IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁙ Impact Factor 7.39 ⁙ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11784

Inverse Cooking: Recipe Generation from Food Images

Ashwini Hegde[1], Rabia Ishrath[2], Ranjitha Y[3], Yashaswini N[4], Sandesh R[5]

Department Of CSE, VVIET, Mysore, India [1-5]

Abstract: The advances in the classification of individual cooking ingredients are sparse. The problem is that there are almost no public edited records available. This work deals with the problem of automated recognition of a photographed cooking dish and the subsequent output of the appropriate recipe. The distinction between the difficulty of the chosen problem and previous supervised classification problems is that there are large overlaps in food dishes (aka high intra-class similarity), as dishes of different categories may look very similar only in terms of image information. The combination of object recognition or cooking court recognition using Convolutional Neural Networks (short CNN) and the search for the nearest neighbours (Next-Neighbour Classification).

Keywords: Inverse cooking, Image processing, Food recognition, Deep learning, Text generation

I. INTRODUCTION

The Image recognition of food items would be a good solution to food recording. Taking a picture would then be a sufficient record. However, we know that there is a wide diversity of types of food. Even within the same food category, there is considerable diversity.

Therefore, despite the attempts at food item recognition, recognition performance is not yet satisfactory. Once the food is identified, proper recipe can be found accordingly. 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 third parties (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.

However, when comparing to natural image understanding, food recognition poses additional challenges, since food and its components have high intra class 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 colours, 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.

II. LITERATURE SURVEY

[1] "Using Deep Learning for Food and Beverage Image Recognition", In this paper authors describe their deep learning contributions to the field: NutriNet, a novel deep learning architecture, and a pixel-level classification solution for images of fake food. NutriNet was trained on a food image dataset of a larger size and containing more food classes than previous works, and was the first to recognize beverage images.

The work on fake-food image recognition includes the first automatic system for recognizing images of fake food, while the visual similarity of fake and real food makes it useful for fake-food experiments as well as real food recognition.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁂ Impact Factor 7.39 ⁂ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11784

[2] "Food Recognition by Combined Bags of Color Features and Texture Features", propose a discriminated food image representation that can perform effective identification of food images in this paper. The conventional image representation mainly includes color and texture distributions (histogram), which are the statistical information based on uniformly quantized color or texture levels. However, these conventional techniques using uniform quantization of the on-hand color and texture in the image lead much information loss for reliably constructing the image.

Therefore, this study proposes to characterize the color and texture information by incorporating the strategy of patch-based bag of features model. This technique can adaptively learn the representative color or texture (prototypes) from the food images for food recognition, and it is possible to recover a more reliable image using the learned prototypes. The experiments using the proposed approaches show that the recognition rate can be greatly improved compared with the conventional method.

[3] "Auto-Recognition of Food Images Using SPIN Feature for Food-Log System", propose to extract rotation invariant features using circle-segmentation called SPIN for food recognition, and construct a Food-Log system, which records the contents of food menu, calories and nutritional value for management of the dietary life.

[4] "Sparse Model in Hierarchic Spatial Structure for Food Image Recognition", Propose to apply a sparse model for coding a local descriptor extracted from the food images. Sparse coding: an extension of vector quantization for local descriptors, which is popularly used in Bag-of-Features (BoF) for image representation in generic object recognition, can represent the local descriptors more efficient, and then obtain more discriminant feature for food image representation.

Moreover, in order to introduce spatial information, a hierarchic spatial structure is explored to extract the feature based sparse model. Experiments validate that the proposed strategy can greatly improve the recognition rates compared with the conventional BOF model on two databases: the constructed RFID and the public PFID. Pittsburgh Fast-Food Image Dataset (PFID) consists of various fast-food images which include 61 categories and 18 images in each category.

Due to the better recognition performance, the nonlinear SVM classifier, which has high computational cost, is applied in BoF model. However, the proposed strategy can further improve the recognition performance even with linear SVM.

III. PROPOSED SYSTEM

The inverse cooking system generates cooking instructions, title, and ingredients form the food images. We are using neural joint embedding. In embedding, the paired (recipe and image) data in order to learn a common embedding space. Embedding make it easier to do machine learning on large inputs like sparse vectors representing words. Cuisine classification is classified cuisine from food ingredients generated by Inverse cooking system. It will classify the food by its types.

There are many kinds of cuisine like Indian cuisine, Italian cuisine, Mexican cuisine, French food, Thai food, Chinese food etc. it will help people to know the type of cuisine by from taking the picture of the food images. It is very useful for the people who don't know any knowledge about the food.

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.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁑ Impact Factor 7.39 ⁑ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11784

IV. PROPOSED SYSTEM

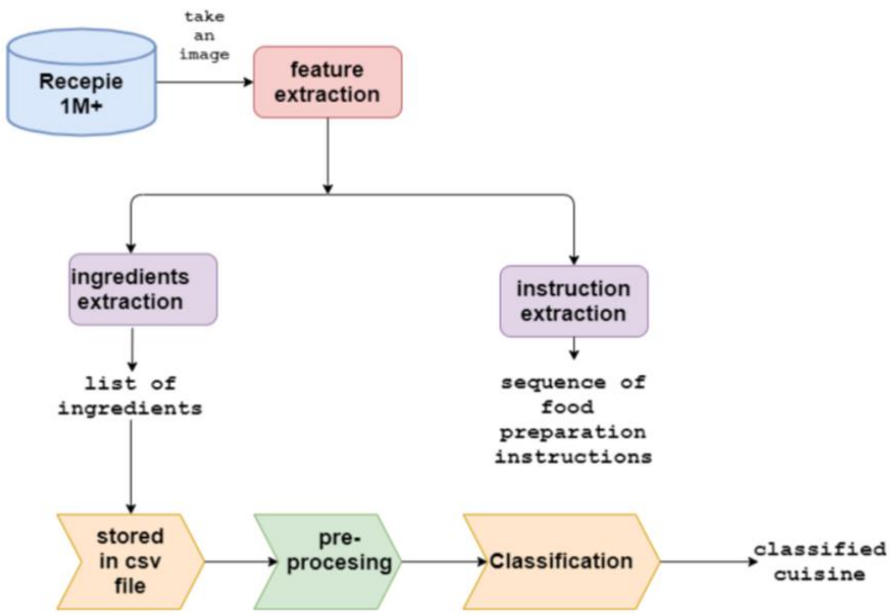

Fig. 1. Block Diagram

V. RESULTS

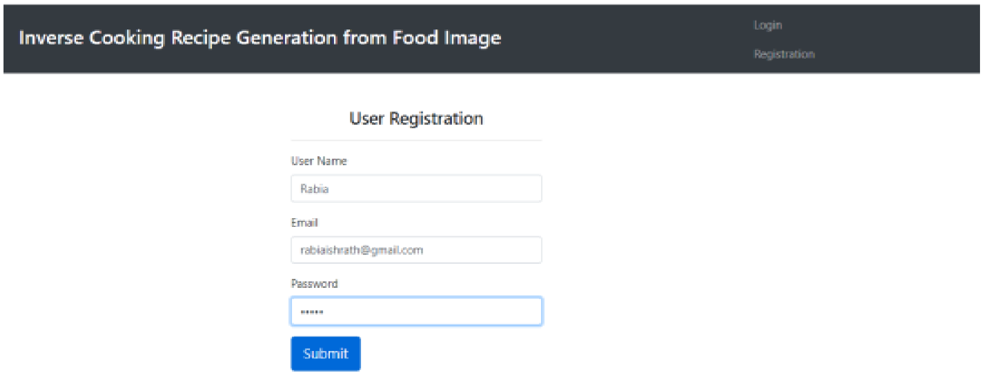

Fig. 2 New User Registration

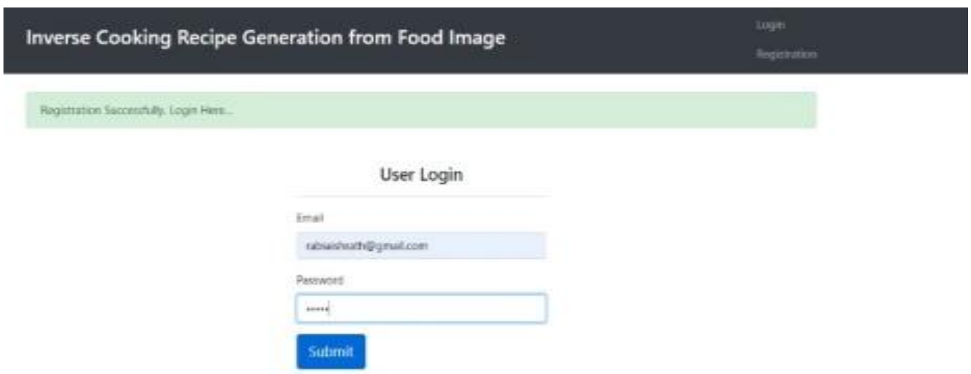

Fig. 3 User Login

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁑ Impact Factor 7.39 ⁑ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11784

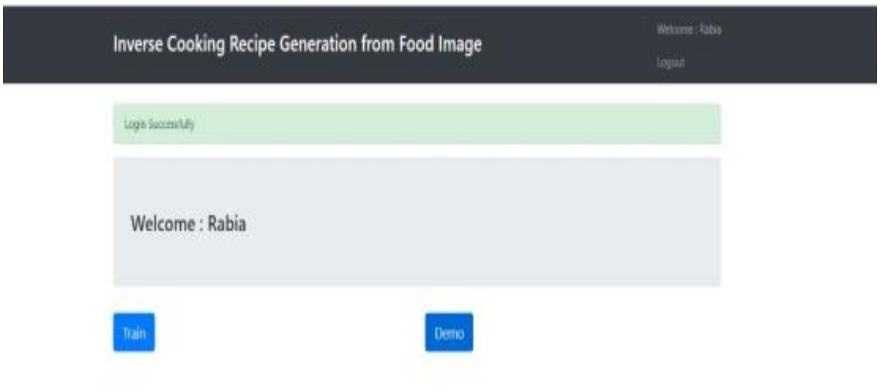

Fig. 4 Welcome Page

Fig. 5 Upload food Image

Fig. 6 Food Prediction

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ჯ Impact Factor 7.39 ჯ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11784

Fig. 7 Food Origin

Fig. 8 Ingredients Required

Fig. 9 Recipe & Youtube Link

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11784

VI. CONCLUSION

In this paper, 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.

ACKNOWLEDGMENT

We wish to thank everyone who anonymously participated in the user studies.

REFERENCES

[1] Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. Food-101–mining discriminative components with random forests. In ECCV, 2014.
[2] Micael Carvalho, R´emi Cad`ene, David Picard, Laure Soulier, Nicolas Thome, and Matthieu Cord. Cross-modal retrieval in the cooking context: Learning semantic text-image embedding. 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. Chinesefoodnet: 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 (Inter- national Social Science Council), 1988.
[11] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N. Dauphin. Convolutional sequence to sequence learning. CoRR, abs/1705.03122, 2017.
[12] Yunchao Gong, Yangqing Jia, Thomas Leung, Alexander Toshev, and Sergey Ioffe. Deep convolutional ranking for multilabel image annotation. CoRR, abs/1312.4894, 2013.
[13] Kristian J. Hammond. CHEF: A model of case-based planning. In AAAI, 1986.
[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In CVPR, 2015.
[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, 2016.
[16] Luis Herranz, Shuqiang Jiang, and Ruihan Xu. Modeling restaurant context for food recognition. IEEE Transactions on Multimedia, 2017.
[17] Shota Horiguchi, Sosuke Amano, Makoto Ogawa, and Kiyoharu Aizawa. Personalized classifier for food image recognition. IEEE Transactions on Multimedia, 2018.

