Food Image To Recipe Converter

[1] K.Arhant , [2] G.Naveen , [3] S.Aasritha Reddy
dept. Artificial Intelligence
Amrita School of Engineering
Bengaluru, Karnataka

[1] bl.en.u4aie22123@bl.students.amrita.edu, [2]bl.en.u4aie22115@bl.students.amrita.edu,
[3]bl.en.u4aie22144@bl.students.amrita.edu.

***Abstract*—The project aims to generate food recipe by classifying the image present in the dataset and along with that shall also give calorie, degree of the food, based on the image using multiple machine learning models**

Keywords*— *food recipe, degree of food, calorie

I. INTRODUCTION

Food is indeed a very important aspect in the life of a human being. It helps us gain energy, help sustain living and also provides rich nutrients to help us do our day- to-day activities. Food is thus an essential component in a human's life. Tasty food adds more nutrients and interest to children especially in their growing ages. But due to lack of finding the proper recipe or taking the worst case not knowing the dish's name can cause mother's and many people limit their food exploration capacity, making them limit to their daily food, which would have bored them. When children and even teenagers lose interest in eating food, they tend to lose their appetite and thus effects their health and doesn't let them enjoy food anymore, thus it is a sheer necessity to eat proper food. To make proper food, modern age people seek for recipes in the internet which do not actually tell them the quantity to enjoy quality food. Health freaks who are very much concerned about their health often tend to analyze their daily intake of food based on the number of calories and the protein thus giving a health-conscious recipe. Thus, we have are introducing Recepto. Recpeto is a image classification model that gives the statistics of food and also its recipe for a healthy diet through machine learning models. We take different parameters since we are very much health conscious and believe in the ideology of "a fitter world, makes a better world", thus we use different machine learning models to classify the image of the food and give the recipe based on that.

II. LITERATURE SURVEY

Food computing is becoming a major area of research, with the ultimate goal of creating machine learning systems that can automatically generate recipes based on food images. Current methods rely on finding similar recipes in a database, which limits their success. This paper proposes approaches, which uses powerful language models to directly generate recipes from images. The classification works in three major steps : (1)**BLIP Model** : Bootstrapping Language-Image Pre-training, is a state-of-the-art model used in the FIRE system for generating titles for food images. Here's a breakdown of its key aspects: BLIP is a multimodal model, meaning it can process both visual information (images) and textual information (language).It's specifically trained on a large dataset of images and their corresponding captions. (2) **Vision transformer with a decoder** : To extract the proper ingredients that are required for the creation of the recipe we need a tool that extracts the ingredients. There are multiple approaches but the best approach is through Vision transformer that is specified to bring out the ingredients that are necessary for the food. (3) **T5 Model** : Well last but not the least, once we are ready with the ingredients it is time to cook the food, thus we use the T5 model that is used in generating the recipe of the food and make it feel tasty.**[1]**

We tend to create a web and a mobile application platform that can be useful in making the model. Thus, to implement it as a dual service we have planned to blend in machine learning models that shall suffice both the sources. To come up with that we have used the following approaches : (1)**Advanced Search**: Forget keyword limitations! Feast In's innovative search algorithm helps you refine your recipe search based on available ingredients, dietary restrictions, cooking time, and even cuisine. No more sifting through irrelevant recipes – find the perfect fit for your needs. (2)**Search by Image**: Ever seen a mouthwatering dish but don't know what it is or how to make it? Feast In's powerful image recognition technology comes to the rescue. Simply upload a picture, and Feast In will suggest similar recipes, unlocking a world of culinary possibilities.**[2]**

This paper, by Obaid et al. (2020), delves into the world of deep learning models for image classification. As the authors highlight, the "big data age" has fueled the development of increasingly complex deep learning algorithms, capable of surpassing traditional machine learning methods in feature learning and expression. This has led to significant advancements in image classification tasks within the field of computer vision. The paper provides a comprehensive overview, starting with an introduction to deep learning itself. It then delves into various deep learning models used for image classification, including: **Convolutional Neural Networks (CNNs):** These are the workhorses of image classification, adept at extracting spatial features from images. **Recurrent Neural Networks (RNNs):** While less common, RNNs can be useful for handling sequential data like image sequences. **Transformers:** Newer architectures like transformers are gaining traction, offering advantages in handling long-range dependencies within images. The authors compare and contrast these models based on their strengths and weaknesses, highlighting factors like accuracy, computational efficiency, and suitability for specific tasks. They also discuss the impact of factors like dataset size and diversity on model performance. Finally, the paper focuses on two popular benchmark datasets for image classification: CIFAR-10 and CIFAR-100. The authors compare the performance of various deep learning models on these datasets, providing insights into their relative effectiveness. Overall, this paper offers a valuable resource for anyone interested in understanding the current landscape of deep learning models for image classification. It provides a clear and concise overview of the key concepts and models, along with insights into their strengths, weaknesses, and practical considerations. **[3]**

This research, by Gulzar (2023), investigates the use of MobileNetV2 with deep transfer learning for classifying fruit images. This approach leverages pre-trained models to improve accuracy and efficiency in tasks with limited data. The study focuses on a dataset of 40 different fruits and utilizes MobileNetV2, a lightweight and efficient CNN architecture. By applying transfer learning, the pre-trained weights of MobileNetV2 are fine-tuned on the fruit image dataset, allowing the model to learn fruit-specific features. **[4]**

A new study uses machine learning to predict the degree of processing in any food, revealing that over 73% of the US food supply is ultra-processed. This has concerning health implications, with increased reliance on ultra-processed food linked to higher risks of metabolic syndrome, diabetes, and other issues. Replacing these foods with less processed alternatives could significantly improve health outcomes, highlighting the importance of better informing consumers about processing levels.**[5]**

This paper provides a conceptual and technical foundation for building a food image to recipe converter using machine learning. This paper, mentions that the system outputs the recipe names of the food, ingredients, and cooking procedures using machine learning datasets. This paper, mentions using the convolutional Neutral networks (CNN) to categorize food images into various categories and output matched recipes. this paper helps our project to how to offering insights into model architecture using CNN, dataset considerations, and output structures, enhancing the practical applicability and performance of our project**.[6]**

This paper is on automatic food recognition and nutrient estimation from food images using some datasets and computer versions which is relevant to our food image to recipe converter project. It tells the state-of-the-art methods for processing food images, including classification, segmentation, and volume estimation. The systematic review of these methods, along with insights into the strengths and limitations, can inform the design of our model. By this paper, we can enhance the accuracy and effectiveness of our food image to the recipe converter project**.[7]**

This paper provides the use of a subset of the 1M+ dataset and it involves using a mobile application with an approach to search encodings of food images generated through a DenseNet-121 CNN. This approach simplifies the model and uses CNN directly for mapping, it calculates the similarity index for input food image and image in the dataset. It uses DenseNet-121, a deep CNN for processing the food images. It also tells the use of distance metrics and the KNN algorithm to find the closest matching result. Finally, this paper informs and guides various aspects of our project covering dataset selection, model architecture, nutritional value integration, and similarity index calculation. **[8]**

This paper provides the limitations of existing food-logging tools and aims to improve precision and convenience. this paper uses leveraging advanced machine learning techniques for food recognition to potentially improve the accuracy of recipe conversion. In their paper, these optimized models are integrated into an Android app named Food Insight. finally, this paper can assist our food image to recipe converter project by providing insights into addressing challenges, leveraging advanced machine learning techniques, integrating knowledge bases, handling contextual information, and offering a practical example of Android app implementation**. [9]**

This paper provides creating standardized recipe datasets for machine learning is tough because recipes come in various formats and languages. In this study, we collected recipe datasets available publicly and made them consistent using dictionary and rule-based methods. We also used specific resources to convert measurements. This gave us two sets of data—one with ingredient embeddings and the other with recipe embeddings. When we tested a machine learning system to predict nutrients using these datasets, we found that combining embeddings using domain knowledge worked better than the usual methods**.[10]**

III. METHODOLOGY

A1. Confusion Matrix

The code A1 provides the details of the K-Nearest Neighbors Classifier on our project dataset food image to recipe converter to predict the food titles on the basis of cleaned ingredient text, and finds the accuracy, precision, recall and f1 score and forms a confusion matrix.

A2. MSE ,RMSE, MAPE calculation

The code A2 reads the dataset of purchase data from the excel file and finds the mean squared error, root mean squared error, mean absolute percentage error and R-squared score for provided pair of actual and predicted columns in the dataset. The usage of Numpy library in Python made it easy for us to compute and also make the code efficient.

A3. Scatter Plot classifying two classes

The code A3 generates a random dataset with X and Y values and assign the points to 2 different classes where the representation of classes in done through blue and red. Class 0 is indicated using "blue color" and class-1 is indicated using "bed color". Finally, it creates a scatter plot of the training data and colors the points as for their class color blue or red.

A4. Using KNN classification (for $k = 3$)

The code A4 test the dataset with values of X and Y varying between 0 and 1 with increment the values of 0.1 and classifying these points with training data using K-Nearest Neighbors Classifier with K=3. finally making a scatter plot of the test data output with test points colored where class-0 in blue color and class-1 in red color.

A5. *KNN classification for different values of k*

The code A5 test the dataset with values of X and Y varying between 0 and 1 with increment the values of 0.1 same as A4 question and classifying these points with training dataset using K-Nearest Neighbors Classifier with different K values for (k=1, k=3, k=5 , k=7). finally making a scatter plot of the test data output with test points colored where class-0 in blue color and class-1 in red color.

A6. *Considering two features and classes from dataset*

The code A6 reads our project dataset food image to recipe converter and considering any two features and classes from our dataset we are repeating question A3, A4 and A5. where it assigns the points to 2 different classes where class-0 in blue color and class-1 in red color. Finally, it creates a scatter plot of the training data where in x-axis Ingredients lengths and y-axis Instructions length are present.

A7. *Finding the ideal k value*

After the usage of the Random-Search CV() function we came to a conclusion that the ideal value of K can be 5. The reason for that is that because it provides a good balance between the nearest neighbors for generalization. It can also maintain flexibility and can give us the accurate output which we require for our project. Keeping in mind that we are not using all the classes in our dataset we have come up choosing the ideal value as 5 and this may vary as we consider more and more classes from the dataset.

IV. RESULT

The project is currently under development and we are still under exploration to blend in as much as functionality required to make it a user-friendly application. The insight on machine learning is indeed a very helpful aspect in the complete project since that shall help users determine what they shall be eating and exploring. The project aims to provide an insight of the food quality and quantity of the user. It helps the user give detailed insights of what ingredients were used for making thefood.

Below, we can see the results or output we can get by running the codes A1 to A9 in the codes we have done the testing and training and we calculate the accuracy, f1 score, recall, and precision for the question asked accordingly using our dataset which we are going to use in our project. we printed the confusion matrix and classification report as asked in some of the questions.

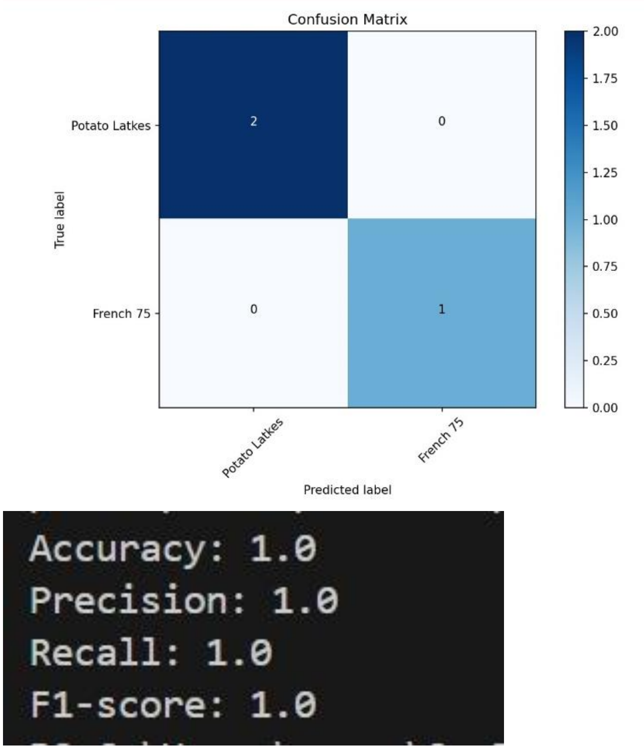
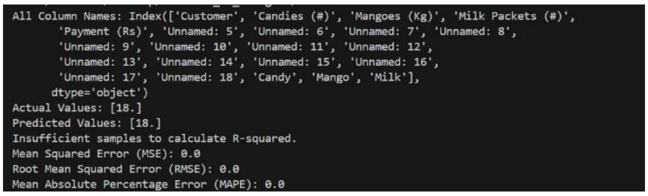


Fig A1



```
All Column Names: Index(['Customer', 'Candies (#)', 'Mangoes (Kg)', 'Milk Packets (#)',
       'Payment (Rs)', 'Unnamed: 5', 'Unnamed: 6', 'Unnamed: 7', 'Unnamed: 8',
       'Unnamed: 9', 'Unnamed: 10', 'Unnamed: 11', 'Unnamed: 12',
       'Unnamed: 13', 'Unnamed: 14', 'Unnamed: 15', 'Unnamed: 16',
       'Unnamed: 17', 'Unnamed: 18', 'Candy', 'Mango', 'Milk'],
      dtype='object')
Actual Values: [18.]
Predicted Values: [18.]
Insufficient samples to calculate R-squared.
Mean Squared Error (MSE): 0.0
Root Mean Squared Error (RMSE): 0.0
Mean Absolute Percentage Error (MAPE): 0.0
```

Fig A2

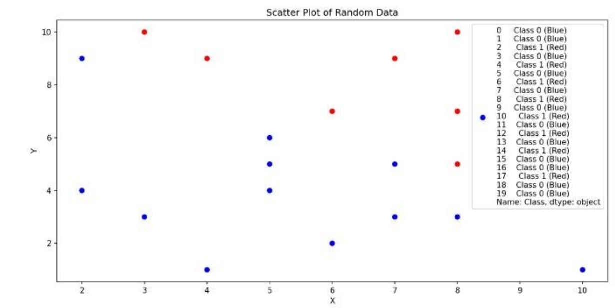


Fig A3

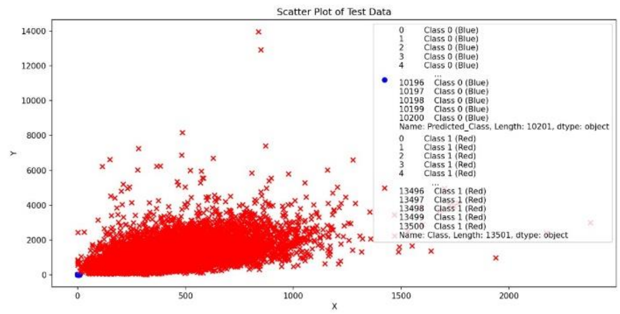


Fig A4

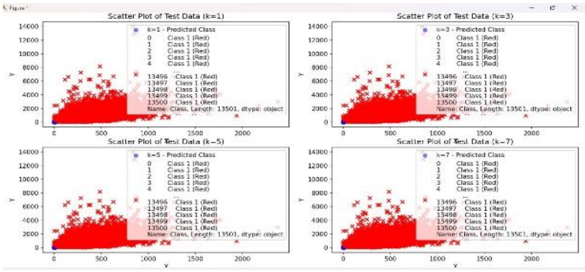


Fig A5

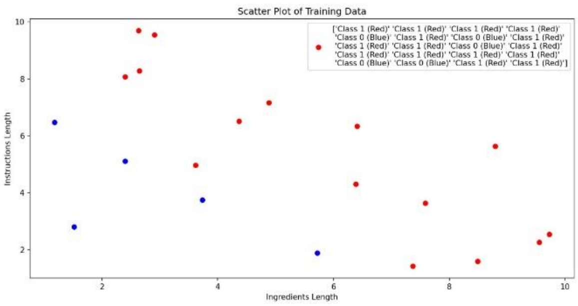


Fig A6

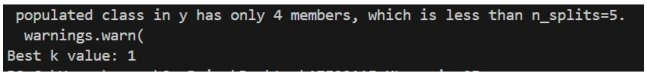


Fig A7

Observations: If we take into consideration the confusion matrix, we observe the separation of the dataset and thus we can classify whether it is well separated or not. From the confusion matrix, we can assess that the dataset is well separated. This classification is since the confusion matrix indicates clear separation and gives high values for the two classes that we had considered there.

Not only that, but we have also observed for the 5th question that as the value of K is increased there are some of changes that take place. When we have the value of K as 1, then the scatterplot seems to be a little rough and discrete, which leads to overfitting. This suggests that our model is sensitive to small variations in the data. But as we increase the value of k 3,5,7 we notice that the decision boundary becomes more and more smoother which indicates that it requires more neighbors. This helps avoid overfitting.

This leads to conclusion that KNN classifier is a reasonable classifier since the classifications are to the point and also are helping us make quicker decisions. However, the increase of K value must not be taken as the only factor into consideration while working with such model and we must also focus on the dimensionality and many other factors that help in the development of the project.

The code generally takes a KNN classifier by our dataset food image to recipe converter and assess the model's fit by finding its accuracy on both the training set and also testing set, utilizing train-test and grid-search for hyperparameter tuning.

The overfitting scenario typically occurs when the value of k is too small. Here as we observed when the value of k was 1 we noticed overfitting but as we increased the value of K the overfitting condition reduced and made it much more reliable.

References

[1] Chhikara, P., Chaurasia, D., Jiang, Y., Masur, O. and Ilievski, F., 2024. FIRE: Food Image to REcipe generation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 8184-8194).J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Ann, E.T.L., Hao, N.S., Wei, G.W. and Hee, K.C., 2021. Feast In: A Machine Learning Image Recognition Model of Recipe and Lifestyle Applications. In MATEC Web of Conferences (Vol. 335, p. 04006). EDP Sciences.

[3] Obaid, K.B., Zeebaree, S. and Ahmed, O.M., 2020. Deep learning models based on image classification: a review. International Journal of Science and Business, 4(11), pp.75-81.

[4] Gulzar, Y., 2023. Fruit image classification model based on MobileNetV2 with deep transfer learning technique. Sustainability, 15(3), p.1906.

[5] Menichetti, G., Ravandi, B., Mozaffarian, D. and Barabási, A.L., 2023. Machine learning prediction of the degree of food processing. Nature Communications, 14(1), p.2312.

[6] A. Channam, B. R. Swarup and S. G. Rao, "Extraction of Recipes from Food Images by Using CNN Algorithm," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1308-1315, doi: 10.1109/I-SMAC52330.2021.9640741.

[7] F. S. Konstantakopoulos, E. I. Georga and D. I. Fotiadis, "A Review of Image-Based Food Recognition and Volume Estimation Artificial Intelligence Systems," in IEEE Reviews in Biomedical Engineering, vol. 17, pp. 136-152, 2024, doi: 10.1109/RBME.2023.3283149

[8] Mane, A., Bawankar, M., Goradia, M., Chaudhari, S. (2023). Image to Recipe and Nutritional Value Generator Using Deep Learning. In: Ranganathan, G., Fernando, X., Rocha, Á. (eds) Inventive Communication and Computational Technologies. Lecture Notes in Networks and Systems, vol 383. Springer, Singapore.

[9] Z. Hakguder, Y. He, W. Chai and J. Cui, "Smart Diet Management through Food Image and Cooking Recipe Analysis," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 2603-2610, doi: 10.1109/BIBM55620.2022.9994888.

[10] G. Ispirova, T. Eftimov and B. K. Seljak, "Predefined domain specific embeddings of food concepts and recipes: A case study on heterogeneous recipe datasets," 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 2022, pp. 4074-4083, doi: 10.1109/BigData55660.2022.10020698.