Recipe retrieval from food image classification

Adarsh Singh Kushwah
adarsh21111@iiitd.ac.in
MT21111

Akash Rawat
akash21005@iiitd.ac.in
MT21005

Charisha Phirani
charisha21117@iiitd.ac.in
MT21117

Niharika
niharika21132@iiitd.ac.in
MT21132

Shubham Rana
shubham21092@iiitd.ac.in
MT21092

1 ABSTRACT

The recipe retrieval system from food image classification focuses on the categorization of food images into their respective classes and retrieving the recipes from the food image provided by the user. The work also aims to retrieve the recipes from the ingredients provided by the user. If the user is short of some ingredients, the model gives the best recipe to be cooked from those ingredients. The system provides an interactive user interface for the recipe retrieval from food images as well as from the ingredients provided by the user.

2 INTRODUCTION

Eating habits affect the lifestyle. The technology has moved from text based classification to image based classification. We are proposing a recipe retrieval system via food image classification and also generation of food recipes from ingredients provided as the text based inputs. In the Information Retrieval System designed here the user needs to upload the food image for which he wants to get the recipe or he can also take the food image from the camera and upload it, the system will provide the recipe of the food image provided. The user can also enter the ingredients and number of recipes he wants to have with those ingredients. The system will provide the recipes generated from those ingredients. The model is designed in a way that it can be used by a naive user.

3 LITERATURE SURVEY

Fotios S. Konstantakopoulos et al. proposed the deep learning-based food image recognition model on the MedGRFood dataset that contains 42,880 food images belonging to 132 different food classes. The author used the EfficientNetB2 deep learning model, pre-trained on the ImageNet dataset. After doing data augmentation on the dataset and fine-tuning the pre-trained model, got top-1 accuracy of 83.4% and top-5 accuracy of 97.8%.

It is difficult to maintain a healthy lifestyle with the right food choices. The paper presents the food ingredients retrieval from the food image datasets using CNN. Marc Bolanos et al. used datasets like Food101(101,000 images), Recipes5k(4,826 recipes), Ingredients101(446 ingredients). The researchers have used a model which assigns binary values for showing the presence or absence of ingredients in the food images using the sigmoid function. They have used InceptionV3 and ResNet50 with a precision of 53.43% and 58.55% respectively.

The food preferences affect the health and lifestyle. This paper presents the various classification techniques and deep learning models on various food image datasets for image based food recognition. Ghalib Tahir et al. used some datasets like UECFOOD-256(256 different food images), UECFOOD-100(100 different kinds of food images), Pakistani Food Dataset(4928 food images), Food-101(101000 food images) and PFID (1098 food images) and applied feature extraction over these using ResNet-50, DenseNet201 and InceptionResNet-V2. For food category classification the author has used Adaptive Reduced Class Incremental Kernel Extreme Learning Machine (ARCIKELM). The accuracy obtained for the model on dataset Food-101 is 87.3%, UECFOOD-100 is 88.7%, UECFOOD-256 is 76.51%, PFID is 100% and Pakistani Food is 74.8%.

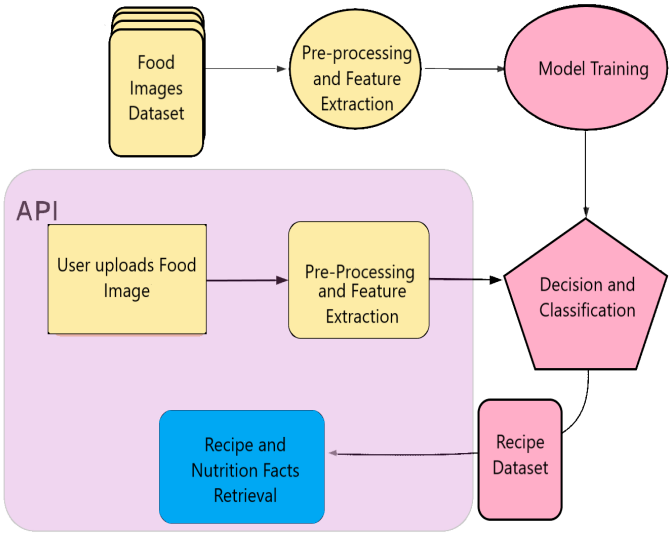

Figure 1: Recipe and nutritional facts retrieval from food image.

Malina Jiang proposed the food image classification with CNN. Malina experimented in the Food-101 dataset, which composed of 101,000 food images from 101 classes of foods, each type having 1000 images. The loss function used was the cross-entropy loss for all the models. The baseline models performed severely in terms of accuracy as the pictures had noise, and the number of classes was huge. Later, Malina used transfer learning using ResNet50, VGG16, and InceptionV3 models pre-trained on the ImageNet dataset. After fine-tuning the models in terms of custom preprocessing the images, hyper-parameter, and weights, it was concluded that the InceptionV3 model outperformed all other models with Top-1 accuracy of 61.4% and Top-5 accuracy of 85.2%.

Yuzhen Lu et al. proposed an optimal method to deal with the image classification problems, specifically in food. He has used the

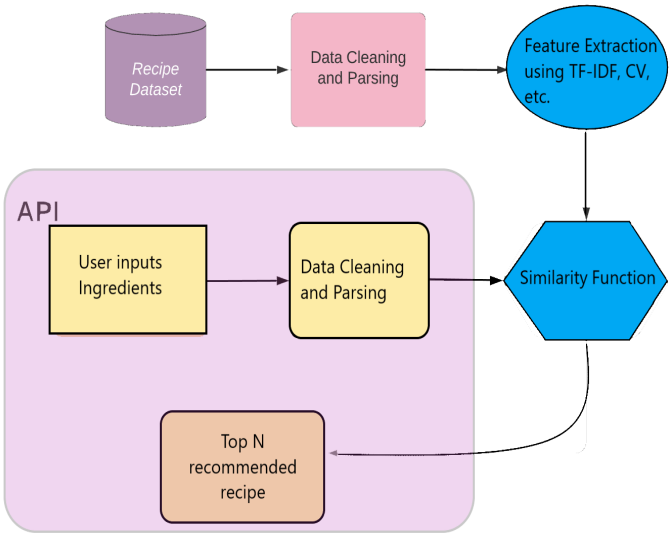

Figure 2: Recipe recommendation from ingredients

ImageNet dataset for performing the research. However, he has taken 5822 images from the large dataset. The author presented three methods, such as BoF with SVM model, pre-trained CNN model with five layers, and CNN with the expansion of image data technique. Upon evaluating the models, the best working model with 74% accuracy on the test dataset is CNN with the data expansion technique.

Jay Baxter proposed a system on identifying food using baseline models like Bag of SIFT and RGB Histogram. The author has compared the baseline RGB Histogram using the 7 Category data accuracy with ingredient-level features STF(Semantic Texton Forests), where the baseline has 71.2% accuracy and overall STF methods combine to give 81.2% accuracy, showing a 10% improvement in accuracy, telling us that every feature captures some information that another one doesn't.

4 METHODOLOGY

We created the Indian food image dataset that consists of more than 40 classes of food, each class having more than 200 different food images. After that, we used the data augmentation technique to increase the size of the training dataset by slightly modifying the original images of food using techniques such as flipping, zooming, cropping, and rotation. We have applied the Keras EfficientNetB2 deep learning model and transfer learning method to fine-tune the existing pre-trained model on the dataset. To fine-tune the EfficientNetB2 model on the dataset, we have used the following hyperparameters such as weights="ImageNet", pooling="MAX", regularizer="L2",activation="Softmax", loss="catagorical_crossentropy", matrices="accuracy", epoches="40". Dataset has been split into training_set=80%, testin_set=10% and validating_set=10%. Food Image Classification Result best accuracy on epoch=9 with validation accuracy=89.048% and loss=0.496. We have also proposed a method to generate the recipes from the list of ingredients provided as an input by the user. The method gives the most suitable recipes that can be cooked from those provided ingredients.The dataset we used

Figure 3: Food Image Classification Result

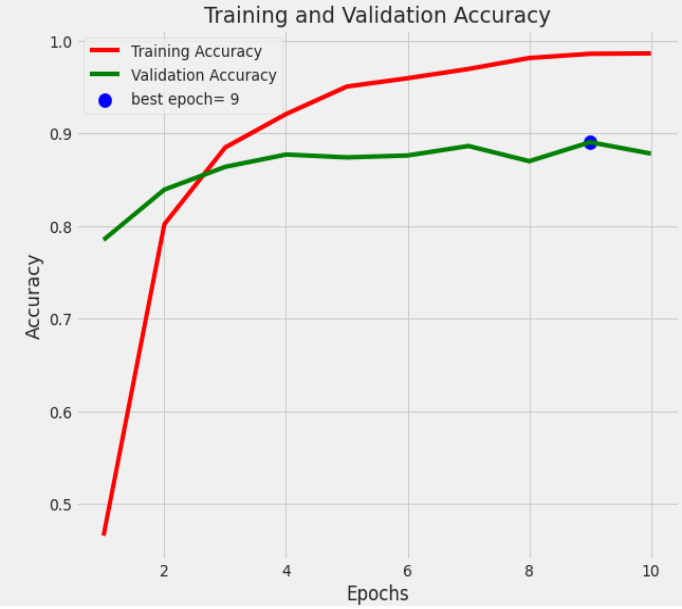

Figure 4: Training Validation Accuracy

is Indian food 101 which contains 255 food items with their ingredients. Using the count vectorization method, we evaluated the similarity between the ingredients list provided by the user and the list of ingredients of each recipe present in the dataset. There is also an interactive user interface that has been designed in which the

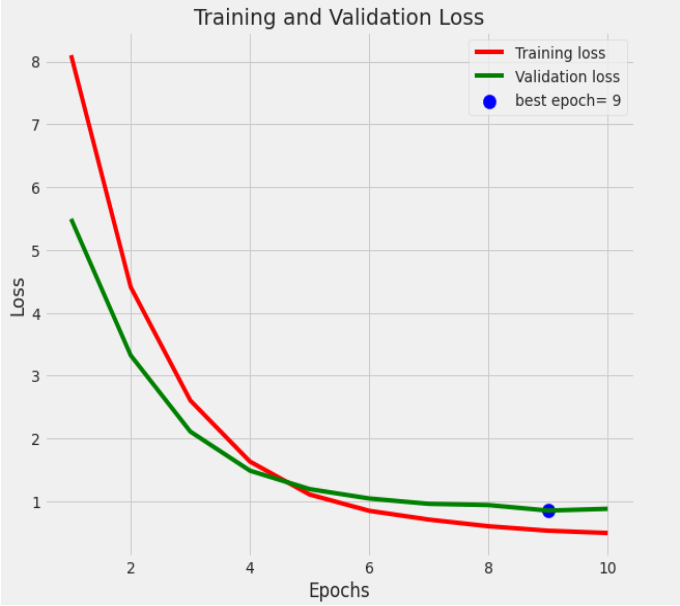

Figure 5: Training Validation Loss

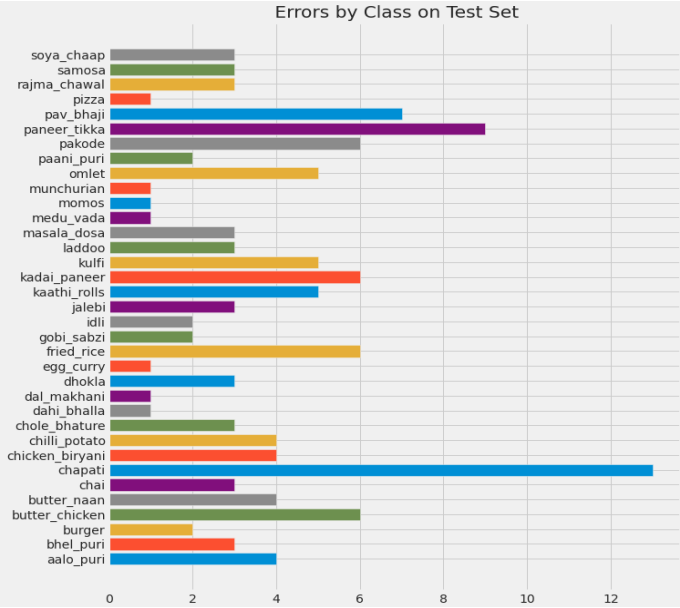

Figure 6: Error By Class

user needs to upload the food images and the model will be classifying the food image first and predicting its recipe. Another function provided by the model if the user enters the list of ingredients and the number of recipes he wants to retrieve from those ingredients, then the interface provides the list of recipes from those ingredients provided by the user. Using the count vectorization method, we evaluated the similarity between the ingredients list provided by

Input Ingredients	N (Number of Recipes)	Recommended Recipes
ghee wheat sugar	4	Boondi Gajar ka halwa Ghevar Gulab jamun
Maida flour, sugar	2	Balu shahi Shankarpali
Flour, ghee, kewra, milk	1	Ghevar

Figure 7: Recipe retrieval based on ingredients

the user and the list of ingredients of each recipe present in the dataset.

5 EVALUATION

After training the model entirely, we have saved the model. We have created a function which will take a food image as an input and, using the trained model, will evaluate the food image. It will give the probability of the image belonging to the particular class. The class having the most probability will be the output for that specific image. In Figure 3, food image classification is shown with the class name along with its probability belonging to the class from the model.

While training the model, we have seen the comparatively best result in epoch 9 with the least error and maximum accuracy on the validation dataset. Figure 4 and 5 shows that the best epoch was 9, with least validation loss of less than 1 and best validation accuracy close to 90%. In the initial dataset, some food classes have fewer images than other classes. To resolve this issue, we have augmented these classes to make the dataset balanced. After training the model and analyzing it, we have seen that still, some classes have more errors in the test set than others. Figure 6 shows the errors of each class on the test set.

We have used the pywebIO python library for the interactive user interface to make a web page on localhost. The web application will ask the user to either enter the food image or the ingredients for the recipe retrieval. In the case of food image as an input, our system will predict the food class and give the recipe and nutritional facts of that food class. In the case of food ingredients, our system will provide the top-k recipes as per the value of k, which a user gives. Figure 7 depicts that we have provided a random list of ingredients to the model for generating the most suitable top N recipes, where N is the number of recipes a user wants.

6 CONCLUSION

The recipe retrieval system from food image classification aimed to build a self compiled dataset containing 40 different classes of food with each class having more than 200 images of food. The model is created using Keras EfficientNetB2 deep learning model and transfer learning method to fine-tune the existing pre-trained model on the dataset. We also used a count vectorization method

for retrieval of recipes from the ingredients. We aim to explore more in the information retrieval domain and build more efficient methods for retrieving recipes after food image classification. We also aim to improve the accuracy of the model by tuning the hyper parameters.

REFERENCES

[1] Ghalib Tahir, Chu Kiong Loo.2021. *A Comprehensive Survey of Image-Based Food Recognition and Volume Estimation Methods for Dietary Assessment.* URL: https://arxiv.org/pdf/2106.11776.pdf

[2] Malina Jiang. 2019. *Food Image Classification with Convolutional Neural Networks.* URL:https://cs230.stanford.edu/projects_fall_2019/reports/26233496.pdf

[3] Yuzhen Lu. 2019. *Food Image Recognition by Using Convolutional Neural Networks (CNNs).*URL: https://arxiv.org/pdf/1612.00983.pdf

[4] Jay Baxter. 2017. *Food Recognition using Ingredient-Level Features.* URL: http://jaybaxter.net/6869_food_project.pdf

[5] Fotios S. Konstantakopoulos. 2021. *Mediterranean Food Image Recognition Using Deep Convolutional Networks* . URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=9630481

[6] Marc Bolaños. 2017. *Food Ingredients Recognition through Multi-label Learning.* URL: https://arxiv.org/abs/1707.08816