Food Recognition System And Nutrition Tracking Using Convolutional Neural Network Model

Mrs.P.Indhumathi
Department of Artificial Intelligence and Data Science
K.S.Rangasamy College of Technology
TamilNadu,India.
indhumathip@ksrct.ac.in

Jayakumar T
Department of Artificial Intelligence and Data Science
K.S.Rangasamy College of Technology
TamilNadu,India.
jayakumar20810@gmail.com

Ragul K
Department of Artificial Intelligence and Data Science
K.S.Rangasamy College of Technology
TamilNadu,India.
ragulkt143@gmail.com

Saravana Bharathi S
Department of Artificial Intelligence and Data Science
K.S.Rangasamy College of Technology
TamilNadu,India.
saravanasbr15@gmail.com

2024 Asia Pacific Conference on Innovation in Technology (APCIT) | 979-8-3503-6153-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/APCIT62007.2024.10673543

***Abstract*— A cutting-edge technical advancement that blends artificial intelligence and computer vision is the Food Recognition System (FRS). Food Recognition System is an effective tool for automating the identification and classification of different food products using image analysis in a world with a great deal of culinary diversity. We suggest using the very accurate deep learning model VGG30 for food recognition. Convolutional neural networks (CNNs) like VGG30 are made especially for classifying images. It can accurately identify a variety of food items because it was trained on a dataset of food photos. In order to use VGG30 for food recognition, a collection of food photos must first be produced, with each image labelled with the type of food it depicts. The preprocessed dataset can then be used to train the VGG30 model. Because of its deep architecture, it can capture the fine details of many food products and perform well even in difficult situations with changing backgrounds, lighting, and angles. Second, because of its scalability, VGG30 can manage big datasets with ease, which makes it appropriate for applications where there may be a significant volume and variety of food photos. Furthermore, the model gains from transfer learning, which minimizes the requirement for labelled food data and expedites development. Its interpretability makes it easier to comprehend how it identifies food items, which improves openness and confidence. Furthermore, VGG30 is a dependable option for precise food recognition across a variety of domains because to its broad community support and cutting-edge performance. Lastly, because of its versatility, it can be adjusted to particular jobs or preferences, which makes it a perfect fit for the Food Recognition System.**

Keywords*— *Continual Learning, Long-Tailed Distribution, Food Recognition

I. Introduction

A. Continual Learing

In the domains of artificial intelligence and machine learning, the concept of continuous learning, also referred to as lifelong learning or incremental learning, presents a paradigm shift. Its main goal is to tackle the challenge of adapting models to changing data distributions over time. Unlike conventional machine learning methods that assume a fixed dataset, continual learning aims to empower models to continuously learn from new information while retaining knowledge gained from past experiences. This dynamic framework is especially vital in practical applications where data is non-stationary, as it enables models to stay pertinent and efficient in the midst of evolving circumstances.

B. Long-Tailed Distribution

A long-tailed distribution is a statistical concept that shows a notable accumulation of events in the tail section of the distribution. This is where unusual events or extreme values are more common compared to a standard distribution. Unlike a normal distribution, where data points are concentrated around the mean and taper off slowly, a long-tailed distribution displays an extended and frequently dense tail, signifying a greater occurrence of rare events. This distribution pattern is widespread in different practical situations, such as income distribution, web traffic, and the prevalence of uncommon diseases.

C. Food Recognition

Within the field of computer vision and artificial intelligence, food recognition is a fascinating and quickly developing field. It entails creating models and algorithms that can recognize and classify different foods that are portrayed in pictures or movies. Because social media and smartphones are so widely used, there is an increasing need for food-related material to be shared, which makes automatic food recognition both a technological difficulty and a useful requirement. This cutting-edge field aims to mimic human visual perception and allow computers to discriminate between different dishes, ingredients, and culinary traditions through the use of state-of-the-art machine learning techniques like deep neural networks.

II. LITERATURE REVIEW

Research by Jianping He [1] et al. shown the exceptional efficacy of deep learning methods in a range of picture-based nutrition assessment applications. Food portion proportions and food categorization are two examples of these uses. But current approaches focus on single activities, which causes problems when numerous tasks need to be completed in parallel in real-world circumstances. The authors used a multi-task learning strategy to address this issue, training both the classification and regression tasks at the same time through the use of soft parameter sharing based on L2-norm. To improve food portion size estimation accuracy, the authors also suggested combining cross-domain feature adaption with normalization. The authors' results outperform the current approaches for portion assessment in terms of mean absolute error and classification accuracy, showing great promise for the advancement of picture-based nutritional evaluation.

Xinyue [2] Food item categorization, according to Dish and colleagues, is critical to picture-based dietary assessment since it allows for the measurement of nutrient intake from photographed food. All of the present food classification research, however, is more concerned with defining "food types" than it is with offering precise nutritional content information. This restriction results from differences in nutrition databases, which are in charge of connecting every "food item" to the pertinent data. Thus, categorizing food products using the nutrition database is the study's goal. In order to do this, the authors present the VFN-nutrient dataset, which contains information on the nutritional makeup of every food image in VFN. Since food items are classified more thoroughly than food categories, the dataset has a hierarchical structure as a result.

A approach that uses the teacher network's softmax output as additional advice for training the student network during the early phases of Knowledge Distillation (KD) was introduced by Guo-Hua Wang and colleagues [3]. On the other hand, it has been noted that the result of a high-capacity network does not always match the ground truth labels. Moreover, the performance of the student model is negatively impacted by the classifier layer since the softmax output is less informative than the representation in the penultimate layer. KD has problems when it comes to distilling instructor models from unsupervised or self-directed learning. Lately, there has been a greater emphasis on feature distillation, with less attention paid to directly replicating the teacher's traits in the penultimate layer.

The problem with class imbalanced data in this paradigm, according to Seulki Park [4] et al., is that the lack of data from minority classes reduces the prediction accuracy of the classifier. In this study, we offer a novel solution to this issue by putting forth a minority oversampling methodology that enhances various minority samples by using backdrop photos that provide abundant context supplied by majority classes. Using rich-context photographs from a majority class as background images overlaid with images from a minority class is our primary method for diversifying the minority samples. This approach is simple to use and may be combined with other well-known recognition methods without any problems. Our proposed oversampling technique is empirically validated by ablation studies and large-scale trials.

The Jiangpeng He [5] and his associates have put up a paradigm in which the first step towards image-based dietary assessment is food grouping. Predicting the kinds of food in each input image is the aim. In real-world situations, on the other hand, food kinds are usually investigated over long periods of time to determine which kinds are ingested more frequently than others. This leads to a serious issue of class disparity, which undermines the framework's overall efficacy. Furthermore, since none of the current long-tailed classification algorithms are especially designed to handle food data, there are other issues arising from the similarities between different food classes and the variability within each class. This research study presents Food101-LT and VFN-LT, two new benchmark datasets for long-tailed food classification, in an effort to address these issues. The VFN-LT dataset is very important since it shows that a real long-tailed food distribution system exists. As a result, a unique two-phase strategy is put out to deal with the problem of class inequality.

III. EXISTING SYSTEM

The utilization of meal photos has greatly enhanced the capabilities of deep learning-based food recognition in distinguishing various food types. But there are two main obstacles that prevent its real-world implementation. First, a trained model needs to be able to learn new foods while retaining its memory of previously recognized ones in order to prevent catastrophic forgetfulness. Secondly, the uneven distribution of food images in real-world scenarios, with certain popular food types being more prevalent, poses a challenge. It is imperative to learn from imbalanced data to improve the system's ability to recognize less common food classes. Our project aims to tackle these challenges through long-tailed continuous learning.

IV. PROPOSED SYSTEM

Food Recognition Systems (FRS) represent a cutting-edge solution to the complex and widespread issues associated with food identification, marking a significant advancement in the realm of technological innovation. In a time where visual content reigns supreme in our digital world, FRS harnesses the power of computer vision and artificial intelligence to interpret and classify a vast array of culinary options. Convolutional neural network (CNN) used in the proposed food recognition system is trained on a dataset of food images using VGG30. CNN is able to recognize foods in fresh images by applying its understanding of the distinctive qualities of many dishes. The initial stage of the system's process involves preparing the input image, which may involve resizing it, standardizing its pixel values, and converting it to a format compatible with VGG30. The preprocessed image is then fed into the VGG30 model. The sort of food depicted in the image is predicted by the model by using features that it has extracted from the image. The output of the model is a forecast regarding the kind of food that is seen in the picture. This prediction can be used to recognize the food or to perform other tasks, such as recommending recipes or providing additional information about the food.

A. Load Input Data

During the initial phase of developing a food recognition system with VGG30, the input data is collected. This dataset consists of various images showing different food items, which will be used for training and testing the neural network. The foundation for the categorization procedure is laid by the fact that every image in the collection is connected to a certain food group. The preparation of the input data involves organizing and preparing the dataset to be fed into the neural network for the upcoming training stage.

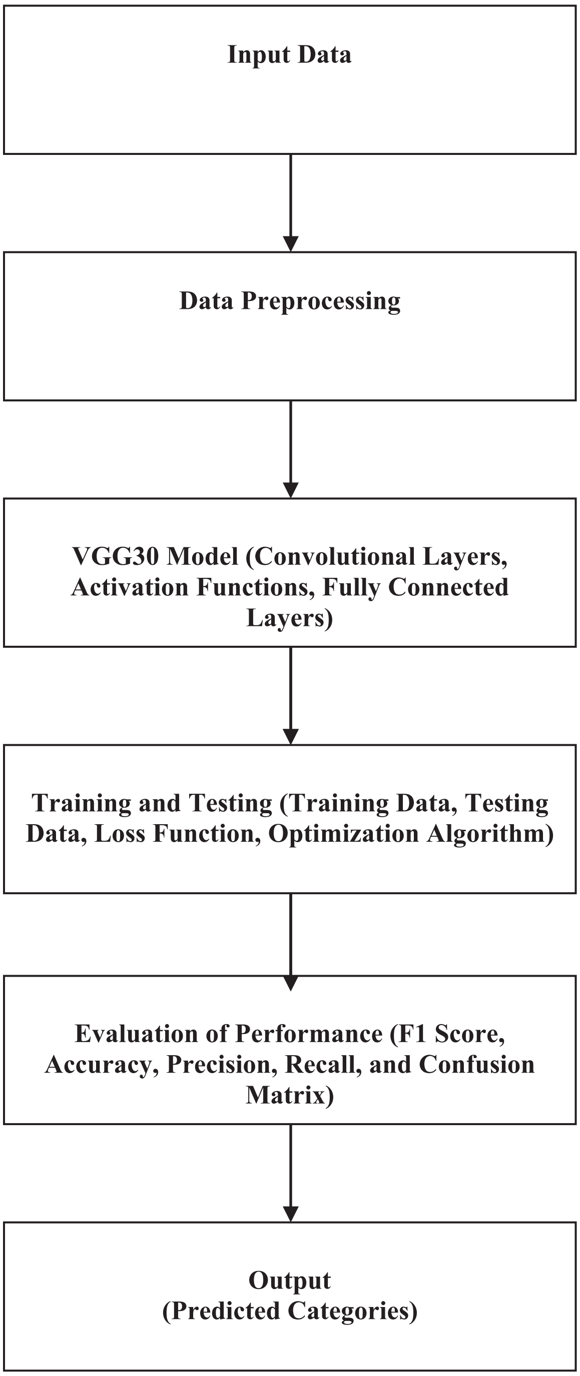

Fig. 1. flow diagram

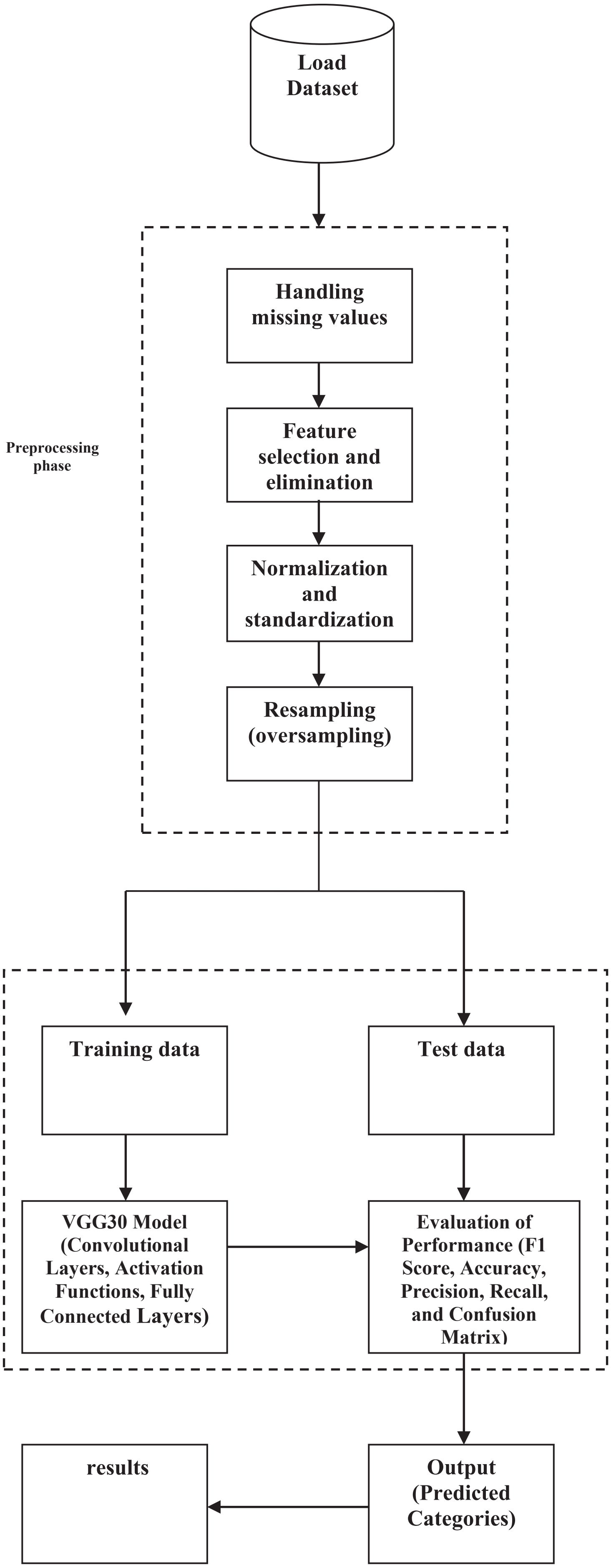

Fig. 2. Block diagram

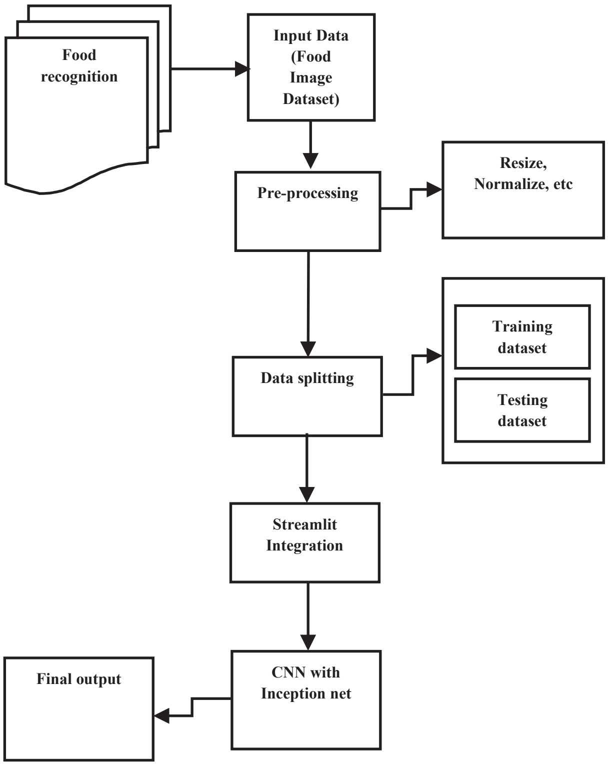

Fig. 3. Proposed flow diagram

B. Data Preprocessing

Upon loading, the data is preprocessed through various stages to ensure its suitability for training. These steps include uniformly scaling the images, normalizing the pixel values to a standard range, and potentially augmenting the data by rotating or flipping it. Data preparation considerably boosts the model's capacity to generalize patterns from the training set to new, unforeseen circumstances.

C. Feature Selection

The utilization of convolutional neural networks (CNNs), such as VGG30, and deep learning in this course involves the application of "feature selection" to identify important components within images. However, it is worth noting that deep learning models like VGG possess the ability to autonomously learn significant features during the training process. This eliminates the necessity for explicit feature selection, as seen in traditional machine learning approaches. The convolutional layers of VGG are specifically engineered to automatically recognize hierarchical features within images.

D. Traning And Testing

Once the data has undergone preprocessing, the VGG30 model undergoes training using a specific portion of the data referred to as the training set. The model learns to identify traits and patterns in the photos that correlate to various food groups during this training phase. The testing set is a distinct subset of the dataset that is used to assess how well the model generalizes to new and unobserved occurrences. It is significant to remember that this testing set is not available to the model throughout the training process. The training process involves adjusting the model's weights based on the calculated mistake or loss at each iteration.

E. Performance Evaluation

After going through training and testing, the efficacy of the VGG30 food recognition system is assessed using a variety of metrics, including accuracy, precision, recall, and F1 score. These metrics provide important information about how well the system classifies food items. A high accuracy level indicates the system's proficiency in categorizing most cases, while precision and recall metrics shed light on its capability to identify all positive instances and make precise positive predictions. In order to assess the feasibility of the food identification system and identify potential areas for enhancement, performance evaluation plays a critical role.

V. ALGORITHM DETAILS

The abstract introduces the Food Recognition System (FRS), a state-of-the-art technological advancement that merges computer vision and artificial intelligence to revolutionize the way individuals engage with and comprehend food. In a global landscape rich in culinary variety, FRS emerges as a robust mechanism for streamlining the identification and classification of numerous food items through image analysis. The proposal advocates for the utilization of VGG30, an exceptionally precise deep learning model, for food recognition assignments. VGG30 stands as a convolutional neural network (CNN) meticulously crafted for image categorization. To ensure diversity among a wide array of food products, an assortment of food images is initially compiled, with each image categorized based on its respective food group. Subsequently, the dataset undergoes pre-processing by resizing all images to a standardized dimension and normalizing pixel values to facilitate model training. Subsets of the dataset are then created for training and testing in order to assess the model's efficacy. The VGG30 model is trained by feeding training images into it and utilizing backpropagation to adjust its parameters using the training dataset. Until convergence or a predefined number of epochs is reached, this iterative process is carried out. Lastly, the model's performance is assessed using the testing dataset, which also yields measures like F1-score, precision, and recall that provide insight into the model's ability to classify food images.

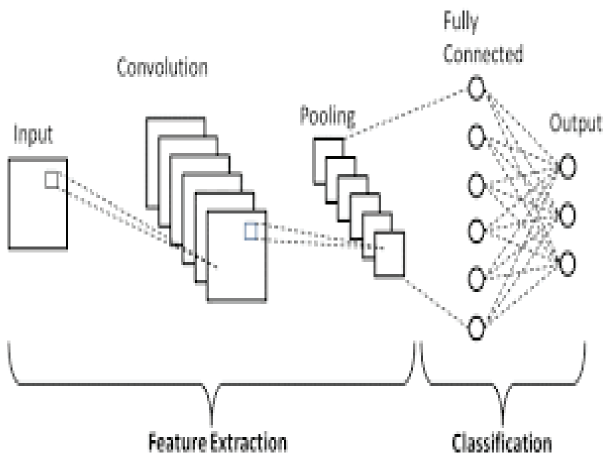

Fig. 4. Architecture Convolutional

VI. RESULT ANALYSIS

Upon training the Food Recognition System (FRS) using the VGG30 model on a dataset of food images and evaluating its performance on a distinct test set, the results reveal its high accuracy in identifying various food items. The assessment criteria, which include accuracy, recall, and F1-score, show how well the FRS can accurately classify food items using picture analysis. Excellent performance metrics in terms of precision, recall, and F1-score are shown by the Convolutional Neural Network (CNN) with the Inception architecture for both classes. The model successfully classifies each class with precision, scoring 0.87 for class 0 and 0.9 for class 1. Furthermore, the model's capacity to capture the majority of cases within each class is indicated by recall scores of 0.87 and 0.9. Notably, the F1-scores for classes 0 and 1 are notably high at 0.89 and 0.88, respectively, striking a balance between recall and precision. The model's overall accuracy of 0.88 signifies its capability to accurately classify examples in both classes. This well-rounded performance is further validated by the macro and weighted average metrics, which also yield favorable outcomes in recall, precision, and F1-score. Consequently, our findings underscore the CNN with Inception architecture's effectiveness in distinguishing between the two classes, hinting at its potential application in real-world scenarios necessitating food recognition. One commonly used metric to evaluate the effectiveness of classification is accuracy. The calculation involves ascertaining the proportion of accurately categorized samples relative to the total number of samples.

Conversely, precision quantifies the proportion of positive class forecasts that actually belong to the positive class. To find this, utilize the formula below:

TABLE I. (TP + FP) / TP = PRECISION

algorithm	Precision	Recall	F score	Accuracy
CNN-VGG 30	0.73	0.7	0.74	0.71
VFN-INSULIN	0.8	0.81	0.84	0.8

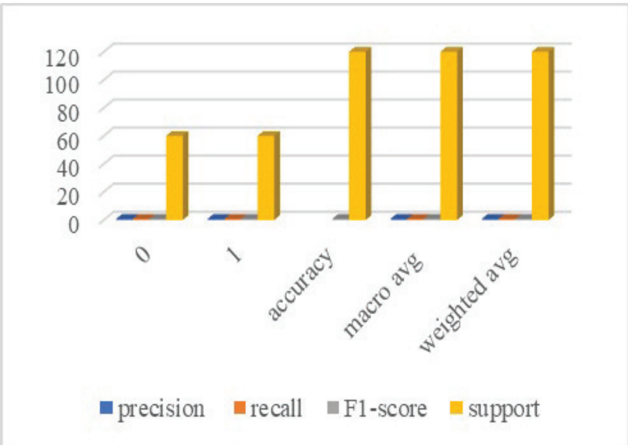

Fig. 5. Comparison Graph

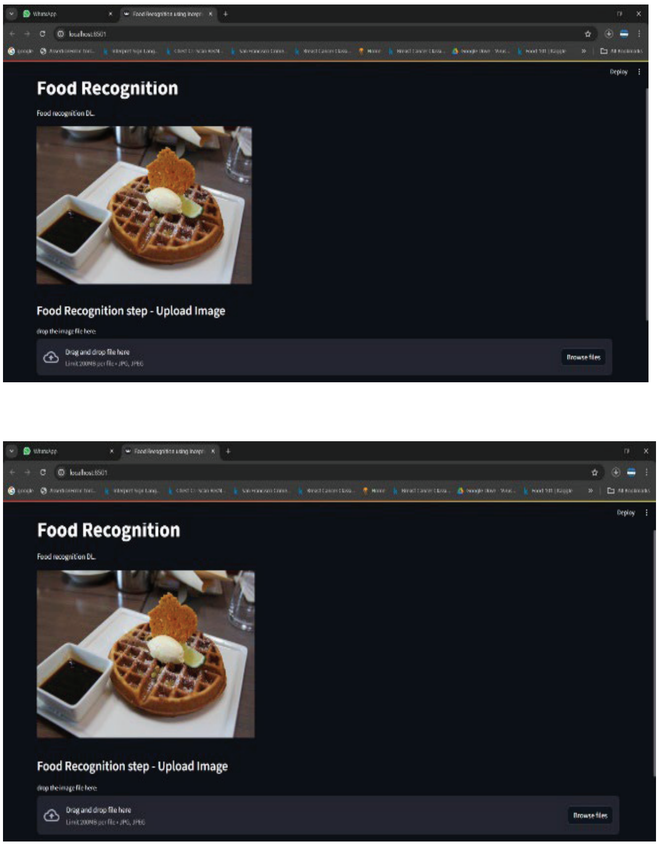

Fig. 6. Output images

VII. CONCLUSION

Ultimately, the suggested system establishes a strong food recognition system through the combination of the reliable and precise Inception Net architecture with the user-friendly Streamlit interface. The process commences with a range of data collection and preparation techniques, guaranteeing a substantial dataset for training purposes. The system's capacity to differentiate between different food categories and display styles is improved through iterative model training alongside the utilization of Inception Net for feature extraction. The incorporation of Streamlit enables a smooth user interaction, simplifying the process of uploading images and obtaining immediate predictions.

VIII. FUTURE WORK

Food Recognition Systems (FRSs) can revolutionize the way people perceive and interact with food. These systems play a crucial role in cultural and culinary education, aiding individuals with dietary restrictions, automating meal tracking, providing nutritional information, and supporting those with food allergies. Among the various benefits of FRSs, one of the most promising areas is assisting individuals with dietary limitations in identifying safe food options.

Enhancing the accuracy and reliability of the system can be achieved by utilizing a broader range of food image datasets, implementing advanced feature extraction techniques, and employing more sophisticated machine learning models. Improving system scalability and efficiency can be accomplished by optimizing the code, utilizing lightweight models, and incorporating distributed computing methods. To enhance user experience, designing user-friendly interfaces and integrating the system with existing applications can promote a more accessible and user-friendly system.

REFERENCES

[1] The creators of this paper are Jiangpeng He, Zeman Shao, Janine Wright, Deborah Kerr, Hymn Boushey, and Fengqing Zhu. The title of the paper is "Perform various tasks picture based dietary appraisal for food acknowledgment and part size assessment." It was introduced at the 2020 IEEE Gathering on Media Data Handling and Recovery, and it tends to be found on pages 49-54 of the meeting procedures.

[2] Runyu Mao, Jiangpeng He, Zeman Shao, Sri KalyanYarlagadda, and Fengqing Zhu are the authors of this paper. "Visual aware hierarchy based food recognition" is the title of the paper. It was introduced at the Global Gathering on Example Acknowledgment Studio, and it tends to be found on pages 571-598 of the meeting procedures.

[3] The paper was published in the IEEE Transactions on Pattern Analysis and Machine Intelligence under the title "Distilling knowledge by mimicking features." The creators are not referenced in the given text. It is found in volume 44, issue 11, and ranges pages 8183-8195.

[4] Seulki Park, Youngkyu Hong, ByeonghoHeo, Sangdoo Yun, and Jin Young Choi are the authors of this paper. The title of the paper is "The greater part can help the minority: Setting rich minority oversampling for long-followed arrangement." It was presented in 2022 at the IEEE/CVF Conference on Computer Vision and Pattern Recognition, and the proceedings of that conference include pages 6887-6896.

[5] The creators of this paper are Zeman Shao, Yue Han, Jiangpeng He, Runyu Mao, Janine Wright, Deborah Kerr, Ditty Jo Boushey, and Fengqing Zhu. The title of the paper is "A coordinated framework for versatile picture based dietary appraisal." It was introduced at the third Studio on AIxFood in 2021, and it very well may be tracked down on pages 19-23 of the studio procedures.

[6] Chao and Hass, "Choice-based user interface design of a smart healthy food recommender system for nudging eating behavior of older adult patients with newly diagnosed type II diabetes," in Proc. Int. Conf. Hum.-Comput. Interact. Cham, Switzerland: Springer, 2020, 201-234.

[7] Data Science and Big Data Analytics released a report by Vairale and Shukla titled "Recommendation framework for diet and exercise based on clinical data: A systematic study." Springer, Singapore, 2019. pages 346–333.

[8] "Health-aware food recommender system," M. Ge, F. Ricci, and D. Massimo, Proc. 9th ACM Conf. Recommender Syst., Sep. 2020, pp. 333–334.

[9] "Differences in dietary intake exist among us adults by diabetic status using NHANES 2009-2016," Luotao Lin, Fengqing Zhu, Edward J. Delp, and Heather AEicher-Miller.

[10] M. Melchiori, N. De Franceschi, V. De Antonellis, D. Bianchini, "PREFer: A prescription-based food recommender system," Computer Standards and Interfaces, vol. 54, pp. 64-75, November 2019.

[11] Tran et al., "Hospital recommender systems: Current and future directions," Journal of Intell. Inf. Syst., vol. 57, no. 1, pp. 171-201, August 2021.

[12] "Improving dietary assessment via integrated hierarchy food classification," IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP), 2021 IEEE, pp. 1–6, Runyu Mao, Jiangpeng He, Luotao Lin, Zeman Shao, Heather A. EicherMiller, and Fengqing Zhu.

[13] "Online continual learning for visual food classification," Jiangpeng He and Fengqing Zhu, Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, October 2021, pp. 2337–2346.

[14] "An open-ended continual learning for food recognition using class incremental extreme learning machines," Ghalib Ahmed Tahir and Chu Kiong Loo, IEEE Access, vol. 8, pp. 82328–82346, 2020.

[15] Shuicheng Yan, Jiashi Feng, Bryan Hooi, Yifan Zhang, and Bingyi Kang. "A comprehensive analysis of deep long-tailed learning," arXiv preprint arXiv:2110.04596, 2021.

[16] "Conversational agents for recipe recommendation," Barko-Sherif, Elsweiler, and Harvey, Proc. Conf. Hum. Inf. Interact. Retr., Mar. 2020, pp. 73-82.

[17] M. Melchiori, N. De Franceschi, V. De Antonellis, D. Bianchini, "PREFer: A prescription-based food recommender system," Computer Standards and Interfaces, vol. 54, pp. 64-75, November 2019.

[18] "Towards user-oriented privacy for recommender system data: A personalized approach to gender obfuscation for user profiles," by Slokom, Hanjalic, and Larson. Article number. 102722, Inf. Process. Manage., vol. 58, no. 6, Nov. 2021.

[19] In October 2019, Yu and colleagues published "A cross-domain collaborative filtering algorithm with growing user and item features via the latent factor space of auxiliary domains" in Pattern Recognition, vol. 94, pp. 96-109.

[20] [A. Kale and N. Auti, "Automated menu planning algorithm for children: Dietary management system employing ID3 for Indian food database,' Proc. Comput. Sci., vol. 50, pp. 197-202, January 2020].