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 MOTIVATION

The food eating habits affect one's health. Many a times people are stuck with ingredients without knowing the recipe to be cooked with those. Since there is a wide variety of food recipes available from various ingredients, so the nutritious food recommendation from available ingredients can help adapt healthy eating habits leading to a healthy lifestyle. There are also the cases while viewing some food images people are eager to know how to cook those and are unable to get that particular food's recipe. There are also the cases when there are limited food ingredients available and one needs to get the best recipe that can be cooked out of those ingredients. This will also help to monitor one's food habits and cook the best food out of the available ingredients. So this motivated us to create a recipe retrieval system based on food image classification.

2 UPDATED PROBLEM STATEMENT

Recipe retrieval using food image classification aims to classify the food images using various deep learning models and recommend recipes along with the nutritional facts of the food image provided by the user. It can also recommend the best recipes from the food ingredients available using some similarity matrices. The system provides an interactive graphical interface so that users can get the recipes out of the food images or the ingredients they provide to the system. The dataset has been prepared by extracting various food image and recipes datasets available on the google images via using web scraping tool: selenium (an automated tool) and python script for saving the images.

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.

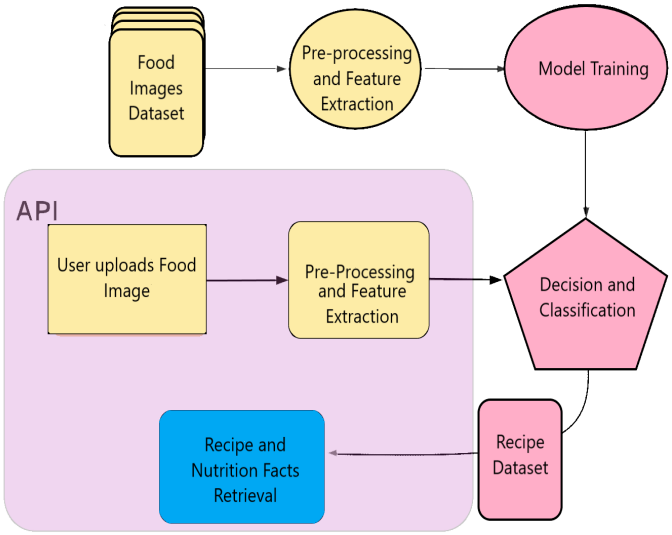

Figure 1: Recipe and nutritional facts retrieval from food image.

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%.

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

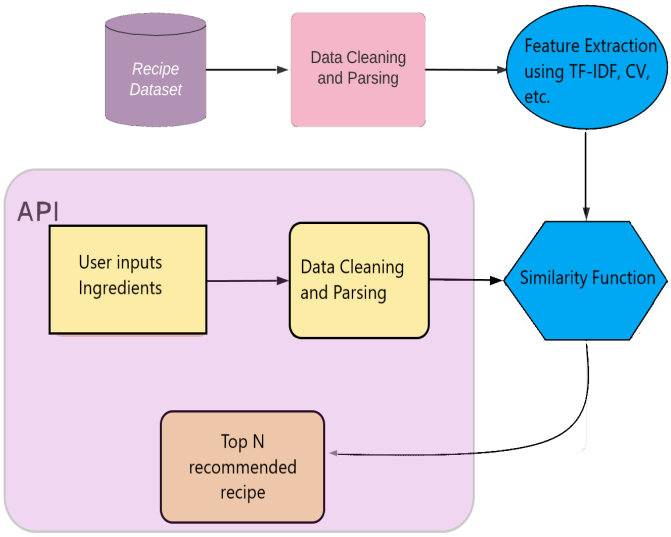

Figure 2: Recipe recommendation from ingredients

Figure 3: Food Image Classification Result

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 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 PROPOSED METHOD

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%, testing_set=10% and validating_set=10%. We got the best accuracy on epoch=9 with validation accuracy=89.048% and loss=0.496.

We propose 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 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.

5 BASELINE RESULTS

In Figure 3, food image classification is shown with the class name along with its probability belonging to the class from the model.

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%. Figure 6 shows the errors of each class on the test set.

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 PLAN OF WORK

6.1 Before Midterm review (Completed)

- Gathering and preparing datasets using some web scraping tools.
- Preparing some baseline models for food image classification.
- Recommendation of recipes using food images or ingredients.

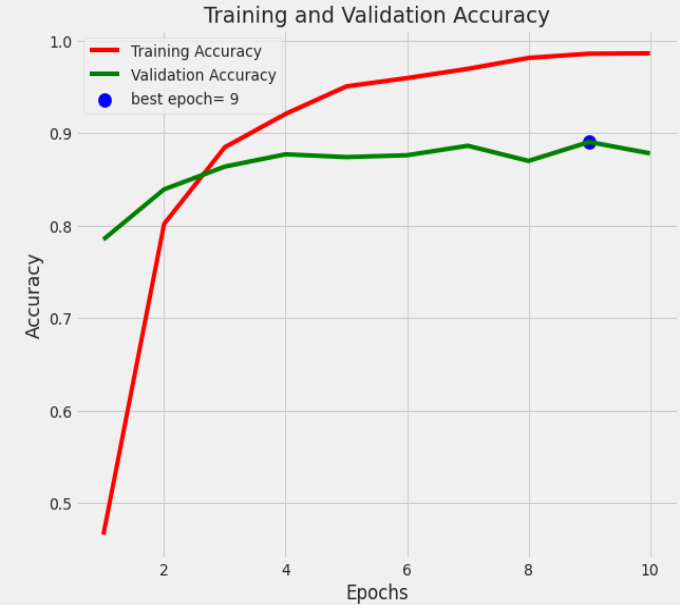

Figure 4: Training Validation Accuracy

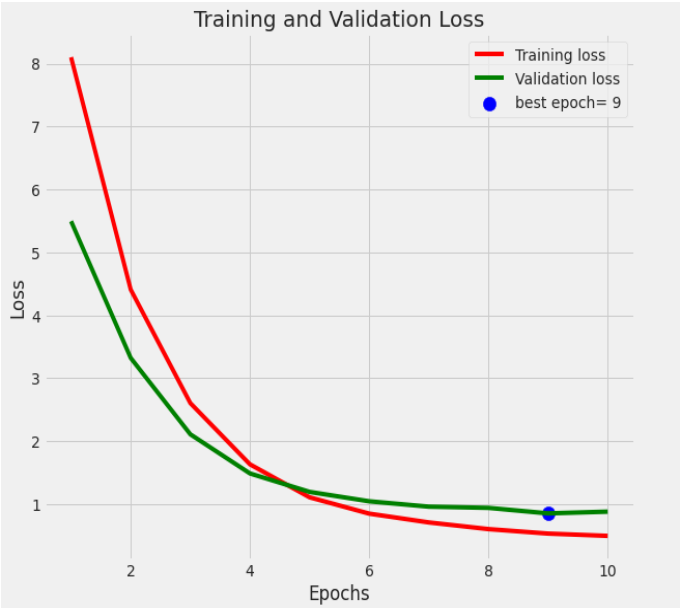

Figure 5: Training Validation Loss

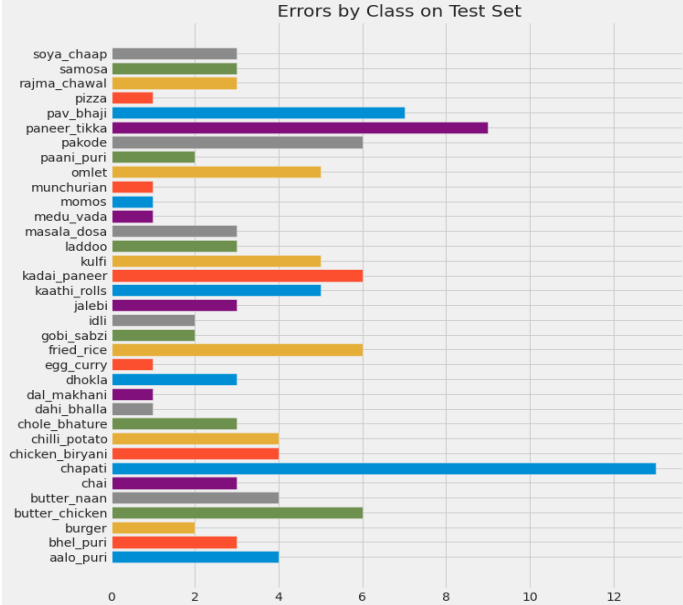

Figure 6: Error By Class

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

6.2 After Midterm

- Development of user interface.
- Deployment of model on cloud.

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