

Calories Estimation of Food and Beverage Using Deep Learning-Based Image Analysis

E. Mahammad Irfan¹, B. Priyanka², C. Uday Kiran³, R. Vinod Kumar⁴,
K. Mahesh⁵

^{1,2,3,4,5}Department of CSE, Tadipatri Engineering College, Tadipatri.

Abstract

Accurate monitoring of caloric consumption is crucial for sustaining a healthy lifestyle and mitigating the risk of lifestyle-related ailments, including obesity and diabetes. Conventional methods for estimating caloric intake necessitate manual food logging, a process that is both time-intensive and frequently imprecise. To overcome this limitation, this study introduces a deep learning-based methodology for the automated classification of food items and subsequent caloric estimation, leveraging food images. The proposed system employs Convolutional Neural Networks (CNNs) in conjunction with image processing techniques to analyse food photographs taken under diverse conditions. Initially, a comprehensive dataset, encompassing images of food and beverages from various cuisines and presentation styles, is gathered and subjected to preprocessing, including resizing, normalisation, and segmentation techniques. A deep convolutional neural network (CNN) model is utilised to classify food items by extracting hierarchical visual features from raw pixel data. Subsequently, a regression model is employed to compute calorie values, utilising the features derived from the trained CNN. The results from the experiment show that the method proposed has a very high accuracy in food classification and a very good consistency in calorie prediction. In addition, the system allows for the tracking of weekly calorie intake the user's way, so the user can very well keep an eye on their eating habits. As a result, the suggested solution offers a practical and automated tool for nutritional evaluation, aiding individuals, dietitians, and healthcare professionals in promoting healthier eating habits.

Keywords: Food Image Classification, Calorie Estimation, Deep Learning, Convolutional Neural Networks, Image Processing, Dietary Monitoring

INTRODUCTION

The study in question looks into the problem of health issues linked to the lifestyle changes, and the most common offenders being obesity, diabetes, and heart-related ailments, among others. Bad eating habits plus a lack of knowledge about nutrition are the factors that make it worse. Overeating is one of the main reasons for this problem, and unfortunately, it is usually done without or with little oversight. Thus, very careful measurements apart from constant monitoring of the daily intake of calories are vital for one to live a healthy lifestyle and for the health problems to be less likely if not totally avoided. Traditional calorie-counting methods where users were required to log everything including foods and even the size of portions have been generally regarded as very slow, burdensome, and also very inaccurate because of human mistakes. The growth of AI and computer vision has led to the possibility of executing dietary assessments automatically by looking at food pictures. As a result, the practicality of other techniques for

picture-based dietary assessment is meanwhile coming up. The vast majority of the population has smartphones that come with top-notch cameras, therefore the taking of meal images is fairly straightforward, and building on this, spontaneous analysis of the food items is truly possible. In this regard, one of the most sought-after deep learning model types, CNNs, has reached extraordinary levels in numerous image recognition and categorization tasks by self-learning the hierarchical visual features (like texture, shape, and color) from raw pixel data. It suggests a deep learning-based architecture that can classify food and beverages along with estimating their caloric content from an image. The framework combines CNN-powered food acknowledgment with caloric prediction based on regression for the purpose of determining the caloric content of goods from pictures. Furthermore, the framework adds a weekly tracking module to monitor the intake pattern through time.

LITERATURE SURVEY

Deep learning has been the cause of great improvements in the performance of image-based detection and recognition systems. As the main technique for visual feature extraction, Convolutional Neural Networks (CNNs) are characterized by their ability to develop spatial hierarchies from raw image data as well as to uncover the best features. An example of the integration of CNN-based number plate recognition with an IoT vehicle registration system for real-time application is given by Jawale et al. Their study brought out the strengths of deep learning against the variation in lighting and backgrounds which is very crucial for image analysis-based systems.

Selmi et al. delivered the DELP-DAR framework, and it was an intelligent combination of detection and recognition stages that resulted in a unified pipeline for license plate recognition. Their method focused on preprocessing and segmentation's vital role in improving recognition accuracy. Likewise, Huang and others brought forth the development of a single deep neural network that could detect mixed-style license plates and this reduced the reliance on multiple specialized models while at the same time improving the overall computational efficiency. All of these studies indicate that end-to-end CNN architectures are the ones that offer better performance than the traditional image processing techniques. Being able to get higher recognition accuracy has been attributed to the architectural enhancements. Pattanaik and Balabantaray have tried out the conventional Xception architecture along with the Mish activation function, obtaining considerable improvements in recognition performance. Henry and others have taken CNN-based recognition to international license plates thus resolving issues connected to character diversity and layout variations. Li et al. have presented a deep learning framework that jointly performs license plate detection and recognition in an end-to-end manner. The detection and recognition speed for real-time applications has been the main subject of several studies. Zou et al. used YOLOv3 in combination with a deep recognition network to make the license plate detection and recognition process more efficient. Kaur et al. demonstrated that the CNN-based systems are capable of automating the whole process of vehicle monitoring whereas Kulkarni et al. carried out a detailed review of the traditional and modern recognition techniques, exposing the common issues such as illumination variation, occlusion, and diversity of datasets. New methodologies based on multi-tasking and edge techniques have been introduced in recent research as well. Alim et al. proposed an edge system that, at the same time, recognizes license plates and detects faces which results in reduced latency and increased scalability. Agarwal et al. coupled license plate detection with facial feature recognition to facilitate the identification of suspicious vehicles, which is indicative of the trend towards integrated intelligent systems. The importance of robust feature extraction and optimization in visual recognition tasks is further highlighted

by the techniques, such as preprocessing under varying illumination conditions and optimization strategies for text detection, discussed by Anila et al. and Kohli et al. respectively. Even though the current literature makes it very clear that CNN-based models are very effective for tasks related to detection and recognition, the majority of such studies still only consider categorical identification and do not address the issue of estimating quantitative attributes from images. Moreover, the existing systems often only support single-task recognition and are not equipped with mechanisms for continuous data aggregation, trend analysis, and user-centric insights. The lack of integrated regression-based estimation and long-term monitoring severely restricts their use in the health and lifestyle areas.

PROPOSED SYSTEM

The system suggested will identify a given food item, automatically using an image and determine the calorie content of a food item simply and efficiently. Rather than entering by hand what we are eating, the user is given an opportunity to capture a picture of their meal on a mobile device. This image is processed by the system which individually cleans the image and makes the necessary adjustments to it (resizing and normalizing) so that it can be analyzed appropriately. The type of food is then identified by a Convolutional Neural Network (CNN) model due to the ability of this model to learn patterns like features of the food (shape, color and texture).

The system, after identifying the food item in question, estimates its calorie content, by way of a trained regression model, which is operational based on the set of features detected by the CNN. The results of the calculations are presented in form of calories to the user. Moreover, this data is stored in the system, and users can monitor the number of calories they eat each day or each week to learn about their consumption patterns. All in all, this program offers an effective, fast, and fair method to keep track of food intake and calorie-only that are easy to use without any type of manual labor.

SYSTEM ARCHITECTURE

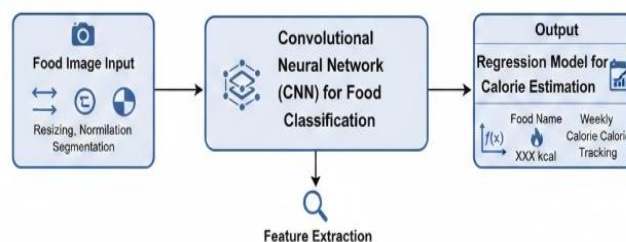


FIG. 1. SYSTEM ARCHITECTURE OF THE PROPOSED FOOD CALORIE ESTIMATION SYSTEM

A. Image Acquisition

System's first phase is such as acquiring food and drink photos either using a mobile phone camera or importing photos from a local hard disk. The images might differ a lot with respect to light, background clutter, and viewing angle. However, their presentation styles may also vary. The system is purposely made to withstand such variations and to be able to generalize nicely across different types of cuisines and foods thereby guaranteeing its robustness

B. Image Preprocessing

- Image preprocessing is the process which plays the most critical role in the enhancement of classification accuracy and performance of calorie predictions. Every input image is processed through the following steps:
- Image Resizing: All the images are brought down to a fixed resolution for uniformity and compatibility with the CNN input layer.
- Normalization: Pixel values are set to a common scale so that variations caused by lighting will have less effect and the model will converge faster.
- Noise Reduction: Audio and visual signals that come but are not wanted are filtered out thus improving the quality of the video signal.
- Image Segmentation: Special methods are used to divide the food part from the back of the picture, thus allowing the model to pay attention to the most important features related to food.
- Data Augmentation: To make the dataset more diverse and to avoid overfitting, techniques like rotation, horizontal flipping, zooming, and translation are done.

C. Food Classification Using Convolutional Neural Networks

The image, after going through preprocessing, is passed through a deep Convolutional Neural Network (CNN) that has been specifically trained to classify foods. Convolutional Neural Network(CNN) is built with multiple nodes that are convolutedly, being followed by pooling and finally connected layers.

- The low-level and high-level characteristics such as edges, textures, shapes, and color patterns are extracted by the convolutional layers.
- The pooling layers keep the significant data and lessen the spatial dimensions.
- The fully connected layers execute the last classification by matching the features extracted to the already defined categories of food.

The model of Convolutional Neural Network (CNN) is trained with food images that are labeled, and proper loss function and optimizer are applied to get classification accuracy as high as possible.

D. Feature Extraction

After training the CNN model, from the last convolution or intermediate layers, deep feature vectors are extracted. These are very good and unique visual characteristics of the food item. Not only that, the reuse of these extracted features as inputs of the calorie estimation module, thus prediction is made to be efficient and accurate.

E. Calorie Estimation Using Regression Model

Calorie estimation is treated as a regression problem. The deep features which are drawn from CNN are given to a regression model that estimates the approximate calorie contents of the food item. The model of regression establishes the connection of different visual attributes with the varying calorie contents, from the training data. This method does away with the need for manual input of portion size and thus, provides a solution that's easily applicable and is effective in calorie prediction across a wide range of food categories

F. Calorie Tracking and Weekly Analysis

TIn a database, along with timestamps, the expected calorie values are saved. Total daily and weekly

calorie consumption is calculated by the system using individual food entries. Weekly calorie consumption reports are created for the users to make them aware of their consumption habits and to help them spot areas of unhealthy diets. The system is capable of giving health-related insights based on the established thresholds that are mainly aimed at promoting the development of good eating habits.

G. Algorithm:

CNN-Based Food Classification and Calorie Estimation

- Submit image of food or drink
- Resize and normalize the image
- Extract food area using segmentation
- Augment image data if needed
- Use trained CNN to identify the food item
- Obtain deep feature vectors from CNN
- Use regression model to predict calories
- Save the predicted calorie amount
- Calculate daily and weekly calorie intake
- Show classification and calorie estimation results

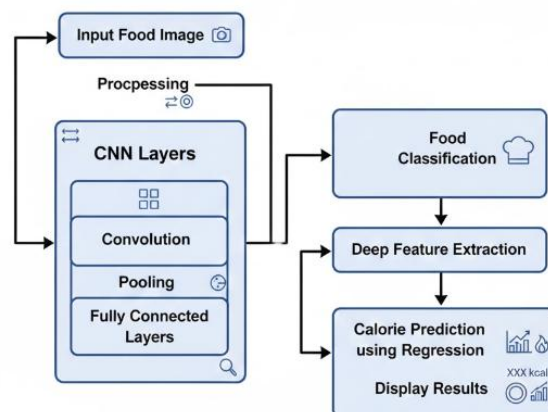


Fig. 2. Workflow of CNN-based food classification and calorie estimation

RESULTS AND DISCUSSION

The proposed system has been tested with a wide range of food images that are taken in a variety of real-life factors, such as the change of light, angles, and backgrounds. The Convolutional Neural Network (CNN) has been shown to perform well in the accurate recognition of food items by learning effectively visual features of the item i.e. color, texture, and shape. The calorie estimation aspect also gave a good and consistent results in comparison with actual calorie value. Altogether, the system demonstrated a high level of accuracy in the food classification sphere and acceptable calorie prediction performance, which is why the system can be used in practice.

Nevertheless, the system demonstrated that there were slight constraints to processes of images containing more than one item of food, unclear lighting environments, and objects that were partially obstructed that compromised marginally the level of prediction. Regardless of the challenges, the system was stable and efficient that gave fast results and allowed users to follow their daily calorie intake as well as weekly

calorie intake. The feature will assist the users learn more about their eating habits and make healthier lifestyle decisions. Finally, the system is an efficient and easy-to-use automated solution to recognizing food and estimating calories, which can be improved in the future to work with complex situations.

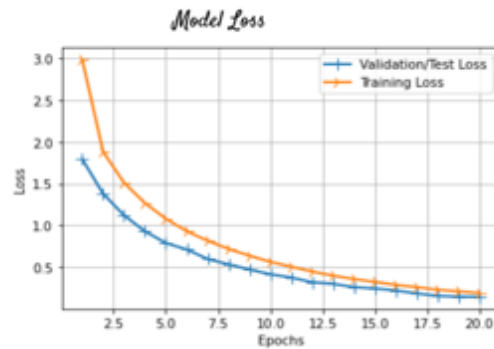


FIG 2. MODEL LOSS ROC GRAPH

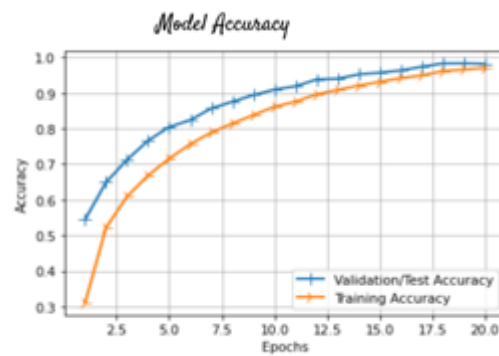


FIG 3. MODEL ACCURACY ROC GRAPH

Nutrition Report

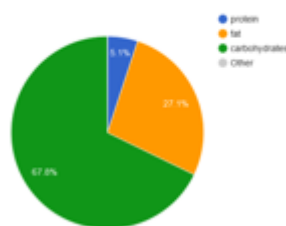


FIG 4. PIE CHART

Nutrition Report

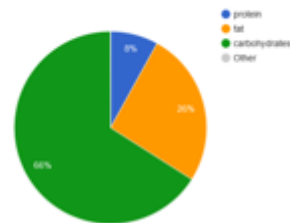


FIG 5. PIE CHART

SCREENSHOTS



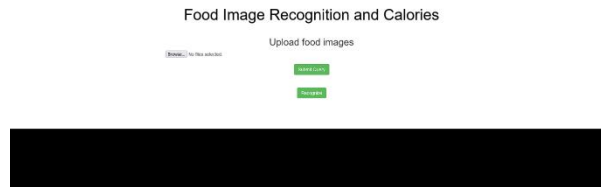
FIG 6. INDEX PAGE



Food Image Recognition and Calories

Login

FIG 7. LOGIN PAGE

**FIG 8.LOGIN PAGE**

Food Image Recognition and Calories
Result



Apple pie- Estimate Calories 237
For 100 Grams Quantity

FIG 9. RESULT PAGE

CONCLUSION AND FUTURE WORK

This research highlighted an automated deep learning model for the classification of food and beverage and the eventually, the calorie estimation of the same through image analysis. The usage of Convolutional Neural Networks for food identification along with regression-based calorie prediction made the diet assessment very precise and there was absolutely no need for calorie manual logging. In their results, the promoted system was shown to be dependable in terms of food classification accuracy and calorie estimation precision, thus, being a suitable device for dietary control and health management. On the other hand, the present-day system, while effective, still has some limitations. One of the major reasons behind this is the varied estimation of calories due to being dependent on one image interpretation, which means that portion size, food arrangement, and occlusion would also play a role in the final accuracy of estimation. Besides, the model effectiveness is also directly related to the training dataset's quality and diversity. In the next phase of the project, the researchers plan to add portion size estimation through depth sensors or multi-views images, user-specific dietary profiles, and widen the dataset by including more cuisines. Further, the system is to be improved even more by introducing it as a real-time mobile app that comes with personalized nutritional guidance.

REFERENCES

1. Jawale, M. A., P. William, A. B. Pawar, and Nikhil Marriwala. "Implementation of Number Plate Detection System for Vehicle Registration Using IoT and Recognition Using CNN." *Measurement: Sensors* 27 (2023): 100761. <https://doi.org/10.1016/j.measen.2023.100761>.

2. Selmi, Zied, Mohamed Ben Halima, Umпада Pal, and M. Adel Alimi. “DELP-DAR System for License Plate Detection and Recognition.” *Pattern Recognition Letters* 129 (2020): 213–223. <https://doi.org/10.1016/j.patrec.2019.11.012>.
3. Huang, Qiuying, Zhanchuan Cai, and Ting Lan. “A Single Neural Network for Mixed Style License Plate Detection and Recognition.” *IEEE Access* 9 (2021): 21777–21785. <https://doi.org/10.1109/ACCESS.2021.3055721>.
4. Pattanaik, Anmol, and Rakesh Chandra Balabantaray. “Enhancement of License Plate Recognition Performance Using Xception with Mish Activation Function.” *Multimedia Tools and Applications* 82, no. 11 (2023): 16793–16815. <https://doi.org/10.1007/s11042-022-13684-4>.
5. Henry, Chris, Sung Yoon Ahn, and Sang-Woong Lee. “Multinational License Plate Recognition Using Generalized Character Sequence Detection.” *IEEE Access* 8 (2020): 35185–35199. <https://doi.org/10.1109/ACCESS.2020.2974702>.
6. Li, Hui, Peng Wang, and Chunhua Shen. “Toward End-to-End Car License Plate Detection and Recognition with Deep Neural Networks.” *IEEE Transactions on Intelligent Transportation Systems* 20, no. 3 (2019): 1126–1136. <https://doi.org/10.1109/TITS.2018.2847291>.
7. Kaur, Parneet, Yogesh Kumar, Shakeel Ahmed, Abdulaziz Alhumam, Ruchi Singla, and Muhammad Fazal Ijaz. “Automatic License Plate Recognition System for Vehicles Using a CNN.” *Computers, Materials & Continua* 71, no. 1 (2022): 759–775. <https://doi.org/10.32604/cmc.2022.020276>.
8. Zou, Yongjie, Yongjun Zhang, Jun Yan, Xiaoxu Jiang, Tengjie Huang, Haisheng Fan, and Zhongwei Cui. “License Plate Detection and Recognition Based on YOLOv3 and ILPRNET.” *Signal, Image and Video Processing* 16, no. 2 (2022): 473–480. <https://doi.org/10.1007/s11760-021-02014-4>.
9. Kulkarni, Parag, Amit Khandebharad, Dattatray Khope, and Pramod U. Chavan. “License Plate Recognition: A Review.” In *Proceedings of the Fourth International Conference on Advanced Computing (ICoAC)*, 1–8. IEEE, 2012. <https://doi.org/10.1109/ICoAC.2012.6416850>.
10. Quraishi, Abdul Awal, Farid Feyzi, and Asadollah Shahbahrami. “Detection and Recognition of Vehicle Licence Plates Using Deep Learning in Challenging Conditions: A Systematic Review.” *International Journal of Intelligent Systems Technologies and Applications* 22, no. 2 (2024): 105–150. <https://doi.org/10.1504/IJISTA.2024.135274>.
11. Alim, Fikret, Enes Kavakli, Sefa Burak Okcu, Ertugrul Dogan, and Cevahir Cigla. “Simultaneous License Plate Recognition and Face Detection Application at the Edge.” In *AI and Optical Data Sciences IV*, vol. 12438, 232–243. SPIE, 2023. <https://doi.org/10.1117/12.2650253>.
12. Agarwal, Vrinda, Aaron George Pichappa, Manideep Ramisetty, and BalaMurugan M. S. “Suspicious Vehicle Detection Using Licence Plate Detection and Facial Feature Recognition.” *arXiv preprint arXiv:2304.14507* (2023). <https://arxiv.org/abs/2304.14507>.
13. Anila, Satish, and N. Devarajan. “Preprocessing Technique for Face Recognition Applications under Varying Illumination Conditions.” *Global Journal of Computer Science and Technology* 12, no. 11 (2012): 13–18.
14. Kohli, Himani, Jyoti Agarwal, and Manoj Kumar. “An Improved Method for Text Detection Using Adam Optimization Algorithm.” *Global Transitions Proceedings* 3, no. 1 (2022): 230–234. <https://doi.org/10.1016/j.gltp.2022.03.029>.
15. Murthy, G. S. N., Shanmukha Rao Allu, Bhargavi Andhavarapu, Mounika Bagadi, and Mounika Belusonti. “Text-Based Sentiment Analysis Using LSTM.” *International Journal of Engineering Research & Technology* 9, no. 5 (2020): 1204–1208