

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/379937519>

A Deep Learning Neural Network-based System for Food Recognition and Calorie Estimation

Conference Paper · December 2023

DOI: 10.1109/ICIMIA60377.2023.10426548

CITATION

1

READS

274

6 authors, including:



[Pallavi Yarde](#)

Sri Balaji University Pune

8 PUBLICATIONS 49 CITATIONS

[SEE PROFILE](#)



[Vipul Vekariya](#)

Parul University

61 PUBLICATIONS 138 CITATIONS

[SEE PROFILE](#)



[Natrayan L](#)

Saveetha University

516 PUBLICATIONS 12,949 CITATIONS

[SEE PROFILE](#)

A Deep Learning Neural Network-based System for Food Recognition and Calorie Estimation

Pallavi Yarde¹, Dibyhash Bordoloi², Rahul Mohan Chavan³, Vipul Vekariya⁴, Harshal Patil⁵, Natrayan L⁶

¹ Sr. Assistant Professor, Department of Technology & Management, Balaji Institute of Technology & Management, Sri Balaji University, Pune, Maharashtra, 411 033, India

pallavi.yarde@bitmpune.edu.in

² Associate Professor, Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India, 248002

dibyhashbordoloi@geu.ac.in

³ HOD, Department of Physical Education and Sports, Kholeshwar College Ambajogai, Maharashtra, 431517, India

rahulchavan045@gmail.com

⁴ Professor, Department of Computer Science and Engineering, Parul Institute of Engineering and Technology, Parul University, Post Limda 391760, Waghodia, Gujarat, India

vipul.vekariya18435@paruluniversity.ac.in

⁵ Associate Professor, Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, 412 115 India

harshal.patil@sitpune.edu.in

⁶ Assistant Professor, Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, 602105, India

natrayanphd@gmail.com

Abstract— The global increase in health awareness reflects a rising desire among individuals to pursue healthy and normal lives. Obesity has emerged as an important health issue in the modern world due to its rapid pace. Obesity is defined by medical authorities as having a Body Mass Index (BMI) of more than 30 kg/m², which can lead to a variety of health difficulties like liver failure, cholesterol, respiratory disorders, diabetes, cardiac complications, and, in serious cases, cancer. Maintaining a diet rich in nutrients and low in calories is critical in the fight against obesity. Individuals, on the other hand, frequently struggle with food decisions. In response to this difficulty, this research presents a system that offers calorific calculation followed by food recognition to aid in preventing obesity. The goal of this research is to identify types of food and estimate their calorific values. For food recognition, a novel Attention-based Dilated Convolutional Neural Network Logistic Regression (A-CNNLR) model is introduced and compared to a traditional CNN model. The proposed A-CNNLR model demonstrates a recognition rate of 99%, surpassing the CNN model's 97%. The study determines the calories in the food according to their volume using data from the ECUST Food Dataset. This study provides a novel approach to the fight against obesity by using sophisticated algorithms for precise food detection and calorie estimation.

Keywords— Food, Calorie, Deep Learning, Neural Network, Augmentation, Accuracy.

I. INTRODUCTION

Obesity is a medical condition that causes major health concerns due to the accumulation of excess body fat [1]. Obesity is becoming more widespread, particularly among children and young adults. As a result of embracing a Western lifestyle and the pervasive influence of the internet

and television ads, children and teenagers are eating more fast food and unhealthy snacks [2]. A change in diet can cause chronic ailments such as heart disease, stroke, diabetes, liver problems, and kidney damage, to name a few. Food journals, which help people keep track of their calorie intake, are becoming more popular as a way to address these health concerns. The emergence of apps for smartphones and cameras that can recognize foods and estimate their calorie content has greatly simplified this chore. However, this strategy raises concerns due to the wide range of food images from various cuisines around the world.

Current methods frequently fail to recognize individual food items accurately, much alone generating a credible estimate of total calorie content [3]. Our investigation is motivated by the need to overcome these obstacles. We provide a hybrid A-CNNLR model that was created to detect precise food from images and calculate calorie value based on the mass and volume of the food item. This unique method enhances accuracy and provides a comprehensive solution to the challenges that current systems confront in the areas of food identification and calorie estimation.

II. LITERATURE SURVEY

This section reviews some of the recent works on food detection and calorie calculation. The study [4] proposes a calorie estimation method for an Android mobile app. The initial stage in the estimate process is to take a picture of the meal with the phone's camera and use the depth image from the augmented reality core library. They can split out the edible space using a Thai cuisine image dataset and a finely calibrated Mask R-CNN. Following that,

they estimate the quantity of calories in each dish using K-Nearest Neighbour, Support Vector Regression, Linear Regression, and Deep Neural Network (DNN). The DNN technique is the most efficient, yielding forecasts with the highest accuracy, and minimum error rate. In the paper [5], the authors present a novel wearable sensor in the shape of a necklace fitted with a piezoelectric sensor for identifying skin mobility in the lower trachea during eating. This system exclusively utilizes temporal amplitude-varying signals for feature extraction, enabling more precise food categorization in comparison to previous systems that relied on spectral attributes. According to the results of a statistical analysis, the best results can be achieved with a frame length of 30 samples and swallowing occurring at the frame's finish. Therefore, the significance of the chewing pattern in the categorization of food intake is highlighted. In addition, a novel method is provided for calculating the calorie content of meals by calculating the mass of solid food from the number of swallows. The technology relies on an app for mobile devices and helps users keep a healthy lifestyle by delivering instantaneous feedback on the types, quantities, and calories of the food they consume.

In the study [6], the performance of several Machine Learning (ML) algorithms is evaluated and compared in terms of calorie estimation through the processing of food pictures. The paper investigates well-known food image datasets and calorie calculation methods, drawing on research from the last five years to propose a dependable mix of methodology and data for an effective food image processing system. The reliable approaches for analyzing food images, according to the results, are Convolutional Neural Network (CNN) and support vector machine (SVM). The public datasets are employed for food recognition. Furthermore, the study underlines the effectiveness of the mathematical model technique for calorie estimation, providing developers with important insights into upgrading existing food picture processing apps by utilizing reliable algorithms, datasets, and computation strategies. In the study [7], two separate picture classification systems are presented: a CNN and a Residual Neural Network (ResNet), both of which are trained to recognize six different types of food based on color attributes. Both forms of categorization are subjected to a thorough performance assessment. To assess the effectiveness of each categorization approach, datasets including 400 images for each food type are compiled from a number of sources. CNNs have four layers in the proposed system, while

ResNets have fifty. Notably, RGB values (Red, Green, Blue) play an important role in the color feature extraction technique for determining the food category. CNN outperforms ResNet in identifying different types of food, according to the testing results.

The study [8] presents a collection of DL models for regression from food photographs to nutrient estimation utilizing vast amounts of data using deep learning (DL) approaches. In this study, the ChinaMartFood-109 food database was built and published, containing 10,921 images with 23 nutritional components spanning 18 important food types. The optimization of Inception V3 in conjunction with other cutting-edge deep CNN results in top-1 and top-5 accuracy of up to 78% and 94%, respectively. The study also compares three alternative nutritional estimate algorithms and discovers that normalization has the best regression coefficient. The paper [9] presents a system for creating and implementing an ImageAI-based food calorie prediction system. In addition to recognizing foods, this technology may provide a description of their ingredients as well as a calorie count. The proposed method estimates both the food category and the quantity of calories from pictures using the ImageAI and RetinaNet feature extraction object identification models. After capturing the food image and segmenting it into its component components, calories are calculated using nutritional information tables. To assess a food's nutritional value, first identify the food type, then the ingredients, and then combine and compute the calories for the entire item. In the study [10], they created and tested a DL model with OpenCV to identify Indian foods and assess their calorie content. The photographs used to create the Indian Cuisine dataset were gathered from the internet and then meticulously pre-processed. The model categorizes Indian cuisine using a CNN and then utilizes image processing techniques to compute how many calories are in each dish. When applied to the task of identifying Indian cuisine, the developed model is highly accurate, with a maximum detection rate on the training and testing data. When compared to the actual calories in the food, calorie calculations frequently have a margin of error of roughly 10%.

III. METHODOLOGY

Figure 1 depicts the methods used for recognizing food and estimating its calorie value.

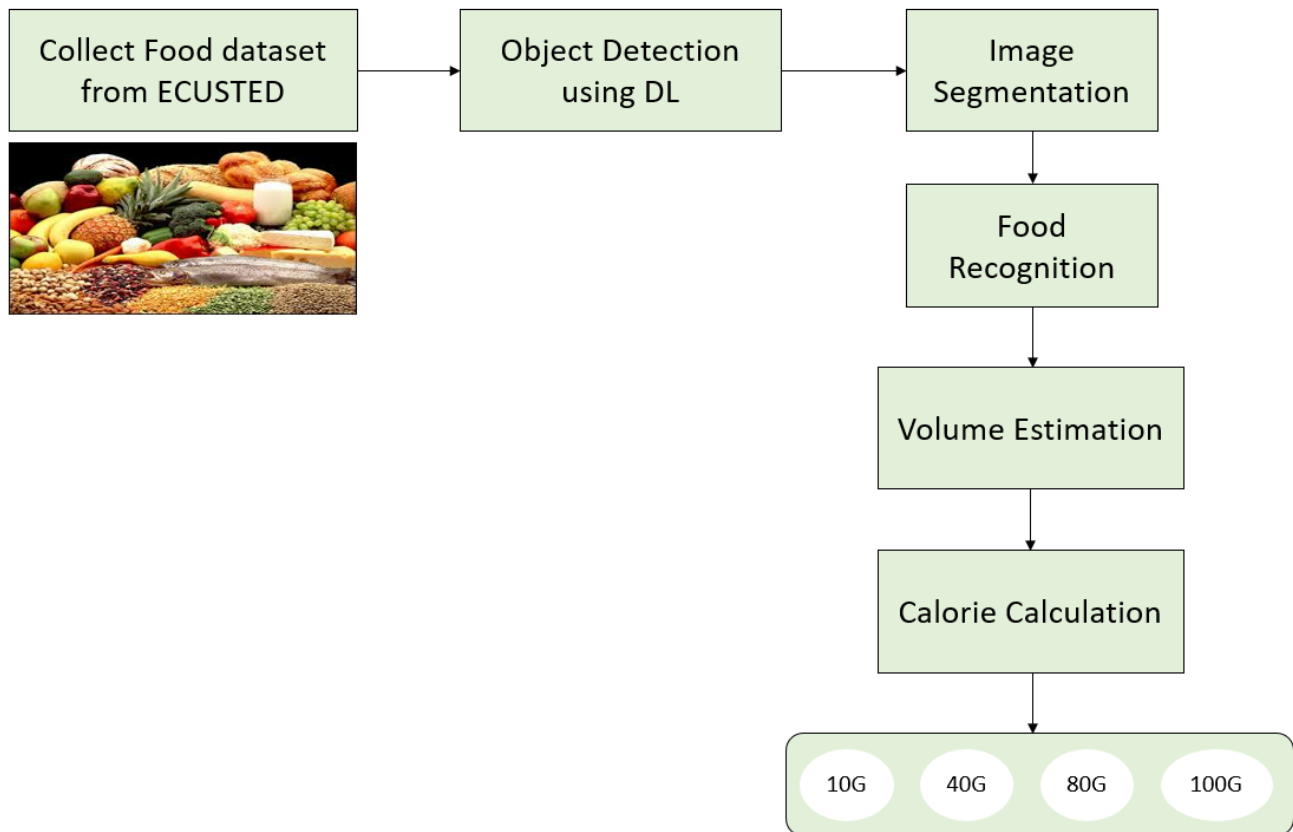


Fig. 1. Proposed Method Workflow

The acquisition of images from the ECUSTED database is the first step in the process. Following that, a DL model created specifically for this task will be suggested to recognize and segment the identified food items inside these images. The suggested DL model recognizes and segments food successfully, utilizing new approaches to improve accuracy. Following the completion of the food recognition stage, the calorie content is calculated. This is accomplished by determining the volume of identified food items. Volume estimation is critical in assessing nutritional content, allowing for a precise estimation of the calorie value of each food item.

A. Data and Processing

It is crucial to assess calorie estimation algorithms using a modern food dataset today. To cover this gap, we used the ECUSTFD data set [11]. Figure 1 illustrates that ECUSTFD contains 2978 food photographs divided into 19 categories. On cell phones, many sets of images were shot, each of which comprised the same dish of food from different perspectives (both above and below). Each image contains no more than two foods, and each image has a One Yuan coin for calibration reasons. Researchers can pick between two versions of the dataset: one that includes the original photos and one that has the images resized. Each scaled image in the collection has a resolution of less than 1000*1000 pixels. In addition to the images, the collection provides the following information for each one:

- Each object in each image has a bounding box (annotation is only available for shrunk photos due to the difficulty of annotating original images with resolutions that exceed screen dimensions).
- The mass of each food item is determined using an electronic scale.
- The dataset contains volume information as a reference because extracting volume from food images is more difficult than extracting volume from mass data.
- Data on food density and energy are provided to aid with calorie estimation. The measured volume and mass are used to calculate density in ECUSTFD. Calorie counts are determined by inspecting each food's nutrition label.

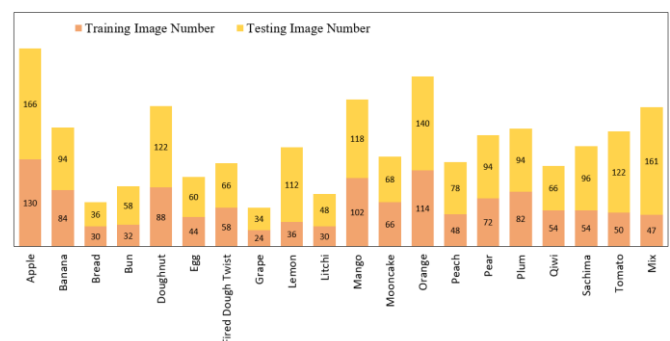


Fig. 2. Data Distribution

We selected a dataset with a wide variety of images of food, however, it's important to note that the sample wasn't randomly distributed. The data distribution of the ECUSTFD database is given in Figure 2. We applied geometric tactics like as rotation, flipping, and contrast enhancement to compensate for this difference [12]. The dataset should now provide a more realistic representation of the full spectrum of available foods as a result of these additions. We used techniques like rotation and flipping to change the orientation of the food images, and color contrast modifications added to the visual depth. This type of data augmentation improves the generality and robustness of the model by exposing it to more situations and variables within the dataset.

B. Food Recognition

We proposed a hybrid model (A-CNNLR) that can detect objects in low-light conditions. To classify objects, the model employs a Dilated CNN (DCNN) [13] and a novel hierarchical attention network [14]. The initial stage is to get the dataset, followed by preparation. A unique vector routing strategy designed for retrieving feature information from the network's deep layers is used to send the cleaned and prepared input to the dilated CNN's attention layer. When it comes time to categorize the obtained features, the LR Classifier is employed.

DCNN Layer: At this stage, we consider specific feature variables like color features that can visually determine the object. The DCNN's ability to learn unique features rapidly and accurately is enhanced by several hidden layers. DL, as opposed to traditional ML, can automatically learn and classify leaves with ailments, indicating its efficacy in this situation. This method involves a procedure that moves from human-driven requests to automatic understanding through expertise, enabling a computer to understand big ideas by reducing things into smaller portions. Feature extraction and selection are profoundly impacted by the outcomes of the DCNN's numerous layers and the attention layer. This is a crucial step in the A-CNNLR architecture. The Dilated Convolution Layer (DCL) explores deep to find composition feature data by searching for hierarchical, granular features. The feature findings are pooled and forwarded to a DCNN layer to provide DCV output, as opposed to the instantaneous dilated convolution processes performed by typical CNNs. equation (1) describes the variables produced by the deep CNN layer.

$$dcv_i = dconv_1, dconv_2, dconv_3, \dots, dconv_n \in R^{n \times d^*} \quad [1]$$

The input sequence is represented by d , while the outcome of the DCL is represented by dcv .

$$DCV^1 = [dcv_1^l, dcv_2^l, \dots, dcv_n^l] \in R^{n \times d^*}, l \in (1, L) \quad [2]$$

The convolutional box as a whole is denoted by L in equation (2), whereas the block filter is denoted by k . Take a look at the l -th block in particular.

$$W^1 \in R^{k \times w \times k}, W^l \in R^{k \times w \times d^*} \quad [3]$$

Using the weight vector w , the filtering matrix W in equation (4) operates over time indexed by k . The connection between two neighboring blocks is described in Equation (4).

$$DCV = F(W^1, DCV^{l-1}) \quad [4]$$

In this case, f represents a linear algebra formula, and w represents the length of the sliding window used to filter the input. The typical approach for finding $dcv_1^l \in DCV^l$ is shown in equation (5).

$$dcv_t^l = ReLU(W^1 \oplus [dcv_{t+1r}^{l-1}]_{t=0}^{w-1}) \quad [5]$$

The symbol \oplus represents convolution, and the letter r represents the depth of the dilation's deep layer. The overall length is $(w - 1)2^{L-1}$ when the ReLU activation is used on each individual block. Unlike the typical deep convolution layer, which grows network layer parameter weights exponentially, this hierarchical method grows weights linearly. Finally, we generate a hierarchical map of $DCV^1, DCV^2, \dots, DCV^L$ using the correlation between the coupling coefficients of the upper and lower levels. The SoftMax function is used to determine the values of the set b_{io} . We currently have $DCV^1 = [dcv_1^l, dcv_2^l, \dots, dcv_n^l]$ in $R^{n \times k^*}, l \in (1, L)$, where each k filter operation gives a dcv output in $R^{n \times k^*}, l \in (1, L)$. The COV_{io} value is the output characteristic. Now compute $COV^1 = [COV_1^l, COV_2^l, \dots, COV_n^l] \in R^{M \times dv}$, where dv is the size of the convolution term and M is the number of final convolutions. The route from DCV^l to COV^l is performed for the purpose of final feature extraction and data production. In Equation (6), multiplying dcv_i by W_j yields the projected vector \widehat{dcv}_{jl} , which represents the transformation of the raw vector features.

$$\widehat{dcv}_{jl} = dcv_i * W_j \quad [6]$$

Small vectors can be normalized, while large vectors can be converted to unit vectors, both of which increase information transmission efficiency in the complex routing process. We employ a multilayered DCL iterative layered routing strategy to get from step one to step two. To express how dcv has changed sr_{fij} , we refer to the softmax routing function as sr_{fij} and the agreement between i and j as a_{ij} .

$$a_{ij} = cov_i * \widehat{dcv}_{jl} + sr_{fij} \quad [7]$$

The dilated convolution technique is widely acknowledged for improving the efficiency and scalability of the convolution routing method. We compute the last convolution layer on our own in this stage, using the formula $COV^1 = [COV_1^l, COV_2^l, \dots, COV_n^l] \in R^{M \times dv}$. The final convolution, as shown in Equation (8), produces a unified COV .

$$COV^1 = [COV^l, COV^l, \dots, COV^l] \quad [8]$$

The action will be sent down through the hierarchical layers. The retrieved features of the dilated convolution $[COV^l, COV^l, \dots, COV^l]$ will be allocated.

Hierarchical Attention Layer: This essential layer receives as input each target convolution and outputs a single actual variable that has been specified and aggregated for attention. For each target convolution $cov_i \in R^{dv}$ in COV , we rank the attention a_{ij} that is created and used in the classification layer. The attention task can be calculated using equations (9) and (10).

$$e_i = a(q, a_{ij}) \quad [9]$$

$$a_{ij} = \frac{\exp(e_i)}{\sum_k \exp(e_k)} \quad [10]$$

This is the probability of the *COV* of the convolution pool over the entire pool, where q is a training pattern vector and k is the likelihood of the *COV* of the convolution pool within the entire pool. Following the extraction of visual features, the aggregated total is then added to the overall target DCL of the downstream pattern, leading to a fixed attention aggregation variable. Lastly, Equation (11) is used for calculating the attention-based dilated convolutional retrieved features.

$$ACOV_i = q^T a_{ij} \quad [11]$$

Using Equation (11), the retrieved features of the attention mechanism $ACOV_1, ACOV_2, \dots, ACOV_n$ are translated to F_1, F_2, \dots, F_n . These characteristics are then used to train an LR classifier.

Classification Layer: Using the LR method, the A-CNNLR model classifies images using the retrieved features. To validate the proposed model, various metrics for monitoring performance and comparisons of sickness images are used.

The goal of this layer is to calculate the probabilistic model by employing the equation $p(y|S)$, where y is the predicted category. The LR algorithm for fixed-length and care-oriented aggregation is employed to provide input for the multi-layer categorization using the vector o . This technique predicts object classes based on raw data. First, we obtain and pre-process an image of the object. The attentive DCNN then sends the completely extracted feature to the attentive hierarchical layer. The properties of the attentive DCNN are employed in the last layer to produce LR predictions for 19 different types of objects.

Begin by gathering the pre-processed features F_i of the CNN. Set the LR classifier's starting values. Determine the sigmoid function of the LR classifier ($\theta^T x$) and run it. The result o can then be calculated using the computational function indicated in Equations (12) and (13).

$$o = \text{sigmoid}(\theta^T x) \quad [12]$$

$$o = \frac{1}{(1 + e^{\theta^T x})} \quad [13]$$

Figure 3 demonstrates our concept in action. This figure depicts how each layer works, from the source phrase to the classification model. The generated output is fed into the next processing layer's input in one way.

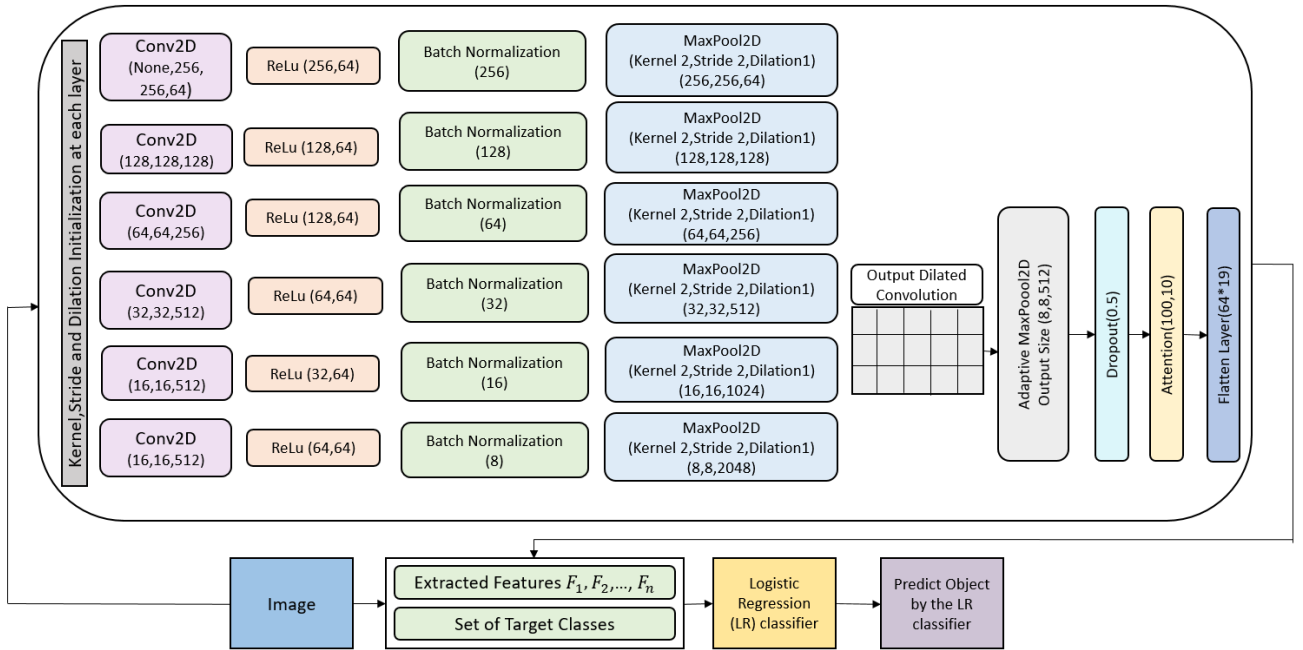


Fig. 3. A-CNNLR Architecture

C. Calorie Estimation

Calibration items are used to generate scale factors, which are subsequently utilised to compute volume. They use a 1 CNY coin as an example to demonstrate the volume computation technique. Because the coin's diameter is 2.5 cm, we can use Equation 1 to calculate the scale factor for the side view.

$$\alpha_s = \frac{2.5}{(W_s + H_s)/2} \quad [15]$$

In this equation, the width of the bounding box is indicated by W_s and the height by H_s . Similarly, Equation 2 can be used to get the scale of the top view.

$$\alpha_T = \frac{2.5}{(W_T + H_T)/2} \quad [16]$$

Following that, ellipse, column, and irregular shaped foods are distinguished. Equation 3 shows how we modify the volume estimating methods to account for different types of foods. Height is marked by H_s in the side view, and the number of foreground pixels in row k ($k \in 1, 2, \dots, H_s$) is denoted by L_s^k . L_s^{MAX} saves the maximum

number of pixels in the foreground. The compensation factor (β) is set to 1.0 by default, therefore each meal type has its unique value.

$$v = \begin{cases} \beta * \frac{\pi}{4} * \sum_{k=1}^{H_s} (L_s^k)^2 * \alpha_s^3, & Shape = Ellipse \\ \beta * (s_T * \alpha_T^2) * (H_s * \alpha_s), & Shape = Column \\ \beta * (s_T * \alpha_T^2) * \sum_{k=1}^{H_s} \left(\frac{L_s^k}{L_s^{MAX}} \right)^2 * \alpha_s, & Shape = Irregular \end{cases} \quad [17]$$

Following that, using Equation 4, calculate the mass of each food item based on the volume estimate. Here, v (cm^3) indicates the volume of the present food, while ρ (g/cm^3) indicates its density.

$$m = \rho * v \quad [18]$$

After that, we use Equation 5 to calculate the number of calories in the dish.

$$C = c * m \quad [19]$$

Here, m (g) is the actual food's mass, and c (Kcal/g) is the number of calories it contains.

IV. RESULT AND DISCUSSION

In this section, we look at the results of the food recognition and calorie estimation processes.

A. Food Recognition

The novel A-CNNLR approach for food recognition is introduced, and its performance is compared to that of the conventional CNN approach. Both accuracy and loss measurements are used to validate these models. Figures 4 and 5 show the CNN model's accuracy and loss plots, which show a maximum accuracy of 0.97 and a loss score of 0.15. Figures 6 and 7 show the accuracy and loss plots for the A-CNNLR model, which shows improved performance with a maximum accuracy of 0.99 and a significantly lower loss score of 0.02. These findings strongly show that the A-CNNLR model outperforms the CNN model in food recognition.

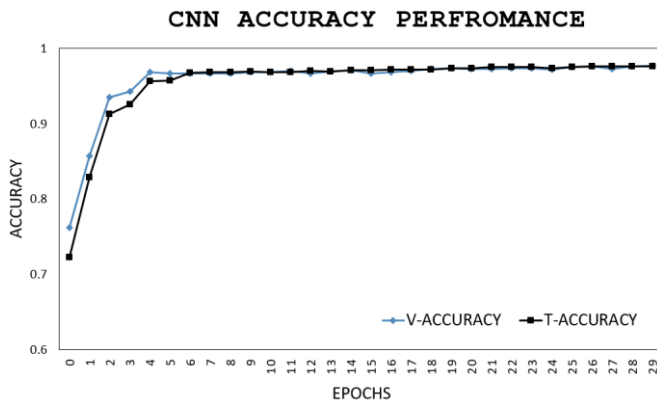


Fig. 4. CNN Accuracy Plot

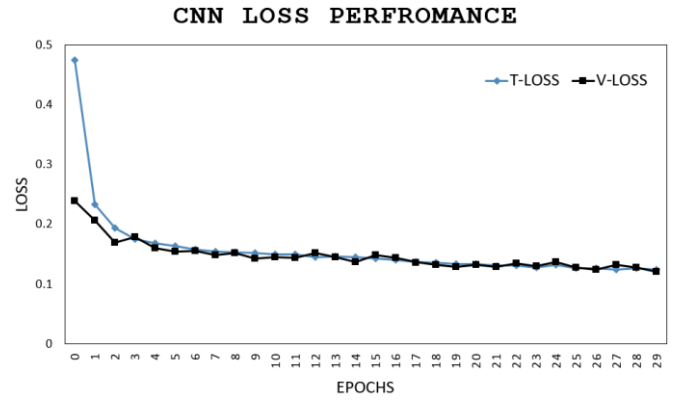


Fig. 5. CNN Loss Plot

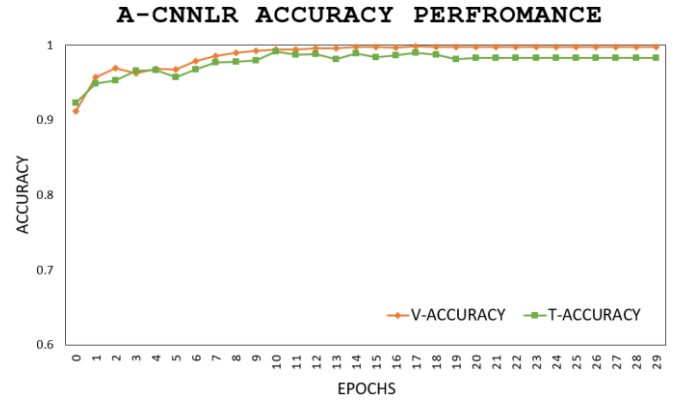


Fig. 6. A-CNNLR Accuracy Plot

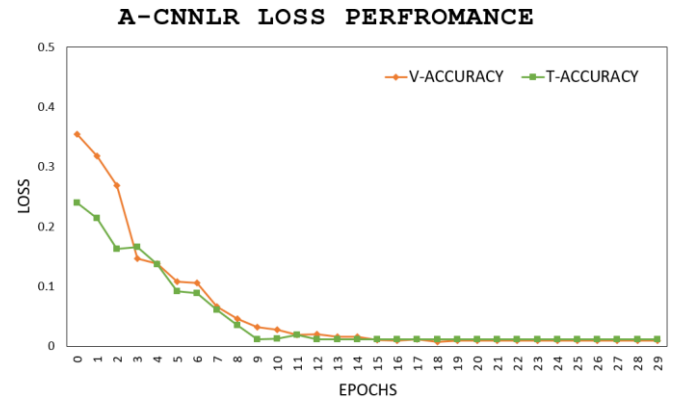


Fig. 7. A-CNNLR Loss Plot

B. Calorie Estimation

Various foods such as apple, orange, tomato, egg, and banana are evaluated in the calorie estimation phase. Figure 8 shows the comparison of actual and predicted calories of the food items. Table 1 compares the actual calorie content to the projected calorie values in great detail. This table is an important measure for determining the system's accuracy. Notably, the percentages of accuracy for each food item range from 95.65% to 99.07%. This indicates the system's effectiveness in estimating the calorie content of various foods and highlights the practical utility of the proposed A-CNNLR model in the field of food recognition and calorie calculation. These findings highlight the potential influence of advanced models in improving the accuracy and reliability of such systems.

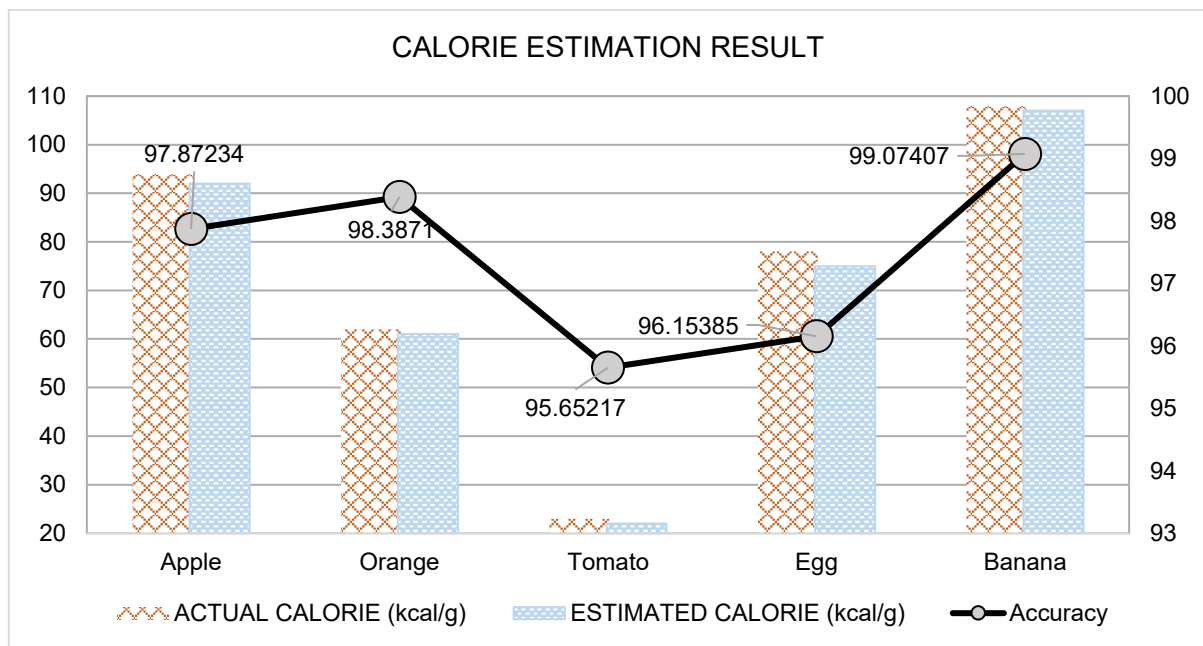


Fig. 8. Outcome of calorie estimation process

Table 1. Calorie estimation

Food	Actual Calorie (kcal/g)	Estimated Calorie (kcal/g)	Accuracy (%)
Apple	94	92	97.87234
Orange	62	61	98.3871
Tomato	23	22	95.65217
Egg	78	75	96.15385
Banana	108	107	99.07407

V. CONCLUSION

This study has investigated the effectiveness of the suggested model in food image recognition. To begin, a large collection of food images was collected, and data samples were distributed equally using augmentation approaches to improve model resilience. Following that, the proposed A-CNNLR model for food recognition was presented and its performance was compared to the classic CNN model. The results of this comparison research demonstrated that the proposed A-CNNLR model had greater accuracy, with a recognition rate of 99%. Furthermore, the study investigated determined the calorie content of recognized food items based on their volume. The calorie estimation accuracy was verified, yielding good findings and confirming the model's capacity to reliably recognize food items and their calories. It's worth noting that the project's scope was purposefully limited to fruits and vegetables, as estimating the calories of fast food and other cooked dishes requires a thorough understanding of ingredients and proportions. The limitations of the project open the way for future improvements, pointing to opportunities for development such as including a wider variety of food categories, such as fast food and cooked products. Segmentation

REFERENCES

- [1]. Iqbal, Rana Khalid, Farah Masood, and Sidra Ikram. "How Obesity Affects Our Health." *National Journal of Health Sciences* 4, no. 3 (2019): 113-118.
- [2]. Skinner, Asheley C., Eliana M. Perrin, Leslie A. Moss, and Joseph A. Skelton. "Cardiometabolic risks and severity of obesity in children and young adults." *New England Journal of Medicine* 373, no. 14 (2015): 1307-1317.
- [3]. Zhou, Jun, Dane Bell, Sabrina Nurat, Melanie Hingle, Mihai Surdeanu, and Stephen Kobourov. "Calorie estimation from pictures of food: Crowdsourcing study." *Interactive journal of medical research* 7, no. 2 (2018): e9359.
- [4]. Sombutkaew, Rattikorn, and Orachat Chitsobhuk. "Image-based Thai Food Recognition and Calorie Estimation using Machine Learning Techniques." In *2023 20th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pp. 1-4. IEEE, 2023.
- [5]. Hussain, Ghulam, Bander Ali Saleh Al-rimy, Saddam Hussain, Abdullah M. Albarrak, Sultan Noman Qasem, and Zeeshan Ali. "Smart piezoelectric-based wearable system for calorie intake estimation using machine learning." *Applied Sciences* 12, no. 12 (2022): 6135.
- [6]. B. Alshujaa, F. AlNuaimi, W. Mansoor and S. Atalla, "Comparison of Food Calorie Measurement Using Image Processing and Machine Learning Techniques," *2022 5th International Conference on Signal Processing and Information Security (ICSPIS)*, Dubai, United Arab Emirates, 2022, pp. 123-128, doi: 10.1109/ICSPIS57063.2022.10002488.
- [7]. A. N. M. Zulfikri, F. Y. A. Rahman, S. Shabuddin and R. Mohamad, "Food Recognition based on Deep Learning Algorithms," *2022 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, Langkawi Island, Malaysia, 2022, pp. 1-4, doi: 10.1109/ISIEA54517.2022.9873669.
- [8]. Ma, Peihua, Chun Pong Lau, Ning Yu, An Li, and Jiping Sheng. "Application of deep learning for image-based Chinese market food nutrients estimation." *Food Chemistry* 373 (2022): 130994.
- [9]. Kumar, G. Kiran, D. Malathi Rani, K. Neeraja, and Jeethu Philip. "Food Calorie Estimation System Using ImageAI with RetinaNet

- Feature Extraction." In *Advanced Techniques for IoT Applications: Proceedings of EAIT 2020*, pp. 93-102. Springer Singapore, 2022.
- [10]. Sathish, Suriyakrishnan, S. Ashwin, Md Abdul Quadir, and L. K. Pavithra. "Analysis of Convolutional Neural Networks on Indian food detection and estimation of calories." *Materials Today: Proceedings* 62 (2022): 4665-4670.
- [11]. Liang, Yanchao, and Jianhua Li. "Computer vision-based food calorie estimation: dataset, method, and experiment." *arXiv preprint arXiv:1705.07632* (2017).
- [12]. Alomar, Khaled, Halil Ibrahim Aysel, and Xiaohao Cai. "Data augmentation in classification and segmentation: A survey and new strategies." *Journal of Imaging* 9, no. 2 (2023): 46.
- [13]. Lei, Xinyu, Hongguang Pan, and Xiangdong Huang. "A dilated CNN model for image classification." *IEEE Access* 7 (2019): 124087-124095.
- [14]. Feng, Jiashi, Huan Xu, Shie Mannor, and Shuicheng Yan. "Robust logistic regression and classification." *Advances in neural information processing systems* 27 (2014).