IJARST

International Journal of Advanced Research in Science and Technology

journal homepage: www.ijarst.com

ISSN 2319 – 1783 (Print)

ISSN 2320 – 1126 (Online)

Recipe Generation from Food Images

V.Varun, Y. Sai Pranay, B. Athirath, Mr.CH. Srinath Reddy

Sreenidhi Institute of Science and Technology Hyderabad, India.

ARTICLEINFO

Article history:

Received	08 May 2023
Accepted	18 May 2023
Available online	12 June 2023

Keywords:

Inverse cooking,
Recipe 1M dataset,
CNN,
retrieval-based approaches.

ABSTRACT

Many people enjoy food photography because it highlights the beauty of food. Images of food do not, however, provide any information regarding the method of preparation or the difficulty of the recipe used to create each dish. Convolutional Neural Network (CNN) is used to create an inverse cooking system that creates cooking instructions from food photographs. Without prescribing any order, the system uses a novel architecture to forecast elements and their dependencies. Then, while concurrently taking into account the image and implied ingredients, it provides cooking directions. On the Recipe 1M dataset, the system's performance was carefully assessed, and the results showed that ingredient prediction was more accurate than with earlier techniques. By utilising both the image and the inferred ingredients, the system was also able to generate high-quality recipes. Human review revealed that these recipes were more interesting than those produced by retrieval-based methods.

1. INTRODUCTION

Food is a vital part of human life because it not only gives us energy but also shapes our identity and culture.We spend a lot of time each day engaging in food-related activities including cooking, eating, and conversing, and the adage "We are what we eat" accurately captures the significance of food in determining who we are. Social media's introduction has increased the popularity of food culture, with users posting images of their meals online with the hashtags #food and #foodie. This pattern highlights how important food is to our culture. What you want altered should go here. After that, click the button below. It's that simple! In addition, how we prepare and eat food has changed over time. While most individuals used to prepare their own food at home in the past, today we frequently buy food from outside sources like restaurants and takeaways. As a result, it can be difficult to find out specifics about the ingredients and cooking methods applied to our food. In order to determine the components and cooking directions from a prepared meal, inverse cooking systems arerequired.

Visual recognition tasks like object detection, semantic segmentation, and natural picture classification have made substantial strides in recent years. However, due to the substantial intraclass variability and deformations that take place while cooking, food recognition poses more difficulties than natural image interpretation. Ingredients in cooked foods frequently appear in a variety of hues, shapes, 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 descriptions

2. METHODOLOGY

Previously, food understanding efforts have primarily focused on categorizing food and ingredients. However, a comprehensive visual food recognition system should not only recognize the type of food or its ingredients but also comprehend its preparation process.The image-to-recipe problem has typically been treated as a retrieval task, where a recipe is retrieved from a fixed dataset based on the image similarity score in an embedding space. The effectiveness of these systems largely depends on the size and diversity of the dataset and the quality of the learned embedding. As a result, these systems may fail when a matching recipe for the image query is not present in the static data.

In the current methodology, we are training CNN with recipe details and images, and this model can be used to predict recipes by uploading related images. We used a dataset of one million recipes, and we selected 1000 recipes from it because training CNN with the entire dataset of images would consume a lot of memory and take hours.

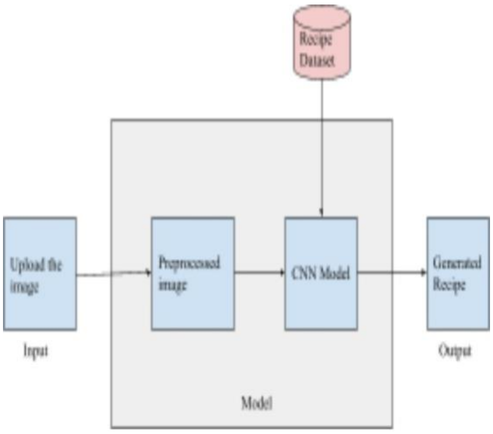
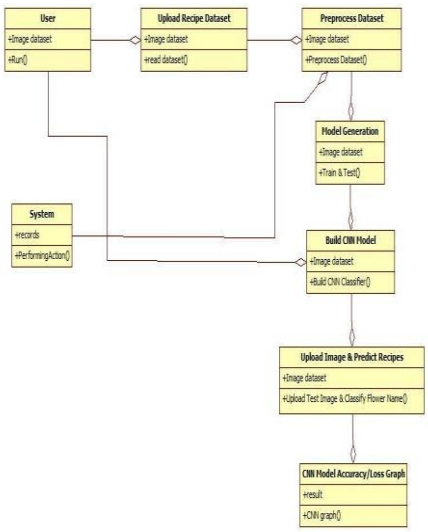


Fig: 3.1 System Architecture

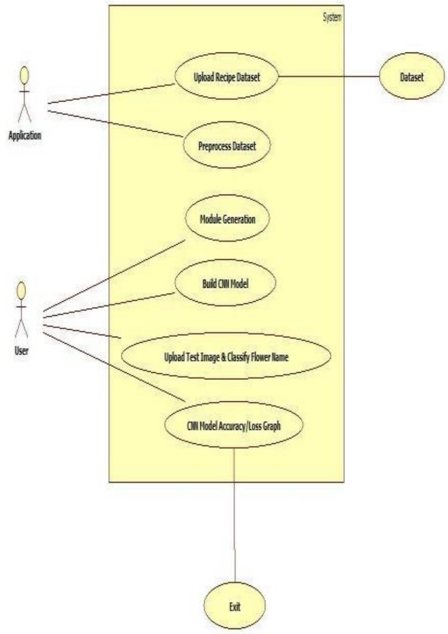
The methodology involves the following steps:

1. Data Collection: Collecting high-quality food images from a diverse set of sources is critical for building an accurate and robust recipe generation model. It's important to ensure that the collected images cover a wide range of cuisines, ingredients, and cookingstyles.
2. Image Preprocessing: Preprocessing the food images using computer vision techniques can help to extract useful features and improve the accuracy of the recipe generation model. This can involve techniques such as resizing, normalization, and feature extraction using CNNs.
3. Recipe Generation: Generating high-quality and diverse recipes that match the input food image is a challenging task that requires a combination of deep learning and optimization techniques. It's important to ensure that the generated recipes are both feasible and appealing to the user.

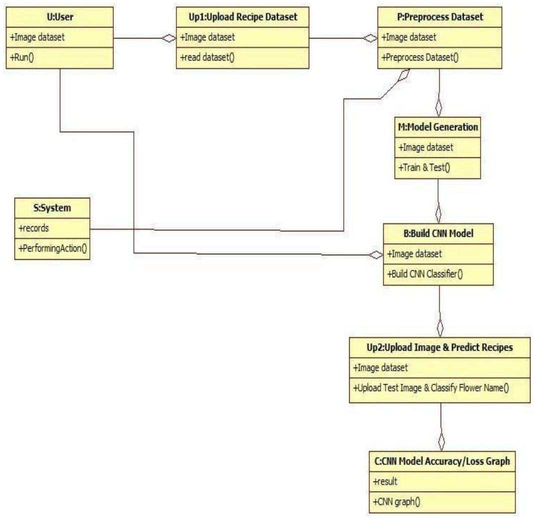
4. UML DIAGRA Class Diagram



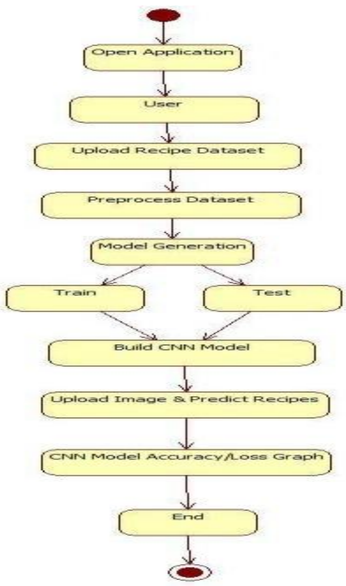
USE CASE DIAGRAM



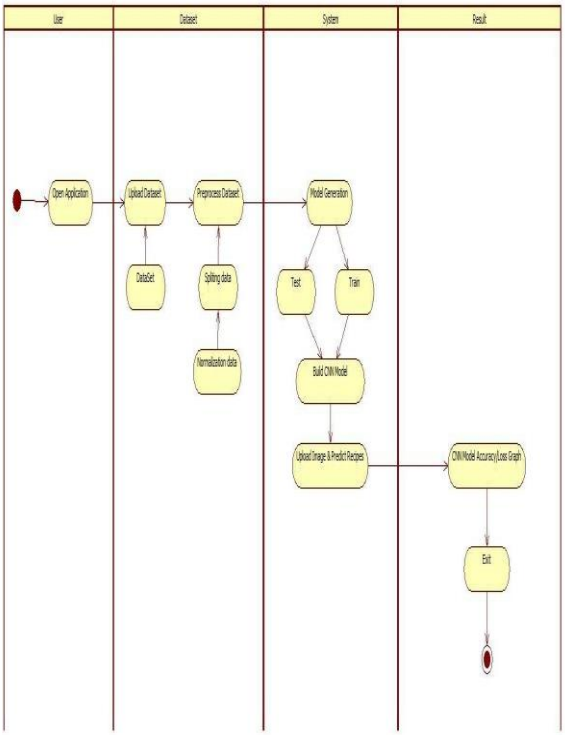
OBJECT DIAGRAM



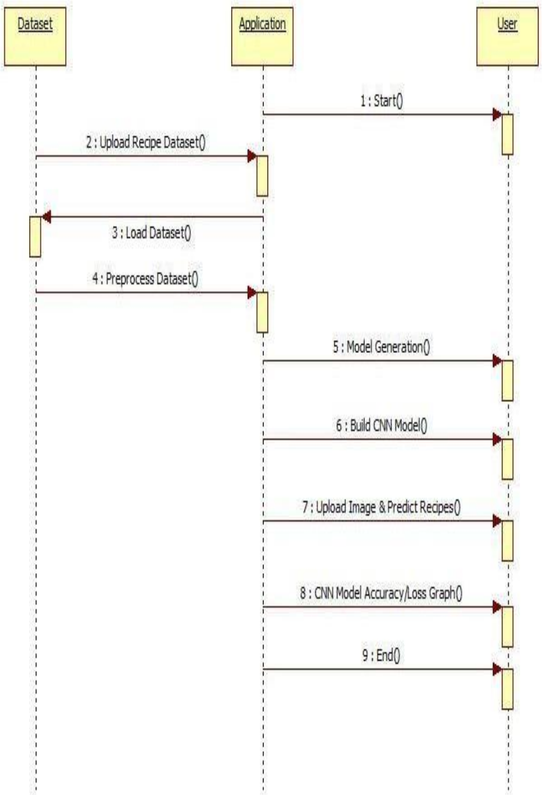
STATE DIAGRAM



ACTIVITY DIAGRAM

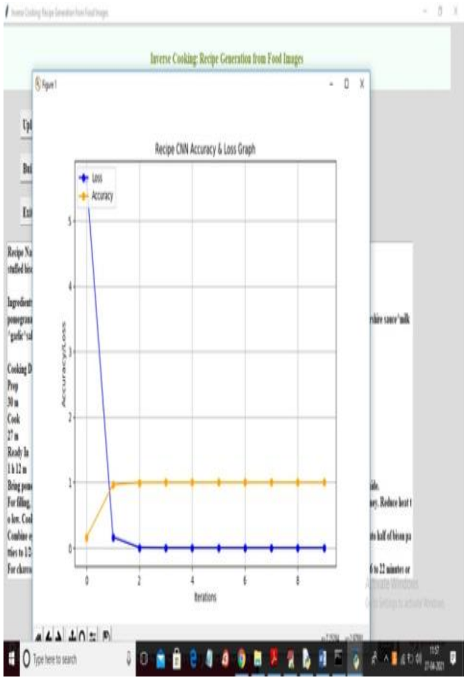


SEQUENCE DIAGRAM



5. RESULTS AND DISCUSSIONS

The result of the project is mainly based on the performance of the trained deep learning models and the evaluation metrics used. It involves measuring the accuracy of the generated recipes, which is often evaluated based on the similarity of the generated recipe to the original recipe, which is usually sourced from the online recipe databases. Wehave used more than 1000 images from the Recipe 1M dataset to train the model and got an accuracy of 99.662% in predicting the recipe name, ingredients and recipe preparation.



In the graph above, the x-axis depicts epochs, the y-axis the accuracy/loss value, and the orange line the accuracy. With each passing epoch, the accuracy grew to 1 (100%) and the loss reduced to 0. Any CNN model with a high accuracy and low loss will be regarded as efficient.

6. CONCLUSION

The goal of our project was to create an image-to-recipe generating system that could create a recipe from a food photograph, complete with a title, ingredients, and cooking directions. First, by identifying groups of elements from culinary photographs, we showed the need of modelling dependencies. Second, welooked into instruction generation that depends on both images and inferred elements, highlighting the necessity of taking into account both modalities at once. Finally, we validated the difficulty of the task based on the findings of a user research, and we established that our system outperforms current image-to-recipe retrieval techniques.

7.REFERENCES

1. Luc Van Gool, Matthieu Guillaumin, and Lukas Bossard. Using random forests to mine discriminative components for food 101.2014, ECCV.

2.Laure Soulier, Nicolas Thome, David Picard, Matthieu

Cord, Micael Carvalho, and Remi Cadene. Learning semantic text-image embeddings for cross-modal retrieval in the context of cookery. 2018 in SIGIR

3.Jing-Jing Chen andChong-Wah Ngo. Deep-based ingredient recognition for cooking recipe retrieval. In ACM Multimedia. ACM, 2016.

4.Jing-Jing Chen, Chong-Wah Ngo, and Tat-Seng Chua.Cross-modal recipe retrieval with rich food attributes. InACM Multimedia. ACM, 2017.

5.Mei-Yun Chen, Yung-Hsiang Yang, Chia-Ju Ho, Shih-Han Wang, Eugene Chang,Che-Hua Yeh, and Ming Ouhyoung are some of the other participants. Chinese food recognition and amount estimation automatically.2012's SIGGRAPH AsiaTechnical Briefs.

6.Xin Chen, Hua Zhou, and Liang Diao. Chinese Food Net: A large-scale image dataset for Chinese food recognition. CoRR, abs/1705.02743, 2017.

7.Bo Dai, Dahua Lin, Raquel Urtasun, and Sanja Fidler. Towards diverse and natural image descriptions via a conditional gan. ICCV, 2017.

8 Krzysztof Dembczy´nski, Weiwei Cheng, and Eyke H¨ullermeier. Bayes optimal multilabel classification via probabilistic classifier chains.In ICML, 2010.

9 Angela Fan, Mike Lewis, and Yann Dauphin. Hierarchical neural story generation. In ACL, 2018.

10 Claude Fischler. Food, self and identity. Information (International Social Science Council), 1988.