FoodScan: Food Image Detection and Analysis

Parisha Agrawal[1], Saurabh Kumar[1], Annu Kumari[1]

Abstract

This report details the development of FoodScan, a web application designed to empower users with information about their food choices by combining machine learning with data-driven analysis. The system classifies dishes from uploaded images, retrieves relevant recipes and provides detailed nutritional information, making informed dietary decisions simple and intuitive. Using the Food-101 dataset, an InceptionV3-based model was trained and deployed for food detection. The application integrates APIs for recipe retrieval and nutrient comparison while offering a user-friendly web interface for navigation and interaction. This project demonstrates the practical integration of advanced technologies to bridge the gap between food recognition and actionable dietary insights.

1. Introduction

Making informed dietary choices can be challenging. Often, individuals lack knowledge about the nutritional content of their meals. FoodScan aims to simplify this process by providing a user-friendly platform and bridges the gap between the food we consume and an understanding of its nutritional value. With the growing importance of maintaining a balanced diet, FoodScan serves as a tool to aid users in making informed dietary decisions. By leveraging machine learning, recipe, and nutrient data, the platform not only identifies dishes from images but also provides recipes and nutritional breakdowns. The primary objectives of the project include dish classification from uploaded images, retrieval of recipes and detailed nutritional information, comparison of dish nutritional values with recommended daily intake guidelines, and the provision of features such as nutrient comparison and diet planning.

2. Literature Review

The intersection of computer vision and food science has led to significant advancements in food image analysis. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have been widely adopted for food image classification and recognition[1].

Transfer Learning: Pre-trained CNN models like InceptionV3 and ResNet have been successfully adapted for food image classification[2]. These models, initially trained on large-scale image datasets, can be fine-tuned on smaller, food-specific datasets to achieve high accuracy.

Custom Architectures: Researchers have proposed custom CNN architectures tailored for food image analysis. These architectures incorporate specialized layers and techniques to capture the unique visual characteristics of food items.

Challenges and Future Directions:

Data Quality and Quantity: The availability of large, high-quality food image datasets remains a significant challenge.

Handling Food Variations: Food items can exhibit significant variations in appearance due to cooking methods, presentation styles, and cultural differences.

Real-time Processing: Developing efficient and real-time food image analysis systems is essential for practical applications like mobile food recognition apps.

Multimodal Learning: Integrating information from multiple modalities, such as text and sensory data, can further enhance food image analysis.

3. Dataset

The Food-101 and Indian Food Images Dataset was utilized for this project. Food-101 dataset, comprises 101 food categories with over 100,000 labeled images, is widely recognized for its diversity and utility in food classification tasks. And Indian Food Images Dataset, das 4000 Indian Food Images in 80 different categories or classes. Images were resized to match the input dimensions required by the InceptionV3 model. Preprocessing involved augmentation techniques such as rescaling, shear, zoom, and horizontal flipping to improve the model's generalization capabilities during training.

4. Methodology

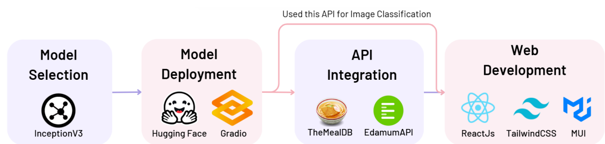

Figure 1: Workflow

4.1. Model Selection and Training

The InceptionV3 pre-trained model was utilized for feature extraction due to its robust performance in image classification tasks [2]. The model was further modified with custom layers, including global average pooling for dimensionality reduction, dense layers for classification, and a softmax output layer. Data augmentation techniques like rescaling, shear, zoom, and horizontal flipping were employed to enhance the model's generalization capabilities. L2 kernel regularization and dropout layers were added to prevent overfitting during training. Stochastic Gradient Descent (SGD) with a learning rate of 0.0001 and momentum of 0.9 served as the optimization algorithm. Callbacks like ModelCheckpoint and CSVLogger were used to monitor and save the best-performing model iterations.

4.2. Model Deployment

The trained model was deployed as a backend API using the Hugging Face platform and further integrated with Gradio for a user-friendly interface.

4.3. API Integration

The FoodScan application utilizes two external APIs. TheMealDB API allows for retrieving relevant recipes based on the identified dish [7]. Additionally, RapidAPI's Edamum API provides detailed nutritional information for various ingredients.

4.4. Web Development

A user-friendly web interface was built using React framework, Tailwind CSS for styling, and Material-UI components for interactive elements. This enables users to upload food images, view classification results, and access detailed nutritional information, recipe suggestions, and diet planning tools.

5. Result

The InceptionV3 model achieved a classification accuracy of 78% on the Food-101 and Indian Food Images Dataset, demonstrating its effectiveness in food recognition tasks. The web application successfully integrates features such as dish classification, recipe retrieval, and nutritional analysis. Users can upload images of dishes and receive accurate classification results along with relevant recipes and detailed nutrient breakdowns. Nutrient comparison tools provide actionable insights, while the diet planning features offer personalized recommendations based on individual health profiles.

Plots such as training and validation accuracy/loss curves demonstrated the model's convergence during training.
A confusion matrix was generated to visualize classification performance across the food categories. Nutritional values were presented graphically, allowing users to compare dish compositions against recommended daily intake values.

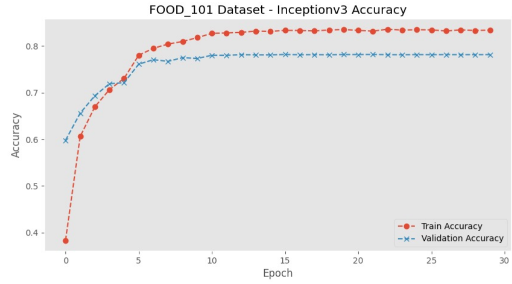

Figure 2: InceptionV3 Accuracy

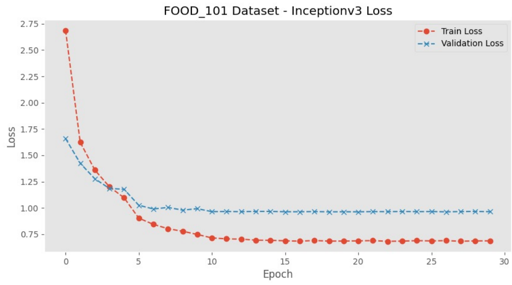

Figure 3: InceptionV3 Loss

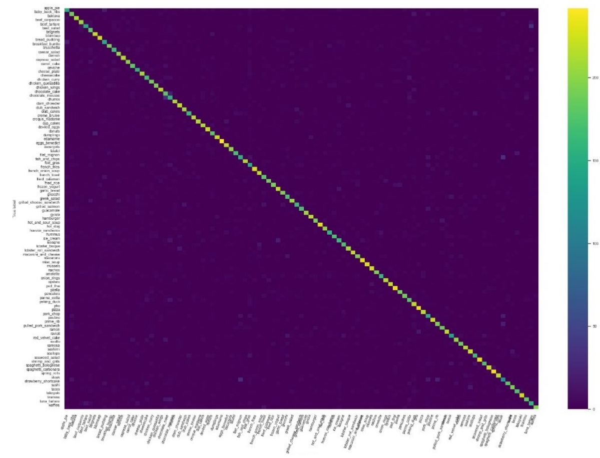

Figure 4: Confusion Matrix

Figure 5: Predictions

6. Webserver

FoodScan's web application integrates both the backend API and frontend for seamless functionality. The backend, deployed on Hugging Face, hosts the trained model FlaskAPI. Gradio was used for testing and demonstration purposes. The frontend, built using React, offers an interactive and responsive interface. Features such as image upload, result visualization, and navigation to recipes and nutritional analysis are easily accessible to users, ensuring a smooth experience. Website

7. Future Work

Future developments could focus on enhancing model accuracy by training on a larger and more diverse dataset. Expanding the dataset to include more food categories and regional dishes would improve the application's versatility. Real-time food detection through a mobile application could provide users with instant recognition capabilities. Additional features such as a calorie tracker and meal history logs would further enrich the user experience. Integration with wearable devices for personalized dietary suggestions and fitness tracking could be explored, making FoodScan a comprehensive health and nutrition platform.

References

[1] Salim, S., Aini, N. A. (2021). Food Detection and Recognition Using Convolutional Neural Network. Journal of Physics: Conference Series, 1963(012001), 1-6. https://www.researchgate.net/publication/305513345$_{F}ood_{D}etection_{a}nd_{R}ecognition_{U}sing_{C}onvolutional_{N}eural_{N}etwork$

[2] Dhawan, S., Singh, K., Arora, M. (2021). Cervix Image Classification for Prognosis of Cervical Cancer using Deep Neural Network with Transfer Learning. EAI Endorsed Transactions on Pervasive Health and Technology, 7(27), e5. https://doi.org/10.4108/eai.12-4-2021.169183

[3] Kokare, S., Kandekar, P. (2024). Food Detection and Nutrition Analysis by Image. International Journal of Research in Applied Sciences and Engineering Technology, 11(6), 4257-4261

[4] Food Images (Food-101). (2018, May 15). Kaggle. https://www.kaggle.com/datasets/kmader/food41

[5] Free meal API — TheMealDB.com. (n.d.). https://www.themealdb.com/api.php

[6] Nutrition Analysis API – Documentation – Edamam. (n.d.). https://developer.edamam.com/edamam-docs-nutrition-api

[7] Rapid. (n.d.). section-1. https://rapidapi.com/