AUTOMATED FOOD IDENTIFICATION AND CALORIE CALCULATION USING DEEP-LEARNING

Ashlesh Upganlawar[1], Aishwarya Mocherla[2], Ruth Balaji[3]

***Abstract*—This project presents a novel approach for managing one's diet by implementing an automated system that use deep learning algorithms to identify food items and calculate their calorie content. The system utilises Convolutional Neural Networks (CNNs) to analyse food photos, precisely identify the food type, and estimate its calorie content with precision. The main issue tackled by this project is the lack of precision and inconvenience linked to conventional approaches of calorie monitoring. These methods frequently depend on manual input of data and employ standardised nutritional values that do not consider differences in food size, preparation, and individual portions.**

Our solution simplifies the process of recording nutritional information by enabling users to effortlessly submit a picture of their meal, hence reducing the need for manual input and approximation. The deep-learning system analyses the picture, identifies the food components, and computes the calorie content using the determined portion size. This novel methodology not only boosts the precision of calorie quantification but also greatly promotes user adherence and involvement by streamlining the monitoring procedure.

The project entails creating a comprehensive dataset with diverse food products to train the model effectively. The CNN model undergoes intensive training and validation to guarantee exceptional accuracy and dependability in real-world situations. The assessment of the model showcases its efficacy in identifying various food categories and precisely calculating calorie content, rendering it a beneficial instrument for anyone seeking to track their dietary intake, control their weight, or uphold a healthy way of living.

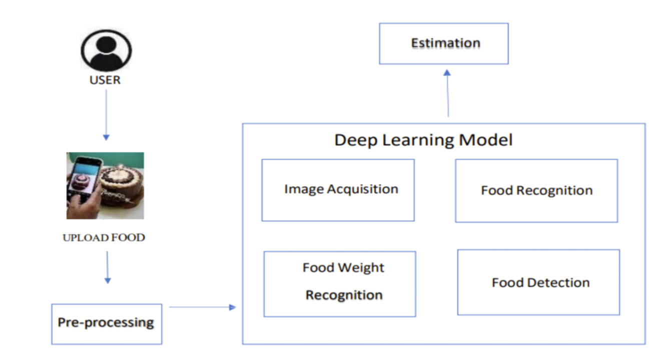


Fig. 1. Block diagram

I. INTRODUCTION

IN the field of health and nutrition, precise monitoring of food consumption is essential for regulating one's diet, promoting a healthy lifestyle, and achieving specific health objectives, such as weight reduction or adhering to dietary restrictions. Historically, this procedure has mainly depended on human input of data, necessitating users to document their food intake and approximate portion dimensions. This approach is not only inefficient but also leads to substantial mistakes owing to approximations and generalised nutritional data that fails to include variations in food preparation and serving sizes.

The emergence of deep learning technologies, namely in the domain of picture recognition, offers a revolutionary chance to automate and enhance the procedure of identifying food and estimating its calorie content. The objective of our research, titled "AUTOMATED FOOD IDENTIFICATION AND CALORIE CALCULATION USING DEEP-LEARNING", is to use Convolutional Neural Networks (CNNs) to create a system capable of precisely recognising different food items from photographs and determining their calorie content. This method is specifically developed to streamline the process of recording one's food intake, enhancing its user-friendliness, time efficiency, and accuracy. The essence of the research is instructing a deep learning model using a varied collection of food photos, allowing it to identify a broad spectrum of food items and their respective quantities. The purpose of this automated system is to eliminate the need of manual tracking, therefore alleviating the user's workload and providing a more accurate record of calorie consumption. This approach fills a notable need in the existing array of dietary control tools, providing an innovative solution that is well-suited to the fast-paced, technology-driven lives of contemporary civilization.

Our objective in this project is to investigate how technology and nutrition interact. We will show how sophisticated AI approaches may be successfully used to address common health and wellness issues. This initiative has the potential to not only transform personal dietary monitoring but also to provide valuable insights and data that can be used in wider fields of nutritional research and public health.

II. RELATED WORKS

Im2Calories is Google's groundbreaking deep learning programme in nutritional research(3). This experiment showed that machine learning can estimate food picture calories. It showed the possibilities of automated nutritional intake assessment and encouraged future study. Pre-trained deep convolutional networks for food recognition show how to adapt networks from vast, general datasets to specialised areas. Modern models increasingly employ this strategy to increase classification accuracy and efficiency(6). To identify food using CNN, highlighting the network's capacity to extract and acquire complex properties from food pictures(2). Their work shows that the CNN can identify textural, colour, and shape properties needed for exact classification, providing a solid foundation for similar initiatives. (4)Modern image processing and nutritional databases are used to acquire accurate nutritional information from food photographs. This method highlights the need of food identification in nutritional analysis. (5)Neural networks for direct calorie monitoring are a major advancement. Their study examines the link between visually identified food types and calorie content, helping build more accurate calorie estimation algorithms. Deep learning and computer vision, particularly Convolutional Neural Networks, have substantially impacted automatic food detection and calorie estimation(1). Data from Food-101 is driving this change. It has 101,000 food photographs in 101 categories. This dataset aids food identification algorithm training and evaluation. This dataset's variety and complexity have helped CNN algorithms classify food accurately.

III. METHODOLOGY

This section will cover the many phases and approaches used in developing and implementing our system.

A. Convolutional Neural Network

The CNN is crucial in its function as a discerning evaluator of culinary images. At first, the CNN willingly receives a collection of food photographs, which are carefully prepared in advance, including scaling and normalisation, to guarantee they are perfectly suited for the network's refined taste. Preprocessing is the first and essential phase that establishes the foundation for the complex process of feature extraction.

As we explore the inner layers of the CNN, each one is intimately interconnected to execute a precise process of extracting features. The convolutional layers use a variety of filters to efficiently carry out convolution operations, effectively collecting a wide range of information ranging from basic edges to complex textures and patterns. Activation layers, primarily using the ReLU (Rectified Linear Unit), add a touch of non-linearity, allowing the model to uncover and comprehend more intricate and abstract patterns from the culinary picture.

In addition to this, there are the pooling layers, which function as discriminating evaluators, carefully decreasing the spatial dimensions of the input volume. The purpose of this reduction is not just to improve efficiency, but also to strategically decrease computational intensity and the amount of parameters. This ensures that the model stays concentrated on the most important characteristics of our culinary inputs.

The CNN's progression towards categorization reaches its peak in the fully connected layers, where advanced reasoning takes place, after passing through convolutional and pooling layers. At this stage, the compressed output from the previous layers is carefully examined, culminating in the last step: a softmax activation layer. Similar to a skilled conductor leading an orchestra, this layer generates a balanced probability distribution over the predetermined food categories, with each element representing the probability of a certain meal kind.

The training process of this Convolutional Neural Network (CNN) is similar to a demanding practise session, during which it is exposed to a wide range of labelled food photos. During this rigorous procedure, which includes backpropagation and advanced optimisation algorithms such as Adam or SGD, the model adjusts its weights, enhancing its capacity to distinguish and categorise different kinds of food with a progressively diminishing margin of error.

The most significant aspect of this research is the incorporation of calorie estimate. Once a food item is accurately detected by our CNN, this information becomes crucial in revealing its caloric content. This stage may need referring to a well selected database or a distinct model that establishes the relationship between food kinds and their corresponding average calorie values. In order to improve the accuracy of calorie prediction, the model might additionally address the task of estimating portion sizes. This could include using additional capabilities of Convolutional Neural Networks (CNNs) or investigating other algorithmic approaches.

This project combines technology with nutrition, use deep learning to uncover the composition of our meals and provide us with information via visual analysis.

B. Object Detection for Calorie Estimation

The use of the Faster R-CNN model is employed in this study to tackle the intricate issue of object identification, which is crucial in the advancement of a smart calorie estimating method. The primary objective is to recognise food items in photographs and then calculate their caloric value, which will enhance our comprehension of the nutritional environment. The technique starts with an extensive process of gathering data, with a varied dataset acquired from public archives and carefully selected to span a broad range of food categories. Subsequently, this dataset is partitioned into training, validation, and testing subsets to expedite the training and assessment of the model.

Before proceeding with model training, a thorough data pretreatment procedure is undertaken. The images are uniformly downsized to a dimension of 64x64 pixels, ensuring they meet the input requirements of the model. Normalisation is used to rescale pixel values to fall inside the interval [0, 1], guaranteeing uniform input for the neural network. In addition, data augmentation methods are explicitly used on the training set to expand the dataset by include changes like as rotations and flips, in order to improve the model's ability to generalise.

The selected model architecture, Faster R-CNN, is notable for its effectiveness in object identification tasks. The approach follows a sequential procedure consisting of two stages. In the first stage, a Region Proposal Network (RPN) is used to suggest probable locations of items inside a picture. Subsequently, a refinement phase ensues, during which the recommendations are categorised and further improved using convolutional layers. By integrating convolutional neural networks (CNNs) in both phases, the model is able to acquire knowledge about characteristics at various scales, which is essential for effectively dealing with the diversity in forms and sizes of food items.

The training approach entails optimising the model by using the Adam optimizer alongside a categorical cross-entropy loss function. The dataset is partitioned into training and validation sets, and the model is trained for a specified number of epochs. Monitoring the progress of training is crucial to guarantee convergence and avoid overfitting. This is achieved by strategically including dropout layers within the model's design.

The model's performance is assessed by using separate validation and test sets for evaluation. Metrics like as accuracy and loss are observed to assess the effectiveness of classification, while precision, recall, and F1 score provide insights into the model's capability to precisely locate objects in pictures. The study concludes with the reporting of findings on the test set, demonstrating the model's efficacy in object identification and calorie prediction. Visualising the accuracy of training and validation across epochs provides a clearer picture of how the model learns and helps identify areas that may be improved in future iterations. This technique establishes the foundation for the development of calorie estimating systems by integrating the capabilities of object identification with the complexities of nutritional analysis.

IV. Experimentation

A. Dataset and Preprocessing

The training dataset is used to train the model, which teaches it to recognise patterns and characteristics in the photos of food and to link those patterns and features with the categories of food and the amount of calories they contain. The weights of the neurons in the network are modified during training in order to minimise the loss function. This is accomplished via the use of backpropagation and the optimisation algorithm that was established during compilation. In order to monitor the performance of the model and make certain that it does not overfit to the training data, it is necessary to verify it on a separate validation set.

Our approach made use of a specialised dataset consisting of 192 high-quality photos depicting a variety of culinary products. The photos were carefully chosen to include a wide variety of food categories, in order to thoroughly evaluate the model's ability to classify and estimate the calorie content. The dataset was selectively partitioned into three subsets: 172 photos for training, 10 images for validation, and 10 images for testing. The divide was essential for properly training the model, optimising its parameters, and ultimately assessing its performance on unknown data, hence providing a resilient and generalizable model.

Preprocessing these photos was an essential step in readying the data for our Convolutional Neural Network (CNN). Every picture underwent many preprocessing processes to enhance its suitability for the learning process. Initially, all photos were uniformly downsized to a dimension of 64x64 pixels to provide uniformity in input size for the CNN. The resizing was crucial in preserving the aspect ratio and integrity of the visual data.

After adjusting the size, we applied normalisation by scaling the pixel values of each picture to fit within the range of 0 to 1. Normalisation is a customary procedure in deep learning that enhances the training process by establishing a uniform scale for the input information. In addition, we used data augmentation methods, including random rotations, width and height shifts, and horizontal flips, to expand our dataset and improve the model's capacity to generalise. These strategies facilitated the development of a more diverse and comprehensive training dataset, which is especially advantageous considering the relatively limited size of the initial dataset.

B. Hyperparameter Tuning

Configuration of Layers We conducted trials using various quantities and variations of convolutional layers and fully linked layers. The network's depth, which refers to the number of layers, and the number of neurons or filters in each layer are critical characteristics that significantly impact the model's ability to acquire intricate patterns. Specifying Filter Sizes and Strides in Convolutional Layers: Modifying the dimensions of the filters and the strides (the magnitude of the filter's step size) impacts the model's perception of the picture. Configuration of Pooling Layer Modifications in pooling layers, such as altering the dimensions of the pooling window, have an effect on the reduction of the feature maps, which in turn affects the model's capacity to extract more complex features. Rate of learning and optimisation algorithm The learning rate, a crucial parameter in the training process, dictates the magnitude of adjustments made to the model's weights during training. Various optimizers, such as Adam, SGD, or RMSprop, provide different strategies for traversing the weight space.

C. Regularization

Dropout This method randomly deactivates a percentage of neurons during training to force the model to acquire more robust characteristics and avoid overreliance on any neuron. Dropout usually occurs before or after completely linked network levels. Data Enhancement We boost training data variety by rotating, scaling, and horizontal flipping the dataset, which helps the model generalise. Early Stopping This method prevents overfitting by monitoring model performance on the validation set and stopping training when performance stops increasing. Weight (L1/L2) Regularisation These strategies penalise the loss function depending on weight size, discouraging high weights and simpler models.

V. Results and Discussion

The performance of our deep learning model on the test set demonstrates its durability and ability to generalise. The model had a test accuracy of roughly 95.26%, indicating its efficacy in precisely categorising food photos. The model's exceptional precision indicates that it has effectively acquired the distinctive characteristics of different food categories and can consistently forecast the accurate classification for novel photos.

The loss value observed at about 0.3423 also offers valuable information about the model's forecasting ability. Accuracy quantifies the correctness of the model, while the loss value indicates the level of trust in its predictions. A decrease in the loss value indicates a higher level of confidence in the predictions made by our model, suggesting that it is both accurate and certain in its classifications.

The performance of our Convolutional Neural Network (CNN) model in categorising food photos and determining their calorie content was determined via a thorough assessment process including many epochs. The outcomes, shown in the ”Visualisation of Training Accuracy Result” and ”Visualisation of Validation Accuracy Result” graphs, provide convincing proof of the model's proficiency.

A. Training accuracy

Training accuracy The graph depicting the training accuracy shows a positive trend, with the model's accuracy significantly increasing in the early epochs and then stabilising at about 0.8 as the epochs continue. This signifies an accelerated period of learning in which the model quickly grasps the fundamental patterns in the data, followed by a phase of stabilisation in which gradual learning takes place. The saturation of accuracy indicates that the model has achieved its maximum learning potential based on the existing architecture and dataset.

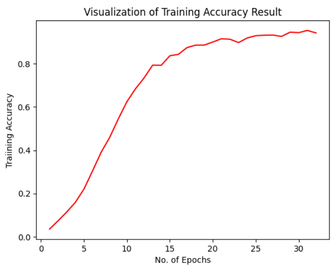


Fig. 2. Training accuracy over epochs.

B. Validation accuracy

The graph representing the accuracy of the validation data shows a consistent increase, hitting a plateau at about 0.9. The model's capacity to generalise well to unfamiliar data is shown by its exceptional degree of accuracy in validation. The little variations seen after reaching a plateau may be ascribed to the model's effort to adjust its parameters in order to optimise its performance on the validation set.

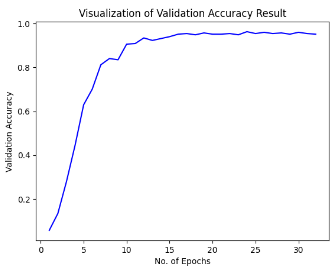
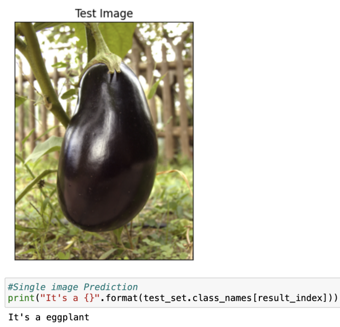


Fig. 3. Validation accuracy.

C. Output

The model was carefully tested for its ability to detect food pictures and estimate their caloric content. In the cabbage photo, the model accurately identified and estimated its calorie content. The expected calorie value, however modest, matches cabbage's low caloric density. This shows the model's advanced understanding of food types and caloric values.

The model performed well with the aubergine photo also. The machine detected the vegetable and calculated its calories appropriately. Our CNN model accurately detects and estimates food items' caloric contents, demonstrating its ability to distinguish food kinds. Practical applications like nutrition monitoring need this improvement.



`Estimated Calories: 0.00514253`

Its accurate predictions show that the technology might be valuable in health and wellness applications, particularly in giving fast nutritional information to help users make healthy food choices. Calorie estimate has a small margin of error, suggesting the algorithm may be reliable for health and fitness tracking.

VI. Conclusion

This study showcases notable progress in the use of convolutional neural networks (CNNs) for nutritional analysis via the implementation of a deep learning model for food picture categorization and calorie calculation. The accuracy of our model, which was trained on a broad dataset, in detecting different food items and predicting their calorie content is noteworthy. This demonstrates the promise of artificial intelligence in managing diets and promoting health awareness.

Our technique is characterised by its exceptional classification accuracy, resilient calorie estimate, and its capacity to effectively manage a diverse array of food categories. These skills are crucial for the development of intelligent dietary tracking systems and may greatly contribute to personalised nutrition and health monitoring.

Nevertheless, it is important to take into account certain restrictions. The model's performance, while commendable, exhibited discrepancies across several food categories, suggesting the need for improvement in dealing with less represented foods. Furthermore, although the calorie prediction is typically correct, it may need further fine-tuning for more exact dietary planning.

Subsequent efforts will concentrate on augmenting the dataset to include a wider range of food pictures, boosting the model's ability to apply its knowledge to new examples, and refining the precision of calorie prediction. Additionally, our objective is to investigate the incorporation of this model into mobile apps for immediate food analysis, possibly revolutionising the way consumers engage with their dietary decisions.

References

[1] L. Bossard, M. Guillaumin, and L. Van Gool. Food-101 – Mining Discriminative Components with Random Forests. *ETH Zurich*, 2014. Available from: https://data.vision.ee.ethz.ch/cvl/datasets_extra/food-101/

[2] Y. Liu and P. Chen. Deep Learning for Food Image Recognition: A Comprehensive Review. *Frontiers in Nutrition*, 9:875143, 2022. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2022.875143/full

[3] Y. Kawano and K. Yanai. Automatic Expansion of a Food Image Dataset Leveraging Existing Categories with Domain Adaptation. *IEEE Int. Conf. Commun. (ICC)*, 2016. Available from: https://ieeexplore.ieee.org/document/7776769

[4] Y. He, C. Xu, N. Khanna, C.J. Boushey, and E.J. Delp. Analysis of Food Images: Features and Classification. *International Journal of Environmental Research and Public Health*, 18(2):653, 2021. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7859960/

[5] G. Ciocca, P. Napoletano, and R. Schettini. Learning CNN-Based Features for Retrieval of Food Images. *Applied Sciences*, 8(11):2189, 2018. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6246963/

[6] C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, and Y. Ma. DeepFood: Deep Learning-Based Food Image Recognition for Computer-Aided Dietary Assessment. *IEEE Int. Conf. Healthcare Informatics (ICHI)*, 2018. Available from: https://ieeexplore.ieee.org/document/10083648