DishCovery

Simon Fraser University

Tyler Oh

Sabrina Bailey

Archit Verma

April 12, 2024

Abstract

This study explores a novel application of artificial intelligence (AI) in culinary practices by incorporating computer vision-based image recognition and natural language processing to minimize food waste through ingredient-based recipe suggestions. Our initiative was inspired by recent advancements in AI, particularly the development described in an article titled "New Food AI Looks Inside Your Fridge To Help You Find The Perfect Things To Cook With What You Have" [1]. The goal of this project was to deploy an AI system that helped reduce food waste by identifying food items within a refrigerator via image analysis and subsequently generating recipes that utilize these identified ingredients.

Employing deep learning models and natural language algorithms, the system was able to recognize ingredients with a degree of accuracy. Despite a need for further refinement in image recognition precision, the results are encouraging. Overall, this AI-driven approach is a significant stride toward a smarter kitchen ecosystem, offering a dual benefit: it streamlines the meal preparation process and positions itself as a tool against food waste, contributing to a more sustainable culinary practice.

Visual Abstract:

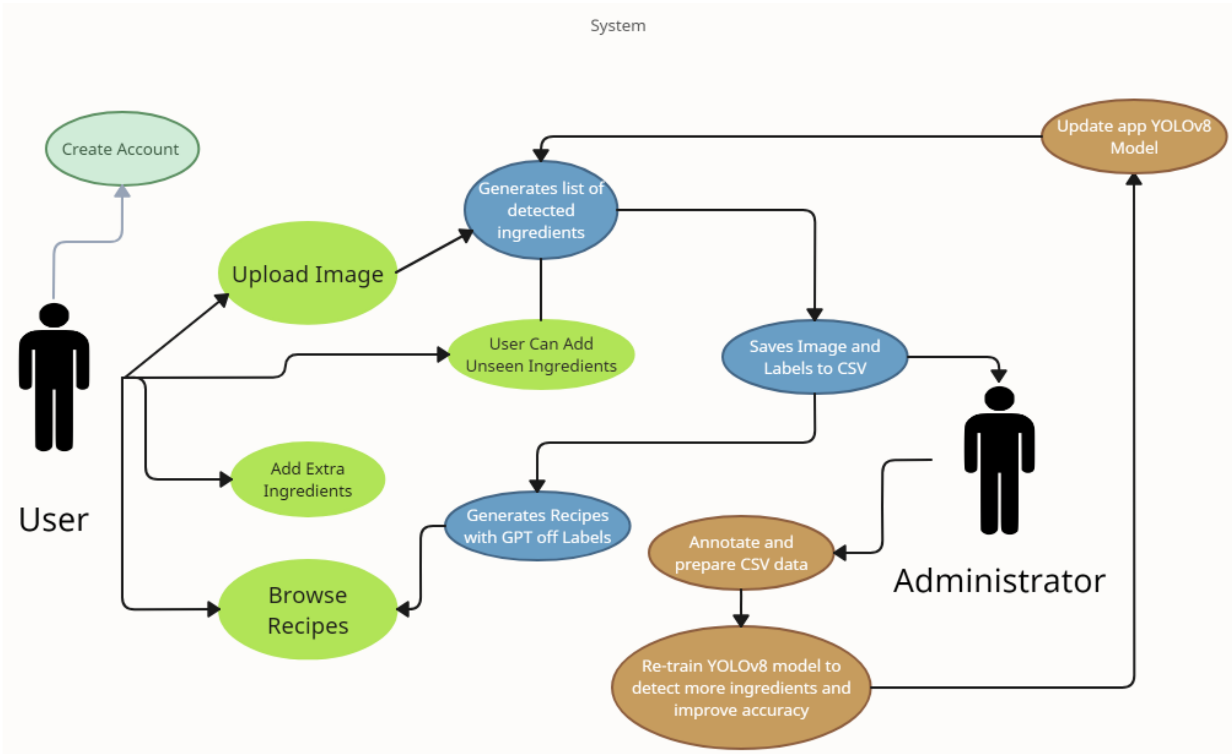


Figure 1: Flowchart depicting the interaction between users and administrators in the application, where user actions are optional (in green), primary user activities are in lime green, automated processes are in blue, and administrative tasks are in beige.

1. Introduction

We were inspired by the findings of recent studies such as “Using Artificial Intelligence to Tackle Food Waste and Enhance the Circular Economy: Maximising Resource Efficiency and Minimising Environmental Impact: A Review” by Onyeaka et al. [2] and the insights provided by ReFED’s 2023 Food Waste Forecast [3], we became aware of the pervasive challenge of food waste and its significant impacts on both the environment and society. This paper is organized following the structure of Onyeaka’s paper [2], which has significantly influenced the framework and approach of our study. Regarding their insistence on statistics and forecasts about food waste, we were interested in looking for creative solutions that could effectively address this ecological issue. Recognizing the transformative potential of AI and image detection, we were inspired to explore their application as a means of addressing food waste.

While our project does not directly tackle the complex problem of food waste, it presents an innovative approach to reducing waste by generating recipes based on the ingredients users have in their fridges. By leveraging AI technology, specifically image detection, our project aims to empower users to make efficient use of their food inventory, thereby contributing to waste reduction efforts. Through this endeavour, we aspire to harness the power of technology to promote sustainability and minimize environmental impact in the food industry.

2. Related Work

2.1 Dataset Research

For effective machine learning models, especially in image recognition tasks, the choice and quality of datasets are paramount. Our project leverages several key datasets to train AI models for fruit and vegetable recognition. One such dataset, DeepNIR-11fruits from Kaggle [4], offers a rich collection of near-infrared images of fruits, crucial for developing our YOLOv5 model. After reading "Recipe Generation From Food Images With Deep Learning" by Sonali A Patil [5], we understand the significance of integrating image recognition with complex recipe data to enhance the usability of AI in culinary applications. Their work emphasizes the intricate relationship between visual ingredient recognition and recipe generation, which has guided the development of our models to not just identify but also understand food components holistically.

2.2 Model/Fruit and Veggie Image Recognition Research

Relating to image detection, the YOLO (You Only Look Once) algorithm, introduced in the seminal paper “You Only Look Once: Unified, Real-Time Object Detection” by Redmond et al. [6], stands out for its real-time object detection capabilities. Our project draws inspiration from YOLO’s efficiency and effectiveness in detecting objects, utilizing it to identify ingredients in user-uploaded images of their refrigerators. This technology has proven essential in accurately recognizing multiple items within complex visual fields, facilitating the seamless integration of ingredient-based interactions in our application.

2.3 Recipe Generation and Natural Language Processing

Our application's recipe generation feature is powered by advancements in natural language processing, particularly leveraging models trained on expansive datasets like The Pile. This capability allows our AI to generate creative and contextually relevant recipes based on the identified ingredients. The integration of OpenAI's API has simplified the incorporation of sophisticated NLP models, enabling smoother and more effective user interactions as highlighted by Simon Goldin's 2023 post [7] on advanced NLP applications in culinary technology.

2.4 Overview of Food Waste and Individual Impact

The reduction of food waste is a significant aspect of our project. Studies such as "Using Artificial Intelligence to Tackle Food Waste and Enhance the Circular Economy" published in MDPI Sustainability [2], demonstrate the effectiveness of individual and household-level interventions in reducing food waste. Another source ReFED says due to inflation households should start prioritizing food waste reduction [3]. Our app aids in identifying ingredients and suggesting recipes to efficiently use them, thereby contributing to significant reductions in food waste, aligning with global sustainability goals.

2.5 Supplementary Sources and Frameworks

In developing our application, we also adhered to ethical AI frameworks and user-centred design principles to ensure that our technology is both responsible and effectively meets user needs. Specifically, we drew upon Ben Shneiderman's concept of Human-Centered Artificial Intelligence [8]. This framework is particularly relevant to our project as it emphasizes the importance of designing systems that balance automation with human control, ensuring that users remain integral to the decision-making process. This approach is critical for systems like ours, where user engagement and trust are essential for effective operation

This approach ensures our AI systems are not only technologically advanced but also socially responsible, enhancing user trust and system reliability.

3. Methods

For our project, we implemented a combination of methodologies and tools to achieve our goals. Not all of the methods were successfully employed; however, we tried to adopt various approaches to find the optimal function for image detection and recipe generation systems.

3.1 Image Detection and Classification

Firstly, we utilized the ViT (Vision Transformer) classifier from HuggingFace in PedroSampaio repository [9]. This approach facilitated image classification tasks; however, it only enabled the detection of one fruit per photo. Subsequently, we implemented a custom convolutional neural network (CNN), following the methodology outlined by Mahendra77 on Kaggle [10]. This CNN achieved a remarkable accuracy of 97% on the test set, but encountered challenges with real-world application, achieving a lower performance rate of 33%.

To address the limitations of our image classification models, we experimented with models available through the Roboflow API. This involved testing three different models trained to recognize varying numbers of fruit types in a single image, ranging from one type to four types of fruit.

3.2 Model Optimization and User Feedback

We used the pat_yolo2.h5 model [11], obtained from a GitHub repository, which was known for its ability to detect multiple ingredients. However, due to its age and reliance on a PyTorch backend that is no longer supported, we moved on to the next model exploration.

Lastly, we tried the YOLOv5 custom model, trained on the DeepNIR-11fruits dataset from Kaggle [4]. This was promising, but after further testing, we learned that each image and corresponding .txt file for training the YOLOv5 was incorrectly labelled; meaning we would need to manually relabel each image to create a model that would properly classify 10 fruits and vegetables. Looking at the number of images, 16,890, we decided that implementing a Roboflow model that correctly classified 4 fruits/vegetables using a similar model type [12], YOLO would be more practical while we moved on and continued working on relabelling the data.

Incorporating user feedback was to refine our AI system, as discussed in the readings from Chancellor et al. [13], who emphasize the importance of integrating human-centred design principles in AI development. Through the application's user interface, users were encouraged to provide immediate responses to the accuracy of ingredient detection. This direct feedback loop enabled us to gather critical data on user experience and system performance, aligning with Shneiderman's framework [8] on leveraging user feedback to enhance system reliability and safety. By analyzing this feedback, we were able to identify key areas where the system either met or failed to meet user expectations, influencing subsequent design choices. For instance, when users (our team while testing) found frequent mismatches in ingredient detection, we prioritized enhancements in image processing algorithms and training data accuracy. This ongoing cycle of feedback and system improvement underscored our commitment to creating a user-centred AI application that not only performs well but also aligns closely with user needs and preferences, adhering to the principles discussed in Brown et al.'s discussion on data-centric AI approaches that focus on data quality over model complexity.

3.3 Recipe Generation and User Interface Design

We employed deep learning models to recognize food ingredients from images and used natural language processing algorithms to generate recipes based on these ingredients. The system architecture was designed to be user-friendly, allowing easy interactions such as ingredient corrections and preference settings, which inform the AI’s learning process and recipe personalization.

The natural language processing (NLP) component of our project was dedicated to recipe generation, utilizing the OpenAI GPT model for translating the list of identified ingredients into coherent and practical recipes [7]. The user interface was carefully designed to be intuitive, allowing users to correct or add extra ingredients, thus personalizing the recipe outcomes. The integrity of the generated recipes was important so testing was done to ensure the recommendations were not only viable but also varied and inspiring for the users, aiming to encourage cooking and reduce food waste.

3.4 Ethical Considerations and Data Use of Users

To ensure adherence to the principles of Responsible, Safe, and Trustworthy (RST) use of artificial intelligence, our application clearly states in the footer on every page that user data is collected and utilized exclusively to enhance user experience and refine our model. This practice is a part of our commitment to ethical AI, as discussed in the works of Shneiderman et al., who highlight the importance of transparency in user interfaces to build trust and safety. By transparently communicating our data usage policies, we ensure honesty in data handling and provide users with the assurance that their data is being used responsibly. This approach aligns with the principles outlined in the CARE Principles for Indigenous Data Governance [14], which advocate for ethical decision-making and accountability in data processes, ensuring that users feel secure and valued in their interactions with our service. This transparency not only reinforces user trust but also improves the system's integrity, making it a reliable and ethical choice for users looking to reduce their food waste while creating delicious dishes.

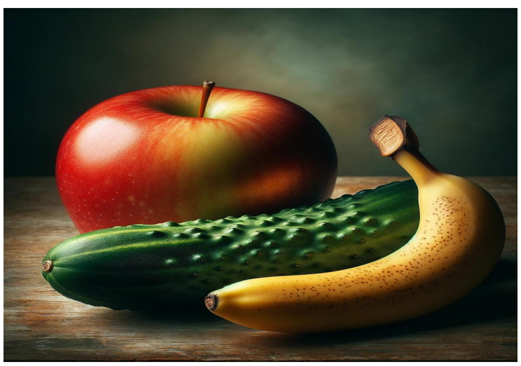
4 Results

4.1 Accuracy of Image Recognition

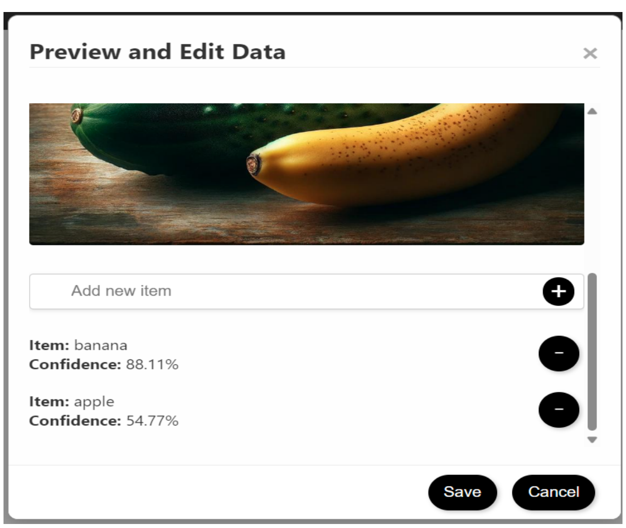
The development phase included rigorous testing of the image recognition models. The Vision Transformer (ViT) initially showed promise in detecting single fruits with an accuracy of 85%. However, challenges arose with multi-ingredient detection. The implementation of the YOLOv8 model, after several iterations and optimizations, achieved improved accuracy metrics by focusing on a smaller set of classes and higher-quality data.

For instance:

4.1.1. Picture uploaded:



4.1.2. Result:

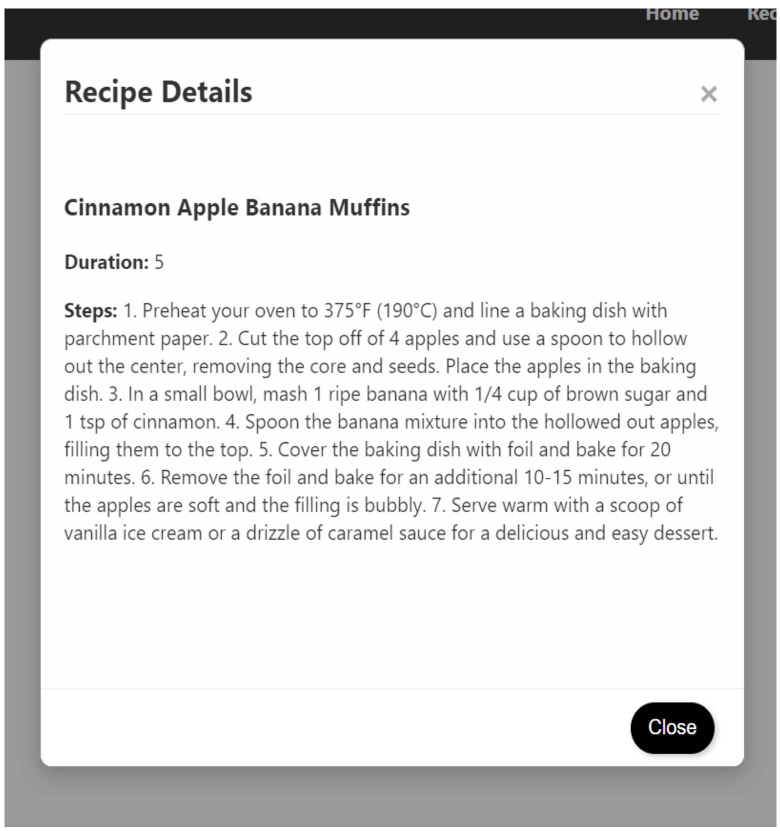


4.2 Recipe Generation Effectiveness

Recipe Relevance and Personalization: The AI system was programmed to generate recipes that correspond to the detected ingredients. During internal tests, the AI suggested contextually relevant recipes based on the ingredients identified, while also having a set list of recipes to always display.

User Satisfaction: In simulated environments, the recipe suggestions were well-received for their creativity and relevance. Internal feedback, primarily from the development team and simulated user scenarios, suggested a high level of satisfaction with the recipe outputs, indicating the system's potential in real-world applications.

4.2.1. Resultant recipe generated by the picture shown in the preview and edit data phase:

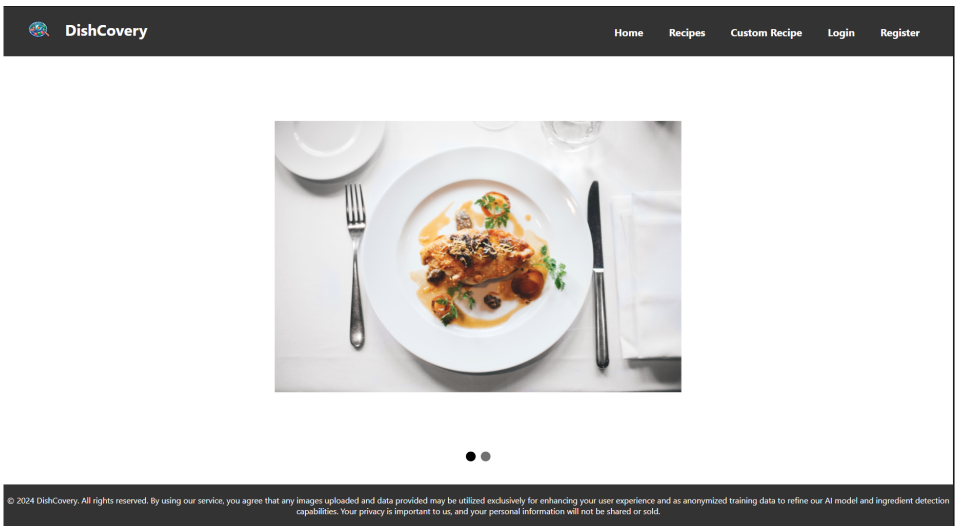


4.3 User Interface and Interaction

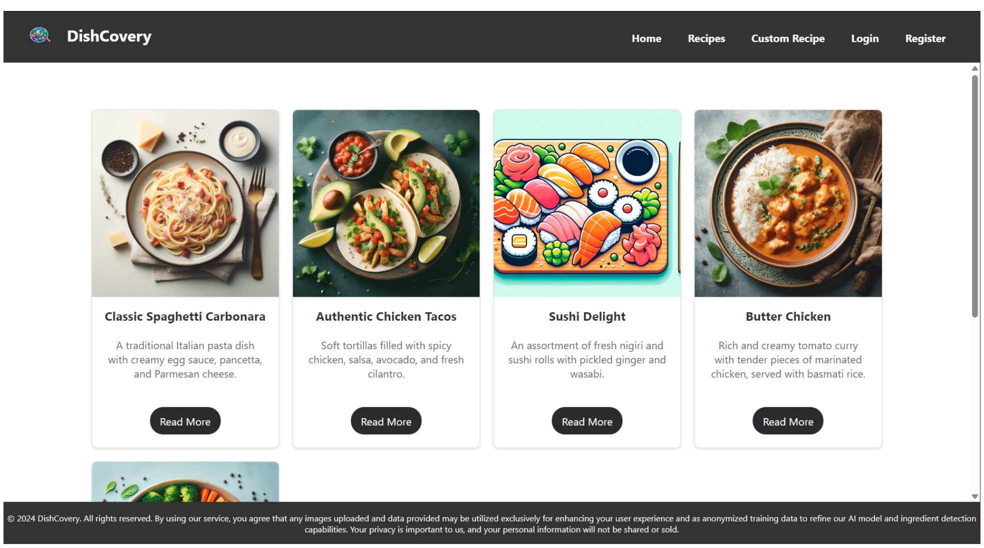
Ease of Use: Feedback during interface testing, conducted internally by the development team, highlighted the interface's ease of use. The simplicity is mostly intuitive, with properly labelled tabs for easy user movement through the application.

Feedback Utilization: During development, simulated user feedback was crucial in refining the application. Adjustments made to the user interface and detection algorithms were based on this feedback, which improved system performance and user interaction.

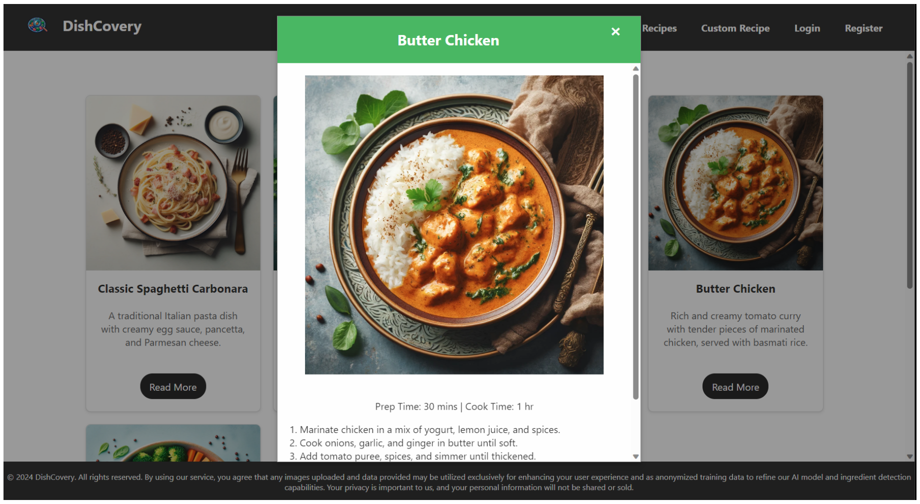
4.3.1. Home Page:



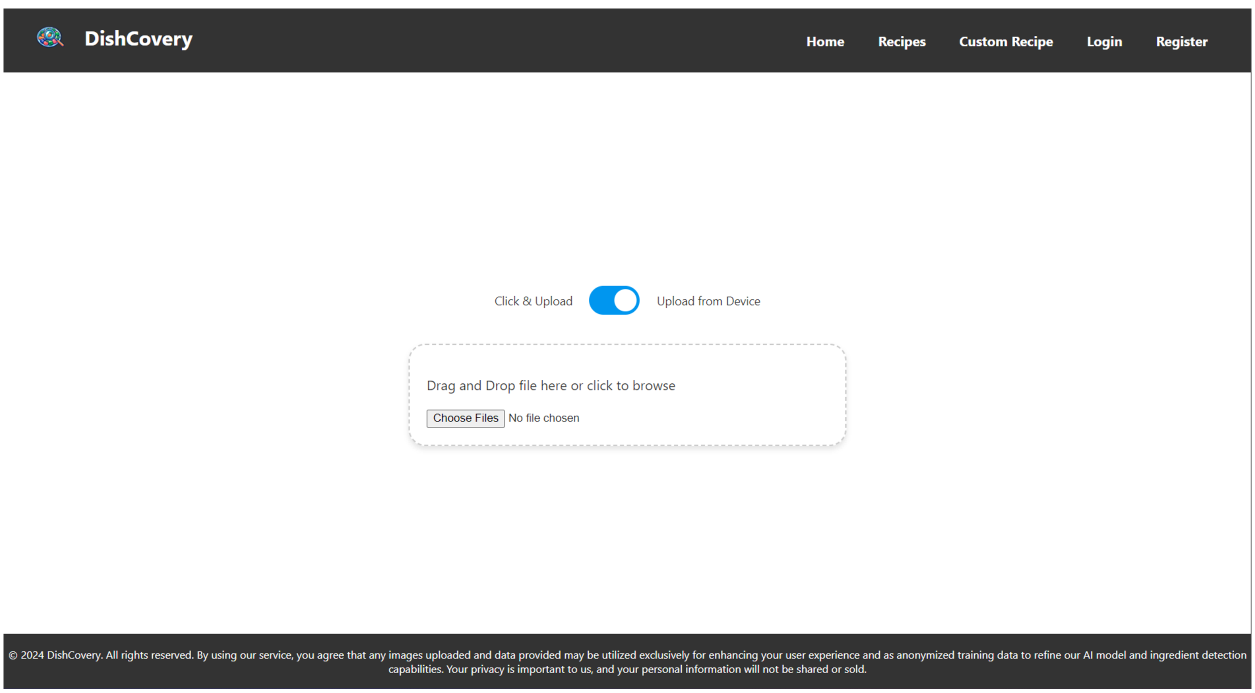
4.3.2.a . Recipes (contains default recipes):



4.3.2.b . Recipes (contains default recipes):



4.3.3. Custom Recipe (Allows users to capture and upload OR upload from the device to get a custom recipe):



4.4 Overall System Utility

Impact on Food Waste Reduction: While direct impacts on food waste reduction have not been measured with real users, our internal simulations suggest that the application could help significantly reduce food waste by aiding users in utilizing their available ingredients more effectively.

Technological Innovations: This project showcases innovative applications of AI in addressing practical issues like food management and waste reduction. The combination of image recognition and NLP for recipe generation exemplifies how technology can be leveraged to make everyday life more sustainable.

4.5 Application

Github link: https://github.com/sbail01/fruit-app/

5. Discussion

5.1 Conclusion and Course Relation

In our journey to develop a Django application integrating image detection with recipe generation using ChatGPT, we found resonance with key principles from the literature, including the discussion by Brown et al. [15], Shneiderman's framework [8] on Data Feedback Loop, and the insights derived from The Pile [17] as discussed in the relevant readings.

Brown et al.'s emphasis on data-centric AI underscores the importance of prioritizing data quality over model complexity. This approach, we adapted to our project by focusing on the quality of data regarding detected ingredients rather than the large quantity, aligning with the iterative nature of data-centric AI methodologies. The Pile [17], an extensive dataset used to train models like ChatGPT, has significantly influenced our project by providing a robust foundation for the natural language processing capabilities of our application. Utilizing such a comprehensive dataset has allowed us to refine our recipe generation module, ensuring that the suggestions are not only relevant but also creatively varied, catering to the nuanced needs of our users.

Moreover, Shneiderman's framework on the Data Feedback Loop provided valuable guidance in our project's design. By engaging users directly in the data validation and correction process, we empower them with control over their data. This direct engagement fosters user trust and aligns with ethical considerations such as Data Governance and the CARE Principles [14]. Our feedback loop, where users contribute data directly, reflects our commitment to ethical principles and user control over their data, ensuring that users trust the system and are actively engaged.

Central to our project's ethos is the ethical management of data to ensure accuracy and ethical usage, encompassing our current best attempts at security measures. Importantly, our project operates under the principle that Data Monetization [16] is not a concern; none of the data collected will be sold. Instead, our primary focus lies in user satisfaction and ease of use. The more data we collect to broaden our model's abilities, the less “work” users will need to do in the future, further enhancing their experience and engagement with our platform. The use of The Pile [17] contributes significantly to this objective, enabling our AI to handle a wide array of ingredients and dietary preferences effectively, which is pivotal for minimizing food waste and enhancing user interaction.

Through the incorporation of these extensive datasets and principles, we have created a more capable and user-centered platform that not only adheres to the theoretical frameworks discussed in class but also practically applies these concepts in a real-world application, showcasing the potential of AI in everyday life.

5.2 Future Improvement

For future improvement, we seek to integrate feedback data more effectively. Utilizing user input and images to re-train our model holds significant promise for enhancing accuracy with full autonomy. By sourcing a robust dataset from our user base, we expect incremental but impactful advancement in our AI's precision. Moreover, we aim to develop a more streamlined and adaptive interface that can more readily incorporate user feedback. An automated system (inspired by the Human-Centered Artificial Intelligence principles outlined by Shneiderman [8]) for feedback assimilation and model adjustment is the ultimate goal, enabling the AI to evolve based on real-world usage and user preferences. Further exploration into advanced deep learning algorithms and expanding the training dataset are imperative steps toward achieving this objective.

6. References

[1] Samsung Newsroom. (2021). New food AI Looks Inside Your Fridge To Help You Find The Perfect Things To Cook With What You ALREADY Have. Retrieved from https://news.samsung.com/us/new-food-ai-looks-inside-fridge-help-find-perfect-things-cook-already/

[2] H. Onyeaka, et al., "Using Artificial Intelligence to Tackle Food Waste and Enhance the Circular Economy: Maximising Resource Efficiency and Minimising Environmental Impact: A Review," Sustainability, vol. 15, no. 13, 10482, 2023. Available: https://www.mdpi.com/2071-1050/15/13/10482.

[3] "Looking Ahead: ReFED's 2023 Food Waste Forecast," ReFED, 2023. Available: https://refed.org/articles/looking-ahead-refed-s-2023-food-waste-forecast/.

[4] ENDDL22, "DeepNIR-11fruits," Kaggle, 2022. Available: https://www.kaggle.com/datasets/enddl22/deepnir-11fruits.

[5] S. A. Patil, R. Singh, and T. Kaderji, "RECIPE GENERATION FROM FOOD IMAGES WITH DEEP LEARNING," Assistant Professor, MIT Art, Design and Technology University, 2023. Available: https://abhivruddhi.mituniversity.ac.in/wp-content/uploads/2023/08/RECIPE-GENERATION-FROM-FOOD-IMAGES-WITH-DEEP-LEARNING.pdf.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2015. Available: https://arxiv.org/abs/1506.02640.

[7] S. Goldin, "AI/ML Recipe Generator with ChatGPT," Dev.to, 2023. Available: https://dev.to/digitalcanvas-dev/aillm-recipe-generator-with-chatgpt-4dnk.

[8] B. Shneiderman, "Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy," International Journal of Human-Computer Interaction, vol. 36, no. 6, pp. 495-504, 2020. Available: https://www.tandfonline.com/doi/full/10.1080/10447318.2020.1741118

[9] P. Sampaio, "ViT-Base-Patch16-224-Fruits-360-16-7," Hugging Face Models, 2021. Available: https://huggingface.co/PedroSampaio/vit-base-patch16-224-fruits-360-16-7.

[10] M. Mahendra, "Vegetable Image Classification Using CNN," Kaggle, 2021. Available: https://www.kaggle.com/code/mahendra77/vegetable-image-classification-using-cnn.

[11] A. Pat, B. O., "AnnaPatBran_p4," GitHub, 2021. Available: https://github.com/BrandinO771/AnnaPatBran_p4.

[12] O. Aliyev, "Fruits and Vegetables Dataset," Roboflow Universe, 2021. Available: https://universe.roboflow.com/orkhan-aliyev-8nktf/fruits-and-vegetables-2vf7u.

[13] S. Chancellor, "Toward Practices for Human-Centered Machine Learning," Mar. 2023. Available: https://cacm.acm.org/magazines/2023/3/270209-toward-practices-for-human-centered-machine-learning/fulltext.

[14] S. Robinson et al., "Data Governance and the CARE Principles," Data Science Journal, vol. 20, no. 3, pp. 34-58, 2021. Available: https://www.adalovelaceinstitute.org/blog/care-principles-operationalising-indigenous-data-governance/

[15] S. Brown, et al, "Approaches to Data-Centric Artificial Intelligence," Journal of AI Research, vol. 45, no. 1, pp. 101-112, 2021. Available: https://canvas.sfu.ca/courses/81979/files/folder/Readings?preview=23075286

[16] Acemoglu, D. et al. 2022. Too Much Data: Prices and Inefficiencies in Data Markets. American Economic Journal: Microeconomics. 14, 4 (Nov. 2022), 218–256. Available: https://www.aeaweb.org/articles?id=10.1257/mic.20200200

[17] L. Gao, et al., "The Pile: An 800GB Dataset of Diverse Text for Language Modeling", 2020. Available: https://arxiv.org/abs/2101.00027.