

South Indian Recipe Recommendation from Ingredient Image

Praguna Manvi
Undergraduate Student,
Department of CSE, B.M.S. College
of Engineering, Bangalore, India

Vishnu M. P.
Undergraduate Student
Department of CSE, B.M.S. College
of Engineering,
Bangalore, India

Nitish Vivian Maximus
Undergraduate Student,
Department of CSE, B.M.S. College
of Engineering, Bangalore, India

Parva Chauhan
Undergraduate Student
Department of CSE, B.M.S. College
of Engineering, Bangalore, India

Umadevi V.
Associate Professor and Head
Department of CSE,
B.M.S. College of Engineering,
Bangalore, India

Muzammil Hussain
Founder, Tika Data Services Pvt.
Ltd., Bangalore, India

ABSTRACT

Applications involving ingredient recognition are very limited and most do not work in less ideal conditions like the ones faced in a typical kitchen. The main reason for this is the dataset that the existing models are based on. These datasets do not account for real-world factors like noise, blur, etc. in the input image since they are trained on images obtained from controlled and nearly idealistic environments. For these reasons, a new dataset was created, consisting of real-world images which represent the scenarios users are most likely to face during daily use. A simple and robust system was developed that aims to address this issue. A multi-label classification model was built to identify multiple ingredients present in a single image. A personalized recommendation system that recommends a list of South Indian recipes to users based on the ingredients identified was also developed.

General Terms

Image Recognition, Recommendation, Machine Learning, Deep Learning.

Keywords

Ingredients, recipe, multi-ingredient dataset, recommendation, multi-label classification, Recipe Recommendation, KC47 Kitchen dataset, South Indian.

1. INTRODUCTION

Deciding which dish to prepare with available ingredients is a problem that everybody can relate to. The work presented here aims to alleviate that problem and expedite the cooking process by providing a simple and robust solution which helps the user by providing a list of South Indian dishes along with their recipes based on the image of ingredients available.

The majority of research on ingredient identification is usually conducted in environments in which backgrounds, lighting and even framing of the ingredients in the image is controlled and these result in idealistic scenarios, which is not always the case. The most common scenarios, however do not possess many of the traits of ideal scenarios and result in input images which are blurry, low resolution or even distorted.

This paper presents a multiple ingredient recognition solution

using a multi-label classification network and a recommendation system to recommend users a list of South Indian recipes based on the identified ingredients from the input image.

To aid the process of training and execution of this task, a multi-class (vegetables, fruits, spices, etc.) image dataset was constructed.

The workflow of the proposed system is shown in Fig 1.

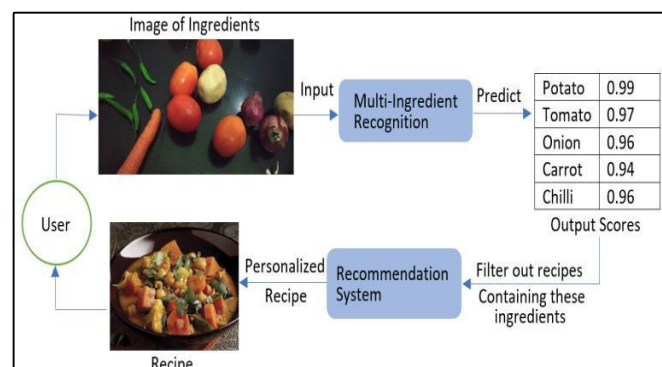


Fig 1. Recipe Recommendation from Ingredient image The work is summarized as follows:

- Showcased a new multi-class ingredient image dataset captured in a kitchen environment and South Indian recipe and reviews dataset to enable more research.
- Developed a multi ingredient classification system based on the dataset constructed.
- Developed a personalized recipe recommendation system for South Indian recipes based on ratings collected.

2. RELATED WORK

2.1 Multi-ingredient Image Recognition

This paper [1] proposes a system which can identify multiple ingredients present in an image and also recommend recipes based on the results of identification. They make use of Spatial Regularization Networks for ingredient recognition. The dataset is constructed in a supermarket environment. Recommendation uses Neural collaborative filtering method.

2.2 Multi-Label Image Classification

The approach used in this paper [2] is among the first Deep Neural Network based label embedding frameworks for multi-label classification. This paper makes use of an approach called Canonical Correlated AutoEncoder to achieve the task of multi-label classification with more accuracy through better relating of features with label domain data. The model is trained to learn from an image label space and then predict labels accordingly. The proposed method also works with training data that has labels missing.

2.3 Recommendation of Recipes

This paper [3] proposes a unique method of recommending recipes using Machine learning models and Bayesian Optimization, rather than Collaborative Filtering which is a widespread approach. This method is unique and fast in that it does not wait to get real world data and uses simulated data for training the model. This saves a lot of time and an accurate model can be developed quickly and with lesser human resources.

2.4 Ingredient Recognition

The approach in [4] compares several image descriptors and then chooses a suitable combination. It makes use of feature fusion which combines more than one feature to provide a feature vector that can be better used for recognition. Preprocessing of images includes segmentation of the object in the picture from the background. This approach uses Convolutional Neural Networks to build a model that can identify the fruits or vegetables in the given input image.

3. INGREDIENT RECOGNITION

3.1 KC47 Kitchen Dataset

The ingredient datasets that are currently available contain images of only a single class. Hence, the ingredient classification models generated by these datasets cannot handle real world scenarios, such as kitchen environments, with much accuracy. Therefore, a Multi-Class image dataset consisting of 47 classes that include vegetables, fruits, spices and pulses, set in a kitchen environment was constructed to meet the requirements.

The following are some unique aspects of this dataset:

- Images in this dataset were captured in varying conditions and backgrounds, with ingredients kept on counter tops, wooden table, white sheet and table cloth in both low and high light exposures.
- There are a varying number of classes for each image instance.
- Images have resolutions ranging from 1500 x 1500 to 4600 x 3500.
- Images were captured using 3 different smartphones.

3228 train images and 933 test images were captured through the collection procedure. Fig 2 represents a sample of images from the KC47 dataset and shows an example of the results of the multi ingredient recognition model and the frequency distribution of the 47 ingredients is shown in Fig 3.

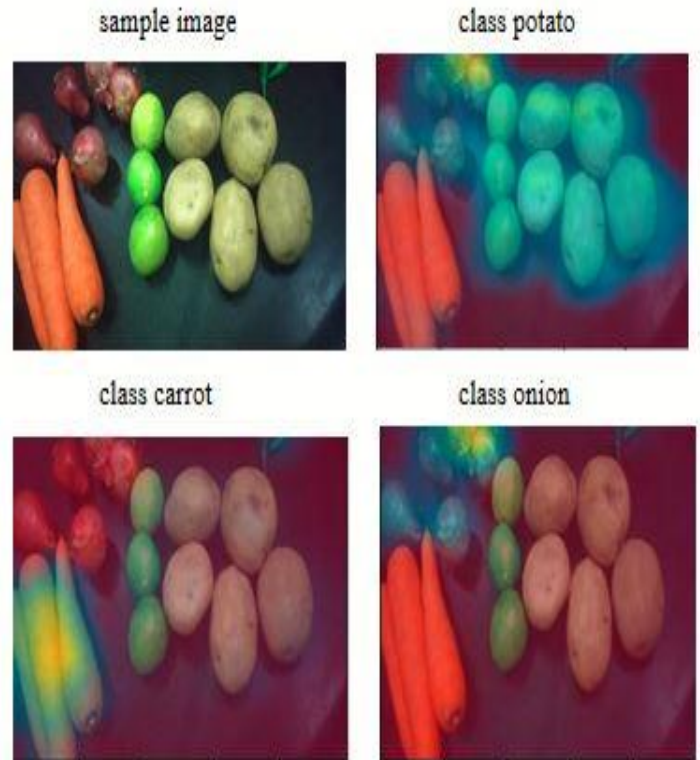


Fig 2. Sample Image from KC47 dataset with identified classes

3.2 Multi-Ingredient Recognition

Multi-Ingredient recognition is a multi-class classification problem with the ingredient images captured in a regular kitchen environment. The classification models are trained on the KC47 dataset.

Various state of the art approaches were experimented on the dataset. The approaches used in this work are: Resnet-101[5], a deep network, which overcomes vanishing gradient problem using residual blocks. InceptionV3 [6], a computationally efficient approach with aggressive regularization. InceptionResnet [7], combines the above two approaches. DenseNet [8], a computational and memory efficient approach consisting of connected dense blocks.

All the experimented approaches were transfer learned on imagenet weights. The varying input images while training was resized to 1000 x 1000 aspect ratio. Keras and TensorFlow were used to implement the models. In order to avoid overtraining, data augmentation was performed. The loss function binary cross entropy was used while training. Adam optimizer and optimized learning rates had been employed.

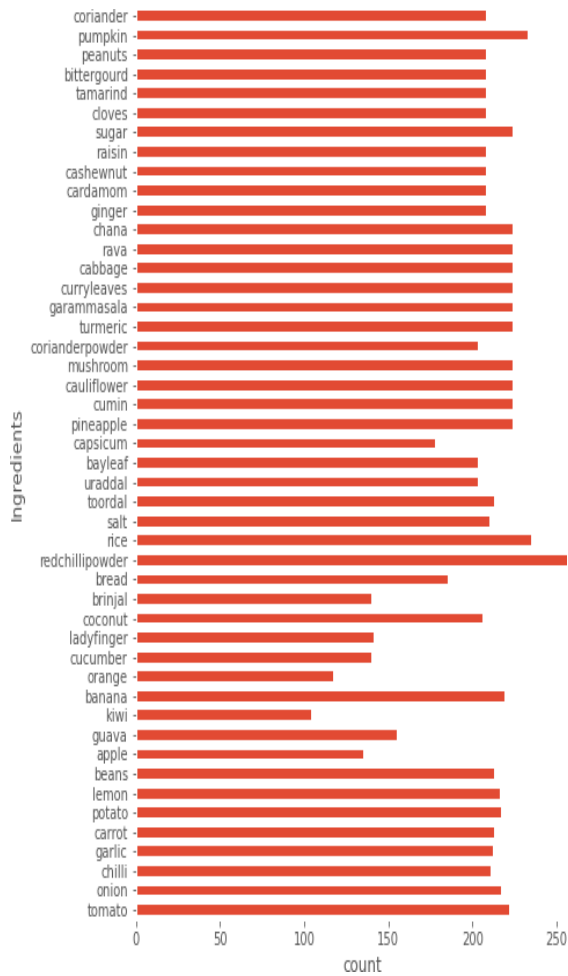


Fig 3. List of ingredients

4. RECOMMENDATION

4.1 South Indian Recipes and Reviews

Dataset

A South Indian recipe dataset was constructed by taking recipes of dishes from various websites. The dataset currently has 126 South Indian recipes.

The rating dataset was constructed by collecting user responses through a web application [9]. It was developed using the MERN (MongoDB Express ReactJS NodeJS) stack. During the data collection process, duplicate responses with the same IP address were allowed. The respondents rated the dishes from the South Indian Recipe Dataset on the scale of 1 to 5.

A total of 1502 responses was received. Each respondent was asked to rate at least five dishes of their choice while submitting their responses. Figure 4 shows the distribution of the number of ratings in the dataset. It can be observed that the majority of respondents have rated 5 to 7 dishes.

4.2 Recommendation Model

The recommendation model is aimed at understanding the user's preference of dishes and to provide them with recommendations. The recipes are filtered out using the recognized ingredients. The recommendation is performed from these recipes.

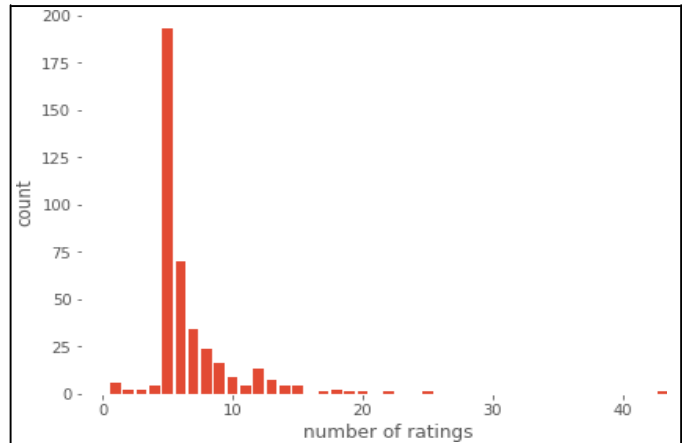


Fig 4. Distribution of number of ratings per respondent

Content-based and collaborative filtering (CF) [10] are common solutions in the recommendation domain. However, since there is a dearth of interaction, the implicit connection that exists among users and recipes is difficult to uncover when using the traditional CF method.

The two main types of CF are: (i) UserCF: measures how similar two users are and recommends items that these similar users liked. (ii) ItemCF: measures the similarity of ratings of different items and recommend items that are similar. There are a few issues that result when using this approach. The main issue that arises is the scalability, computation and sparsity of the user-item matrix.

Deep neural networks were used to model the interactions between users and items and this circumvented the issues that arise when implementing CF traditionally.

A Neural Matrix Factorization (NeuMF) [11] model was implemented, which is a combination of Generalized Matrix Factorization (GMF) [12] model and Multilayer Perceptron model (MLP). This combination results in taking advantage of linearity and non-linearity [13] of the models respectively.

The MLP network is a densely connected network, with four dense hidden layers with Rectified Linear Unit (ReLU) activation function, using binary cross entropy loss function and Adam optimizer.

The GMF network was constructed on similar lines, with output layer being densely connected with a single neuron and uses binary cross entropy loss function with Adam optimizer. It is a point wise matrix factorization to approximate factorization of a matrix into two matrices.

5. EXPERIMENTATION

5.1 Ingredient Identification

Accuracy is the number of correct predictions made with respect to all the predictions made. Precision is expressed as the proportion of true positives to all instances classified as positive. Recall is the proportion of true positives to all the instances that are actually positive. F1 score is the harmonic mean of precision and recall. Mean Average Precision (mAP) is a metric used to evaluate object identifiers. It is the mean of average precision that is calculated for all classes. Metrics used were macro and micro precision (P-C and P-O), macro and micro recall (R-C and R-O), macro and micro accuracy (A-C and A-O), macro and micro F1 measure (F1-C and F1-

O) for performance comparison. These metrics were calculated on predictions and true values of 900 test images and InceptionV3 yielded superior results as seen in table 1.

Table 1. Quantitative comparison for Multi Label Classification

Approach	mAP	P-C	P-O	R-C	R-O
DenseNet	86.86	82.16	81.15	71.3	71.26
Inception Resnet	90.49	89.82	88.08	72.48	72.55
Resnet101	90.55	81.03	78.45	89.53	90.14
InceptionV3	91.1	86.45	85.53	80.53	81.1

Approach	A-C	A-O	F1-C	F1-O
DenseNet	97.18	97.18	74.96	75.88
Inception Resnet	97.68	97.68	76.5	79.56
Resnet101	97.84	97.84	83.77	83.89
InceptionV3	97.97	97.97	82.28	83.26

5.2 South Indian Recipe Recommendation

Hit Ratio (HR@10) and Normalized Discounted Cumulative Gain (NDCG@10) were used as the evaluation metrics for the recommendation system. The Hit Ratio (HR@10) measures the number of hits in the top ten recommended results for each user. Normalized Discounted Cumulative Gain (NDCG@10) prioritizes by the hit position logarithmically and performs normalization, hence the weightage is logarithmically decreased moving down the ranks. These metrics were calculated on test data, constructed by taking one item for each user present in the training data.

Item-based and user-based collaborative filtering (ItemCF) and (UserCF) were implemented as the baseline models. The results improved by employing GMF, MLP and NeuMF approaches. GMF approach yields superior results as seen in table 2.

Table 2. Quantitative comparison for Recommendation

Approach	HR@10	NDCG@10
UserCF	6.5312	2.8716
ItemCF	7.0439	3.1669
NeuMF	23.72	12.23
MLP	23.86	13.042
GMF	27.005	13.994

6. CONCLUSION

This work presented a solution for the task of multi ingredient recognition and subsequent recommendation of South Indian recipes. This was achieved using a combination of deep learning approaches - Multi-label Classification and Recommendation systems.

Many approaches to multi-ingredient recognition on the dataset were explored. A personalized South Indian recipe recommendation system was also developed. To enable further research on this topic, a multi-ingredient image dataset and South Indian recipes and user ratings dataset was constructed.

7. ACKNOWLEDGEMENT

The authors would like to thank B.M.S. College of Engineering, Bangalore and Tika Data Services Pvt. Ltd., for providing support and guidance.

8. REFERENCES

- [1] Zhang, Lin, Jianbo Zhao, Si Li, Boxin Shi, and Ling-Yu Duan. "From Market to Dish: Multi-ingredient Image Recognition for Personalized Recipe Recommendation." In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1252-1257. IEEE, 2019.
- [2] Yeh, Chih-Kuan, Wei-Chieh Wu, Wei-Jen Ko, and Yu-Chiang Frank Wang. "Learning deep latent space for multi-label classification." In *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.
- [3] Garrido-Merchán, Eduardo C., and Alejandro Albarca-Molina. "Suggesting cooking recipes through simulation and bayesian optimization." In *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 277-284. Springer, Cham, 2018.
- [4] Rocha, Anderson, Daniel C. Hauagege, Jacques Wainer, and Siome Goldenstein. "Automatic fruit and vegetable classification from images." *Computers and Electronics in Agriculture* 70, no. 1 (2010): 96-104.
- [5] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
- [6] Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-2826. 2016.
- [7] Nguyen, Long D., Dongyun Lin, Zhiping Lin, and Jiuwen Cao. "Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation." In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1-5. IEEE, 2018.
- [8] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708. 2017.
- [9] ratingcollector.net, Website created for collecting ratings from users. URL: www.ratingcollector.net [Last accessed: July 2020]
- [10] Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl. "Item-based collaborative filtering recommendation algorithms." In *Proceedings of the 10th international conference on World Wide Web*, pp. 285-295. 2001.
- [11] He, Xiangnan, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. "Neural collaborative filtering." In *Proceedings of the 26th international conference on world wide web*, pp. 173-182. 2017.
- [12] Dziugaite, Gintare Karolina, and Daniel M. Roy. "Neural network matrix factorization." *arXiv preprint arXiv:1511.06443* (2015).
- [13] Krishna, Vaibhav, Tian Guo, and Nino Antulov-Fantulin. "Is Simple Better? Revisiting Non-linear Matrix Factorization for Learning Incomplete Ratings." In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 1289-1293. IEEE, 2018.