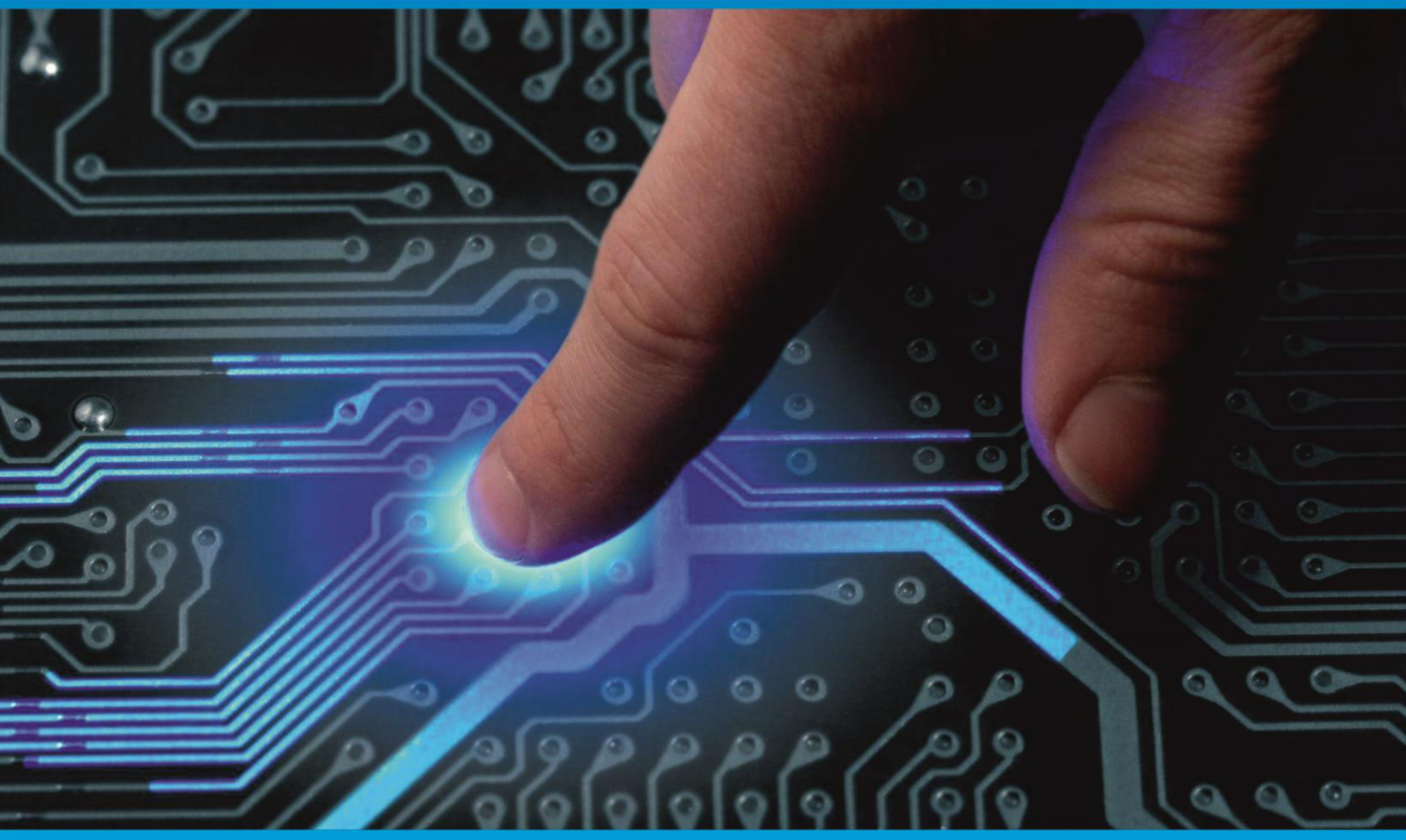
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 4, April 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0904110 |

A Survey on Counter Cooking: Generating Recipe from Food Images using techniques of Machine Learning

Prof. Swati Patil[1], Aparna Shitole[2], Madhu Basrani[3], Rutuja Jadhav[4] and Swaraj Gore[5]

Assistant Professor, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India[1]

UG Student, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, Maharashtra, India[2]

UG Student, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, Maharashtra, India[3]

UG Student, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, Maharashtra, India[4]

UG Student, Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, Maharashtra, India[5]

ABSTRACT: We introduced an image-to-recipe generation method, which takes a food image and offers a recipe consisting of a title, ingredients, and a bunch of cooking directions. We first predicted sets of ingredients from food images, showing that modeling dependencies matter. Then, we explored instruction generation conditioned on images and gathered ingredients, highlighting the importance of reasoning about both modalities at the same time. Lastly, user study results validate the difficulty of the task and illustrate the superiority of our system against state-of-the-art image-to-recipe retrieval methods.

KEYWORDS: Text Generation, Image-to-Text, Image Encoder, Ingredient Decoder

I. INTRODUCTION

Food is key to human life. In addition to the fact it provides us with energy it additionally characterizes our personality and culture . As the well-known old saying goes, we are what we eat, and food related exercises like cooking, eating and discussing it take a significant portion of our everyday life. Food culture has been spreading like never before in the current advanced period, with numerous individuals sharing pictures of food they are eating across social media. Questioning Instagram for #food prompts at any rate 300M posts; also, looking for #foodie results in at any rate 100M posts, featuring the certain worth that food has in our society. Additionally, eating patterns and cooking culture have been advancing over time. Previously, food was for the most part set up at home, yet these days we often devour food arranged by third parties (for example takeaways, providing food and eateries).

Subsequently, the access to detailed information about prepared food is limited and, as an outcome, it is difficult to know definitely what we eat. In this manner, we argue that there is a requirement for recipe generation system, 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 with natural image understanding, food recognition represents extra difficulties, since food and its component have high intraclass variability and present substantial deformations that happen during the cooking cycle. Ingredients are often blocked in a cooked dish and arrive in a variety of colors, structures and textures. Further, visual ingredient identification requires high level thinking and prior knowledge(for example cake will probably contain sugar and not salt, while croissant will presumably include butter). Previous efforts on food understanding have mainly focused on food and ingredient categorization. However, a system for comprehensive visual image recognition should not only be able to recognize title or its ingredients, but also gives all possible recipes.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

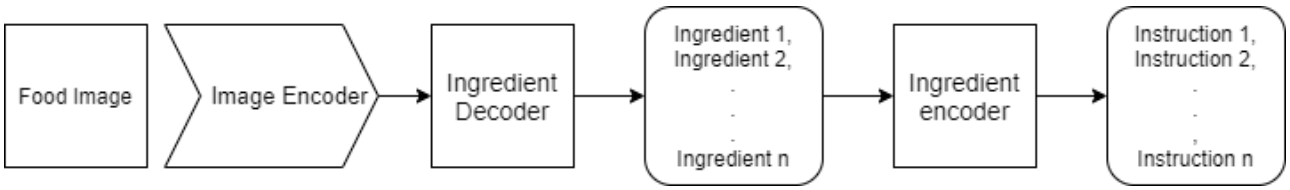
| DOI: 10.15680/IJIRCCE.2021.0904110 |

Traditionally, the image-to-recipe issue has been figured as a recovery task , where a recipe is retrieved from a fixed dataset dependent on the image similarity score in an embedding space. The performance of such system highly depends upon the dataset size and variety, as well as on the quality of learned embeddings. Not surprisingly these systems fails when a matching recipe for the image query does not exist in the static dataset. A choice to defeat the dataset limitations of recovery systems is to formulate the image-to-recipe issue as a conditional generation one. Therefore , in this paper, we present a system that creates a cooking recipe containing a title, ingredients and cooking instructions directly from a image.

II. RELATED WORK

A. *Generating recipes from images*

Generating a recipe (title, ingredients and instructions)from an image is a challenging task, which requires a simultaneous understanding of the ingredients composing the dish as well as the transformations they went through, e.g. slicing, blending or mixing with other ingredients. Instead of obtaining the recipe from an image directly, we argue that a recipe generation pipeline would benefit from an intermediate step predicting the ingredients list. The sequence of instructions would then be generated conditioned on both the image and its corresponding list of ingredients, where the interplay between image and ingredients could provide additional insights on how the latter were processed to produce the resulting dish. Above figure illustrates our approach. Our recipe generation system takes a food image as an input and outputs a sequence of cooking instructions, which are generated by means of an instruction decoder that takes as input two embeddings. The first one represents visual features extracted from an image, while the second one encodes the ingredients extracted from the image.



B. *Understanding Food*

Our goal is to extract the ingredients from food, which can be done by using computational methods to analyze food data. There is currently a vast literature in computer vision dealing with a variety of food related tasks, with special focus in image classification. As of today social media is developed a lot and has vast amount of food images which are getting posted every day. Recipe 1M dataset help us establish recipes and food images. It is worth noting that, to the best of our knowledge, is the only work for recipe generation task on Recipe1M dataset. Our DGN approach improves the recipe generation performance by introducing the decomposing idea to the generation process.

C. *Text Generation*

Text generation is a widely researched task, which can take various input types as source information. Machine translation is one of the representative works of text-based generation, in which the decoder takes one language text as the input and outputs another language sentences. Image-based text generation involves both vision and language, such as image captioning, visual question answering To be specific, image captioning is to generate suitable descriptions for the given images, and the goal of visual question answering is to answer questions accompanied with the image and text. Text generation related tasks are accelerated by some new state-of-the-art models like the Transformer and BERT, which are attention-based. Many recent works achieve superior performance with attention-based models. In our work, we compare the results of using the pretrained BERT and normal embedding layer as the ingredient encoder.

III. CONCLUSION

We introduced an image-to-recipe generation 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 modelling 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.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 4, April 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0904110 |

REFERENCES

1. Decomposed Generation Networks with Structure Prediction for Recipe Generation from Food Images: arXiv:2007.13374 .
2. Inverse Cooking: Recipe Generation from Food Images, Salvador, Amaia and Drozdzal, Michal and Giro-i-Nieto, Xavier and

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details