
ISSN: 2321-2152
IJMECE
International Journal of modern
electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

IJMECE

ISSN2321-2152

www.ijmece .com

Vol 12, Issue.2, 2024

AUTOMATED FOOD IMAGE CLASSIFICATION USING DEEP LEARNING APPROACH

Y. PAVANI NAGA DURGA[1], Dr.V.Bhaskara Murthy[2]

[1]MCA Student, B V Raju College, Kovvada, Andhra Pradesh, India.

[2]HOD & Professor, B V Raju College, Kovvada, Andhra Pradesh, India.

ABSTRACT:

People enjoy food photography because they appreciate food. Behind each meal there is a story described in a complex recipe and, unfortunately, by simply looking at a food image we do not have access to its preparation process. Therefore, in this paper we introduce an inverse cooking system that recreates cooking recipes given food images. Our system predicts ingredients as sets by means of a novel architecture, modeling their dependencies without imposing any order, and then generates cooking instructions by attending to both image and its inferred ingredients simultaneously. We extensively evaluate the whole system on the large-scale Recipe1M dataset and show that (1) we improve performance w.r.t. previous baselines for ingredient prediction; (2) we are able to obtain high quality recipes by leveraging both image and ingredients; (3) our system is able to produce more compelling recipes than retrieval-based approaches according to human judgment.

INTRODUCTION

Food is fundamental to human existence. Not only does it provide us with energy—it also defines our identity and culture. As the old saying goes, we are what we eat, and food related activities such as cooking, eating and talking about it take a significant portion of our daily life. Food culture has been spreading more than ever in the current digital era, with many people sharing pictures of food they are eating across social media. Querying Instagram for #food leads to at least 300M posts; similarly, searching for #foodie results in at least 100M posts, highlighting the unquestionable value that food has in our society. Moreover, eating patterns and cooking culture have been evolving over time. In the past, food was mostly prepared at home, but nowadays we frequently consume food prepared by thirdparties (e.g. takeaways, catering and restaurants). Thus, the access to detailed information about prepared food is limited and, as a consequence, it is hard to know precisely what we eat. Therefore, we argue that there is a need for inverse cooking systems, which are able to infer ingredients and cooking instructions from a prepared meal. The last few years have witnessed outstanding improvements in visual recognition tasks such as natural image classification, object detection and semantic segmentation. However, when comparing to natural image understanding, food recognition poses additional challenges, since food and its components have high intraclass variability and present heavy deformations that occur during the cooking process. Ingredients are frequently occluded in a cooked dish and come in a variety of colors, forms and textures. Further, visual ingredient detection requires high level reasoning and prior knowledge (e.g. cake will likely contain sugar and not salt, while croissant will presumably include butter). Hence, food recognition challenges current computer vision systems to go beyond the merely visible, and to incorporate prior knowledge to enable high-quality structured food preparation descriptions.Previous efforts on food understanding have mainly focused on food and ingredient categorization. However, a system for comprehensive visual food recognition should not only be able to recognize the type of meal or its ingredients, but also understand its preparation process. Traditionally, the image-to-recipe problem has been formulated as a retrieval task, where a recipe is retrieved from a fixed dataset based on the image similarity score in an embedding space. The performance of such systems highly depends on the dataset size and diversity, as well as on the quality of the learned embedding. Not surprisingly, these systems fail when a matching recipe for the image query does not exist in the static dataset. An alternative to overcome the dataset constraints of retrieval systems is to formulate the image-to-recipe problem as a conditional generation one.

Reverse cooking is a fascinating subject of research aimed at creating a recipe for a certain dish. Although many people have cooking experience, reverse cooking requires the ability to understand the complex relationships between ingredients, cooking techniques, and flavours. Reverse cooking has many real-world applications, such as creating recipes for meal planning services, personalized recommendations for food delivery, and culinary education. In recent years, there has been an increasing interest in using machine learning techniques to solve the reverse cooking problem. These techniques range from traditional rule-based systems to the most advanced neural network architectures. Although considerable progress has been made, there are still many challenges to be addressed in this area, such as the lack of large-scale data sets, the difficulty of modelling complex cooking techniques, and the need to Create personalized recipes. In this context, reverse cooking offers an exciting opportunity to combine culinary knowledge with machine learning techniques. By doing so, we are able to develop intelligent systems capable of generating high-quality recipes that meet individual preferences and dietary requirements. The field is changing rapidly and this is an exciting time to get involved in reverse cooking system research and development.

SURVEY OF RESEARCH

There are many works that have been carried in the past on recipe generation. Here is a survey of some works which help in understanding the previous techniques and gives a clear view on the challenges on which the researchers have worked and the new things that they have introduced. Lukas Bossard et al. [1] introduced a new dataset for food recognition called Food-101 in 2014. The dataset contains over 100,000 images of 101 food categories, and they proposed a method for mining the discriminative components of the images using random forests. Their method outperforms several state-of-theart algorithms for food recognition on the Food-101 dataset. They also conducted a detailed analysis of the dataset, including the distribution of the categories and the difficulty of the recognition task . They have highlighted the importance of large-scale and diverse datasets in food image analysis. Micael Carvalho et al. [2] proposed a cross-modal retrieval approach for the cooking context, which involved learning semantic text-image embeddings to link cooking recipes and their corresponding food images. They used a multi-modal deep neural network to learn the embeddings, which consisted of a textual embedding and a visual embedding. The textual embedding is learned from the recipe ingredients and instructions, while the visual embedding is learned from the food images. They evaluated their approach on a dataset of 5,000 recipeimage pairs and also conducted a qualitative analysis of the retrieved results and showed that their method is able to retrieve relevant recipes and images. They have provided a novel approach for linking cooking recipes and food images, highlighting the importance of cross-modal retrieval in the food image analysis field. Chong-Wah Ngo et al.[3] introduced a deep-based approach for ingredient recognition in cooking recipes, which is essential for cooking recipe retrieval. They have used a convolutional neural network (CNN) to extract features from food images and then used these features to recognize the ingredients in the corresponding recipes. They have evaluated their approach on a dataset of 600 recipes and also conducted a user study to evaluate the effectiveness of their method for recipe retrieval and showed that their approach improves the retrieval performance compared to a baseline method and highlighted the importance of ingredient recognition in cooking recipe retrieval. Jing-Jing Chen et al. [4] proposed a cross-modal recipe retrieval approach that considers rich food attributes such as taste, cuisine, and occasion, in addition to the ingredients and food images. They have used a deep neural network to extract features from food images and text features from the recipe ingredients and attributes. They then used a multimodal fusion method to combine the features and perform cross-modal retrieval and evaluated their approach on a dataset of 1,000 recipes. They also conducted a user study to evaluate the effectiveness of their method and show that their approach improves the retrieval performance compared to a baseline method. They have highlighted the importance of considering various aspects of food when performing food image analysis tasks.

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 present methodology we are training CNN with recipe details and images and this model can be used to predict recipe by uploading related images and we used 1 million recipe dataset and from this dataset we used 1000 recipes as training the entire dataset with images will take lots of memory and hours of time to train CNN model.

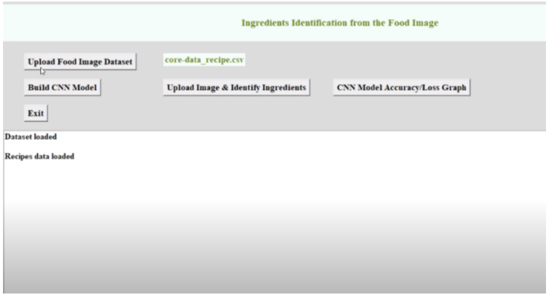
The methodology involves the following steps:

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 cooking styles.
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.

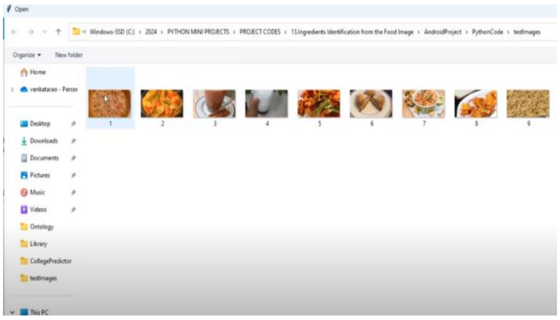
Producing a formula from a picture may be a challenging assignment, which needs a synchronous understanding of the fixings composing the dish as well as the changes they went through, for example, slicing, mixing or blending with other fixings. Instead of getting the formula from a picture straightforwardly, we contend that a formula era pipeline would advantage from a halfway step foreseeing the fixings list. The arrangement of instructions would at that point be created conditioned on both the picture and its comparing list of fixings, where the interaction between picture and fixings seem provide additional experiences on how the last mentioned were prepared to deliver the coming about dish. Figure 4 outlines our approach. Our formula era framework takes a nourishment picture as an input

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2, 2024

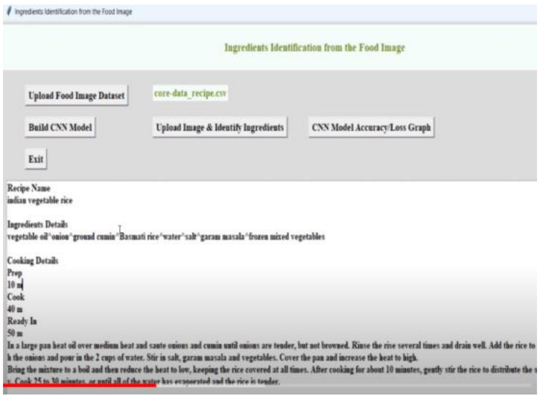
and yields a sequence of cooking informational, which are created by means of an instruction decoder that takes as input two embeddings. The primary one speaks to visual highlights extracted from an picture, whereas the moment one encodes the fixings extricated from the picture.



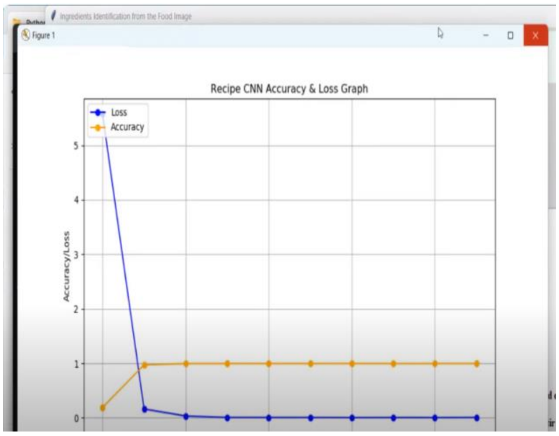
Training images:



Output results:



Accuracy Graph:



CONCLUSION

The aim of our study was to develop an image-to-recipe generation system that can produce a recipe, including a title, ingredients, and cooking instructions, from a food image. Firstly, we demonstrated the significance of modeling dependencies by predicting groups of ingredients from food images. Secondly, we investigated instruction generation that is dependent on both images and inferred ingredients, emphasizing the need to consider both modalities simultaneously. Finally, based on the outcomes of a user study, we verified the complexity of the task and confirmed that our system outperforms existing image-to-recipe retrieval methods.

REFERANCES

[1] LukasBossard, Matthieu Guillaumin, and Luc Van Gool. Food-101–mining discriminative components with random forests. In ECCV, 2014.

ISSN2321-2152
www.ijmece .com
Vol 12, Issue.2, 2024

[2] Micael Carvalho, R´emi Cad`ene, David Picard, Laure Soulier, Nicolas Thome, and Matthieu Cord. Crossmodal retrieval in the cooking context: Learning semantic text-image embeddings. In SIGIR, 2018.

[3] Jing-Jing Chen and Chong-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. In ACM Multimedia. ACM, 2017.

[5] Mei-Yun Chen, Yung-Hsiang Yang, Chia-Ju Ho, Shih-Han Wang, Shane-Ming Liu, Eugene Chang, Che-Hua Yeh, and Ming Ouhyoung. Automatic chinese food identification and quantity estimation. In SIGGRAPH Asia 2012 Technical Briefs, 2012.

[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.