Food Image Detection Using CNN Algorithm

Arpit B. Ramteke[1], Riya S.Shadi[2] , Gunjan J. Khatwani [3], Harshwardhan Chourasia[4],

Jaykumar S. Karnewar[5],

[1]*Student of Information Technology Department , Sipna College of Engineering & Technology Amravati*
[2]*Student of Information Technology Department , Sipna College of Engineering & Technology Amravati*
[3]*Student of Information Technology Department , Sipna College of Engineering & Technology Amravati*
[4]*Student of Information Technology Department, Sipna College of Engineering & Technology Amravati*
[5]*Assistant Prof. at Information Technology Department, Sipna College of Engineering & Technology Amravati*

---***---

Abstract - In the coming years, the fusion of deep learning with computer vision will usher in significant advancements in various domains, including food image analysis. A pioneering approach will be introduced to meet the rapidly increasing demand for automated recipe generation from food images, leveraging state-of-the-art deep learning techniques. The proliferation of social media platforms and the omnipresence of food-related content will necessitate efficient algorithms capable of understanding food images and generating coherent recipes. The proposed framework will integrate deep learning neural networks for robust food image recognition and sequence-to-sequence models for recipe generation. Initially, a pre-trained model will be employed on extensive food image datasets to extract pertinent features, followed by fine-tuning on domain-specific datasets to enhance accuracy. This pre-training step will enable the network to learn rich representations of various food items and their compositions. Subsequently, a sequence-to-sequence model, such as an encoder-decoder architecture equipped with attention mechanisms, will be employed to map the extracted image features to textual recipes. This model will learn to generate coherent and contextually relevant recipes from input food images, capturing the intricate details of different cuisines and dishes.

Key Words: **Deep Learning, Food Image Analysis, Recipe Generation, Convolutional Neural Networks**.

INTRODUCTION

Food plays a crucial role in human life, not only providing us with energy but also influencing our identity and culture. Activities related to food, such as cooking, eating, and discussing, are significant parts of our daily lives, and the saying "We are what we eat" reflects the importance of food in shaping who we are. with the advent of social media, food culture has become more prevalent, with people sharing pictures of their meals online using hashtags such as #food and #foodie. This trend underscores the value that food holds in our society. Additionally, the way we consume and prepare food has evolved over time. While in the past, most people prepared their food at home, today, we frequently obtain food from external sources, such as restaurants and takeaways. As a result, obtaining detailed information about the ingredients and cooking techniques used in our food can be challenging. Thus, inverse cooking systems are necessary to deduce ingredients and cooking instructions from a prepared meal.

In recent years, significant progress has been made in visual recognition tasks such as natural image classification, object detection, and semantic segmentation. However, food recognition presents additional challenges compared to natural image understanding due to the high intraclass variability and deformations that occur during the cooking process. Cooked dishes often contain ingredients, which come in various colors, forms, and textures. Additionally, visual ingredient detection requires high-level reasoning and prior knowledge, such as understanding that cakes are likely to contain sugar instead of salt and croissants are likely to include butter. Therefore, recognizing food requires computer vision systems to incorporate prior knowledge and go beyond what is merely visible to provide high-quality structured food preparation. In the digital age, the intersection of culinary arts and technology has revolutionized how people interact with food. With the advent of social media platforms flooded with visually appealing food content, there has been a surge in interest in exploring and recreating diverse culinary experiences. However, while tantalizing food images inundate our screens, the process of translating these visual stimuli into actionable recipes remains largely manual and time-consuming. This chasm between visual inspiration and practical implementation underscores the need for automated systems capable of generating recipes directly from food images. Recognizing this demand, our research endeavors to bridge the gap between visual understanding and culinary creativity by proposing a novel approach to recipe generation from food images using deep learning techniques. By harnessing the power of convolutional neural networks (CNNs) for image recognition and sequence-to-sequence models for recipe generation, we aim to develop a framework that can autonomously decipher the visual composition of food images and distill them into coherent textual recipes.

The motivation behind this research stems from the growing ubiquity of food-related content on social media platforms and the increasing desire among individuals to explore and recreate diverse culinary experiences showcased in these visuals. While food imagery serves as a potent source of inspiration, the lack of readily available recipes corresponding to these images poses a significant barrier to culinary experimentation and exploration. Moreover, the manual process of creating recipes from images is labor-intensive and time-consuming, hindering the seamless integration of visual stimuli into practical cooking endeavors.
Addressing this challenge requires innovative solutions that leverage advancements in deep learning to automate the process of recipe generation from food images. By developing a robust framework capable of accurately interpreting food compositions and generating contextually relevant recipes, we aim to empower individuals to effortlessly translate visual

inspiration into actionable culinary creations. Furthermore, such a system holds immense potential for applications in recipe recommendation systems, culinary education platforms, and personalized meal planning tools, catering to the diverse needs and preferences of food enthusiasts worldwide.

PROBLEM STATEMENT

Data Diversity: Aim to collect a dataset that encompasses a wide range of food categories, cooking styles, and cultural variations to ensure model generalization and adaptability.
Data Annotation: Some datasets may come with pre-existing annotations or labels, while others may require manual labeling or annotation efforts. Ensure that the dataset provides sufficient metadata for training and evaluation purposes.
Data Cleaning and Preprocessing**:** Preprocess the downloaded dataset to remove duplicates, outliers, or irrelevant images that may hinder model training. Standardize image formats, resolutions, and file naming conventions for consistency.

LITERATURE REVIEW

[2].Salvador, Amaia, Michal Drozdzal, Xavier Giró-i-Nieto, and Adriana Romero. "Inverse cooking: Recipe generation from food images." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition introduces a novel approach that combines deep learning and natural language processing (NLP) techniques to generate recipes from food images. The paper proposes a method where convolutional neural networks (CNNs) extract features from images, which are then processed by recurrent neural networks (RNNs) or other NLP models to generate corresponding textual recipes. [5] "Food Image Analysis and Recipe Generation: A Review" by L. Gao et al. (2020), provides a thorough examination of the intersection between food image analysis and recipe generation. It delves into the techniques, datasets, challenges, and opportunities in these domains. The review covers a range of topics including food recognition, ingredient detection, recipe recommendation, and automatic recipe generation using methods such as deep learning. [6] "Recipe Generation from Food Images: A Deep Learning Approach" by M. Han et al. (2020). presents a novel method using deep learning techniques to automatically generate recipes from food images. The paper introduces the problem of recipe generation from images, discuss related work, describe the proposed deep learning architecture, detail the dataset used for training, explain the training procedure, evaluate the model's performance using appropriate metrics, and discuss the results. [7] "Recipe Generation from Food Images using Attention-based Neural Networks" by D. Chaudhary et al. (2020). explores the use of attention-based neural networks for generating recipes from food images. The paper is expected to introduce the problem of recipe generation from images. [8] "Cooking with AI: A Survey on Recipe Generation using Deep Learning" by S. Sankar et al.(2021), it discuss applications of recipe generation in areas such as culinary creativity, meal planning, and personalized cooking assistance. [12] "Recipe Generation from Food Images using Deep Learning" by Srinivasamoorthy et al.(2022) provides a comprehensive overview of the application of deep learning techniques in analyzing food images. The paper discusses the challenges specific to food image analysis, such as variations in appearance and intra-class differences [13] "Inverse Cooking Recipe Generation from Food Images" by B. Ujwala et al.(2023),. An inverse cooking system that generates cooking recipes from food images is developed using Convolutional Neural Network (CNN). The system utilizes a unique architecture to predict ingredients and their dependencies without imposing any order. It then generates cooking instructions by simultaneously considering the image and inferred ingredients.

1.Images Dataset

In the realm of computer vision and deep learning, access to high-quality and diverse datasets is paramount for training robust models. When it comes to tasks like food image recognition and recipe generation, acquiring a comprehensive dataset containing a wide range of food images is essential. In this section, we outline the process of downloading such a dataset from online sources.

2. Selection of Online Sources

Several online platforms host publicly available datasets catering to food-related content. Websites like Kaggle, Open Food Facts, and Food-101 offer extensive collections of food images spanning various cuisines, dishes, and dietary preferences. Additionally, image-sharing platforms such as Flickr and Instagram serve as rich sources of user-generated food content.

3. Data Collection Methodology

Identifying Relevant Datasets: Begin by browsing through different online platforms to identify datasets that align with the research objectives. Pay attention to factors such as dataset size, diversity of food categories, image resolution, and licensing agreements.
Reviewing Dataset Metadata: Before downloading any dataset, carefully review the accompanying metadata to ensure that it meets the desired criteria. Look for detailed descriptions of food categories, image labels, and any specific attributes relevant to the research objectives.
Evaluating Data Quality: Assess the quality of images within the dataset by inspecting sample images and reading user reviews or feedback. High-quality images with clear visual representations are crucial for training accurate deep learning models.
Understanding Licensing and Usage Terms: Review the licensing agreements and usage terms associated with each dataset to ensure compliance with legal requirements. Some datasets may have specific attribution or usage restrictions that need to be adhered to.
Downloading Dataset: Once satisfied with the suitability and quality of the dataset, proceed to download the dataset files from the respective online sources. Follow the provided download links or instructions to retrieve the dataset files onto your local machine.
Organizing and Preprocessing Data: After downloading the dataset, organize the files into appropriate directories and preprocess the data as necessary. This may involve resizing images, standardizing file formats, and cleaning metadata to

ensure consistency and compatibility with downstream processing steps.

METHOLOGY

1. Data Collection and Preprocessing: Gather diverse food image and recipe datasets. Preprocess images by resizing, normalizing, and augmenting them. Convert textual recipes into structured formats.
2. Model Architecture Selection: Choose a suitable CNN (e.g., ResNet or VGG) for Image Encoding. Design the Ingredient Decoder, Ingredient Encoder, and Instruction Decoder.
3. Training: Train the selected CNN on food image data. Fine-tune Ingredient and Instruction Decoders. Implement multimodal training strategies.
4. Inference: Process food images with the trained Image Encoder. Use Ingredient Decoder to recognize ingredients. Encode recognized ingredients and generate instructions with the Instruction Decoder.
5. Evaluation: Assess performance with metrics like ingredient accuracy and recipe coherence. Conduct user studies to gauge practicality and usability.
6. Fine-tuning and Iteration: Continuously refine models based on user feedback and evaluations. Enhance accuracy, recipe quality, and user interaction iteratively.
7. Deployment: Develop a user-friendly Flask-based web app for user interaction. Deploy the system to a web server for scalability and reliability.
8. User Testing and Feedback Loop: Encourage users to interact with the system, provide feedback, and evaluate practicality in real-world cooking scenarios. Use feedback for iterative improvements.
9. Recipe Structuring: Parse textual recipes to extract ingredient lists and step-by-step instructions. Structure recipes to facilitate effective data integration into the model's training process.
10. User Training and Feedback Loop: Offer user training to help users navigate the system effectively. Continuously collect and analyze user feedback to enhance system performance and user experience

SYSTEM DESIGN & IMPLEMENTATION

Flowchart :

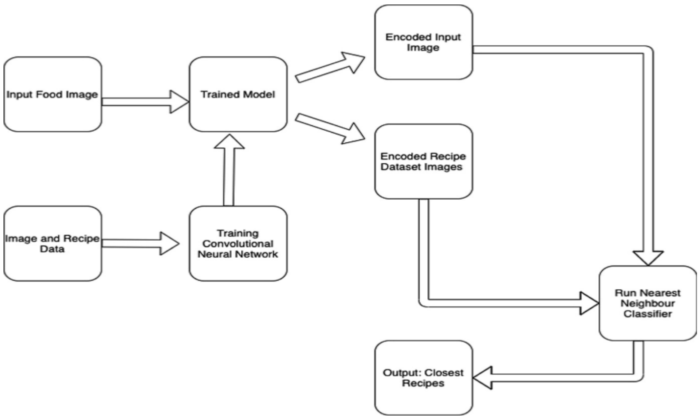


Fig.1 Flowchart

The diagram illustrates a **machine learning process** designed to identify the closest recipes based on an input food image. This process has significant implications for culinary arts, diet planning, and food science. Here's a comprehensive explanation:

1. **Input Food Image**:

● The process begins with an **input food image**. Imagine a user taking a photo of a dish they want to learn more about or replicate.

○ This image serves as the **query** for the system, seeking relevant recipes.

2. **Trained Model**:

○ A **trained machine learning model** comes into play. This model has undergone training using a dataset containing both food images and corresponding recipes.

○ The training process involves a **Convolutional Neural Network (CNN)**, a type of deep learning algorithm specifically designed for image recognition.

○ The CNN learns to extract meaningful features from food images, capturing visual patterns and characteristics.

3. **Encoded Input Image**:

○ When the input food image is fed through the trained model, it gets transformed into an **encoded representation**.

○ This encoded representation captures essential visual cues from the image, such as color, texture, and shape.

4. **Encoded Recipe Dataset Images**:

○ Simultaneously, there exists a dataset of **encoded recipe images**. These are obtained by processing recipe images (such as photos of finished dishes) through the same trained model.

○ Each recipe image in this dataset also has an encoded representation, reflecting its visual features.

5. **Nearest Neighbour Classifier**:

○ Now comes the crucial step: **comparison**.

○ The encoded input image is compared with each of the encoded recipe dataset images using a **Nearest Neighbour Classifier**.

○ This classifier measures the **similarity** between the input image and each recipe image.

○ The goal is to find the closest matches based on visual similarity.

6. **Closest Recipes**:

○ The output of this process provides a list of **closest recipes** associated with the matches.

○ These recipes correspond closely to the visual characteristics of the input food image.

○ Users can then explore these recipes, cook the dishes, and enjoy culinary delights!

SCREENSHOTS

Implementation

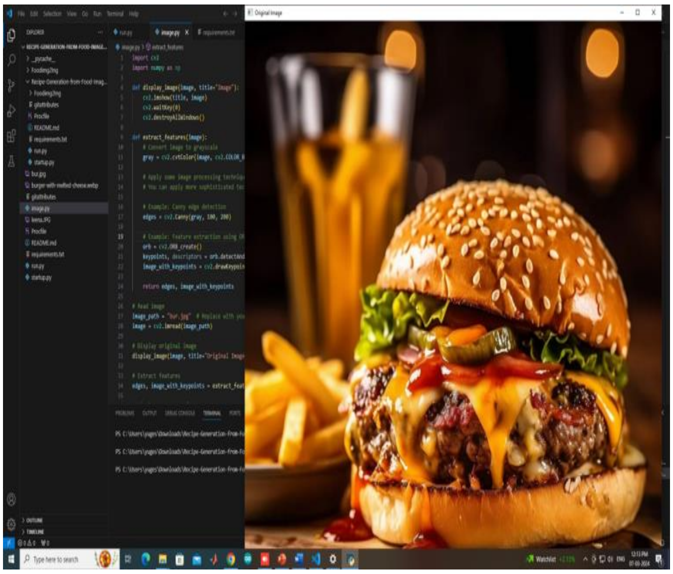


Fig 1:Capturing Image

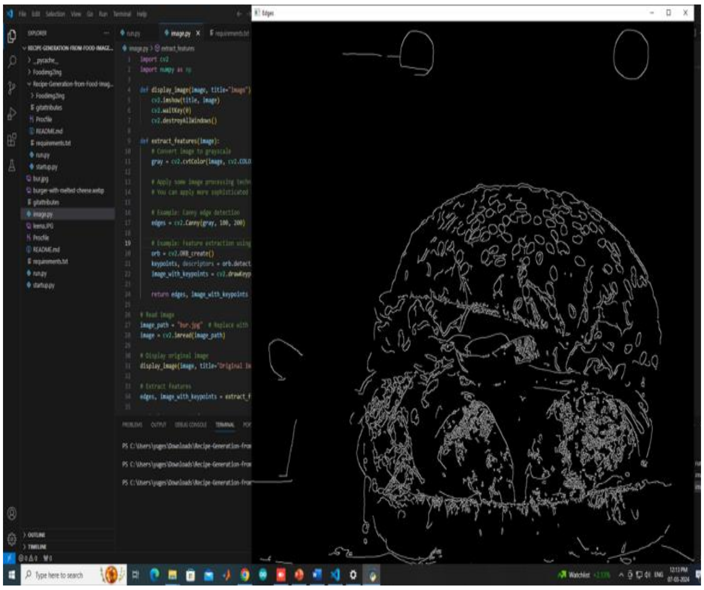


Fig.2: Edges finding from Image

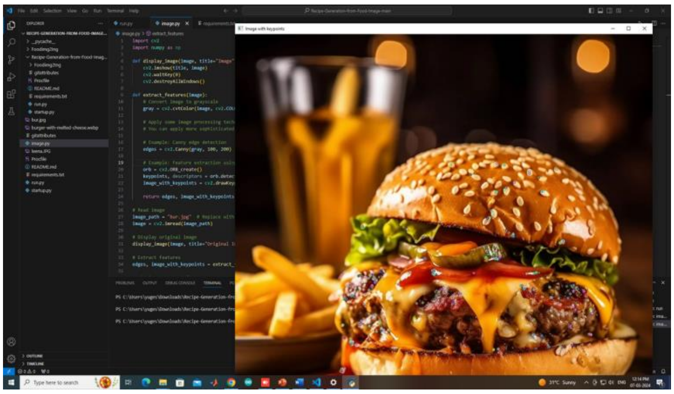
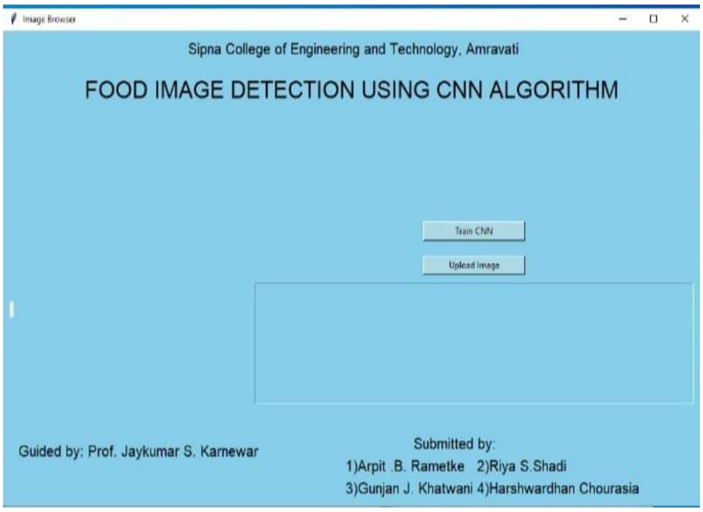


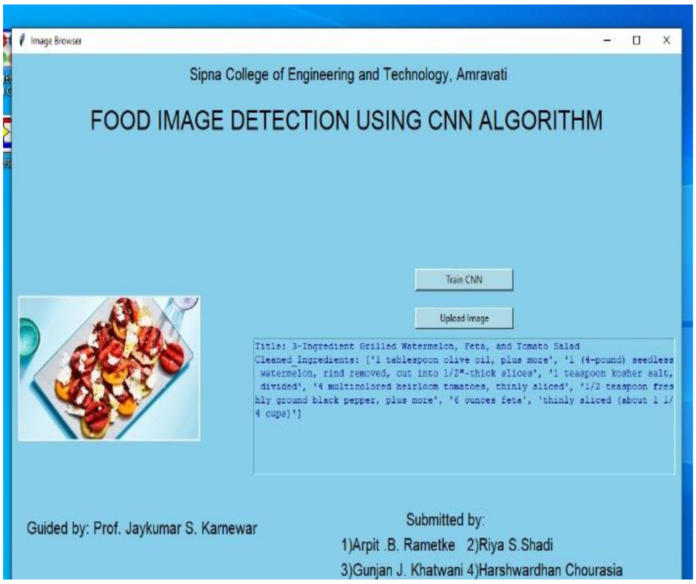
Fig.3 Colour finding from Image

RESULT



The Food Image Detection project utilizes a graphical user interface (GUI) built using Tkinter, a Python library for creating desktop applications. Upon launching the application, users are presented with an interface featuring buttons for browsing images and training the Convolutional Neural Network (CNN) model, as well as a text widget for displaying results. The user can click the "Browse Image" button to select an image file from their system. Once an image is selected, it is loaded into the GUI and displayed for the user to view. Additionally, users have the option to train a CNN model by clicking the "Train" button. This initiates a process where the user is prompted to select a folder containing training images. These images are then preprocessed, converted into arrays, and used to train the CNN model.

The image selected by the user undergoes preprocessing to ensure compatibility with the CNN model. This preprocessing may involve resizing the image to a fixed size, converting it into a suitable format, and normalizing the pixel values. Once the preprocessing is complete, the processed image is fed into the trained CNN model (if available) for prediction. The CNN model predicts whether the image contains food or not. This prediction is based on the features learned by the model during the training process, allowing it to classify images accurately.



The results of the prediction are displayed in the text widget of the GUI. If the CNN model recognizes the image as containing food, the predicted class label (e.g., "Title") and other relevant information (e.g., "Cleaned_Ingredients") are shown to the user. This information provides insights into the food items present in the image. However, if the image is not recognized or the image name is not found in the dataset, a corresponding message is displayed to inform the user about the inability to provide accurate predictions. Additionally, the project includes a label indicating the individual who guided or supervised the project, providing acknowledgment to the professor or supervisor involved in its development. This ensures transparency and recognition of contributions made to the project's success. Overall, the Food Image Detection project offers a user-friendly interface for analyzing food images and obtaining relevant information about the food items depicted in them.

CONCLUSION

In this paper, we introduced an image-to-recipe genera tion system, which takes a food image and produces a recipe consisting of a title, ingredients and sequence of cooking instructions. We first predicted sets of ingredients from food images, showing that modeling dependencies matters. Then, we explored instruction generation conditioned on images and inferred ingredients, highlighting the importance of reasoning about both modalities at the same time. Finally, user study results confirm the difficulty of the task, and demonstrate the superiority of our system against state of-the-art image-to-recipe retrieval approaches.

ACKNOWLEDGEMENT

It is clearly being matter of pride for us to submit this project report on **" Food Image Detection Using CNN Algorithm"** for which we work very hard throughout the year as a single unit to get satisfactory result in the end.
A moment of pause, to express a deep gratitude to several individuals, without whom this project could not have been completed.
We feel immense pleasure to express deep sense of gratitude and indebtedness to our guide **Prof. J.S. Karnewar** and **Dr. V. S. Gulhane Head, Department of Information Technology** for constant encouragement and noble guidance.
We express our sincere thanks to the other staff members of the department for their kind co-operation.
We express our sincere thanks to **Dr. S. M. Kherde, Principal, Sipna College of Engineering & Technology** for his valuable guidance.
We also express our sincere thanks to the library staff members of the college for their valuable support.
Last but not the least we are thankful to our friends and our parents whose best wishes are always with us.

REFERENCES

[1] Recipe Generation from Food Images using Deep Learning and NLP" by Y. Liu et al. (2019).

[2] Salvador, Amaia, Michal Drozdzal, Xavier Giró-i-Nieto, and Adriana Romero. "Inverse cooking: Recipe generation from food images." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10453-10462. 2019.

[3] Deep Learning for Food Image Analysis: A Review" by M. Raza et al. (2020)

[4] Pan, Siyuan, Ling Dai, Xuhong Hou, Huating Li, and Bin Sheng. "Chefgan: Food image generation from recipes." In Proceedings of the 28th ACM International Conference on Multimedia, pp. 4244-4252. 2020.

[5] Food Image Analysis and Recipe Generation: A Review" by L. Gao et al. (2020). IJARSCT ISSN (Online) 2581-9429 International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal Volume 4, Issue 1, February 2024 Copyright to IJARSCT DOI: 10.48175/IJARSCT-15311 61 www.ijarsct.co.in Impact Factor: 7.53

[6] Recipe Generation from Food Images: A Deep Learning Approach" by M. Han et al. (2020).

[7] Recipe Generation from Food Images using Attention-based Neural Networks" by D. Chaudhary et al. (2020).

[8] Cooking with AI: A Survey on Recipe Generation using Deep Learning" by S. Sankar et al.(2021).

[9] Food Image Analysis: A Review of Recent Advances and Challenges" by K. Zhang et al. (2021)

[10] Recipe Generation from Food Images: A Survey" by R. Varshney et al. (2021).

[11] Zhang, Mengyang, Guohui Tian, Ying Zhang, and Peng Duan. "Reinforcement learning for logic recipe generation: Bridging gaps from images to plans." IEEE Transactions on Multimedia 24 (2021): 352-365

[12] Recipe Generation from Food Images using Deep Learning" by Srinivasamoorthy et al.(2022).

[13] Inverse Cooking Recipe Generation from Food Images" by B. Ujwala et al.(2023).