IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.133100

Ingredient Detection and Recipe Recommendation Using Deep Learning

Hency Jostan Dsouza[1], K Sthuthi Nayak[2], Krishii Kirti Karkera[3], Melan Varghese[4], Mr. Shreejith K B[5]

Student, Dept. of Computer Science and Engineering, Mangalore Institute of Technology & Engineering, Moodbidri, Karnataka, India[1-4]

Senior Assitant Professor, Dept. of Computer Science and Engineering, Mangalore Institute of Technology & Engineering, Moodbidri, Karnataka, India[5]

Abstract: In response to the hectic pace of modern life, there's a growing need for a smartphone web app that streamlines meal preparation. Our project aims to address this need by developing a sophisticated recipe recommendation system powered by technologies such as computer vision and machine learning. The primary objective is to simplify the culinary experience for users who often find themselves uncertain about what to cook with the ingredients they have on hand. By leveraging computer vision techniques, our system can accurately identify the ingredients available to the user. This information is then processed using machine learning algorithms to generate tailored recipe suggestions. This approach eliminates the need for extensive meal planning or manual recipe searches, saving users valuable time and effort. To tackle this, we prepared an ingredient dataset containing image 12,558 images across 15 food ingredient classes. The YOLOv8 object detection model was used to detect and classify food ingredients. Additionally, the recommendation system was built using machine learning. In the end, we achieved an accuracy of 96%, which is quite impressive.

Keywords: Object Detection, YOLOv8, FastAPI, TF-IDF, Word2Vec.

I. INTRODUCTION

In a world where convenience frequently takes precedence over creativity in culinary trends, our project seeks to revolutionize home cooking. Our commitment lies in creating an innovative mobile application that tackles a basic kitchen problem: what to make using the ingredients at hand. Our app signifies a change in the way people approach cooking; it goes beyond conventional recipe apps by utilizing object detection technology to evaluate photos of ingredients and provide recipe recommendations based on the detected ingredients.

At its core lies the YOLOv8 model, used in real-time object detection, which serves as the foundation for identifying ingredients within culinary images. Complementing this, the utilizes word embeddings, a technique rooted in natural language processing, to discern semantic similarities among ingredients and curate tailored recipe recommendations.

Powered by a FastAPI backend, the system presents a streamlined interface for seamless interaction, facilitating efficient data transmission and processing. Through meticulously crafted endpoints, it offers a gateway to the underlying functionalities, enabling users to effortlessly harness the system's capabilities.

II. LITERATEURE SURVEY

[1] The paper aims to develop a tool that would help in pairing various ingredients from different cuisines and suggest an alternate ingredient. The goal is to support the creation of novel recipes and, by suggesting substitute ingredients, assist those who are allergic to particular ingredients. The study finds popular ingredient pairs from various cuisines and suggests substitutes using two machine learning models: the vector space model and the Word2Vec model.

[2] This paper aims to develop a system that can recognize food ingredients from images and recommend recipes based on the detected ingredients using deep learning and machine learning techniques. To use a convolutional neural network (CNN) model based on ResNet50 architecture to classify 32 different food ingredients from a custom dataset. Design a recipe recommendation algorithm using a 2D matrix to store the relationship between 19 recipes and 32 ingredients. The paper was able to achieve an accuracy of 94%.

DOI: 10.17148/IJARCCE.2024.133100

[3] The paper proposes a mobile application that allows users to search for Indian cuisine recipes based on the ingredients available at their end, and filter them by course type, and diet type. To use content-based filtering with Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity to measure the relevance of recipes based on the user's input ingredients.

[4] The paper presents a web application called RecipeIS, which can recognize food ingredients from images and recommend recipes based on the recognized ingredients. The paper uses a convolutional neural network model, ResNet-50, to perform image recognition and trains it with a dataset of 36 classes of fruits and vegetables. The paper uses the Edamam API to search for recipes that contain the recognized ingredients and display them to the user, along with nutritional information and cooking instructions.

[5] The paper presented an objective to develop a fruit recognition system based on a modern deep learning technique called EfficientNet, which can classify 25 different types of fruits from images with high accuracy. The authors used the Fruits 360 dataset, which contains 25 categories for their experiment. They trained and tested the EfficientNet-B0 model on this dataset. The paper was able to achieve an accuracy of 95%

III. SCOPE AND METHODOLOGY

Scope

The project's scope involves integrating advanced techniques such as YOLOv8 for object detection, word embeddings using Word2Vec, TF-IDF for semantic analysis, and cosine similarity for recommendation generation. Through the implementation of these methodologies, the system aims to accurately identify ingredients in culinary images and analyze their semantic relationships to provide recipe suggestions. This entails the development of a robust FastAPI backend to process user requests, perform object detection, and compute semantic similarities using Word2Vec, TF-IDF, and cosine similarity metrics. The HTML interface will offer seamless interaction, allowing users to upload images, receive object detection results, and view recommended recipes based on semantic similarities between ingredients.

Methodology

The methodology for this project commences with the meticulous collection of diverse datasets crucial for training and evaluation. Two primary datasets are acquired: one comprising ingredient images and another containing textual recipe data. The ingredient image dataset is sourced from Roboflow, a popular platform for preparing datasets and deploying ML models. An ingredient dataset containing image 12,558 images across 15 food ingredient classes with augmentation like horizontal flipping was prepared. Simultaneously, the recipe dataset is curated from kaggle.com. A dataset that contains 6,303 recipes which include their Sr. No, recipe name, ingredients, and instructions was acquired.

Pre-processing steps ensure that the data is formatted correctly for subsequent processing stages. Object detection plays a crucial role, utilizing the YOLOv8 model to accurately identify ingredients within images uploaded or captured by users. Once the ingredients are detected, they are processed through a recommendation system. This system utilizes advanced techniques such as Word2Vec embeddings and TF-IDF weighting to analyze the ingredients and generate relevant recipe recommendations. By comparing the ingredient profiles of detected items with those in the recipe database, the system suggests top-N recipes that closely match the user's input. The FastAPI backend orchestrates these processes seamlessly, receiving image data from the front end, performing object detection using the YOLOv8 model, and then passing the detected ingredients to the recommendation system. The front end provides an intuitive interface for users to upload images or capture them via their device's camera, facilitating a seamless user experience.

IV. SYSTEM ARCHITECTURE

System design is a crucial phase following the Software Development Life Cycle (SDLC). It starts with requirement analysis and progresses to designing the overall system structure. The main goal of system design is to outline the architecture, modules, their relationships, purposes, and how they integrate. This phase provides a comprehensive overview of the system flow and architecture.

Architectural Design

The system architecture of this project follows a modular and scalable design to accommodate various components seamlessly. At the core lies the backend, implemented using FastAPI, which handles incoming requests from the frontend interface which is made using HTML and orchestrates the interaction between different modules. The backend is responsible for ingredient detection using YOLOv8, semantic analysis employing Word2Vec and TF-IDF with cosine similarity, and recipe recommendation generation.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.133100

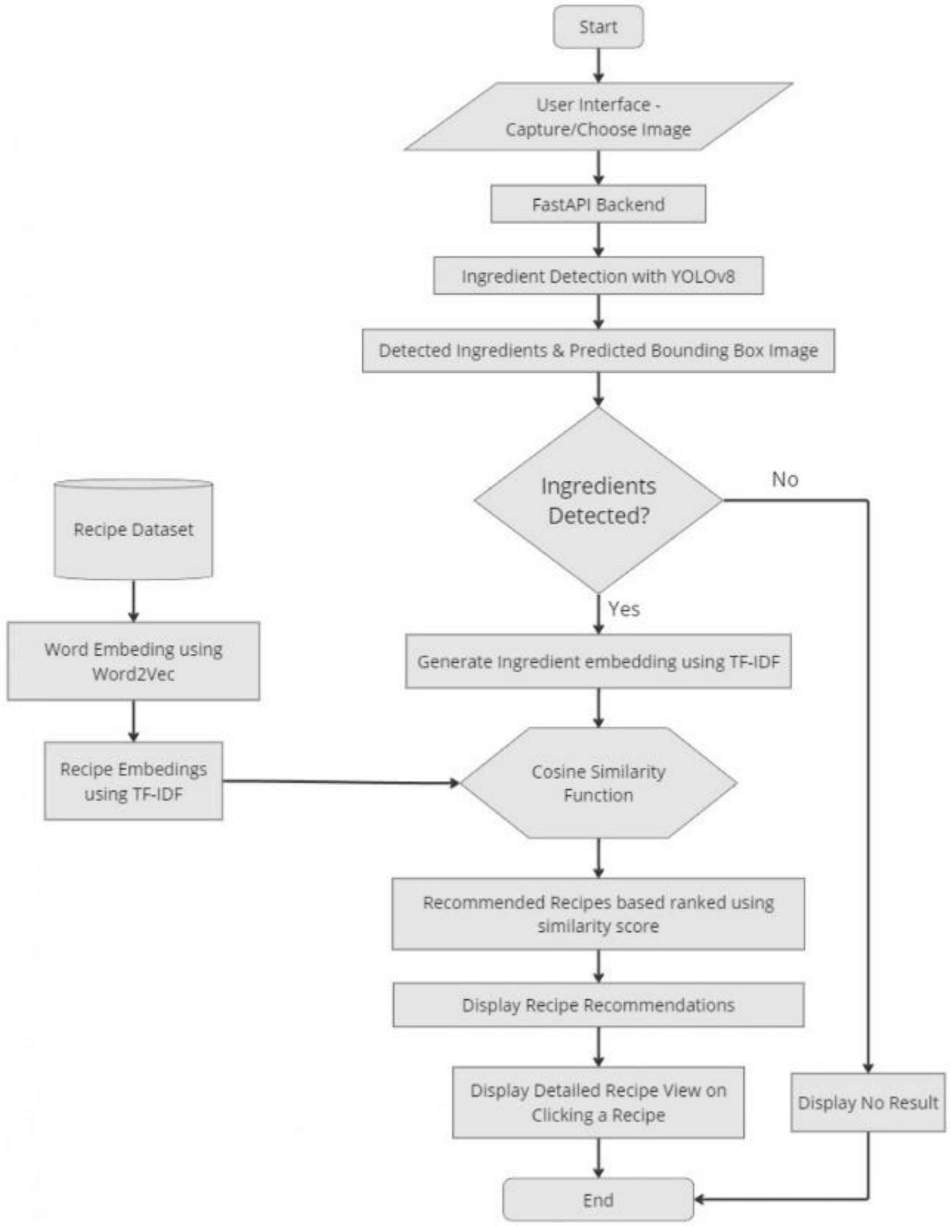

Fig 1 Architecture diagram of Proposed work

YOLOv8 is a convolutional neural network (CNN) architecture for object detection, designed for speed and accuracy. The default input image size is 640x640 pixels, but it supports various resolutions based on specific requirements and hardware constraints. A modified version of the CSPDarknet53 architecture forms the backbone of YOLOv8.

This architecture consists of 53 convolutional layers and employs a Cross-Stage Partial Network (CSPNet) design, which splits the feature maps into two parts at each stage, with one part going through a dense block and the other through a shortcut connection. The backbone network is responsible for extracting features from the input image through a series of convolutional layers and down-sampling operations.

The extracted features are then passed to the neck of YOLOv8, which is a Feature Pyramid Network (FPN). The FPN combines features from different levels of the backbone network to create a feature pyramid, enabling the detection of objects at various scales.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.133100

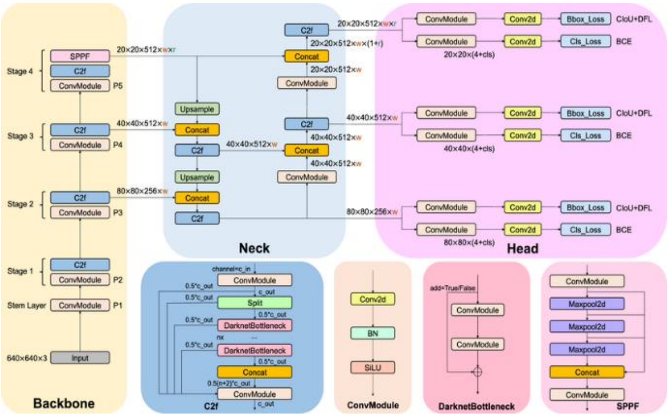

Fig 2 Architecture Diagram of YOLOv8

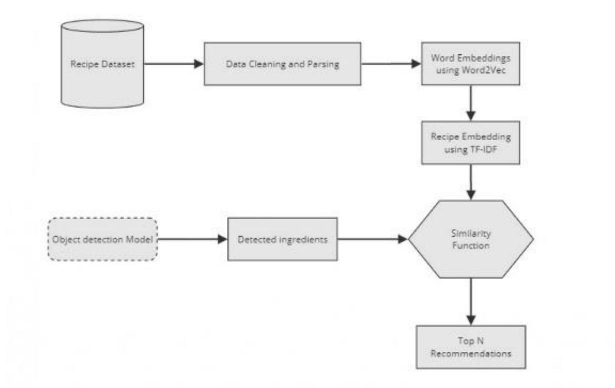

Fig 3 Working of Recommendation System

The recommendation system utilizes Word2Vec which is a 2-layer neural network that produces word embeddings. It is a technique in natural language processing (NLP) that learns word embeddings, which are vector representations of words. These word embeddings capture the semantic relationships between words, meaning that words with similar meanings will have similar vector representations. In this system, the Word2Vec model is trained on the recipe corpus. This allows the system to understand the relationships between input ingredients and recipes. TF-IDF (Term Frequency-Inverse Document Frequency) is a statistical measure used in information retrieval and text mining to evaluate the importance of a word in a document or corpus. It assigns higher weights to words that are more specific to a document and lower weights to words that are common across many documents. In the recommendation system, TF-IDF is used to weight the word embeddings, giving more importance to rare ingredients and less importance to common ingredients. Cosine similarity is then calculated between the input embedding and all the document embeddings (recipe embeddings). Cosine similarity measures the cosine of the angle between two vectors, providing a measure of their similarity, with higher values indicating greater similarity between the input ingredients and the recipe ingredients. The ranking is of the recipes is based on their similarity score with the detected ingredients.

V. RESULT AND CONCLUSION

Result

In this deep learning project, we trained two different models, namely YOLOv8 and Word2Vec for the task of Ingredient Detection and recipe recommendation for images. The training process involved using a dataset of Ingredient images and Recipes and implementing the respective architectures of each model. The YOLOv8 model was trained on a dataset consisting of 15 ingredient classes with about 12,558 images, consisting of 36,159 total annotations. The model was trained for 50 epochs.

The training results showed promising performance for all two models. The accuracy of the model was evaluated. The results are shown below.

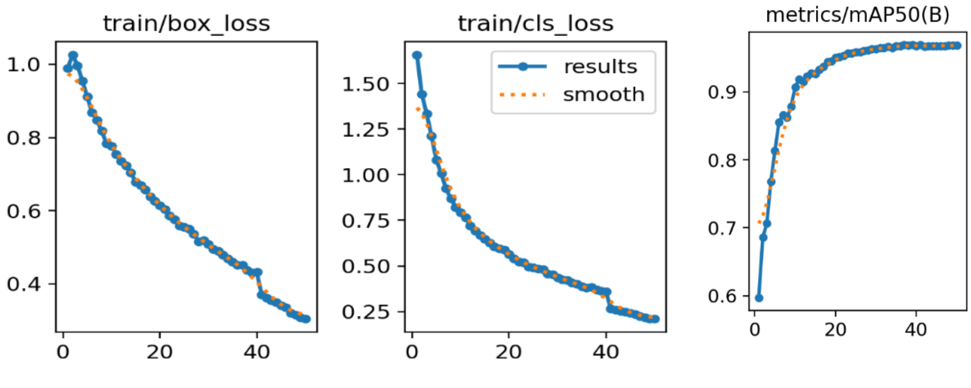

Fig 4 YOLOv8 Model Loss and Accuracy

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.133100

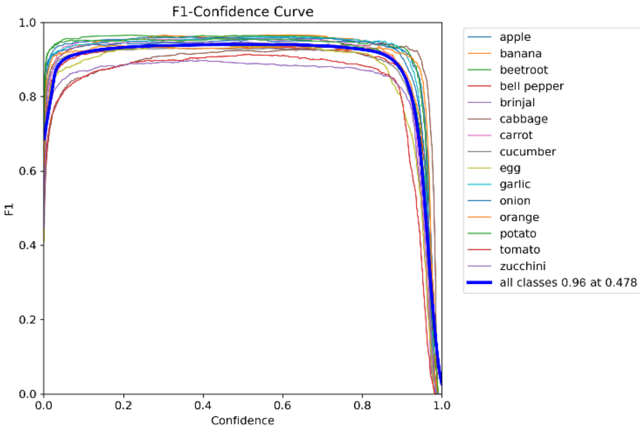

Fig 5 YOLOv8 Model F1-Confidence Curve

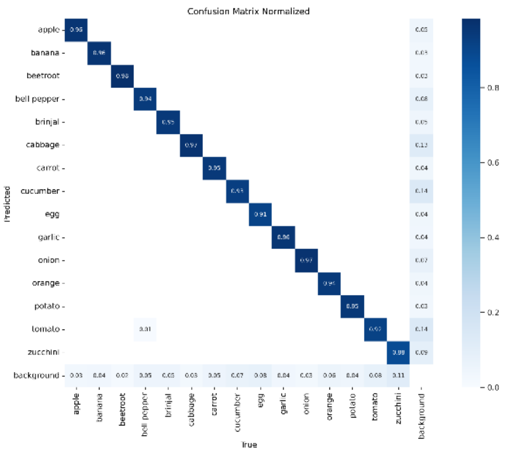

Fig 6 Confusion Matrix for the YOLOv8 Model

TABLE I RECOMMENDATION SYSTEM RESULTS FOR POTATO AND ONION

Sl. No	Recommendations
1	Aloo Matar Masala Khichdi Recipe
2	Chotti Aloo Dum Recipe
3	Nepalese Vegetable Pulao Recipe
4	Uttarakhand Aloo Ke Gutke Recipe
5	Konkani Style Batata Song Recipe
6	Spiced Potato Wedges Recipe

The accuracy of the deployed YOLOv8 object detection model on the Web App is validated through extensive testing and evaluation, demonstrating high accuracy in identifying ingredients from images. The model achieves an impressive accuracy of 96% on the test set. Additionally, the model performed well on a variety of other evaluation criteria, including F1 score, precision, and recall. Semantic analysis utilizing Word2Vec and TF-IDF with cosine similarity effectively computes ingredient similarities between the uploaded images and the recipes in the dataset. While the Word2Vec model can provide recommendations on the detected ingredients, the accuracy of the recommendation system can be evaluated by measuring how well the recommended recipes match the ingredients and how relevant they are to the user's preferences.

Fig 7 YOLOv8 Predictions on the Test Dataset

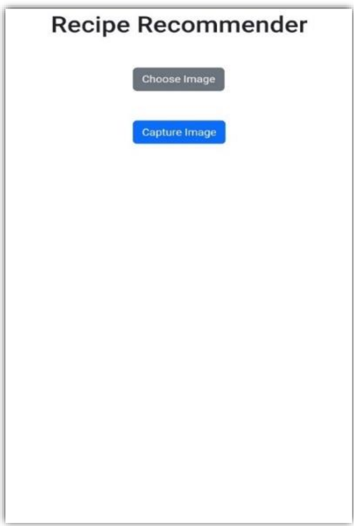

Fig 8 Application Interface

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.133100

Fig 9 Result After Detection of ingredients

Fig 10 Detailed Recipe View

VI. CONCLUSION

In conclusion, the implemented system successfully addresses the need for personalized recipe recommendations based on culinary images. Through the integration of technologies such as YOLOv8 for object detection and Word2Vec with TF-IDF and cosine similarity for semantic analysis, the system achieves accurate and relevant recipe suggestions. The comprehensive dataset collection ensures the quality and diversity of recipe recommendations, enhancing the user experience. Thorough testing and evaluation validate the system's effectiveness, the object detection model achieved an accuracy score of 96% along with high-quality recommendations consistently delivered. Moving forward, continuous refinement and optimization can further improve the system's performance and expand its capabilities. Overall, this project showcases the potential of advanced technologies in revolutionizing culinary exploration and enhancing user satisfaction in discovering new and exciting recipes.

REFERENCES

[1]. Suyash Maheshwari and Manas Chourey, "Recipe Recommendation System using Machine Learning Models", This work was supported in part by Sinhgad College Of Engineering, Pune, India, and in part by the International Research Journal of Engineering and Technology, 2019.
[2]. Dang Thi Phuong Chung and Dinh Van Tai, “A fruits recognition system based on a modern deep learning technique” 2019.
[3]. Md. Shafaat Jamil Rokon, Md Kishor Morol, Ishra Binte Hasan, A. M. Saif, Rafid Hussain Khan, “Food Recipe Recommendation Based on Ingredients Detection Using Deep Learning”, This work was supported in part by American International University-Bangladesh (AIUB), 2022.
[4]. Shubham Chhipa, Vishal Berwal, Tushar Hirapure, and Soumi Banerjee, “Recipe Recommendation System Using TF-IDF”, this work was supported in part by Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, 2022.
[5]. Rodrigues MS, Fidalgo F, Oliveira Â, "RecipeIS—Recipe Recommendation System Based on Recognition of Food Ingredients", Applied Sciences, 2023.