Food Cuisine Analysis using Image Processing and Machine Learning

Prasad Badhan[1], Rhugved Kale[2], Sahil Batham[3], Varun MK[4], Prof. Krishnendu Nair[5]

[1,2,3,4]UG Student, Dept. of Computer Engineering, Pillai College of Engineering, New Panvel, Maharashtra, India
[5]Professor, Dept. of Computer Engineering, Pillai College of Engineering, New Panvel, Maharashtra, India

---

Abstract - *Globalization has added eclectic cuisines to our everyday lives. We constantly find ourselves having a wide variety of food composed of different ingredients and different ways of preparation. It would be good to know the cuisine of the dish we are having and what would be a good fit for us in our future cravings along with its recipe, ingredients, nutritional value, and some other features. This approach if adopted by restaurant networks and food delivery apps would be a good strategy to attract and offer personalization to potential customers. Learning these applications motivated us to come up with this system of predicting and analyzing the cuisine of a plate of food given either its image or its list of ingredients.*

Key Words: **Cuisine Prediction, Food, Recipe, Machine Learning, Image Processing, Ingredients.**

1. INTRODUCTION

Food is the most fundamental and essential part of our everyday lives. It forms an important component of one's culture. Ingredients can be identified to understand a culture's cuisine and they can be based on factors like regional availability, cultural history, climatic conditions, etc. For instance, Mexican cuisine utilizes ingredients like tomatoes, avocado, corn, cheese, beans, chiles, cheese, etc. Indian cuisine uses spices, cumin, turmeric, herbs, vegetables, etc.

Cuisines have different choices of ingredients and even with the same ingredients, can differ in the process of preparation resulting in different tastes and nutritional values. Having said this, it would be a good thing to know the cuisine that we are having so that we can figure out what we would like to have and what we won't. This could be very useful not only to individuals but also for various restaurant networks and food delivery service companies who could incorporate this approach in websites and applications that takes photos and sends them into classification, which then shows the dish in a network in relation to other dishes and ingredients. This can help customers find what cuisines or dishes they are interested in and which places serve the same. Many restaurant owners could use this to find out how the more famous ones are using the same ingredients to create innovative dishes. This could also benefit people with dietary restrictions by helping them figure out the nutritional value of a dish and can also be used in recipe recommendation systems.

Studying these applications motivated us to come up with this system which aims at identifying the cuisine of a plate of food by using two different approaches. The first approach would require feeding in an image of the dish as an input and the second one would be feeding a list of the composing ingredients of the dish. The former approach would simply return the cuisine of the ingredients list in question along with the prediction accuracy while the latter approach would give us the predicted name of the cuisine, its list of composing ingredients, and the instructions/recipe to prepare the dish.

2. LITERATURE SURVEY

In this section, different methodologies employed by researchers have been discussed.

NutriNet, based on deep CNN architecture, employs around 520 food photographs from various food classes and achieves an accuracy of roughly 86.72 percent when tested for 130,517 images, as well as accuracy of 94.47 percent when tested for 130,517 images. The NutriNet identifies food and drink items when presented separately as identifying the two or more food items in a single image is complex. Also, it fails to give an accurate result when images are noisy/blurry.

Then, using Random Forests, they executed the food capture process by mining only the discriminative sections, allowing them to mine all classes at the same time while also sharing information. There are 101 food categories in the dataset utilized in this model, with a total of 101000 food pictures. This proposed model provides a better accuracy (50.76%). The key features of their studies are a free-to-use and huge dataset for food recognition, superpixel-based patch sampling which reduces the number of sliding windows, and food identification through Random Forests Mining.

Another system uses Fuzzy C-means Clustering for Segmentation and Morphological operations to identify and measure the volume, mass, and density of food items in an image. In the MatLab environment, the calorific value is then determined. The process incorporates Fuzzy C-implies Clustering Segmentation which permits every data. Using the Convolutional Neural Network of cuisine detection points to a place with different groups with various degrees of membership. Morphological activities, boundary extraction, expansion, disintegration, and highlight extraction are next performed. This system can recognize raw vegetables and fruits and has nothing to do with cooked food. It also indicates the calorific value of each component separately, rather than providing a total calorific value for the entire dish.

2.1 SUMMARY OF RELATED WORK

The summary of methods used in the literature is given in Table 1.

Table 1: Summary of literature survey

Sr No.	Name Of Paper & Author Name	Algorithm	Success Ratio	Conclusion
1	Cuisine Detection Using the Convolutional Neural Network (2020) Authors: Dipti Pawade, Ashwini Dalvi, Dr. Irfan Siddavatam, Myron Carvalho, Prajwal Kotian, Hima George	Convolutional Neural Network.	73%	This paper aimed at recognizing food items on a plate and calculating the nutritional value of each of them. The user had to feed preprocessed, resized, and grayscale cuisine images and the algorithm used was Convolutional Neural Network-Based.
2	Cuisine classification using recipe's ingredients (2018) Authors: S. Kalajdziski, G. Radevski, I.Ivanoska, K. Trivodaliev and B. Risteska Stojkoska	Naive Bayes, Neural Networks, Support Vector Machines (SVMs).	80.41%	Based on the results we can conclude the following: Naive Bayes models when trained on a proper dataset with an informative feature vector per sample show extremely competitive results.
3	Cuisine Prediction based on Ingredients using Tree Boosting Algorithms (2016) Authors: R. M. Rahul Venkatesh Kumar, M. Anand Kumar, and K. P. Soman	Extreme Boosting algorithm in R Programming and Random forest classifier algorithm using Sklearn python.	80.41%	We can infer that in comparison between the two algorithms Extreme Boost performed well when compared with Random forest. These two algorithms come under ensemble classifiers.
4	What Cuisine? - A Machine Learning Strategy for Multi-label Classification of Food Recipes (2015) Authors: Hendrik Hannes Holste, Maya Nyayapati, Edward Wong	Custom tf-idf scoring model, Random Forest, and Logistic Regression.	77.87%	In this, having used the Logistic Regression model with C being 2.0 and no tf-idf, along with ingredient name preprocessing and bag-of-ingredients count give us a test set error of 0.2213, which isn't much different from reported validations at error.
5	Predicting Cuisine from Ingredients (2015) Authors: Rishikesh Ghewari, Sunil Raiyani	Naive Bayes, Logistic regression, Decision Tree, Random Forest, SVM, and k-Nearest Neighbors.	81.31%	We observed that both Logistic Regression as well as SVM classifiers perform equally well in the prediction task and saw that the bag of words model on ingredients works well in this task.

3. PROPOSED WORK

Classification of cuisine based on either images or ingredients exists but not based on both. It's mostly the images or ingredients that help in finding the particular type of cuisine the food is. With the help of the specific type of ingredients, it's easier for the model to figure out which type of cuisine the particular food falls under. Otherwise, the other way is by image processing to find the cuisine the image falls under for that the acquired dataset needs to be very appropriate and the model should be trained appropriately for getting output with maximum accuracy.

In today's real-world industry there is only research conducted on this cuisine detection but it isn't applied in the current working industries. This will help all the trending food delivery applications to recommend to their customers the food and the restaurants the users are willing according to the cuisine. As nowadays in various countries the particular cuisine such as Korean dishes are getting popular so getting the appropriate suggestion as per the interest of the user will help the food delivery applications to recommend and also it would help them to get the analysis in which place which is cuisine has the most demand this will indirectly help them in the survey for the rest. The existing system is not applied in the real world industry, only research is conducted.

The accuracy which is the most important factor to be considered is not that high in the current model this will lead to a wrong observation for the recommendation system. So, the accuracy needs to be taken into consideration by training the model and also by making a well-defined dataset.

3.1 SYSTEM ARCHITECTURE

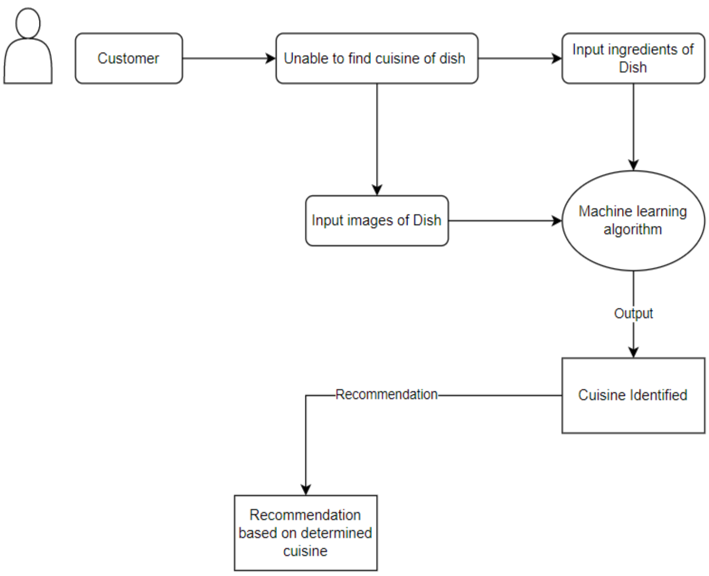

Fig. 1: Proposed system

The proposed system will help to overcome all the loopholes in the cuisine detection system and give a more appropriate accuracy. The proposed system is as follows:

The customer wants to find the cuisine of the dish there are two ways in which the user can do the following

1. Input ingredients of dish: In the following the customer can select the ingredients that are present in the particular dish.

2. Input Images of dish: The Customer has to upload the image of the dish of which the user wants to know under which cuisine the image falls.

Machine Learning Algorithm: After giving input either by selecting the ingredients of the dish or uploading the image of the dish then this algorithm helps to give the most accurate result with the help of the detailed dataset and the trained model. The proposed system needs to have a very accurate and detailed dataset for getting the output very correct which will help in the observations.

Cuisine Identification: The algorithm helps to find out the appropriate output as the cuisine the dish falls under with the help of the image or by ingredients as input.

Recommendation based on determined cuisine: The output of the following will help the food-related applications such as delivery, dietary, restaurants recommendation platforms applications, etc will be in benefit by getting to know their customers much better**.** The output of the proposed system will also help to get a better observation for the survey of a particular area as every place is known for its special cuisine style. By this, the user will be more likely to get recommendations according to their interest and taste in the specific cuisine.

4. REQUIREMENT ANALYSIS

4.1 HARDWARE AND SOFTWARE SPECIFICATIONS

Following are the different hardware and software specifications for the system:

Hardware details

Processor	2 GHz AMD Ryzen
HDD	1 TB
RAM	16 GB

Software details

Operating System	Windows 10
Programming Language	Python
Database	Image repository /Ingredient list

4.2 EVALUATION METRICS

For measuring the accuracy of the system we use certain metrics. These metrics evaluate the quality of the system and give the user a quantifiable output. These systems are typically measured using precision and recall.

Confusion matrix: It is a 2 by 2 matrix used to represent binary classification where one axis represents actual values and the other represents predicted values.

PREDICTION \ ACTUAL	Negative	Positive
Negative	TRUE NEGATIVE	FALSE NEGATIVE
Positive	FALSE POSITIVE	TRUE POSITIVE

Fig. 2: Confusion Matrix

Precision: This metric measures exactness. It calculates the fraction of relevant items out of all the items retrieved. It is the proportion of recommended tv shows that are actually good. The precision is defined in the equation below:

$$P = \frac{TP}{TP + FP}$$

Recall: This metric measures completeness. It calculates the fraction of relevant items out of all relevant items retrieved. It is the proportion of all great shows that were recommended. The recall is defined in the equation below:

$$R = \frac{TP}{TP + FN}$$

F1 Score: The harmonic mean of recall and precision is called the F1 score. It considers both false positives as well as false negatives. This is why it performs exceptionally on an imbalanced dataset. The F1 score is defined in the equation below:

$$F1\ score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

5. IMPLEMENTATION

The implementation detail is given in this section.

5.1 K-NEAREST NEIGHBORS

KNN is an algorithm used for both classification and regression. It is primarily known to be used for classification and is one of the best. Basically, already classified data is plotted using specific features. After that, a new data point is introduced to the algorithm. For finding the class of the new data point, the algorithm will compare this new point to the k closest data point. This algorithm works on the concept of neighbors, initially, we decide the number of neighbors we need. This is basically the number of classes that we need to identify. Next, the Euclidean distance is calculated among the data points. This helps us understand the nearest neighbor to a data point. Eventually, we get different classes based on the data points entered. In this type of algorithm, there is no middle ground, the data point either belongs to class A or class B.

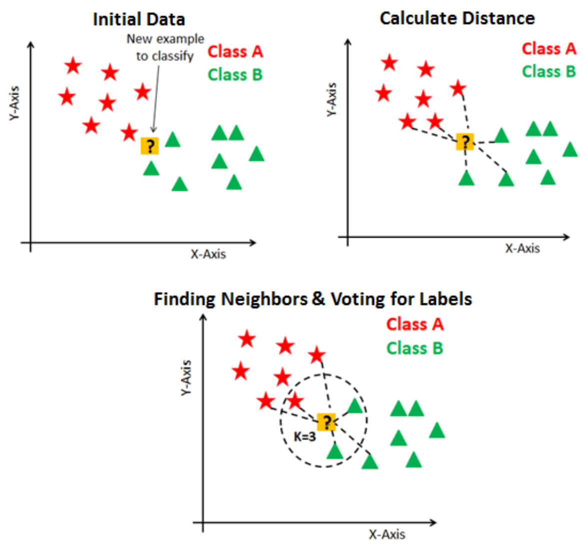

Fig. 3: KNN

5.2 IMAGE RECOGNITION USING PYTORCH

Image recognition involves the usage of neural networks to identify items within images. PyTorch is a built-in library within python that will help us in identifying images.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 04 | April 2022 www.irjet.net p-ISSN: 2395-0072

Neural networks work on the basic structure of our brain. Our brain uses neurons to send messages and complete actions. Similarly, neural networks have nodes which act as neurons, there are multiple layers of neurons depending upon requirement. There is an input layer, an output layer, and a hidden layer. The input layer replicates attributes and the output layer depicts the number of classes to be identified. The hidden layer uses weights and biases to predict the class correctly.

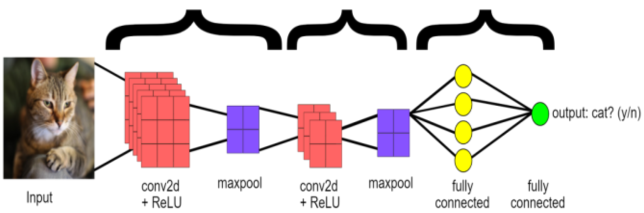

Fig. 4: Image Recognition

5.3 DATASET

The dataset used has been acquired from yummly website. Here, a script was run to collect all the data required to begin the project.

The dataset has about 40,000 rows of different recipes from different cuisines.

Dataset size:

```
yumdf.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 39774 entries, 0 to 39773
Data columns (total 3 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   id           39774 non-null  int64
 1   cuisine      39774 non-null  object
 2   ingredients  39774 non-null  object
dtypes: int64(1), object(2)
memory usage: 932.3+ KB
```

Fig. 5: Dataset

Cuisine distribution:

Italian has the highest number of dishes in our dataset making it the most accurate prediction overall.

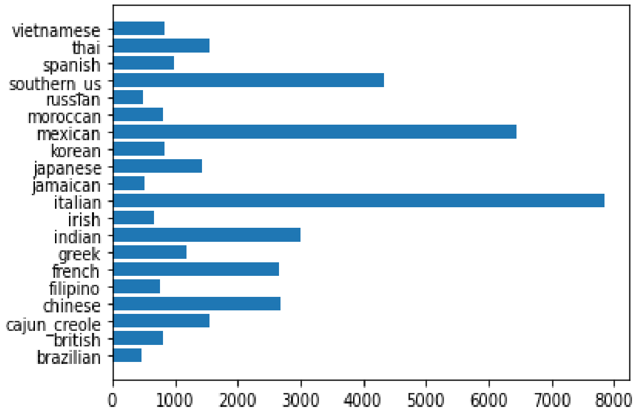

Fig. 6: Cuisine Distribution

6. RESULTS

6.1 CUISINE IDENTIFICATION USING INGREDIENTS LIST

In this approach, the user enters a list of ingredients i.e Rice, Beans, Salsa, Corn, Beef, Tortilla and Lettuce. These items form an essential part of the Mexican cuisine which is what is returned by our algorithm. The algorithm used for prediction is KNN. We calculate the f1-score to measure the accuracy of the model on the user input. As seen, we get an accuracy of 88%.

```
user_input = ["Rice", "Beans", "Salsa", "Corn",
              "Beef", "Tortilla", "Lettuce"]
check_predict = predictCuisine(user_input)
modeldf_subset = modeldf[modeldf['cuisine'] ==
                         check_predict[0]]
print(check_predict + " - f1-score: " +
      str(modeldf_subset['f1-score'].values[0]))
```

```
['mexican - f1-score: 0.88']
```

Fig. 7: Cuisine prediction using list of ingredients as input

6.2 RECOMMENDATION SYSTEM USING IMAGES

In this approach, the user inputs an image (chicken wings) and the algorithm outputs a set of recipes each of which contains the predicted name of the dish along with its list of ingredients and the instructions to prepare it. This approach uses image recognition using PyTorch.

Fig. 8: Input Image

```
RECIPE 1

Title: Chicken wings

Ingredients:
chicken, pepper, salt, butter, oil, paprika

Instructions:
-Preheat oven to 400 degrees f.
-In a large bowl, combine the butter, olive oil, paprika, salt and pepper.
-Add the chicken wings and toss to coat.
-Place wings on a baking sheet and bake for 30 minutes.
-Turn wings over and bake for another 15 minutes.
====================
RECIPE 2

Title: Oven fried chicken wings

Ingredients:
chicken, pepper, salt, butter, oil, paprika

Instructions:
-Preheat the oven to 425.
-Line a baking sheet with parchment paper.
-In a large resealable plastic bag, combine the oil, butter, paprika, salt
-Add the chicken wings and toss to coat.
-Arrange them on the baking sheet in a single layer and roast for 30 minut
====================
```

Fig. 9: Output

7. CONCLUSION

In this project, we implemented a cuisine identification system using two approaches: first one included feeding a list of ingredients as input and the second one demanding image of the dish as an input. The second approach along with returning the list of ingredients, also suggested various ways with which those ingredients could be utitlized to make a recipe. These approaches made use of extensive databases with which we trained our respective model and tested the same with 20% of our dataset as test data along with multiple user provided custom inputs.

The approach was ran through several machine learning algorithms and we found the K-Nearest Neighbours (KNN) approach to be the most suited one for the project which gave us an average accuracy of 71% which was greater than what we got from any other algorithm.

The proposed system can be implemented in various food delivery applications such as Zomato, Swiggy and Uber eats. The system can refer users to restaurants of specific cuisines, this helps the user when they want to try a new type of dish/cuisine. It can also be used for prediction of nutritional value of the food items for the users who are much conscious about the ingredients and nutritions provided in the food items. This system can help the users who are seeking dietary assessment and planning.

ACKNOWLEDGMENT

We feel thankful and would like to express our kind gratitude towards our Principal, Dr. Sandeep Joshi Sir for allowing us to conduct our project 'Food Cuisine Analysis using Image Processing and Machine Learning.'

We respect and thank the HOD of Computer Engineering, Dr. Sharvari Govilkar ma'am for providing us an opportunity and a platform to complete our project duly along with the support provided throughout.

We would like to take this opportunity to express our profound gratitude and deep regards to our project guide Prof. Krishnendu Nair for her exemplary guidance, monitoring, and constant encouragement throughout the course of this project. The blessing, help, and guidance given by her from time to time shall carry us a long way in the journey of life on which we are about to embark.

We owe our gratitude to our mini project coordinator Prof. Payel Thakur who demonstrated an avid interest in our project from the beginning and helped us to keep a check on our progress and made sure we completed each assigned task within the given deadline.

We adore and respect our parents and friends who have undoubtedly been a driving force and a constant motivation in this journey.

REFERENCES

[1] *https://www.kaggle.com/c/whats-cooking/data*: Dataset containing ingredient list for different cuisines.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] X. Li, L. Wang, and E. Sung, "*Multilabel SVM active learning for image classification,*" 2004 International Conference on Image Processing, 2004. ICIP '04., 2004, pp. 2207-2210 Vol. 4, doi: 10.1109/ICIP.2004.1421535.

[3] https://www.kdnuggets.com/2020/04/performance-evaluation-metrics-classification.html: Understanding evaluation metrics.

[4] Y. Deng and B. S. Manjunath. Unsupervised segmentation of color-texture regions in images and video. IEEE

Transactions on Pattern Analysis and Machine Intelligence (PAMI),23(8):800–810

[5] https://www.uplevel.work/blog/feature-predicting-a-cuisine-based-only-on-its-ingredients-using-machine-learning: Predicting Cuisines Based Only On Ingredients In A Dish.

[6] B. Li and M. Wang, "Cuisine classification from ingredients".

[7] H. Su, TW. Lin, CT. Li, MK. Shan, and J. Chang, "Automatic recipe cuisine classification by ingredients," In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, pp. 565–570, 2014.

[8] M. M. Zhang, "Identifying the cuisine of a plate of food," University of California San Diego, Tech. Rep, 2011.