A HEALTH-LIFE RECOMMENDATION SYSTEM: FOOD IMAGE RECOGNITION BASED APPROACH

Hesam Abbaszadi

z5403577

ELEC9771 - Project Report A

Supervisor - Chaojie Li

The University of New South Wales

3/8/2023

Abstract. Healthy lifestyle is pursued by many people in our society. However, our daily diet includes plenty of junk food which may pose a risk to our health. In this research, a food-AI system was developed to recognise the daily diet which can extract nutrition information for customers and recommend some healthy food to them. Food data was collected, and the useful information was already labeled in each picture. At the end, by leveraging Xception, an image recognition system was built for recognizing food images with 74.84% accuracy on testing set. This is important for facilitating a healthy lifestyle for individuals in the future.

Keywords: Food recognition, Xception model, Computer vision, Deep learning, Image classification

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Chaojie Li, for his continuous support, patience, and guidance throughout the course of this project. His knowledge and passion for the subject have been invaluable to my work. I would also like to thank the staff and faculty at UNSW Sydney for providing an encouraging and stimulating environment, and for providing the opportunity to carry out this project. Additionally, my appreciation goes to Google Colab for providing the necessary computing power to train my machine learning models. I am also grateful to the creators and maintainers of the food-101 dataset, which played an essential part in the completion of this project. Lastly, I wish to express my thanks to my family and friends for their understanding, encouragement, and support throughout my academic journey.

Table of Contents

A Health-Life Recommendation System: Food Image Recognition Based Approach

1 Introduction

The pursuit of a healthy lifestyle is a shared aspiration among many, a commitment that remains ever significant in the wake of increasing health concerns globally. Balancing this desire, however, is the widespread prevalence of unhealthy dietary habits, often compounded by the easy availability and appeal of processed and junk foods. These lifestyle choices, though seemingly harmless at first, can over time pose profound health risks. The challenge, therefore, lies not just in raising awareness about healthy foods, but also in facilitating informed dietary decisions.

Amidst these challenges, the technological realm has seen remarkable advances, particularly in the domain of Artificial Intelligence (AI). AI's transformative potential is evident across sectors, and when it comes to healthcare and lifestyle management, its role becomes even more pivotal. Specifically, AI has the potential to assist individuals in making better dietary choices by offering insights driven by data, leading them towards a healthier lifestyle. However, the application of AI in food choices so far has been limited.

This project, which has two main parts, seeks to harness the power of AI for food recognition and health recommendations. In Part A, I delve into the development and fine-tuning of a computer vision model – based on the renowned Xception architecture – to recognize and categorize foods from images. By training this model on the expansive food-101 dataset, I created a robust system that can accurately identify a vast array of food items.

In the subsequent phase, Part B, the aim is to evolve from mere recognition to meaningful recommendations. Using the identified food categories from Part A as a foundation, a recommendation system will be constructed, one that offers nutritious food suggestions tailored to individual needs. Thus, while the immediate goal is food recognition, the overarching vision is to foster healthier food choices by integrating insightful AI-backed recommendations. As I journey through this report, I will shed light on the methodologies, technologies, and strategies that underpin this project, emphasizing the significance of AI in promoting healthier dietary habits.

2 Literature Review

During the literature review phase of this project, the "Stanford University Computer Vision" course available online served as an invaluable resource. This course, comprising 16 lectures, provided an in-depth overview of the core concepts and techniques in computer vision. The topics ranged from fundamental aspects such as image formation, feature detection, and image segmentation, to more advanced subjects like object recognition, scene understanding, and deep learning in vision. Notably, the lectures on convolutional neural networks and deep learning were particularly relevant for this project.

The role of AI in recognizing patterns and making insightful predictions has been vastly explored in various scientific studies. Within this expansive realm of research, the intersection of AI and dietary habits is of particular interest to my study. My approach draws inspiration from the research paper titled "Food-101 – Mining Discriminative Components with Random Forests" by Bossard, Guillaumin, and Van Gool (2014). This study provides significant insights into food recognition using machine learning algorithms.

In their research, Bossard et al. have sought to push the envelope of fine-grained categorization, focusing their efforts on the domain of food recognition. Their work employs the food-101 dataset, comprising 101 food categories with 101,000 images, to train a Random Forests model. The model harnesses the strength of this extensive dataset to mine discriminative components and, ultimately, distinguish between varied food items. This is noteworthy because food recognition poses unique challenges to computer vision, given the diverse presentation and preparation styles of food items.

The core methodology of their research, utilizing a large, diverse dataset to train an algorithm for recognition tasks, aligns with my approach. However, instead of Random Forests, my study leverages the Xception model, a convolutional neural network known for its performance in image classification tasks (Chollet, 2017). The choice to use Xception is driven by its design advantages and its previous successes in large-scale image recognition, such as its pre-training on ImageNet, a large database with more than 14 million annotated images.

While Bossard et al.'s research significantly contributes to food image recognition, it does not extend to providing health recommendations based on the recognized food items. Here, my study aims to bridge the gap. I aspire to build upon the foundation of food recognition and introduce a layer of intelligent, health-conscious food recommendations, thereby adding practical value for individuals seeking healthier dietary choices. My study, therefore, is a blend of the foundational principles established by Bossard et al. and the novel application of the Xception model, coupled with the goal of meaningful health recommendations. In the following sections, I will delve deeper into the methodologies and implementation details that bring this vision to life.

3 Food Recognition System

3.1 Problem Definition

In the digital age, where visual data is prevalent, harnessing the power of image recognition opens a world of possibilities, including in food recognition (Bossard et al., 2014). The ability to accurately identify different food items from an image and extract meaningful nutrition information can significantly contribute to better dietary choices and health management. However, the task is not without challenges. Food items are vastly diverse in appearance, varying greatly in shape, colour, texture, and presentation style. These variables add complexity to the task of image recognition, requiring robust and sophisticated models capable of learning intricate patterns and recognizing subtle differences. The existing models, while making headway in this domain, are yet to reach a level of accuracy that makes them universally reliable and applicable in real-world scenarios. The objective of this part of the study is to develop a food recognition model that can overcome these challenges and deliver high-accuracy results.

3.2 Methodology and Implementaion

To tackle the food recognition problem, the study adopted a two-part approach: data acquisition and model training. Given the need for a large, diverse set of food images, the food-101 dataset was chosen. This dataset, used in the study by Bossard et al., comprises 101,000 images spread across 101 food categories. The images were pre-labelled, which eased the process of supervised learning. For the model, the study chose the Xception architecture. A deep learning model pre-trained on ImageNet with 36 convolutional layers, Xception is known for its robust performance in image classification tasks. The choice to use Xception stemmed from its architectural advantages and proven successes in large-scale image recognition tasks. Training the Xception model on the food-101 dataset, therefore, provided a balanced blend of a sophisticated model and a diverse dataset.

The implementation of the food recognition system took place on Google Colab Pro, leveraging the computational power of NVIDIA A100 GPUs. The Xception model was imported using Python 3 and TensorFlow library. The food-101 dataset was split into a training set and a testing set, with 75,750 images used for training the model and the remaining 25,250 images set aside for testing. To train the model on a larger dataset, extra images were generated from the existing images in training set, using 'ImageDataGenerator' command with a batch size of 100. Given the lack of bounding boxes in the dataset, the model was trained for whole-image classification rather than object detection.

After calling Xception, a few changes were made to the model. Activation functions were changed from ReLU function to Sigmoid function for the last three layers since this found to be effective for model accuracy while training the model. The labels accompanying the images were used as targets in the training process. The model was trained twice. The first time, the model was trained for 60 epochs, with 23 steps for each epoch, with a learning rate of 0.001; after training which took about 33 minutes, the model accuracy started to converge after 51st epoch, settling at 67.04% on 60th epoch. The model was saved. The second time, the saved model was trained again for 60 epochs, with 47 steps per epoch this time and a learning rate of 0.0001; a process that took about 66 minutes to complete and consistently increased the model accuracy, topping up at 82.02% on 60th epoch.

3.3 Results and Evaluation

After training, the model was tested on testing set with a batch size of 32 and achieved an accuracy of 74.84%. This is a promising result, considering the lack of bounding boxes in the labels and the complexity and variability inherent in food images. The accuracy of the model was assessed based on its ability to correctly classify the food item present in an image.

4 Health Recommendation System

4.1 Problem Definition

Recognizing food items from images is only part of the equation in aiding individuals in making healthier dietary choices. The real value lies in offering insightful recommendations based on this recognition. This facet of the project, which will be developed next term, aims to construct a recommender system that provides users with healthier alternatives to their current food choices, thus encouraging a more balanced and nutritious diet. The challenges in building such a system are manifold. Not only does the system need to provide relevant and practical suggestions, but it must also consider a variety of factors such as dietary preferences, allergies, cultural considerations, and nutritional balance. Furthermore, the recommender system must be user-friendly and engaging to encourage sustained use and long-term impact on dietary habits.

4.2 Proposed Methodology and Implementation

Given the challenges, a comprehensive approach will be required to develop the health recommendation system. First, the nutritional information of the food items recognized by the Xception model will need to be extracted. This information can then be analysed to identify healthier alternatives. Moreover, user input will also be considered, such as dietary preferences and restrictions. This user information will be integrated with the recognized food data to tailor the recommendations to individual users. Linear regression algorithms could potentially be used to make the system adaptive, learning from user feedback to improve the relevance and personalization of the recommendations over time.

The health recommendation system will be implemented using Python 3, given its vast library support for AI and machine learning. The detailed implementation will be defined as the project progresses, but a possible approach could include using the Scikit-learn library for developing the recommendation algorithm. For nutritional information, a database with a comprehensive list of food items along with their nutritional details could be employed. Additionally, user interaction components will be designed for data input and recommendation display. The final system aims to be intuitive and user-friendly, guiding the user through the process of capturing a food image, waiting for the system to recognize the food and, finally, receiving a list of healthier food recommendations.

4.3 Anticipated Challenges and Proposed Solutions

Developing a health recommendation system poses several challenges. One primary challenge is ensuring the accuracy and relevance of the recommendations. This could be addressed by integrating a feedback loop into the system, allowing users to rate the relevance of the suggestions, which the system can learn from over time. Privacy is the main concern as the system will handle personal dietary information. Adherence to privacy standards will be a priority during development, ensuring that data is handled securely and ethically. Open sourcing the code would be one reliable solution to this problem. Balancing customization and general health guidelines will also be critical. While the aim is to tailor recommendations to individuals, the system should also align its suggestions with universally accepted nutritional guidelines. The system's engagement factor is also crucial. Strategies to maintain user engagement, such as a friendly user interface and a reward system for following the recommendations, could be employed. These challenges, among others, will be addressed in the next stage of this project.

5 Timeline and Future Work

5.1 Timeline

This project's execution was split into two phases aligned with the academic terms. Term 2 was dedicated to the development of the food recognition system, while the Term 3 is dedicated to the development of the health recommendation system. Here is a summary of the revised timeline:

- Term 2:
  - Weeks 1-2: Project planning and literature review
  - Weeks 3-5: Data acquisition and pre-processing
  - Weeks 6-8: Model training and evaluation
  - Weeks 9-10: Final report writing and project report for Part A
- Term 3:
  - Weeks 1-2: Planning for the health recommendation system
  - Weeks 3-5: System design and implementation
  - Weeks 6-8: Testing, refinement, and user feedback incorporation

- Weeks 9-10: Final report writing and project report for Part B

5.2 Future Work

As discussed, the second part of the project will focus on developing the health recommendation system. This involves extracting nutritional information from recognized food items, implementing a machine learning-based recommendation system, and refining the system based on user feedback. Furthermore, several potential enhancements could be incorporated to improve the computer vision system. These include:

- Expanding the food database: As the current system is trained on the food-101 dataset, future work could involve incorporating additional datasets to increase the system's ability to recognize a wider variety of food items.
- Using bounding boxes: The accuracy of the system is expected to increase drastically if bounding boxes are used in labelling images and training the model.
- User interface: A user interface could be designed and implemented to offer a more personalized and engaging experience.

6 Conclusion

In an era where people are becoming increasingly health-conscious, technology can play a pivotal role in aiding individuals to make better dietary decisions. This project's first phase successfully developed a computer vision system capable of recognizing various food items from images, setting a solid foundation for the second phase. The use of the Xception model and the food-101 dataset proved effective, yielding a model accuracy of 74.84%. This accuracy indicates that the model can correctly identify a significant variety of food items, serving as a critical component in developing a system that can analyse an individual's dietary habits. However, the project's overarching objective extends beyond simple food recognition. The second phase, planned for the upcoming term, aims to develop a health recommendation system. This system will offer healthier alternatives based on the food identified, considering individual preferences and dietary restrictions.

The upcoming challenges include providing accurate, relevant, and personalized recommendations, maintaining user engagement, and handling sensitive dietary information ethically and securely. A timeline has been set to guide the development process, and strategies have been proposed to address these challenges. This project demonstrates the power of artificial intelligence in promoting healthier lifestyles and opens a myriad of possibilities for further enhancements and research. With continued efforts, it is anticipated that this system will not only offer a personalized health assistant to individuals but also contribute significantly to the broader discourse of leveraging technology for public health improvement.

Reference List

Bossard, L., Guillaumin, M., & Van Gool, L. (2014) *Food-101 – Mining Discriminative Components with Random Forests.* European Conference on Computer Vision.

Chollet, F. (2017) 'Xception: Deep Learning with Depthwise Separable Convolutions', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-1258.

Food-101 Dataset (n.d.) Available at: https://www.vision.ee.ethz.ch/datasets_extra/food-101/