IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 787

Research Paper , UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

An Approach to Food Recognition and Nutrition Assessment Based on Machine Learning

Jebaraj Ratna Kumar[1]

Professor, Computer Science and Engineering, J.N.N Institute of Engineering, Chennai, Tamil Nadu, India

S. Senthilnathan[2]**, A. Krishnaveni**[3]**, M. Karthikeyan**[4]**, A. Iyyanarappan** [5]**, K.Philip Vinod**[6]

Assistant professor, Computer Science and Engineering, J.N.N Institute of Engineering, Chennai, Tamil Nadu, India

Abstract

To prevent obesity in the human body, a balanced diet is now required for regular consumption of healthful foods. This research, introduced a novel machine learning-based system that automatically classifies food photos accurately and calculates food qualities. In the training portion of the prototype system, the deep learning model proposed in this paper uses a convolutional neural network to classify food into several categories. Increasing the pre-training model's accuracy is the primary goal of the suggested approach. A client-server model-based prototype system is designed in this paper. An image detection request is sent by the client, and the server handles it. Three primary software components make up the prototype system's design: a server-side module, a text data training module for attribute estimation models, and a pre-trained CNN model training module for classification applications. To attain better categorization accuracy, made experiments with several food categories, each with hundreds of photos, and machine learning training.

Keyterms:

Nutrition estimation, machine learning, Food recognition, image detection

Introduction

Food is necessary for human survival. These days, people are more mindful of what they eat as unhealthy diets can cause a wide range of diseases, including diabetes and obesity. Our ability to identify food items is essential to maintaining our fitness and health. In order to help people plan adequately for their daily calorie intake, the suggested approach will allow both healthy and obese individuals.

Prior studies primarily employed hand-engineered characteristics in conjunction with conventional image processing approaches for food recognition. These techniques include co-occurrence statistics between food items, feature fusion, manifold ranking-based methodology, and relative spatial correlations of local features. These approaches are computationally expensive, have a low recognition rate, or fit poorly to large-scale applications. Accurate recognition is now achieved with the use of different machine learning techniques.

[1] offered a machine learning method for precise attribute estimation and classification of food images. Convolutional neural networks are used by the system to categorise food into different groups. A CNN model, text data training, and a server-side module make up the three primary software components of the client-server prototype system. To increase the accuracy of classification, experiments were carried out using different food groups. [2] Health care providers and dietitians stress how crucial it is to keep a nutritious diet to avoid obesity and other health problems.

[3] addressed that conventional nutritional monitoring systems suffer from poor adherence, underreporting, and imprecision. Modern systems automatically identify food and estimate quantities using machine learning techniques. A convolutional neural network variation is used for ingredient recognition in all research assessed, and 66.7% of them use visual features from deep neural networks for food recognition. Future work should concentrate on employing explainable AI to improve model transparency, learning from unlabeled image datasets, and preventing catastrophic forgetting. [4] For food safety and hygiene inspections, the study shows that deep learning works better than ordinary machine learning algorithms and manual feature extractors.

[5] study combines a new food-matching technology with a validated food-choice research approach. The technique employs natural language processing for food matching and standardisation, and deep learning for the identification of fraudulent food images. [6] It contrasts the advantages and disadvantages of the three primary solutions—platform-based techniques, transfer learning, and design from scratch. The chapter emphasises the significance of comprehending and using these techniques while also providing background information, pertinent statistics, and future directions.

[7] Growing interest in their application in nutrition was discovered in this systematic analysis of 36 studies that used machine learning (ML) algorithms to evaluate food intake in various groups. The meal frequency questionnaire was the most widely used tool, while supervised learning algorithms were the most often employed. R and WEKA, two programmes for data analysis, were also utilised. The application of machine learning (ML) in nutrition is new and difficult; further research is needed for public policies and food reeducation initiatives. [8] With publicly available food photographs from the MyFoodRepo app, a public baseline for automatic food recognition has been set. A dataset of 24,119 photos and 39,325 segmented polygons, divided into 273 classes, was made public by the benchmark. The best-performing models were implemented in production using the MyFoodRepo app, with a mean average precision of 0.568 and recall of 0.885.

[9] suggests using input image classification to estimate food features such as nutritional value and components using a mobile-based approach. It makes use of semantically linked terms from a sizable text corpus and deep learning models. Tests on the Food-101 dataset and its expansion for foods from other continents demonstrate the system's effectiveness, and it is intended for use in medical settings. A mobile app is used to implement the system. [10] presents a mobile-based dietary evaluation system that estimates food items' nutritional contents using deep learning algorithms. Web scraping datasets for several food categories are used by the system. On the other hand, technical problems include analysing inadequate nutrition content information and identifying various food items from a single photograph. For the healthcare industry, the system is developed as an Android application.

[11] examines the effectiveness of deep learning techniques and contrasts existing methods for food recognition and volume/weight estimates. The study comes to the conclusion that, in order to

ISSN PRINT 2319 1775 Online 2320 787

 UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

accurately estimate dietary intake while taking advantage of both opportunities and obstacles in the field, integrated dietary assessment systems paired with various methodologies could be a viable solution.

[12] The Healthy 365 mobile app from Singapore promotes smart consumption and a healthy lifestyle by making food logging easy for more than 100 registered organisations. This paper examines FoodAI, an API service utilised in the app. [13] Food processing efficiency can be increased by machine vision, with image processing being a crucial element. Food types and quality can be identified using machine learning and deep learning models, and tasks like food grading and impurity removal can be handled by follow-up design. An overview of machine vision techniques, including conventional and deep learning methods, as well as present and upcoming trends in food processing are given in this study.

[14] investigates deep learning's potential for food image recognition systems. For protein estimation, the top-5 Arithmetic Mean technique produced the greatest regression coefficient, demonstrating its practical usability. The findings might promote the use of artificial intelligence in estimating food nutrients. [15] Particularly, Deep convolutional neural networks were used to optimise Inception V3, resulting in great accuracy. The study supported artificial intelligence in food analysis by validating the use of three nutritional estimation methods.

Pre – Trained model section

Here, the suggested methodology comes under three distinct sections. The first section deals with CNN models based on transfer learning, the second section deals with text recovery from various sources, and the third section deals with text data training.

Phase of Preparing the Dataset and Per-processing

Generated a dataset comprising hundreds of thousands of images of different foods. Some of the images are relevant to our research study, while others are not, so filtering the data set is important for setting up a model. The Spatial Transform Network to train transformation parameters, which are then applied to the food image and transformed are used. Figure 1 explains the structure bonding of food composition.

Training of Textual Data Models

A machine learning tool called Word2Vec aids in the computation of the vector representations of various words. Because Word2Vec is a more powerful method than the clustering algorithm, it is utilised in place of the latter. Word2Vec is a two-layer neural network. In this work, trained text data using word2vec, skip Gramme, and continuous Bag of words.

ISSN PRINT 2319 1775 Online 2320 787

Research Paper , UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

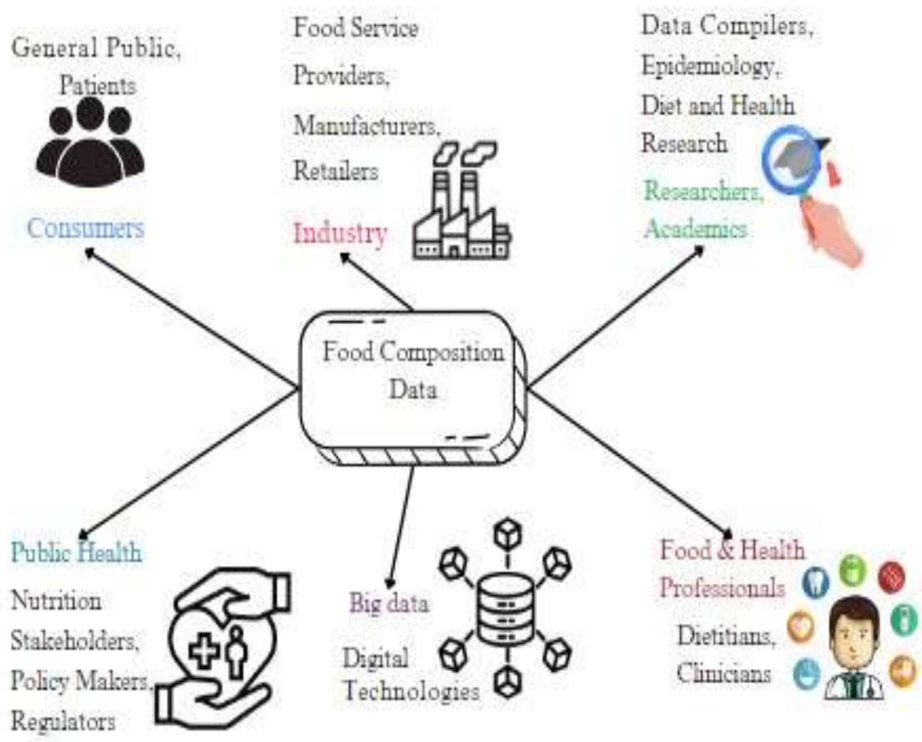


Figure 1: Classification of Food composition data

Design and Implementation

This chapter displays the system's design, flow, and execution along with its outcomes and assessment. The purpose of this demonstration is to shed light on the various parts, processes, and instruments that employed during the system's implementation and illustrate how they cooperated to produce the intended outcomes and functions.

Architectural Overview

This case just employed the server-side architecture in the system that designed for our system. Our system's primary goal is to increase the accuracy of pre-trained models so that developers and architects can utilise it to their advantage by creating custom client-side web- and Android-based applications.

 UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

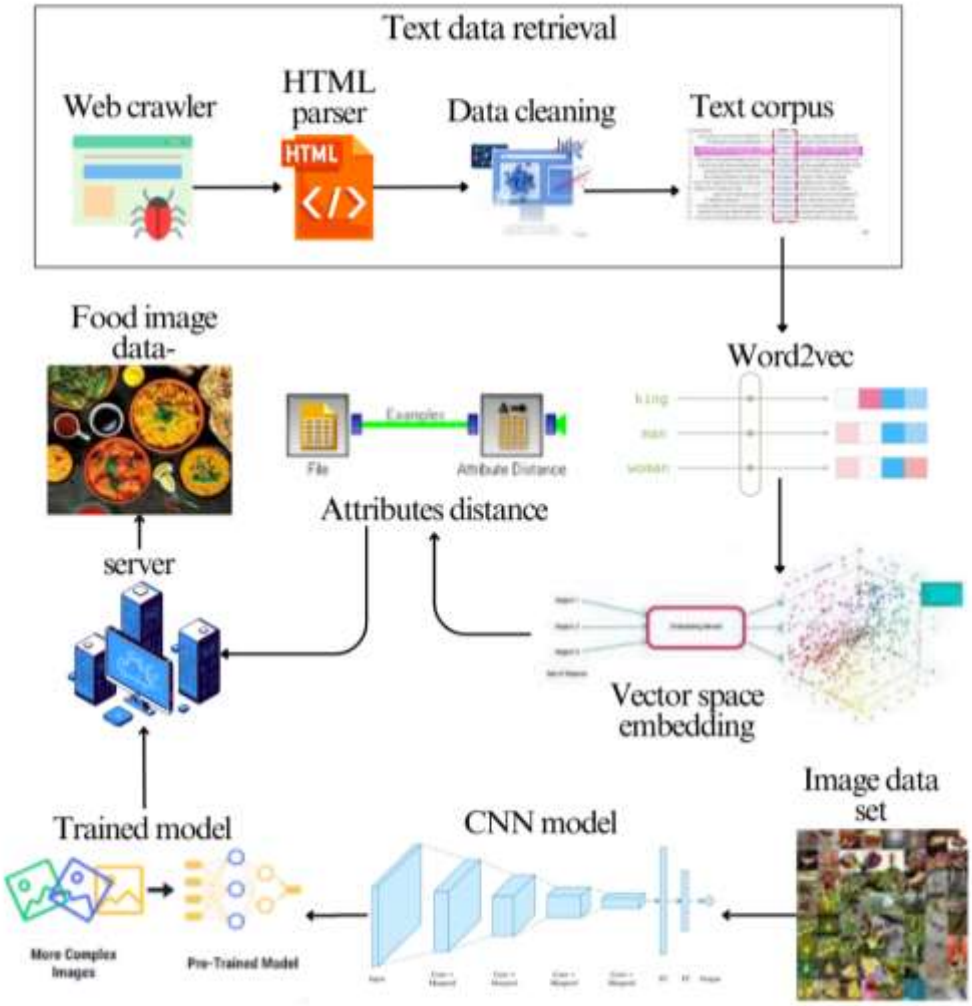


Figure 2: Proposed system's block diagram

Three modules—text data retrieval, text data training, and CNN-based classification model training—are described in Figure 2. Following the completion of the classification process with a pre-trained model, the text data retrieval system receives the food's name. The relevant text will be taken from the Google search and extracted via the URL. After that, HTML will be taken out of websites and stop words will be eliminated using various Python tools and an HTML parser. Following this, it went via the lemmatization and stemming procedure.

A CNN model trained on images of food and related properties is described in Figure 3. The web scraping of around 500 URLs from each food category makes up the text data unit. Text data is also processed by the model. Following text data collection, vector space embeddings based on semantic similarity are obtained by training word2vec on the corpus. These models predict the properties and qualities of every given image input. The suggested system's workflow begins with the generation of textual and picture data and concludes with classification, characteristics, and an ingredients list. Data

ISSN PRINT 2319 1775 Online 2320 787

, UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

augmentation is used to improve and preprocess food photos that have been gathered from various sources. The next steps are to train and fine-tune the pre-trained CNN model. Following the system's successful training and learning phase, the categorization phase begins. Classification is completed upon user input. The attribute estimation model uses the yielded label from the previous phase as its input. The process of creating an attribute model begins with the raw text. Following the preprocessing of textual input, a programme called "Word2Vec" creates vector embeddings made up of distance values. The attribute estimate methodology uses these values as a foundation to generate associated ingredients or attributes.

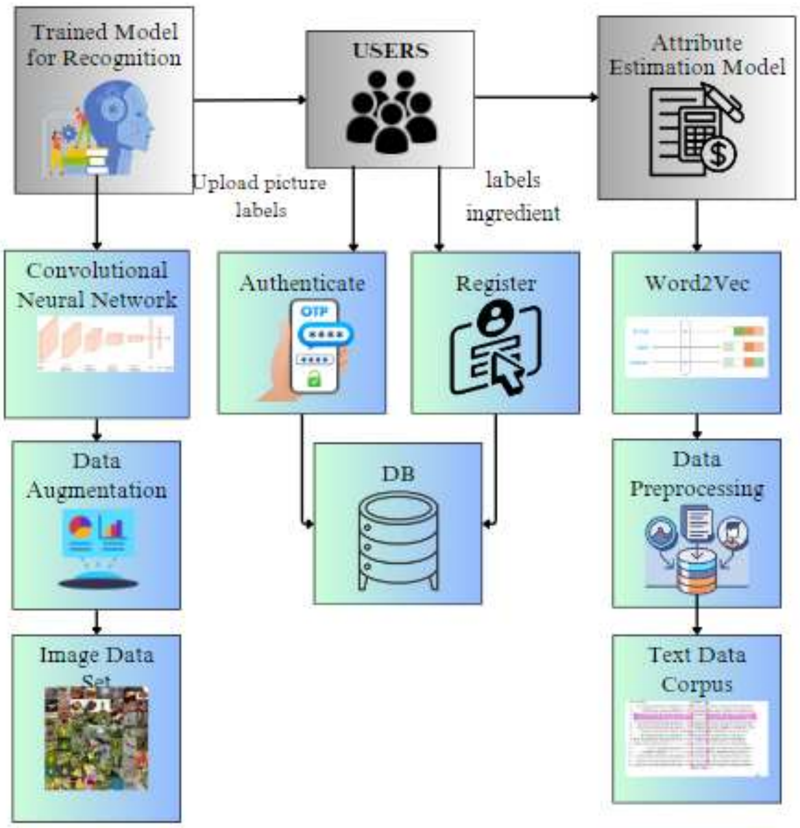


Figure 3: dataflow diagram

Implementing the System

Tensor Flow and the Keras were utilised to implement the model in Python. Computing features from about 50,000 photos involved experiments with several CNN models. With differing accuracies, each model required a different amount of training time, ranging from about 30 hours. With its recommended convolutional and pooling layers, the inception model was categorised with a 91.73% accuracy rate. In essence, the inception model uses ImageNet for pretraining. Additionally, multi-crop evaluation was employed, which involved taking about ten crops at a time during the prediction process. Maximal value derived from anticipated outcomes yields optimal accuracy.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 787

 UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

Table 1: Comparison of Single and Multiple Crops

Model		**Top-1**	**Top-2**	**Top-3**	***Top-5***
	Multiple corpus	98.31%	89.12%	-	-
	multiple corpus	98.56%	91.73%	-	-
	multiple corpus	-	-	-	-
Inception v3	Single corpus	95%	79.8%	91.6%	*87.9%*
Inception v4	single corpus	94.7%	83.8%	92.4%	*89.8%*
V4-101	**single corpus**	**91.2%**	**78.3%**	**88.2%**	***85.4%***

Results and Discussion

Since the CNN model outperforms other models in our suggested issue area, chose it based on Inception-v3 and Inception-v4. To undertake a comparison, these models are refined on both Food-101 datasets and our own datasets. Figure 4 and 5 demonstrates the comparison of accuracy between inception V3 and V4.

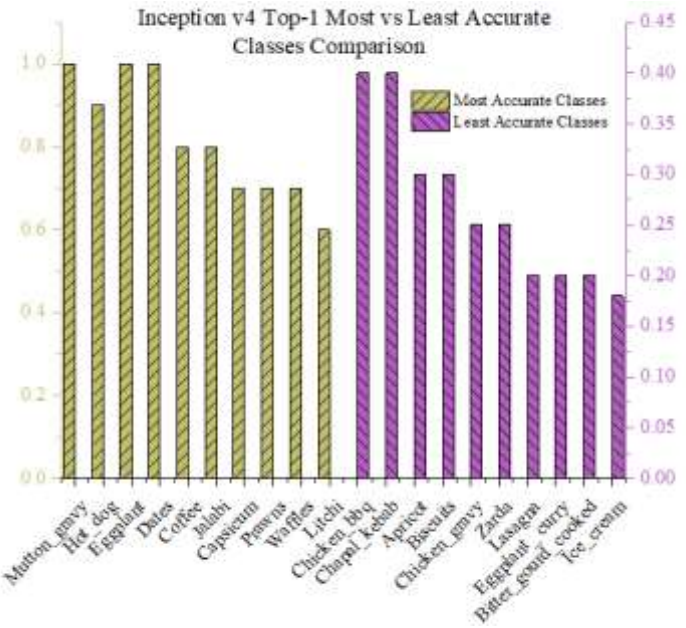
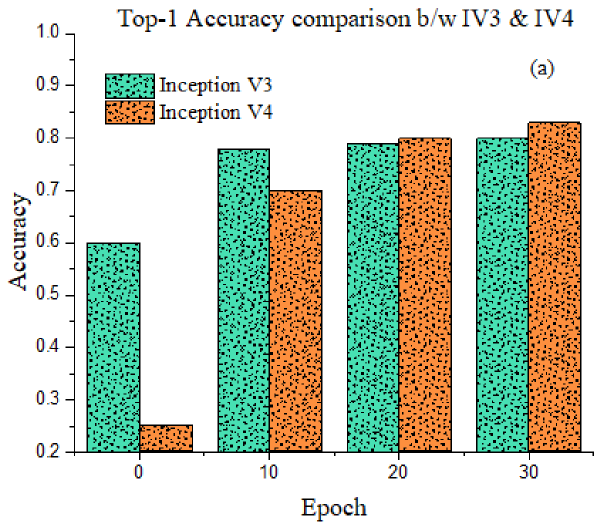


Figure 4. Comparison of top-1 most vs least accurate classes

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 787

, UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023



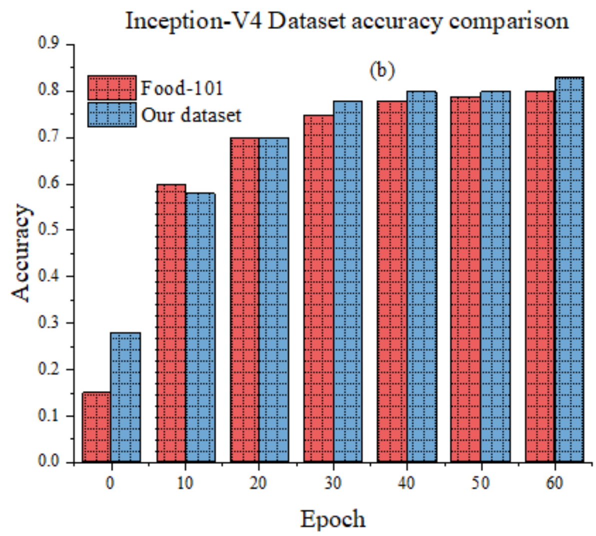


Figure 5. (a) Top-1 accuracy comparison; (b) Inception-v4 model based performance

<u>Network analysis</u>

Following retraining and convergence, the final prediction accuracy is determined by the choice of network models. In this case, the models that used here should be easily adaptable to food recognition. The two top ImageNet models, Inception V3 and Inception-ResNet, are contrasted in Figures 6 and 7.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 787

 UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

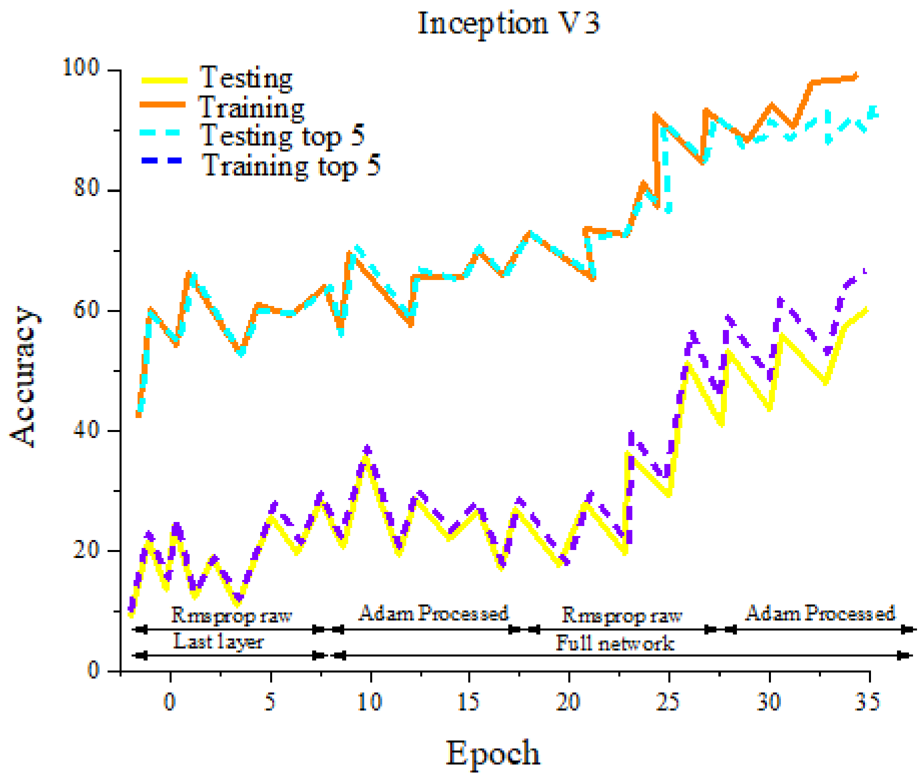


Figure 6. Results of Inception V3: Accuracy of each epoch

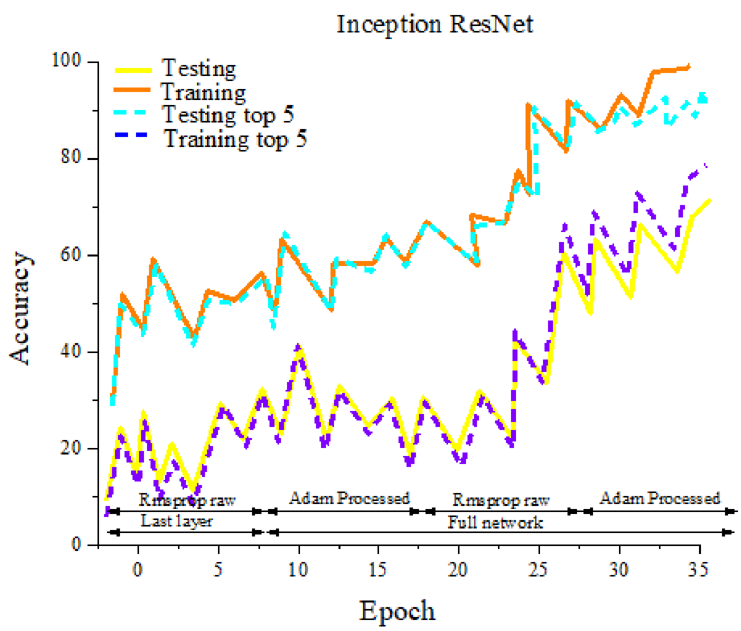


Figure 7. Results of Inception-ResNet: Accuracy of each epoch

First off, just retrain the model's final layer, or softmax layer. This indicates that food recognition can be broadly achieved using the attributes that were taken from ImageNet. As can be seen from the images, Inception-ResNet outperforms Inception V3 in terms of feature extraction after the last layer retraining has converged. Preprocessing does not enhance recognition, as evidenced by the comparison of raw and processed image input, as the earlier layers of those models have already acquired knowledge. Then retrain the model across all levels. Retraining on the complete network yields far better results than retraining solely on the softmax layer, as the Figures demonstrate. This implies that food recognition cannot be accomplished with the attributes that were taken from ImageNet. Additionally, preprocessed input is this time around 3% more accurate than raw input. This is due to the fact that specific preprocessing neurons are freed up to extract more food information, leading to an improved prediction. Furthermore, Inception V3 outperforms Inception-ResNet on entire network training. This implies that while the learning capacity of these two models is similar, Inception-ResNet's features are more adaptable. Finally, overfitting only happens when accuracy rises above 60%; the final overfitting value obtained was 3%.

Identifying different foods

Mixed physical pictures are not sufficiently identified and processed by current technologies. Foods that need to be cooked, liquid foods, and composite foods like salads and sandwiches are not included in them. In subsequent studies, a physical image resembling cooking and a mixed food image are processed by integrating image segmentation algorithms to address the issue of oblique edges in the image or each other causing the recognition detection to fail.

Improvement of Datasets and Systems

Features and data sets have a major impact on the detection's results. The few characteristics present in the current data sets—such as different backgrounds, lighting conditions, and camera angles—make them inadequate. An image processing technique that is faster than lookup tables is combined with a database to store calculated values, food labels, and other factors. Additionally, the architecture of the system and application is optimised.

Conclusion

Nowadays, being overweight is a serious problem in human life. People are observed to be curious to weigh themselves and eat healthily to prevent being overweight. Thus, this research introduces a novel method that provides us with information on the characteristics of the food we eat and its type. After accurately classifying the user-provided image of the food, the system provides information about the dish's properties. Our algorithm has been trained on a dataset comprising our subcontinental food and a typical meal of Food-101. In order to identify food items, have improved the Inception V-3 and V-4 model. Also suggested a way to assess the food's properties using attribute estimation. Multi-crop, data augmentation, and related methods improve the outcomes. Our suggested approach yields an exceptionally high accuracy of 85% for both classification and attribute extraction.

References

1. Shen, Z., Shehzad, A., Chen, S., Sun, H., & Liu, J. (2020). Machine learning based approach on food recognition and nutrition estimation. Procedia Computer Science, 174, 448-453.

2. Subhi, M. A., Ali, S. H., & Mohammed, M. A. (2019). Vision-based approaches for automatic food recognition and dietary assessment: A survey. IEEE Access, 7, 35370-35381.

3. Tahir, G. A., & Loo, C. K. (2021, December). A comprehensive survey of image-based food recognition and volume estimation methods for dietary assessment. In Healthcare (Vol. 9, No. 12, p. 1676). MDPI.

4. Salim, N. O., Zeebaree, S. R., Sadeeq, M. A., Radie, A. H., Shukur, H. M., & Rashid, Z. N. (2021, July). Study for food recognition system using deep learning. In Journal of Physics: Conference Series (Vol. 1963, No. 1, p. 012014). IOP Publishing.

5. Mezgec, S., Eftimov, T., Bucher, T., & Seljak, B. K. (2019). Mixed deep learning and natural language processing method for fake-food image recognition and standardization to help automated dietary assessment. Public health nutrition, 22(7), 1193-1202.

6. Kiourt, C., Pavlidis, G., & Markantonatou, S. (2020). Deep learning approaches in food recognition. Machine Learning Paradigms: Advances in Deep Learning-based Technological Applications, 83-108.

7. Oliveira Chaves, L., Gomes Domingos, A. L., Louzada Fernandes, D., Ribeiro Cerqueira, F., Siqueira-Batista, R., & Bressan, J. (2023). Applicability of machine learning techniques in food intake assessment: A systematic review. Critical Reviews in Food Science and Nutrition, 63(7), 902-919.

8. Mohanty, S. P., Singhal, G., Scuccimarra, E. A., Kebaili, D., Héritier, H., Boulanger, V., & Salathé, M. (2022). The food recognition benchmark: Using deep learning to recognize food in images. Frontiers in Nutrition, 9, 875143.

9. Yunus, R., Arif, O., Afzal, H., Amjad, M. F., Abbas, H., Bokhari, H. N., ... & Nawaz, R. (2018). A framework to estimate the nutritional value of food in real time using deep learning techniques. IEEE Access, 7, 2643-2652.

10. Chaturvedi, A., Waghade, C., Mehta, S., Ghugare, S., & Dandekar, A. (2020). Food Recognition and Nutrition Estimation Using Deep Learning.

11. Lo, F. P. W., Sun, Y., Qiu, J., & Lo, B. (2020). Image-based food classification and volume estimation for dietary assessment: A review. IEEE journal of biomedical and health informatics, 24(7), 1926-1939.

12. Sahoo, D., Hao, W., Ke, S., Xiongwei, W., Le, H., Achananuparp, P., ... & Hoi, S. C. (2019, July). FoodAI: Food image recognition via deep learning for smart food logging. In Proceedings of the

ISSN PRINT 2319 1775 Online 2320 787

, UGC CARE Listed (Group -I) Journal Volume 12, Iss 01 2023

25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2260-2268).Zhu, Lili, et al. "Deep learning and machine vision for food processing: A survey." Current Research in Food Science 4 (2021): 233-249.

13. Zhu, L., Spachos, P., Pensini, E., & Plataniotis, K. N. (2021). Deep learning and machine vision for food processing: A survey. Current Research in Food Science, 4, 233-249.

14. Ma, P., Lau, C. P., Yu, N., Li, A., Liu, P., Wang, Q., & Sheng, J. (2021). Image-based nutrient estimation for Chinese dishes using deep learning. Food Research International, 147, 110437.

15. Ma, P., Lau, C. P., Yu, N., Li, A., & Sheng, J. (2022). Application of deep learning for image-based Chinese market food nutrients estimation. Food Chemistry, 373, 130994.