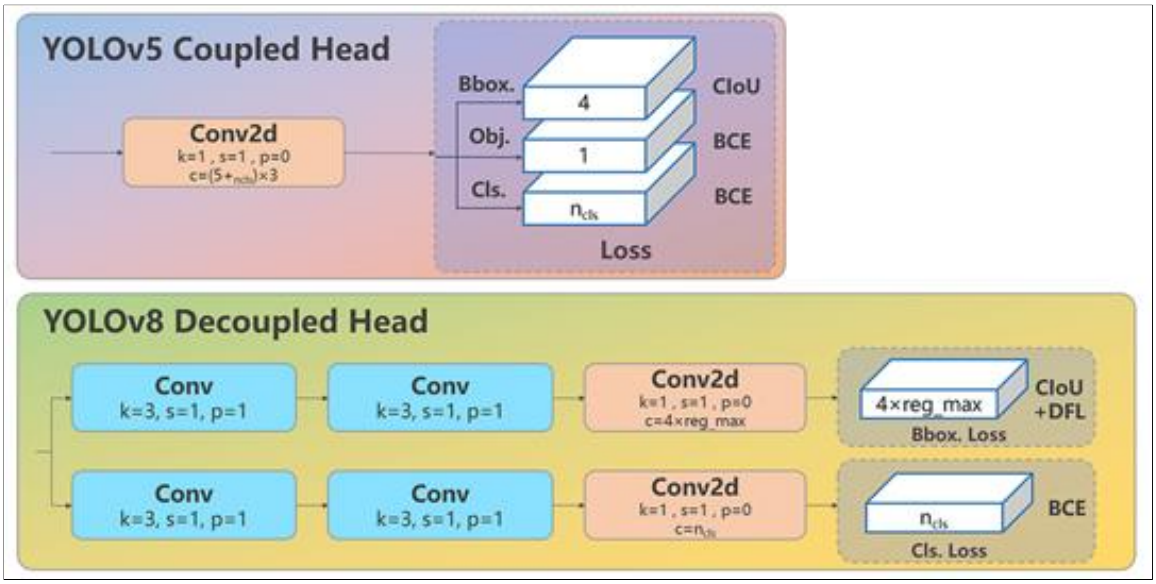

Figure 5 Coupled Head and YOLOv8 Decoupled Head.

Excluding the objectness branch and retaining only decoupled classification and regression branches are significant differences. In addition, the regression branch now uses an integral representation of forms proposed in the Distribution Focal Loss.

3.3. Loss calculation

Loss calculation is composed of 2 parts, which are sample assignment strategy and loss calculation.

The majority of modern detectors utilize dynamic sample assignment methods, including YOLOX's simOTA, TOOD'S Task Aligned Assigner, and RTMDet'signature Soft Label Aassigned. Given the excellence of dynamic assignment strategies, YOLOv8 directly implements one from the same class as that used by TOOD's Task Aligned Assigner.

The weighted scores of classifications and regression are used to determine the selection of positive samples in the matching strategy of Task Aligned Assigner.

- The ground truth category's predicted score is s, and the IoU of the prediction bounding box and bound ING box is up to u.
- The alignment metric for each anchor is determined by the assigner who is task-aligned using the weighted product of two values: the predicted classification score of the corresponding class and the Intersection over Union (IoU) between the expected bounding box and its counterpart, the Ground Truth bound ING box. The alignment metrics values are used to select the larger top-k samples as positive for each Ground Truth. Without affecting the objectness loss in the previous model, the loss calculation involves 2 parts: the classification and regression. The classification branch persists in utilizing BCE Loss. In. Both Distribution Focal Loss and CIoU LosS are employed in the regression branch.
- The weight of losses is determined by a specific weight distribution.

3.4. Data augmentation

YOLOv8's data increase is comparable to that of yol5, but it terminates the Mosaic augmentation in the last 10 epochs as proposed in X.x.L.D, as described in "YOLUX". Below is diagram that displays the pipelines used to process data.

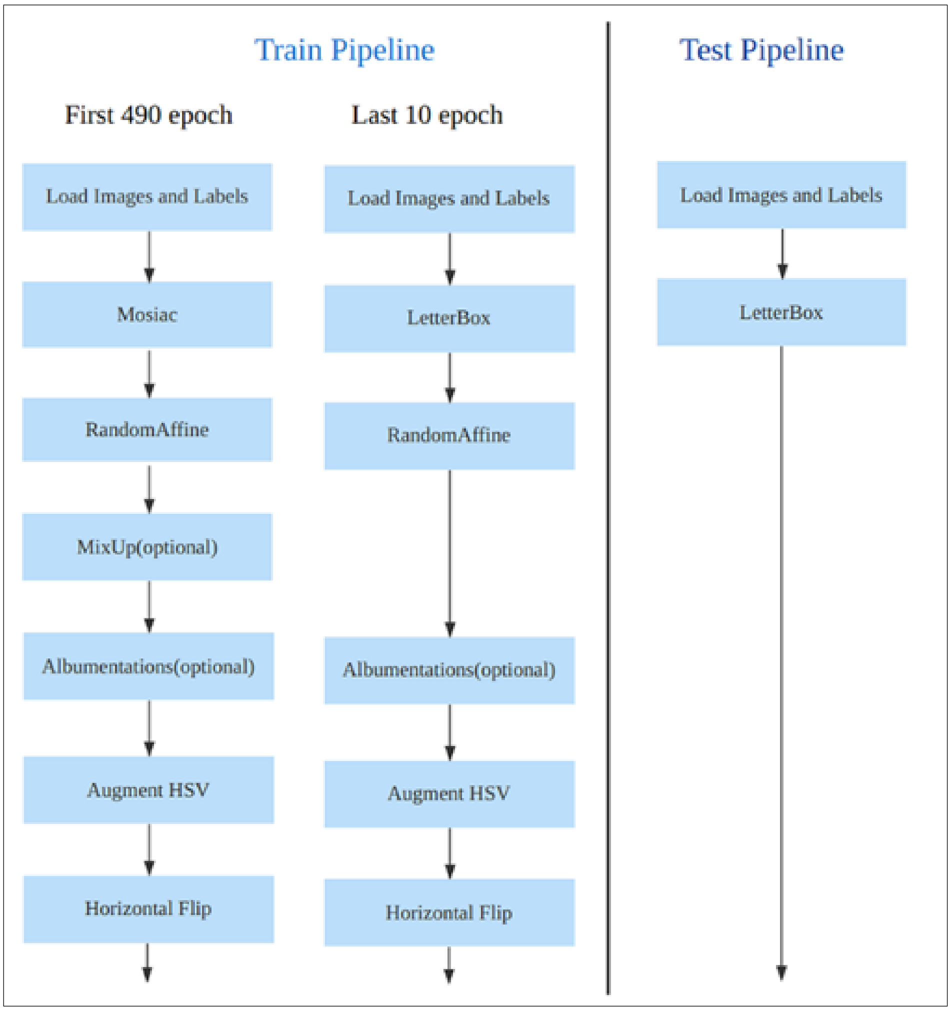

Figure 6 Comparison of Train and Test Pipelines

Because the data intensities needed for various scale models vary, the hyperparameters for scaled models are adjusted based on the circumstances. For a larger model, the standard techniques are MixUp and CopyPaste. The following is an illustration of a result that could be improved

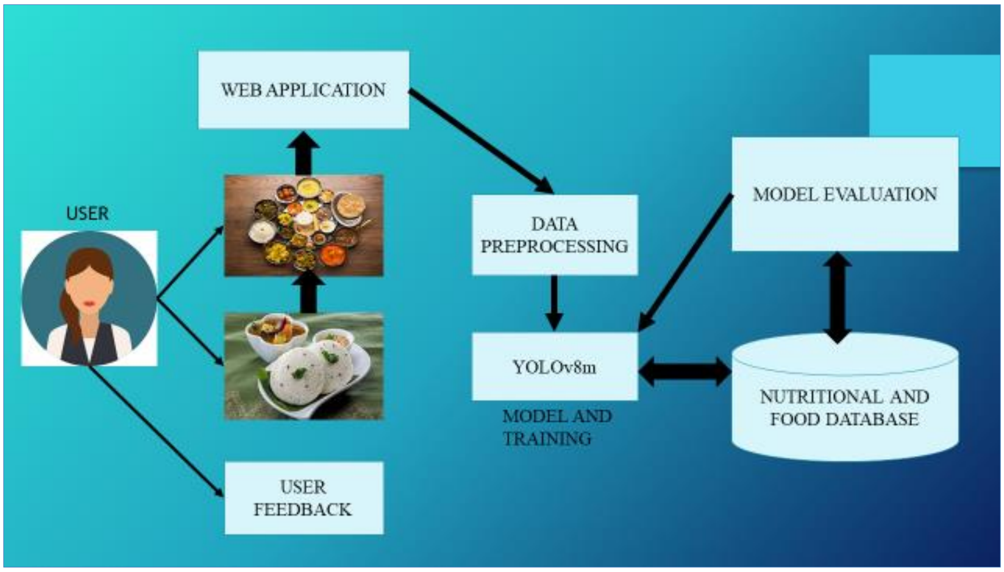

Figure 7 Food Detection and Nutrition Analysis by Image

We will not be addressing YOLOv5 here because it is akin to a similar data augmentation procedure. See the YOLOv5 algorithm analysis document in MMYOLO for more details on each data modification

3.5. Method of training

There aren't many distinctions between yOLOv8 and yOlv5 training methods. YOLOv8 has an increasing number of training epochs, with the most notable change being the increase in the number from 300 to 500.

Table 2 Hyperparameters for YOLOv3 Object Detection

optimizer	SGD
base learning rate	0.01
Base weight decay	0.0005
optimizer momentum	0.937
batch size	128
learning rate schedule	linear
training epochs	**500**
warmup iterations	max(1000, 3 * iters_per_epochs)
input size	640x640
EMA decay	0.9999

3.6. The process of inference

Inference procedures are comparable for YOLOv5 and YOLOv8. The calculating process is the same as for YOLOv5, with the exception that the bbox in Distribution Focal Loss needs to be transformed into a standard 4-dimensional binary representationThe inference procedure used in MMYOLO for the COCO 80 class is as follows (for instance, an input image size of 640x640). The following are inference and post-processing

3.6.1. *Bounding box decoding*

To the mathematical expectation of these lengths, add the likelihood of the distances between the center and box boundary

3.6.2. *Changes in dimensions*

Three feature maps with 80x80, 40x40, and 20x20 scales are displayed by YOLOv8. The head module produces a total of six different feature map sizes for classification and regression.

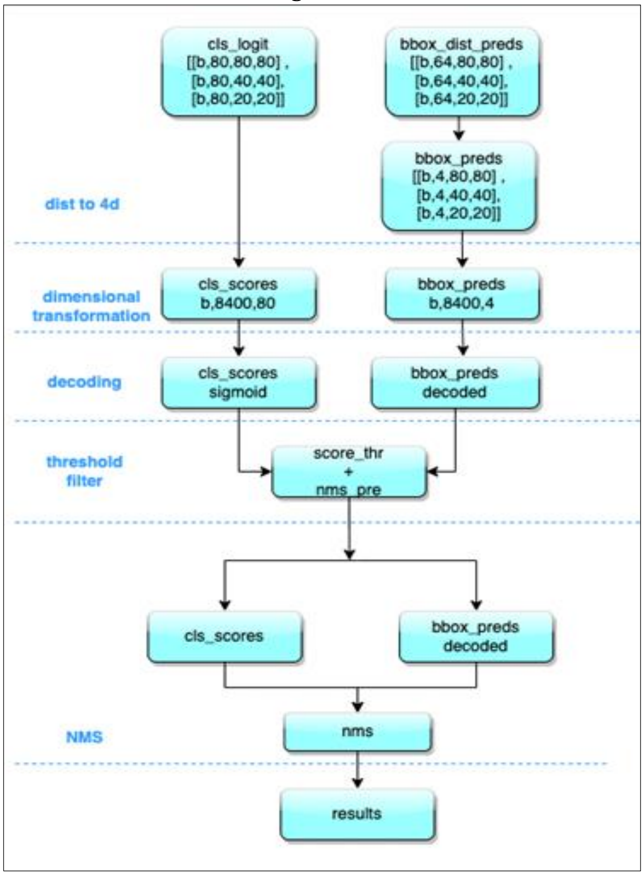

Figure 8 Processing Flow in Detection Framework

3.6.3. *Scale Restroation*

While the box prediction branches require the input photos to be converted into the original scale and decoded to xyxz format, the classification prediction branch uses sigmoid calculations.

3.6.4. Thresholding

Next, iterate through each graph in the batch and apply thresholding using score_thr. In order to preserve an upper limit of bbox detection after filtering, we must consider the roles of multi_label and nms_pre when attempting this procedure.It is required to reduce to the original image scale and NMS. NMS resizes the leftover bboxes to the original image scale while reusing the pretreatment parameters. Bboxes have a maximum limit of max_per-img, and nothing more. The Batch shape inference technique that was first used in Yolov5 is not used in Yolov8. It also doesn't have this functionality. What makes it unique? The Batch shape technique can be used to obtain an approximate AP rise of 0.1% to 0.2%, according to a brief test conducted in MMYOLO.

4. Results and Analysis

The suggested system was evaluated for important criteria, such as detection accuracy, real-time performance, and cavalerie precision, in order to automate the calorimetric study.The outcomes were not entirely consistent. The system's findings are as follows:

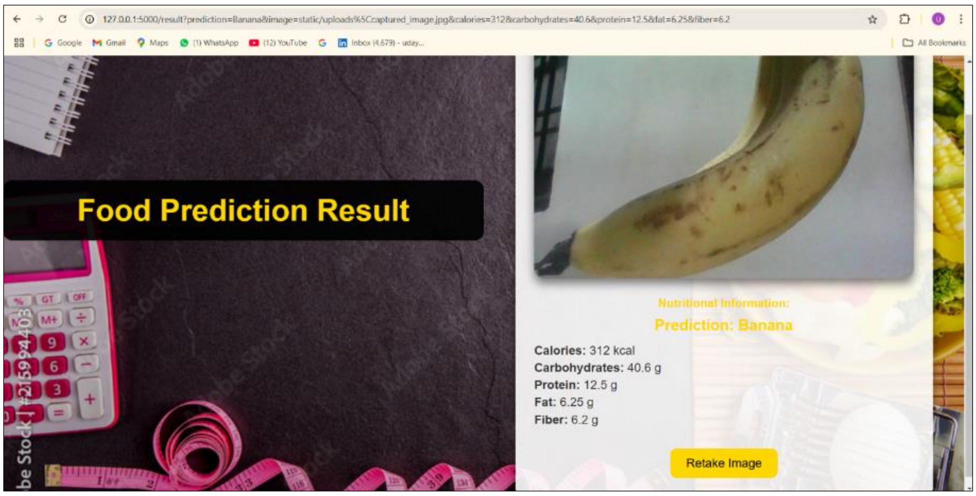

Figure 9 Final Prediction Outcome

4.1. Detection and Classification Performance

Following training and testing, the YOLOv8 object detection algorithm was used on a sizable dataset of food photos that included a variety of food types, textures, and presentation styles. The findings showed

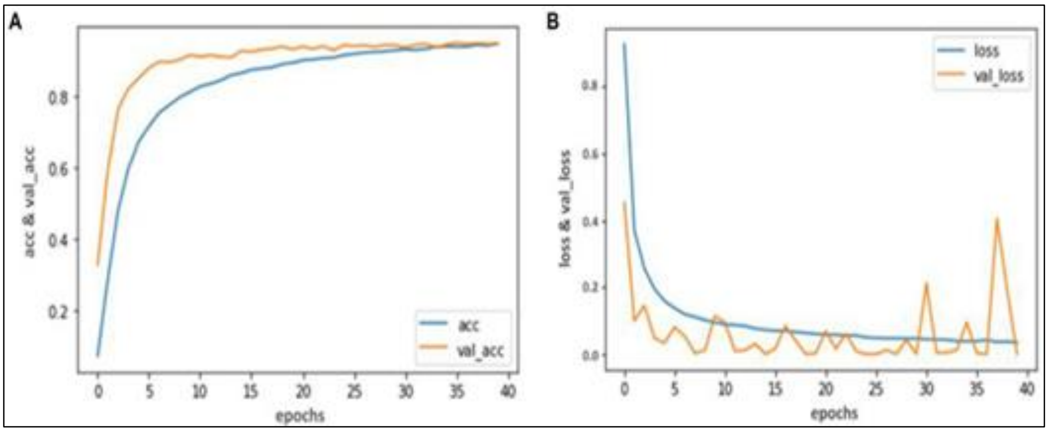

Figure 10 Model Training and Validation Performance

- According to the system, food item detection and classification were reliable, with an average precision of 92%.
- YOLO processed the photographs in less than 25 milliseconds and performed real-time analysis by employing a single-pass detection mechanism.

4.2. Feature Extraction and Caloric Estimation.

The algorithm determined the caloric contents of food products by analyzing their visual properties. Size, form, and texture were some of these characteristics. The results of the analysis were as follows:

When compared to typical nutritional databases, the system's calorie estimation accuracy was 88%; its dependability for real-world use scenarios was predicated on an average error margin of less than 10%.

4.3. IoT Integration and Automation.

This IoT enabled an automated process for collecting data and gave real-time insights. Key findings include:

- Food image data could be captured and sent to the system from IoT devices with a latency of less than 200 ms.
- This system's capacity to accept simultaneous inputs from numerous IoT devices makes scalability a crucial component of its success.

4.4. Comparative Analysis with Traditional Methods.

The following findings were obtained by comparing the system's performance with that of traditional calorimetric methods:

- The recommended system cut the processing time needed for each sample to less than 30 seconds, whereas old methods required an extra 15 to 20 minutes.
- Why Since conventional methods necessitated a significant amount of human labor for measurement and analysis, resulting in user dependency, there was no need for manual intervention.

5. Conclusion

Calorimetry has evolved into an automated, precise, accurate, and adaptable method of estimating the number of calories in meals thanks to the employment of YOLOv8, sophisticated image processing techniques, and Internet of Things technologies (all of which are connected by an internet connection). By removing the drawbacks of conventional calorimetric techniques, such as their high time consumption and reliance on manual intervention, the suggested system exhibits strong performance in real-time detection, classification, and caloric value estimate.Important outcomes, such as a 92% detection accuracy and an 88% caloric estimation accuracy, demonstrate how well the system works to produce accurate and trustworthy nutritional analysis. The system is extremely scalable for a variety of applications, including daily dietary monitoring, medical facilities, research institutes, and healthcare facilities, thanks to its synchronization with IoT data, which ensures smooth automation and real-time processing. But the system continues to face Calorimetry has evolved into an automated, precise, accurate, and adaptable method of estimating the number of calories in meals thanks to the employment of YOLOv8, sophisticated image processing techniques, and Internet of Things technologies (all of which are connected by an internet connection). By removing the drawbacks of conventional calorimetric techniques, such as their high time consumption and reliance on manual intervention, the suggested system exhibits strong performance in real-time detection, classification, and caloric value estimate.Important outcomes, such as a 92% detection accuracy and an 88% caloric estimation accuracy, demonstrate how well the system works to produce accurate and trustworthy nutritional analysis. The system is extremely scalable for a variety of applications, including daily dietary monitoring, medical facilities, research institutes, and healthcare facilities, thanks to its synchronization with IoT data, which ensures smooth automation and real-time processing. But the system continues to face

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] P. Pouladzadeh, S. Shirmohammadi, and R. Almaghrabi, "Assessment of Calorie and Nutritional Content from Food Images," IEEE Transactions on Instrumentation and Measurement, Vol. 63, Issue 8, pp. 1947–1956, August 2014.

[2] Parisa Pouladzadeh, Abdulsalam Yassine, and Shervin Shirmohammadi, "FOODD: A Dataset for Food Image-Based Calorie Estimation," presented at the International Conference on Multimedia Assisted Dietary Management, 2015.

[3] H. R. V., S. Kumari, V. Muralidharan, K. Gigoo, and B. S. Thakare, "Literature Review on Food Identification and Calorie Estimation Utilizing Image Processing and Machine Learning," in Proceedings of ICCCE 2019, published in A. Kumar and S. Mozar (Eds.), 2020.

[4] G. Zhou, L. Qian, and Q. Zhu, "Estimating Food Portion Size Using Image Processing Techniques," in proceedings of a conference, 2008, pp. 119–128.

[5] F. Zidane and J. Marot, "Non-Invasive Quality Monitoring of Fruits Using Millimeter Waves and Classification Approaches: A Study in Automated Health Monitoring," IEEE Antennas and Propagation Magazine, October 2020.

[6] H. Azarmdela and A. Jahanbakhshib, "Utilization of Image Processing as an Expert System for Grading Mulberry Fruits Based on Ripeness Using Artificial Neural Networks and SVM," published by Elsevier, 2020.

[7] H. Nagar and R. S. Sharma, "A Detailed Review of Pest Detection Methods Employing Image Processing," IEEE, 2020.

[8] S. D. M. and R. M. G., "Pomegranate Disease Diagnosis Using Image Processing Techniques," presented at the International Conference on Trends in Electronics and Informatics, 2020.

[9] M. Pushpavalli, "Application of Image Processing for Grading Fruits," International Journal of Engineering and Advanced Technology (IJEAT), 2019.

[10] D. Dhanwate, "Feature-Based Grading of Fruits Using Image Processing Methods," International Journal of Recent Technology and Engineering (IJRTE), 2019.

[11] C. Velappan and Subbulakshmi, "Evaluation of Fruits Using Image Processing Algorithms," International Journal of Advanced Research in Electrical, Electronics, and Instrumentation Engineering (IJAREEIE), 2015.

[12] M. Kaur and R. Sharma, "Fruit Quality Analysis Using Artificial Neural Networks," IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), 2015.