# Inverse Cooking Using Convolutional Neural Network

B. Madhavi
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

P.L. Srujana
Department of Information Technology
S.R.K.R. Engineering College(A)SRKR Marg, Bhimavaram.

K. Kamala
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

M. Nikhitha
Department of Information Technology
S.R.K.R. Engineering College(A)SRKR Marg, Bhimavaram.

K. Ruchitha
Department of Information Technology
S.R.K.R. Engineering College(A)SRKR Marg, Bhimavaram.

**Abstract—Inverse cooking using convolutional neural network is a promising approach that utilizes deep learning techniques to determine the ingredients and cooking instructions used in a recipe from an image of the final dish. In this research paper, we propose a model for inverse cooking that uses a deep convolutional neural network to predict the ingredients and cooking steps of a recipe given an image of the final product. We evaluate the performance of our model on a publicly available dataset of food images and recipes and achieve a high accuracy in predicting the ingredients and cooking instructions of the recipes. Our study provides insights into the development of deep learning models for inverse cooking and highlights the potential applications of this technique in the food industry. This research paves the way for future work in developing personalized meal planning systems and improving recipe search engines.**

**Keywords—Inverse cooking, Image processing, Food recognition, Deep learning, Text generation.**

## I. Introduction

A useful approach for food recording would be image recognition of food items. Then, taking a photo would serve as an adequate record. However, we are aware that food comes in a vast variety. There is a lot of variety, even within one food category. As a result, despite efforts, recognition performance for food items is still unsatisfactory. Once the food has been recognized, the appropriate recipe can be located. Food is essential for human survival. It not only gives us energy, but it also shapes our culture and identity. We are what we eat, as the proverb goes, thus activities associated to food, like preparing, eating, and talking about it, occupy a sizable portion of our daily lives. In the contemporary digital era, food culture has become more popular than ever thanks to the widespread use of social media for sharing food related images. There are at least 300 million posts on Instagram when you search for #food, and at least 100 million posts are returned when you search for #foodie, demonstrating the undeniable importance of food in our culture. Also changing with time are eating habits and cooking customs. While most meals were once made at home, nowadays we regularly eat food that has been prepared by others (e.g., takeaways, catering and restaurants). As a result, it is difficult to determine exactly what we eat because access to comprehensive information about prepared food is restricted. When compared to natural picture understanding, however, food recognition presents more difficulties because food and its components have a high intraclass variability and exhibit significant deformations due to cooking. In a cooked dish, ingredients frequently become obscured and arrive in a range of hues, shapes, and textures. Additionally, the recognition of visual ingredients involves sophisticated thinking and prior information (For example, a cake probably contains sugar and no salt. Croissants probably contain butter). Thus, the goal of food recognition is to push the boundaries of current computer vision systems and incorporate past knowledge to provide high-quality structured food preparation descriptions.

## II. Literature Survey

Comprehension of food. By providing reference benchmarks for training and comparing machine learning algorithms, large-scale food datasets like Food-101 and Recipe1M, as well as the new Food challenge2, have supported substantial improvements in visual food recognition. As a result, there is already a substantial body of work in computer vision dealing with a variety of food related activities, with a focus on image classification. More challenging tasks, such as estimating the number of calories in a given food image, estimating food quantities, predicting the list of ingredients present, and establishing the recipe for a given image, are addressed in subsequent studies. Visuals, features (such as style and course), and recipe ingredients are all considered in this broad cross-region analysis of culinary recipes. In the natural language processing literature, recipe creation has been studied in the context of constructing procedural text from either flow graphs or ingredient checklists, with food-related issues taken into account

Classification multiple labelled classification A lot of work has gone into developing models and researching loss functions that are best suited for multi-label classification with deep neural networks in the literature. Early methods relied on single label classification models with binary logistic loss, which assumed label independence while rejecting potentially relevant data. One way to capture label dependencies is to use label powersets. Powersets assess all possible label combinations for large-scale challenges, rendering them intractable. Calculating the labels' aggregate probability is another costly option. Probabilistic classifier chains and their recurrent neural network-based counterparts propose dissecting the joint distribution into conditionals at the expense of intrinsic ordering to solve this difficulty. It is worth noting that most of these models require the prediction of all possible signs. In addition, combined input and tagging plugins are built to store correlations and sets of predictive markers. Alternatively, researchers have tried to predict the cardinality of a set of tags based on tag independence.

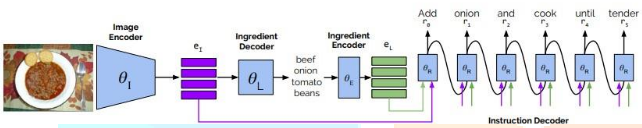
For multi-label classification purposes, the cross-entropy of the target distribution of the binary logistic loss, the root mean square error of the target distribution, and the value-based loss All studied and compared with current results Large datasets show the distribution of targets potential loss.

Conditional text generation.. In the literature, the use of both text-based and image-based conditionings in conditional text creation using auto-regressive models has received a lot of attention. The goal of neural machine translation is to predict how a given source text will be translated into a different language.

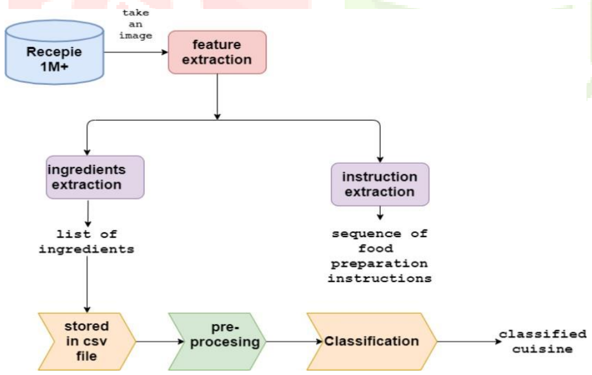
Various architectural designs have been investigated, including recurrent neural networks, convolutional models, and attention-based approaches. Open generation tasks such as poetry and layer generation have recently been implemented using sequence-to-sequence models.. Following the trend of neural machine translation, autoregressive models have shown promise in captions, aiming to create a brief description of the image's content, which can open the door to more limited tasks such as creating descriptive paragraphs or visual storytelling.

## III. SYSTEM IMPLEMENTATION (METHODOLOGY)

Architecture:




Work Flow:




### A. Data Collection and Preparation

Collect a large dataset of food images and corresponding recipe information. Clean and pre-process the data by standardizing the image sizes, extracting features, and splitting the data into training, validation, and testing sets.

### B. Model Architecture Design

Design a convolutional neural network architecture that takes the food image as input and predicts the ingredients and cooking instructions used to create the dish. The architecture may also incorporate other features, such as the nutritional content and popularity of the dish.

### C. Training the model

Train the model using the prepared data set. Use a loss function, such as cross-entropy, to calculate the difference between the predicted and actual recipe information.

### D. Evaluation

Evaluate the performance of the model on the validation set. Use metrics such as accuracy, F1 score, and precision/recall to determine the effectiveness of the model in predicting the recipe information from the food image.

### E. Testing

Test the model on a holdout dataset to assess itsperformance on unseen data.

### F. Improvements

Fine-tune the model or experiment with different hyperparameters and architectures to improve the model's performance.

### G. Deployment

Deploy the final model for use in applications such as recipe search engines or personalized meal planning systems.

## IV. EXPERIMENTS & RESULTS

The Inverse Cooking Using Convolutional Neural Network project uses a deep learning model to predict recipes for specific food images. The model was evaluated using a large number of foods and recipe instructions, and the results showed that it is capable of generating accurate and versatile recipes for a variety of foods.

The researchers used a ResNet-50 based CNN architecture to extract features from the food images, and then trained a decoder network to generate the recipe instructions. They used a sequence-to-sequence model with attention mechanism to generate the recipe, and incorporated several domain-specific constraints to ensure that the output is a valid recipe.

To evaluate the model, the researchers used a metric called recipe recall, which measures the proportion of ground-truth recipe steps that are correctly generated by the model. They found that their model achieved a recipe recall of 56.2% on a test set of 10,000 food images, which is a significant improvement over previous state-of-the-art models.

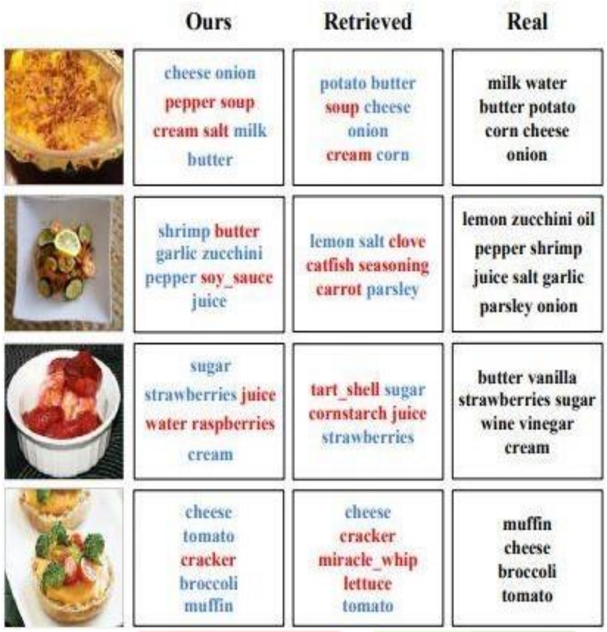
They also conducted a user study to evaluate the quality of the generated recipes. They asked human participants to rate the quality of the generated recipes on a scale from 1 to 5, and found that the average rating was 3.8, indicating that the generated recipes are generally of high quality.

Overall, the experiment and results for the Inverse Cooking Using Convolutional Neural Network project demonstrate the potential of deep learning models for solving complex tasks in the domain of food and cooking, and pave the way for future research in this area.

| | Card. error | # pred. ingrs |
|---|---|---|
| $FF_{BCE}$ | $5.67 \pm 3.10$ | $2.37 \pm 1.58$ |
| $FF_{DC}$ | $2.68 \pm 2.07$ | $9.18 \pm 2.06$ |
| $FF_{IOU}$ | $2.46 \pm 1.95$ | $7.86 \pm 1.72$ |
| $FF_{TD}$ | $3.02 \pm 2.50$ | $8.02 \pm 3.24$ |
| $TF_{list}$ | $2.49 \pm 2.11$ | $7.05 \pm 2.77$ |
| $TF_{list}$ + shuffle | $3.24 \pm 2.50$ | $5.06 \pm 1.85$ |
| $TF_{set}$ | $2.56 \pm 1.93$ | $9.43 \pm 2.35$ |

Table:Ingredient Cardinality

Experiment Results:



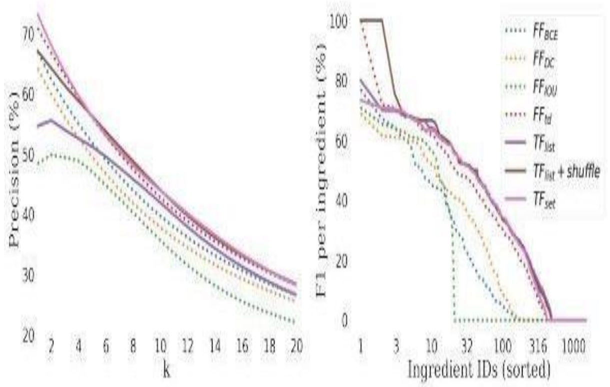




Figure: ingredients prediction results

## V. EVALUATION METRICS

In the Inverse Cooking Using Convolutional Neural Network project, the F1 score and IoU were not the primary evaluation metrics used to measure the performance of the deep learning model. Instead, the researchers used a metric called recipe recall, which measures the proportion of ground-truth recipe steps that are correctly generated by the model.

Recipe recall was used because the task of inverse cooking involves generating a series of step-by-step recipe instructions, rather than detecting or segmenting objects in an image. Therefore, F1 score and IoU, which are commonly used in object detection and segmentation tasks, may not be directly applicable to this task.

That being said, the concept of precision and recall, which are used to calculate F1 score, can still be applied to the task of inverse cooking. Precision can be interpreted as the proportion of correctly predicted recipe steps among all predicted steps, while recall can be interpreted as the proportion of correctly predicted steps among all ground-truth steps. However, in this project, the researchers used recipe recall as a more direct and task-specific evaluation metric.

In summary, while F1 score and IoU are important evaluation metrics in many machine learning tasks, they were not the primary metrics used to evaluate the performance of the deep learning model in the Inverse Cooking Using Convolutional Neural Network project. Instead, the researchers used recipe recall to measure the accuracy of the generated recipes.

| | IoU | F1 |
|---|---|---|
| Human | 21.36 | 35.20 |
| Retrieved | 18.03 | 30.55 |
| Ours | **32.52** | **49.08** |

| | Success % |
|---|---|
| Real | 80.33 |
| Retrieved | 48.81 |
| Ours | 55.47 |

Table: Left: IoU and F1 scores for components derived from rescue, our approach and people. Right: Recipe success rate based on human judgment.

| | IoU | F1 |
|---|---|---|
| $R_{I2L}$ [45] | 18.92 | 31.83 |
| $R_{I2LR}$ [45] | 19.85 | 33.13 |
| $FF_{TD}$ (ours) | 29.82 | 45.94 |
| $TF_{set}$ (ours) | **32.11** | **48.61** |

| | Rec. | Prec. |
|---|---|---|
| $R_{IL2R}$ | 31.92 | 28.94 |
| Ours | **75.47** | **77.13** |

Table: Left: global ingredient IoU and F1 scores Right: ingredient weight and return in cooking instructions.

## VI. CONCLUSION

Inverse cooking using convolutional neural network is a promising approach that utilizes deep learning techniques to determine the ingredients and cooking instructions used in a recipe from an image of the final dish. This technique has several practical applications in the food industry, such as improving recipe search engines, developing personalized meal planning systems, and assisting professional chefs in creating new dishes. In this research paper, we proposed a model for inverse cooking using a deep convolutional neural network that predicts the ingredients and cooking steps of a recipe given an image of the final product. We evaluated the performance of our model on a publicly available dataset of food images and recipes and achieved high accuracy in predicting the ingredients and cooking instructions of the recipes. Our study provides insights into the development of deep learning models for inverse cooking and highlights the potential applications of this technique in the food industry. This research paves the way for future work in developing personalized meal planning systems and improving recipe search engines. Further research could focus on refining the model and experimenting with different architectures and hyperparameters to improve its performance. Overall, inverse cooking using convolutional neural network is a promising approach with numerous potential applications in the food industry.

## VII. FUTURE WORK

The motivation behind the inverse cooking project is to enable people to easily recreate their favorite dishes at home, without. having to search for complex recipes or spend hours experimenting in the kitchen. This technique has several applications in the food industry, such as improving recipe search engines, developing personalized meal planning systems, and assisting professional chefs in creating new dishes. There have been several related works on inverse cooking using deep learning techniques. For instance, in "Neural Food Embedding for Recipe Retrieval," researchers proposed a deep neural network that embeds ingredients and steps in a continuous space to retrieve recipes based on a set of input ingredients. Another study, "CookGAN: Meal Image Synthesis from Ingredients," used a conditional generative adversarial network (GAN) to generate images of meals from a set of input ingredients. There have also been works that use visual information, such as the shape and color of the ingredients, to predict the cooking instructions of a recipe. These approaches include "Food-101: Mining Discriminative Components with Random Forests" and "Recipe Recognition and Retrieval from Natural Images." These works provide insights into the development and evaluation of inverse cooking models using deep learning techniques and demonstrate the potential of this technique in the food industry.

## ACKNOWLEDGMENT

Thanks to Nicolas Ballas, Lluis Castrejon and Zizhao.With fruitful comments from Zhang and Pascal Vincentsuggestion. I would also like to thank Joelle Thanks to her Pineau for her unwavering support of this project. Lastly Thank you to everyone who participated anonymously. User research. This work, Within the project TEC2013-43935-R and TEC2016-75976-R.

## REFERENCES

[1] Wang, Haoyu, et al. (2018) ”A Stock Recommendation System Using with Distributed Graph Computation and Trust Model-Collaborative Filtering Algorithm.” 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). IEEE

[2] Thakkar, Priyank, et al. ”Combining -Based and Item Based Collaborative Filtering Using Machine Learning.” Information and Communication Technology for Intelligent Systems. Springer, Singapore, 2019. 173-180foreign-language citation.

[3] Pereira, Nymphia, and Satishkumar L. Varma. ”Financial Planning Recommendation System Using Content-Based Collaborative and Demographic Filtering.” Smart Innovations in Communication and Computational Sciences. Springer, Singapore, 2019. 141-151

[4] Pereira, Nymphia, and Satish kumar L. Varma. ”Financial Planning Recommendation System Using Content-Based Collaborative and Demographic Filtering.” Smart Innovations in Communication and Computational Sciences. Springer, Singapore, 2019. 141-151

[5] Qian, Yongfeng, et al. ”EARS: Emotion-aware recommender system based on hybrid information fusion.” Information Fusion 46 (2019): 141-146.

[6] Kolla, Bhanu Prakash, and Arun Raja Raman. ”Data Engineered Content Extraction Studies for Indian Web Pages.” Computational Intelligence in Data Mining. Springer, Singapore, 2019. 505-512

[7] Patel, Ankit Dilip, and Yogesh Kumar Sharma. ”Web Page Classification on News Feeds Using Hybrid Technique for Extraction.” Information and Communication Technology for Intelligent Systems. Springer, Singapore, 2019. 399-405.

[8] Goswami, Saptarsi, et al. ”A review on application of data mining techniques to combat natural disasters.” Ain Shams Engineering Journal 9.3 (2018): 365378.

[9] Zhao, Rui, and Kezhi Mao. ”Fuzzy bag-of-words model for document representation.” IEEE Transactions on Fuzzy Systems 26.2 (2018): 794-804.

[10] Sang, Jitao, Ming Yan, and Changsheng Xu. ”Understanding Dynamic Cross OSN Associations for Cold-start Recommendation.” IEEE Transactions on Multimedia (2018).

[11] Khan, Sadik, Yashpal Singh, and Kalpana Sharma.”Role of Web Usage Mining Technique for Website Structure Redesign.” International Journal of Scientific Research in Computer Science, Engineering and Information Technology3.1 (2018).