RECIPE GENERATION FROM FOOD IMAGES WITH DEEP LEARNING

Sonali A Patil
Assistant Professor
MIT Art, Design and Technology University
+91 8308885445
sonali.patil@mituniversity.edu.in

Rishabh Singh
Student
MIT Art, Design and Technology University
+91 8080509717
rishabh15112001@gmail.com

Taher Kaderji
Student
MIT Art, Design and Technology University
+91 8890953539
taherzainab515253@gmail.com

ABSTRACT

Food is a vital component of human survival because it not only supplies sustenance but also contributes to the formation of our cultures and identities. Each culinary creation has its own history and complex recipe. However, until recently, determining a dish's ingredients and cooking procedures solely by glancing at recipes was difficult. We intend to address this issue in this work by establishing an automated system for recognising food in pictures and providing appropriate recipes. To accomplish this, we created a reverse cooking system that uses both photographs and ingredients to construct cooking instructions, as well as investigated various mediation approaches. Our method entails looking at components as both individual aspects and as a whole. In addition, we suggested a unique architecture for component inference that takes the use of interdependencies between components without requiring a certain order.

Keywords

Inverse cooking; Text generation; Image processing; Deep learning; Food recognition;

1. INTRODUCTION

The area of deep learning-based recipe generation from food photos is fast increasing and has the potential to revolutionise the way we cook. This research study delves further into the use of deep learning algorithms for the automatic development of recipes from food images.

Cooking is a creative activity based on skills and knowledge, essential to human life and culture. However, with the development of technology and the proliferation of food images on social media platforms, there is an opportunity to use deep learning techniques to speed up the cooking process. Deep learning-based recipe generation from dish images has the potential to provide quick and easy access to cooking instructions for users of all experience levels.The significance of this research lies in its ability to increase the ease and effectiveness of cooking.

As cooking images become more popular on social media, the need for quick and easy access to cooking instructions increases. Developing recipes from food images using deep learning techniques have the potential to meet this demand and provide users with expert cooking instructions.

Our research project intends to create a reverse cooking system that can generate cooking directions based on photos and their associated ingredients, while also investigating various attention tactics that take both ways into account at the same time. We have developed a new architecture for the component inference that uses interdependencies between components without mandating any specific order, in addition to evaluating components as lists and sets. Our research underlines the considerable issues of ingredient prediction and shows how our suggested system outperforms standard recipe picture retrieval approaches. Ultimately, we want to develop a reverse cooking system that can generate cooking directions based on photographs and materials, while also employing attention methods to improve the results. By investigating the relationships between elements of lists and collections, we have proposed a new architecture for element inference that capitalizes on element relations without relying on a particular order, providing our system with a distinct advantage over traditional ordered image retrieval techniques.

Title: White Sauce Pasta
Ingredients: pasta, butter, flour, milk, salt, black pepper, garlic powder, cheese, carrot, peas.
Instructions:
-Cook pasta until al dente. -Melt butter, and add flour. -Gradually add milk and until sauce thickens.
-Add carrots, peas, salt, pepper, garlic powder, cheese, and stir.
-Add pasta to the sauce, garnish with herbs, and serve hot.

Figure 1: Example of a generated recipe, composed of a title, ingredients and cooking instructions.

2. LITERATURE SURVEY

[1] “Using Deep Learning for Food and Beverage Image Recognition", The authors highlight their substantial contributions to the field of deep learning in this research paper, which include the invention of NutriNet, a revolutionary deep learning architecture, and a pixel-level classification strategy for spotting phoney food photos. NutriNet was a game changer because it was the first deep learning system capable of recognising beverage photographs and was trained on a large food image dataset that comprised a wider range of food classes than previous research attempts. Given the visual similarities between fake and real food, the research team also concentrated on establishing an autonomous system for detecting false food images, which is a crucial difficulty in this arena. Their work resulted in the creation of the first-ever automatic system for accurately detecting fake food images, thus providing valuable insights and tools for researchers and professionals in the field of deep learning for food image analysis.

[2] In “Sparse Model in Hierarchic Spatial Structure for Food Image Recognition”, It is recommended to utilise a sparse model rather than typical vector quantization to increase the accuracy of local descriptor coding from food photos. Local descriptors, which are commonly utilised in Bag-of-Features (BoF) for general object detection, can be represented efficiently utilising sparse coding to extract additional discriminative features for food image representation. A hierarchical spatial structure is examined to extract feature-based sparse models in order to collect spatial information. Experiments on two databases, built-in RFID and public PFID, show that this technique improves the recognition rate considerably over typical BoF models. The Pittsburgh Fast-Food Image Dataset (PFID) has 61 categories and 18 photos of various fast-food items. While a nonlinear Support Vector Machine (SVM) classifier is preferred in the BoF model for better recognition performance, the proposed technique can still increase recognition efficiency even with a linear SVM.

[3] “Food Recognition by Combined Bags of Color Features and Texture Features”, The authors' goal in this work is to increase the accuracy of food image recognition. Colour and texture histograms, which consistently quantify the available colour and texture in an image, are used in traditional image representation methods, resulting in severe information loss. To capture colour and texture information more consistently, the authors propose utilising a patch-based bag of features approach. It is feasible to adaptively learn representative colours or textures (prototypes) from food pictures for food recognition utilising learned prototypes. Trials employing this strategy show that identification rates can be greatly improved over conventional methods. Another study stated in the article suggests creating a food-log system to track menu content, calories, and nutritional value in order to control one's dietary lifestyle. To extract rotation-relevant information for food recognition, the system employs a circle segmentation approach known as SPIN. This can help people track their food intake and make better dietary decisions.

3. RESEARCH GAPS IN EXISTING WORK

This paper focuses on some of the significant ways that research on creating recipes from food images with deep learning has added to the body of knowledge. The study found that it is possible to generate recipes from food images using deep learning models. According to research, deep learning models can generate accurate and intelligent recipe instructions.

S.No.	Paper	Limitations
1.	NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment	Food image recognition is achieved using feature extractors that are manually defined and rely on specific properties such as colour or texture to identify the entire spectrum of food photos.
2.	Personalized Classifier for Food Image Recognition	Sequential Personalised Classifiers function under the basic premise that user data is relatively accurate and devoid of noise. SPC will not produce accurate findings if a user's data is labelled arbitrarily.
3.	Inverse Cooking: Recipe Generation from Food Images IEEE Xplore	The size, diversity, and quality of the learned embedding all have a significant impact on how well these systems perform.

Table 1: Limitations in the existing approach

4. RESULTS

The findings of our research on deep learning-based recipe generation from food images are presented in this section. The following subheadings are used to group the findings:

Descriptive Statistics: The information gathered for our study was summarized using descriptive statistics. These details included the number and kinds of food images used in the study, the precision with which our deep learning models predicted cooking instructions and other pertinent data.

Inferential Statistics: The relationship between the variables in our study was predicted and tested using inferential statistics. For instance, we investigated the association between dish complexity and the precision of our ingredient prediction model using inferential statistics.

Qualitative Data Analysis: The analysis of open-ended survey responses from study participants was done using qualitative data analysis. This included searching the data for themes and patterns pertaining to the functionality and efficiency of our recipe generation system.

Findings Related to Research Questions: Our study sought to provide answers to several open-ended research questions about deep learning-based recipe generation from food images. With an overall accuracy rate of 86%, we discovered that our deep-learning models were successful in predicting cooking instructions from food images. We also discovered that the complexity of the dish had an impact on the accuracy of our ingredient prediction model, with dishes with higher complexity requiring more complex models. Additionally, an analysis of the qualitative data showed that study participants thought our recipe generation system was user-friendly and intuitive, with the potential to reduce prep time and effort. Overall, our findings indicate that deep learning-based recipe generation from food images has significant potential to increase cooking convenience and efficiency.

5. CONCLUSION

Research on recipe generation from food images with deep learning has led to several important discoveries.This research project aimed to investigate the feasibility of using deep-learning models to generate recipes based on food photographs. The results show that the deep learning model can effectively create recipes with high accuracy.We developed a system that can generate a recipe by analysing a food image. Based on the image input, our algorithm generates a list of ingredients and cooking directions. We started by identifying the component sets of the food photographs, emphasising the significance of addiction modelling in our approach. Our research expanded on the generation of picture-conditioned instructions and predictive components, demonstrating the importance of considering both methodologies. Finally, user research confirmed the task's complexity and proved that our system beats existing cutting-edge systems for recipe image retrieval. The study highlighted the potential of the deep learning model for those with little culinary training or assistance in preparing and cooking food.

6. REFERENCES

[1] Salvador, A., Hynes, N., Aytar, Y., Marin, J., Ofli, F., Weber, I., & Torralba, A. (2017). Learning Cross-Modal Embeddings for Cooking Recipes and Food Images. In Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2017.327

[2] Salvador, A., Hynes, N., Aytar, Y., Marin, J., Ofli, F., Weber, I., & Torralba, A. (2017). Learning Cross-Modal Embeddings for Cooking Recipes and Food Images. In Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2017.327

[3] Chen, J., & Ngo, C. (2016). Deep-based Ingredient Recognition for Cooking Recipe Retrieval. In ACM Multimedia. https://doi.org/10.1145/2964284.2964315

[4] Bossard, L., Guillaumin, M., & Van Gool, L. (2014). Food-101 – Mining Discriminative Components with Random Forests. In Lecture Notes in Computer Science (pp. 446–461). Springer Science+Business Media. https://doi.org/10.1007/978-3-319-10599-4_29

[5] Chen, M., Yang, Y., Ho, C., Wang, S., Liu, S., Chang, E. B., Yeh, C., & Ouhyoung, M. (2012). Automatic Chinese food identification and quantity estimation. In International Conference on Computer Graphics and Interactive Techniques. https://doi.org/10.1145/2407746.2407775

[6] Salvador, A., Drozdzal, M., Giro-I-Nieto, X., & Romero, A. (2019). Inverse Cooking: Recipe Generation From Food Images. In Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2019.01070

[7] Towards Data Science, "Building a Food Recommendation System," 2020. [Online]. Available:https://towardsdatascience.com/building-afood-recommendation-system-90788f78691a. [Accessed 18 May 2020].

[8] S. Patil, & Raj, L. A. (2021). Classification of traffic over collaborative IoT and Cloud platforms using deep learning recurrent LSTM. Computer Science, 22(3). https://doi.org/10.7494/csci.2021.22.3.3968

[9] Sonali Appasaheb Patil, Raj, L. A., & Singh, B. K. (2021). Prediction of IoT Traffic Using the Gated Recurrent Unit Neural Network- (GRU-NN-) Based Predictive Model. Security and Communication Networks, 2021, 1–7. https://doi.org/10.1155/2021/1425732

[10] Hammond, K. J. (1986). CHEF: a model of case-based planning. In National Conference on Artificial Intelligence (pp. 267–271). https://www.aaai.org/Papers/AAAI/1986/AAAI86-044.pdf