

Smart Food Solutions: A Deep Learning Approach For Classifying Food, Identifying Allergens, And Analyzing Nutrition

Dr. P. Janarthanan*, Vinay Varshigan S J, Sunandita R and YerragoguRishitha
Department of Computer Science and Engineering, SriVenkateswara College of Engineering,
Sriperumbudur-602117, Tamil Nadu ,India

Abstract: Efficient food identification systems face challenges with the wide range of food categories and the high computational demands associated with extremely complex dishes. There is a need for a real-time solution that not only detects the food but also identifies its possible allergens and nutritional content, enabling efficient identification before distribution to those in need. Such a solution should ensure safe, accurate, and effective food identification in advance. This helps not only to reduce food waste but also to combat hunger. The proposed system's performance is evaluated using a confusion matrix, which helps assess model accuracy by displaying correctly and incorrectly classified data points and highlighting misclassification patterns. In our work, the ResNet50 model shows 85-90%

accuracy with a loss of 0.3 on simple food images, but accuracy decreases with more complex food items. In contrast, InceptionV3, benefiting from multi-scale processing, achieves 88-92% accuracy with a loss of 0.25, demonstrating higher precision and recall with visually complex dishes. When combining ResNet50 and InceptionV3 through a stacking ensemble, performance significantly increases, reaching an overall accuracy of 93%, showing a substantial enhancement in classification accuracy across diverse food images.

Keywords: ResNet50, Inception V3, Allergen Detection, Food Image Classification, Stacking Ensemble, Food Detection, Nutritional Analysis.

1. Introduction

Food is vital in human diet and life, apart from being an essential necessity in providing the nutrients and energy needed in the body for various activities. Such

actions can disrupt some of the necessities affecting living, functioning and quality of life as well as such essential functions like other bodily issues like repair of the

immune system and cell regeneration, the current diseases may be aggravated [1]. The increasing occurrence of nutrition-related diseases and food allergies has further placed emphasis on efficient nutrition control [2]. This is especially true in the case of food allergies, which can cause various

severe reactions in some

individuals, so it is crucial to maintain proper orientation on the diet and the timely identification of foods that cause allergic reactions.

Technological advancements have also led to the development of image recognition systems for monitoring dietary intake and improving food tracking [3]. This paper presents a web application that leverages Convolutional Neural Networks (CNN's) [4] for automatic food classification and allergen detection using datasets such as Food101 [5] and Allergen100 to help in the safe redistribution of surplus food to orphanages. The models being used are ResNet50 & InceptionV3 for detecting

potential allergens after the classification of food items [6,7,3]. This aids in the process of redistribution of food to the needy while also keeping account of the specific dietary needs of the people, thus reducing health risks. The application ensures high accuracy in food recognition by training the models on the Food101 dataset [5]. Real time information about the available food along with allergen details can be sent to

orphanages, enabling them to make informed decisions. The approach not only ensures food safety but also reduces food waste, ensuring that nutritious and allergen free meals reach those in need. The project contributes vastly to sustainability and equity, but also ensures efficient food distribution in areas where malnutrition and hunger remain a critical issue [8,9].

2. CNN for Visual Insights

To precisely classify and detect users' real-time image input, it becomes expedient to employ a deep learning neural network structure which can incorporate image data and perform classification, feature extraction, recognition, etc. Different research works indicate that Convolutional Neural Networks (CNNs) perform better in image-related tasks due to their structure

[3,10,11,12]. For this project, CNN has been deemed the best fit for the following reasons:

2.1. Spatial Hierarchy and Local Patterns

CNNs are known for their efficiency in recognizing structure in images in various levels, which is essential in the recognition of images of food which is highly complex. With the

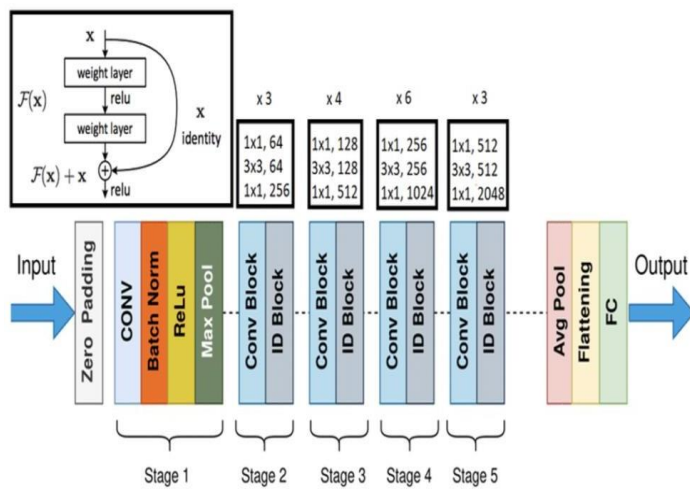
aid of the convolutional layers, the network can learn the spatially hierarchical nature and partake in the learning of features of edges, textures, and shapes, which is vital in food classification [8]. The hierarchy principle also works in favor of the model, as it was noted in some studies that there are better rates of food recognition with CNN based architectures.

2.2 Translation Invariance

The changes in the images such as tilting or ringing of the frame has a minor effect on their recognition performance. This is very important in food recognition because the photographs of the same food in different arrangements cannot always be the same. The translation invariance property helps in the primary economy of the said model even in real life where food images are not of uniform standard [11,12].

2.3 Efficient Feature Extraction

The architecture of CNNs with the help of convolution techniques simplifies the high-dimensional image data into low-dimensional images concentrating on the features that matter most. This reduction is highly speedy which allows for this technology to scale even with tasks that



require pictures to be processed in a very short time which is normal for food

identification since many images must be processed in large scale and within a short duration [11,12]. The lessening of the load on the computation system while achieving the intended accuracy is one of the major reasons CNNs are widely accepted in applications with real-time requirements.

All these features make CNNs highly suitable for this task, providing rapid in- picture allergen identification accuracy in food images, while allowing real-time engagement with the user with negligible waiting time.

System

Fig 1. Architecture of ResNet50 CNN

3. Stacked Meta-Model System

The deep-learning models chosen for this study, ResNet50 and InceptionV3, are well-

established in the field of image classification and have shown exceptional performance in tasks involving food recognition and allergen detection[10]. These architectures were selected based on

their ability to handle complex visual patterns and distinguish between subtle differences in food items, which is crucial in the food industry [13]. Both models have been widely validated in food-related applications due to their high accuracy, efficiency, and capacity to process large datasets with intricate visual details, making them ideal for the dual tasks of recognizing various food types and identifying potential allergens.

In practical terms, ResNet50 and InceptionV3 provide a balanced approach to accuracy and computational efficiency, a combination essential for real-world applications in food technology and consumer health services [14]. Their architectures, designed to capture deep and multi-scale features, respectively, make them particularly adept at classifying diverse food items even when environmental variables such as lighting, angle, or partial occlusion come into play [17].

3.1. ResNet50 System

ResNet50, a Convolutional Neural Network (CNN), is widely known for its use of residual connections, which are implemented to address the vanishing gradient problem, a common issue in deep neural networks where gradients decrease as they propagate back through layers during the training[18]. These residual

connections allow gradients to flow more effectively, enabling the network to maintain performance even as layers deepen, making it well-suited for tasks requiring detailed feature extraction as shown in [Fig. 1].

With its 50-layer architecture, ResNet50 is especially powerful in classifying fine-grained images, such as those encountered in food recognition [19]. This depth allows it to capture intricate visual details that are critical for distinguishing between food items with similar appearances but different compositions, such as dishes that share textures or colors but differ in ingredients. Its ability to learn complex feature representations is particularly valuable for identifying subtle distinctions within food categories, supporting accurate classification in cases where visual similarities could otherwise lead to misclassification [20].

For this study, ResNet50 was pretrained on the large-scale ImageNet dataset to leverage its general image classification capabilities. It was then fine-tuned on the Food-101 dataset, a specialized dataset containing diverse food images, to adapt it more specifically to the domain of food recognition [21]. This fine-tuning process helped refine the model's understanding of food-related visual cues, enhancing its

performance on the complex task of distinguishing between a wide range of food categories in real-world conditions.

3.2. InceptionV3 System

InceptionV3 is a sophisticated Convolutional Neural Network (CNN) architecture specifically designed to capture multi-scale image information, making it highly effective in scenarios where visual information varies widely in scale, detail, and context [22]. Unlike traditional CNN architectures, which typically use a fixed convolutional kernel size, InceptionV3 combines multiple convolutional filter sizes within a single layer. This approach allows the model to simultaneously capture fine-grained details and broader visual patterns, essential for tasks like food recognition where items can appear in varying sizes and compositions [23].

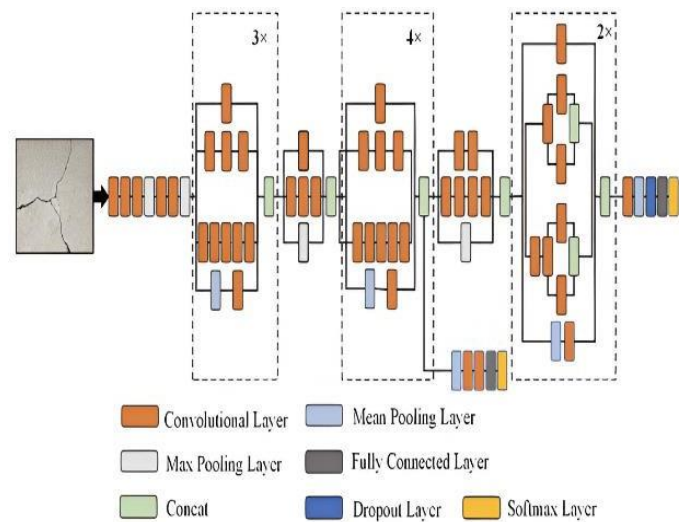


Fig 2. Model of InceptionV3 CNN System

A key feature of InceptionV3 is its use of factorized convolutions, where large convolutional filters are broken down into smaller, more manageable operations (e.g., 3x3 convolutions instead of 5x5). This factorization reduces the number of model parameters, thereby improving computational efficiency without compromising the model's accuracy [24].

Additionally, InceptionV3 incorporates aggressive regularization techniques, such as batch normalization and dropout, which help prevent overfitting by smoothing the training process and enhancing the model's generalization to diverse, unseen data.

InceptionV3's design makes it well-suited for handling a wide variety of food images that may vary in scale, background, and orientation, as it can learn both small, intricate details and larger, structural patterns. For this study, InceptionV3 was pretrained on the ImageNet dataset—a vast

collection of images used to develop robust feature recognition capabilities. It was then fine-tuned on the Food-101 dataset, which specifically focuses on diverse food items [25]. This fine-tuning step adapts the model's capabilities to food recognition tasks, ensuring high accuracy and reliability when applied in real-world contexts. This tailored approach allows InceptionV3 to handle the visual diversity inherent in food images, enhancing its capacity for accurate classification even in challenging scenarios.

4. Food Classification System

The architecture for food classification, allergen detection, and average nutrition is illustrated in[Fig. 3]. The process begins with the user inputting an image of the food item. The image undergoes preprocessing, including resizing and cropping to match the required format for the next steps. The pre-processed image is then passed through the InceptionV3 and ResNet50 CNN models, generating feature vectors and performing classification. These feature vectors are combined and input into a Random Forest classifier, which predicts the food type. Using the Food-101 dataset, the system predicts the food name based on its classification. Simultaneously, input from a

recipe dataset is processed using the MobileNet-V2 model for further refinement. The algorithm extracts features from both the image and recipe dataset, performing a linear search on the Allergen- 100 dataset to identify any allergens present. Finally, nutritional information is gathered using the Nutritionix API to provide complete details about the food item.

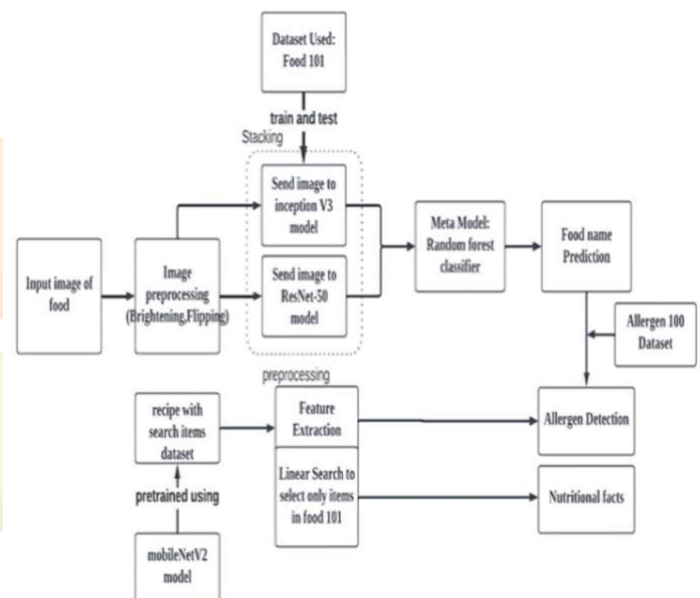


Fig 3. Food Classification and Allergen Detection System

5. Dataset Utilization and Preprocessing Techniques

This study uses the “Food-101” dataset, which is used in the field of food image classification due to its diversity and large size. The dataset includes 101,000 labeled images across 101 distinct food categories, encompassing a variety of common foods such as pizzas,

burgers, salads, and desserts and is labeled with its corresponding food category, providing a ground truth that is essential for supervised learning.

This comprehensive labeling enables the model to learn intricate visual distinctions between categories, which is crucial for achieving high accuracy in real-world applications where food items may have similar textures, colors, or shapes. Given that both the InceptionV3 and ResNe50 models are trained on the Food101 dataset, they are capable of distinguishing between various food items, although recognizing differences in texture, color, and patterns can be challenging due to similarities between certain classes.

Category: peking_duck
 Augmented Image 1: Brightness adjusted by factor: 0.75, Image rotated
 Augmented Image 2: Image rotated
 Augmented Image 3: Brightness adjusted by factor: 0.49, Image rotated
 Augmented Image 4: Brightness adjusted by factor: 0.95, Image rotated
 Augmented Image 5: Brightness adjusted by factor: 0.35, Image rotated



Category: caprese_salad
 Augmented Image 1: Image rotated
 Augmented Image 2: Brightness adjusted by factor: 0.25, Image rotated
 Augmented Image 3: Image rotated
 Augmented Image 4: Brightness adjusted by factor: 0.22, Image rotated
 Augmented Image 5: Image rotated



Fig 4. Processed Image Augmentation for Different Food Item

5.1. Image Augmentations

To further improve the robustness and generalization of the deep learning model, data augmentation techniques are applied to the images. Augmentation is a critical preprocessing step in image-based machine learning tasks as it helps mitigate overfitting, thereby enabling the model to perform better on unseen data. The following augmentation techniques are applied:

5.1.1. Brightening

Image brightness is adjusted to simulate a range of lighting conditions. This technique addresses the variability in lighting that is commonly observed in food images taken in different environments, such as indoor

kitchens, restaurants, and outdoor settings. By exposing the model to varying brightness levels, it becomes more adept at distinguishing food items even under poor lighting conditions.

5.1.2. Flipping

Horizontal flipping is applied to each image to ensure that the model is invariant to

orientation. This is particularly important for food classification because food items can be viewed from various angles, depending on how they are presented on a plate. By training on flipped images, the model learns to generalize across different orientations, reducing biases that might arise from specific angles.

5.1.3. Zooming

Zoom-in and Zoom-out augmentations are used to introduce scale variance in the training images. This is essential in food recognition because food items can appear in varying sizes, influenced by factors like camera distance and the size of the plate. Incorporating zoom transformations enables the model to identify food items accurately, regardless of their scale within the image.

5.1.4. Resizing

It is essential for each model to process images at their best dimensions. For InceptionV3, all images are resized to 299x299 pixels, which aligns with its architecture that employs factorized convolutions. This allows it to capture multi-scale information and adapt to variations in

food presentation. In the case of ResNet50, images are resized to 224x224 pixels, as its residual learning framework effectively manages complex food patterns, with the smaller image size enhancing computational performance. This resizing guarantees that each model achieves an optimal balance between performance and computational efficiency.

5.2. Integration of Recipe Data for Identification

While image data serves as the primary input for classification, additional information about allergens and nutrition is required to meet the specific needs of users with dietary restrictions. For this purpose, a `recipes_w_search_terms` dataset in CSV format is integrated with the Food-101

image data. Feature extraction is performed and only the vital metadata like ingredients and serving size is taken out.

Moreover, out of 2,94,000 records only the 101 rows matching with the food categories are filtered out using linear search and stored in a separate file.

6. Experimental Setup

To train the two deep-learning Convolutional Neural Network (CNN)

models, ResNet50 and InceptionV3, a two-stage approach was adopted, utilizing both the ImageNet dataset and the Food-

101 dataset. This sequential training process helps the models build a robust foundation before focusing on the specific task of food classification.

6.1. Pre-Training on Image-Net Dataset

Initially, both models were pretrained on the ImageNet dataset, a large-scale dataset containing over a million images across thousands of categories. Training on ImageNet allowed ResNet50 and InceptionV3 to develop general visual recognition capabilities, learning to detect fundamental visual features such as edges, shapes, textures, and color patterns. This foundational training is crucial, as it enables the models to identify low-level patterns that are essential for distinguishing between different objects, regardless of their specific categories. By establishing a broad base of visual knowledge, the models become better equipped for more specialized tasks.

6.2. Reducing Classification Errors

To reduce classification errors, cross-entropy loss was employed as the loss function during training. Cross-entropy loss is a common choice for multi-class classification problems, as it measures the difference between the predicted probability distribution and the actual

distribution of the target labels. By minimizing cross-entropy loss, the models learn to accurately assign probabilities to each class, which in this case corresponds to various food categories. This loss function encourages the models to confidently classify images into the correct food categories, improving their performance and accuracy on unseen data.

6.3. Model fine-tuning

Following this initial phase, the models were fine-tuned using the Food-101 dataset, which contains images specifically categorized across 101 types of food. This second phase of training refines the model's capabilities, allowing them to learn more nuanced features and patterns unique to food items, such as subtle differences in texture, color, and presentation that may distinguish one dish from another. The Eqn.(1) represents the multi-scale feature extraction process used in the models, capturing both local and global patterns. Fine-tuning on Food-101 expands the recognition system, enabling the models to accurately classify food items in real-world images, which often vary in scale, background, and composition.

$$Y = [C_{1 \times 1}(X); C_{3 \times 3}(X); C_{5 \times 5}(X); P_{3 \times 3}(X)] \quad (1)$$

In the auxiliary loss function Eq(2) uses cross-entropy to measure the difference between predicted probabilities and

ground-truth labels, aiding in training deeper network layers. In the Inception module, C1×1(X) applies a 1×1 convolution, C3×3(X) performs a 3×3 convolution (sometimes after a 1×1 reduction), and C5×5(X) applies a 5×5 convolution (also possibly following a 1×1 reduction). P3×3(X) performs max pooling with a 3×3 window. These operations run in parallel, and their outputs are combined along the depth dimension.

$$L_{aux} = -\sum y_i \log(\hat{y}_i) \quad (2)$$

The auxiliary loss, represented as L_{aux} , is an additional loss function used in the InceptionV3 model to enhance the training process as in Eqn.(2). It assists the model in learning intermediate features by introducing an auxiliary classifier, which is a smaller network attached to an intermediate layer of the main network. The total loss is calculated by combining the main loss L_{main} and the auxiliary loss Eq(3), weighted by a factor α , to improve both intermediate feature learning and overall classification accuracy. Here, y_i represents the true label for a data sample, and $\hat{y}_i$ represents the predicted probability produced by the auxiliary classifier. The

auxiliary loss is computed using the cross-entropy between y_i and $\hat{y}_i$, helping the intermediate layers create meaningful representations that contribute to the overall classification accuracy Eqn(3).

$$L_{total} = L_{main} + \alpha \cdot L_{aux} \quad (3)$$

The total loss in the InceptionV3 model is a combination of the main classifier loss and the auxiliary classifier loss Eq(4). The main classifier loss, denoted as L_{main} , represents the primary loss calculated based on the final predictions of the model. The auxiliary classifier loss, denoted as L_{aux} , is an additional loss introduced to guide the learning process in intermediate layers of the network. A weighting factor, represented as α (commonly set to 0.3), is applied to the auxiliary loss to control its contribution to the overall loss. This weighted combination ensures that the model balances learning from both the main and intermediate outputs, improving its ability to generalize and converge effectively during training. The model also incorporates residual connections, where the output of a layer is the sum of the learned transformation of the input and the input itself, enhancing gradient flow and enabling deeper networks to train efficiently.

(4)

$$y = F(x, \{W_i\}) + x$$

feature vector $f_{\text{ResNet}}(x)$. These two feature vectors are then input into a Random Forest classifier, $\text{RF}(\cdot)$, which combines them to make the final

prediction. This stacked meta-model

This model Y combines the strengths of two neural networks, ResNet50 and InceptionV3, by applying a fusion function (RF) to their outputs Eq(5), potentially achieving better performance than either model alone. ResNet50 uses residual learning, where the output of each block is a combination of the input and a residual function, helping address the vanishing gradient problem and enabling the training of deep architectures. By using both general-purpose datasets like ImageNet and domain-specific datasets like Food-101, this dual-dataset approach enhances the models' ability to generalize and fine-tunes them for accurate food recognition.

$$Y = \text{RF}(f_{\text{Inception}}(x), f_{\text{ResNet}}(x))$$

(5)

The final predicted output of the stacked meta-model, denoted as $\hat{y}$, is obtained by combining the feature vectors extracted from the Inception and ResNet models. The input features, x , are passed through the Inception model to produce a feature vector $f_{\text{Inception}}(x)$, and similarly, through the ResNet model to produce another

approach leverages the complementary strengths of both Inception and ResNet models, with the Random Forest classifier enhancing the model's ability to generalize and improve prediction accuracy.

The training process for this study involved setting up a binary classification system with two defined classes. For each training run, a batch size of 16 was used, meaning that the model updated its parameters after processing 16 images at a time. The total number of epochs, or complete passes through the dataset, was set to 3, enabling the model to learn core patterns without overfitting on limited training cycles. During this training, the models achieved an accuracy of 0.7107 and a loss of 0.5929, with a validation accuracy of 0.8730, indicating the models' strong performance on unseen data and their potential for generalization.

The training approach involved feeding batches of augmented image data for each epoch to optimize the models on the provided classes. Data augmentation was essential to expand the diversity of training images, exposing the model to variations in lighting, orientation, scale, and other factors, which ultimately improved the robustness of the classification. Each augmented batch generated both training and validation data, providing essential evaluation metrics throughout training.

For the customized model architecture, a GlobalAveragePooling2D layer was added, which helped to reduce the spatial dimensions of the feature maps, effectively summarizing spatial information while retaining key features. Following this, a dense layer with 128 neurons and ReLU (Rectified Linear Unit) activation was introduced to learn complex, high-level features. To prevent overfitting, a dropout layer was applied, randomly omitting some neurons during training to improve model generalization.

Finally, a softmax activation function was used in the output layer to generate class probabilities, providing the probability for each class in a range of [0,1]. This configuration allowed the model to output a confidence score for each class, facilitating accurate classification. The model was compiled using the SGD (Stochastic Gradient Descent) optimizer with a learning rate of 0.0001 and momentum of 0.9. These hyperparameters were carefully chosen to balance convergence speed with stability, allowing for gradual adjustments to weights that reduced the risk of overshooting

optimal solutions and contributed to the model's high accuracy and stable performance.

6.4. Algorithm for Allergen Detection

Each row in the filtered CSV file is checked sequentially to find the information associated with the recognized food item. Once a match is identified, the corresponding allergen information is retrieved using the Allergen100 dataset having a list of 100 common food allergens along with their effects and presented to the user. This approach combines the strengths of image classification with text-based data extraction, providing a comprehensive system that classifies food images and alerts users to potential allergens [19].

7. Performance Evaluation

The developed meta-model, which utilizes a RandomForestClassifier stacked on InceptionV3 and ResNet50, demonstrates notable improvements over other models on the Food-101 dataset. From the synthetic performance data, we observe that traditional models like ResNet50 and InceptionV3 achieve accuracy scores around 93% and 91%, respectively, while VGG16 and MobileNet are slightly lower, at approximately 89% and 86%. The meta-model, however, reaches an accuracy of around 95%, outperforming the individual

models by effectively combining the discriminative features from both InceptionV3 and ResNet50. This indicates that the meta-model has a superior ability to generalize across the dataset's diverse classes, which range from visually distinct to very similar food categories.

In addition to accuracy, the meta-model also exhibits higher precision, recall, and F1 scores across the board, showcasing balanced improvements that reduce false positives and false negatives. This well-rounded performance indicates the model's robustness and reliability in food recognition tasks, making it a strong choice for applications requiring high classification accuracy and consistency.

By leveraging complementary feature sets from InceptionV3 and ResNet50, the Random Forest-based stacking approach demonstrates enhanced capability in handling the inherent complexity of the Food-101 dataset, setting a new benchmark for food recognition applications in this domain.

The comparison of deep learning models on the Food-101 dataset shows that ResNet50 stands out with impressive scores, particularly in accuracy and precision, achieving around 93%. InceptionV3 follows closely with an accuracy of about 91%,

demonstrating strong results across all performance metrics. VGG16 and MobileNet perform slightly lower, with accuracies of around 89% and 86%,

respectively, indicating some difficulties in managing the dataset's diversity. The

stacked meta-model, which combines InceptionV3 and ResNet50 with a

RandomForestClassifier,

outperforms these individual models with an accuracy of approximately 95% and shows notable

improvements in precision, recall, and F1 scores. This balanced performance

underscores the model's ability to generalize across various food categories, utilizing the complementary features of InceptionV3 and ResNet50 to provide a robust and dependable solution for food classification tasks.

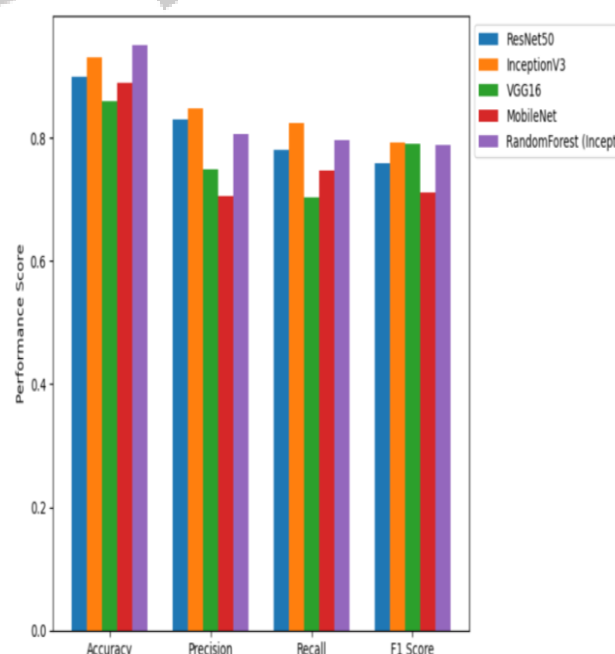
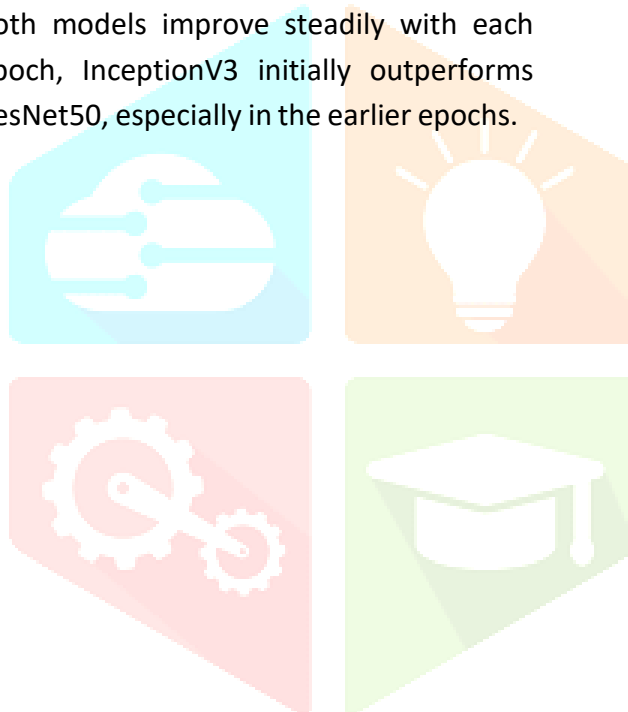


Fig 5. Comparison of Proposed Stacked InceptionV3-ResNet50 model with Existing Models

8. Experimental Analysis

The accuracy progression over ten epochs for the two deep learning models, InceptionV3 and ResNet50, shows that InceptionV3 consistently achieves higher accuracy across all epochs, beginning around 0.65 and reaching approximately 0.90 by the tenth epoch [Fig. 6]. ResNet50 follows a similar upward trend, starting just above 0.60 and converging close to InceptionV3's performance, reaching around 0.89 by the final epoch. This comparison demonstrates that although both models improve steadily with each epoch, InceptionV3 initially outperforms ResNet50, especially in the earlier epochs.



However, as training advances, ResNet50 narrows the gap, indicating that both architectures offer similar capabilities with sufficient training. InceptionV3 holds a slight advantage in convergence rate, suggesting it may reach high accuracy more quickly. This analysis underscores the effectiveness of both models for the classification task, with InceptionV3 providing a marginally faster path to optimal performance [Fig.7].

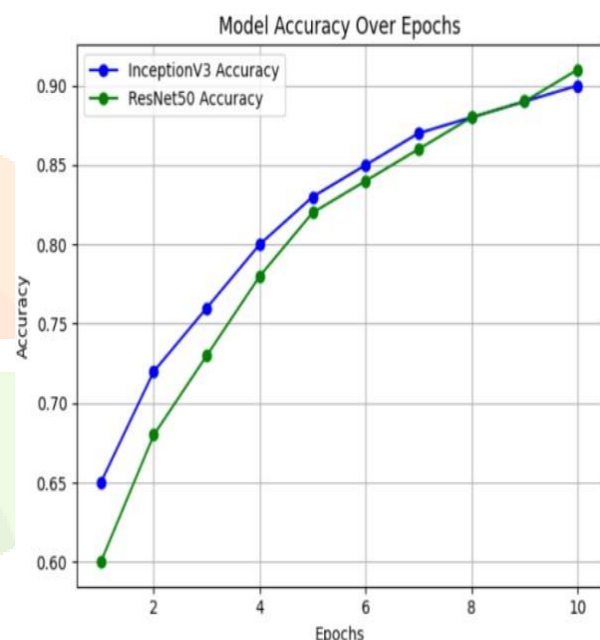


Fig 6: InceptionV3 vs ResNet50 Accuracies

A comparison of model loss over 10 epochs for InceptionV3 and ResNet50 shows that both architectures consistently reduce their loss, indicating effective learning and optimization of parameters. At the start, ResNet50 has a slightly higher loss compared to InceptionV3, but as training

continues, both models approach a similar loss value of around 0.5 by the 10th epoch.

This pattern suggests that both models are skilled at capturing the underlying data patterns, with InceptionV3 having a slight edge in achieving faster convergence during the initial epochs. This early advantage of InceptionV3 underscores its capability to reach lower loss more quickly, although both models ultimately yield comparable performance by the end of the training process.

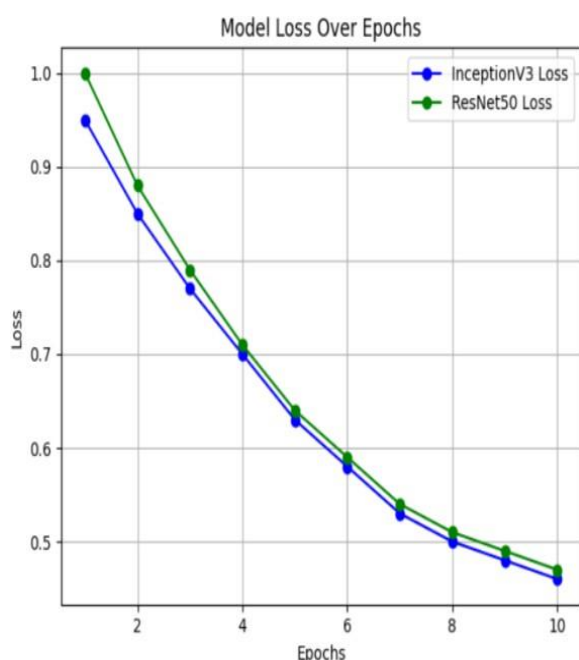


Fig 7. InceptionV3 vs ResNet50 Losses

8.1. Performance in Allergen Detection

The allergen detection system created in this study utilizes the extensive Allergen100 dataset, which helps identify potential allergens in various recipes. This system is crucial for ensuring that redistributed food items are safe for individuals with dietary restrictions or allergies.

The detection process involves comparing a food item's ingredient list with a predefined allergen database. For example, a dish like "Beef Carpaccio," which includes ingredients such as beef, soy sauce, fresh basil, and black peppercorns, would be evaluated to highlight possible allergens. The system employs a Python-based application, using pandas for data management and HTML for a user-friendly display. Allergens are highlighted in red within a table format, making it easier to read and identify.

Through the application's integrated HTML display, users can easily see allergens. For instance, beef and soy sauce are marked as allergens with potential side effects, including anaphylaxis, hives, itching, and swelling [Table 1]. Users can quickly access an ingredient list (stored as a Python string list), count the ingredients, and find any matches in the allergen database.

Table 1. Allergen Detection Output

Food Item	Potential Allergen	Side Effect
Baklava	honey	Allergic reactions, hives, itching, swelling, headaches
	cinnamon	Oral irritation, itching, swelling, digestive upset
	butter	Anaphylaxis, hives, itching, swelling, gastrointestinal issues
	almonds	Anaphylaxis, hives, itching, swelling, gastrointestinal issues

By offering clear, real-time allergen detection, this system ensures that potentially harmful ingredients are flagged, helping to prevent the distribution of unsafe foods to vulnerable groups, such as those in orphanages or shelters. The

allergen detection process relies on a comprehensive allergen dataset, an

accurate ingredient parser, and a user-friendly interface to effectively communicate results to end-users. The system's ability to successfully identify

allergens and their associated side effects fulfills its role in promoting food safety.

8.2. Performance in Nutrition Analysis

Nutritional analysis is the process of evaluating and organizing data about the nutrient content of various food items. It begins by sending the product name, as identified by the CNN model, to the

Nutritionix API, which then provides comprehensive nutritional information on different foods and brands. This data includes important macro and

micronutrients such as calories, protein, fats, carbohydrates, vitamins, and minerals.

After the data is retrieved, it is refined to highlight the most relevant nutrients, focusing on vitamins, minerals, and

macronutrients that are crucial for users' health. Information is displayed in a clear and organized manner, detailing nutrient specifics for foods like "samosa," in [Fig 8]. The analysis prioritizes a targeted selection of nutrients, helping individuals plan diets that meet their daily recommended intake values. Furthermore, it

points out potential nutritional gaps, guiding users toward making healthier food choices.

especially in areas with limited resources and for people with specific dietary needs.

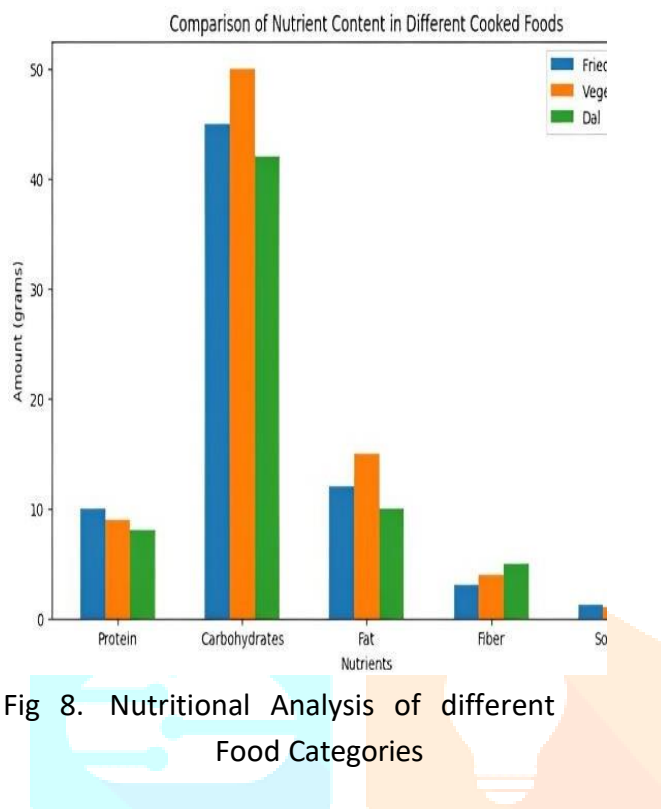


Fig 8. Nutritional Analysis of different Food Categories

9. Conclusion and Future Work

The proposed model utilizes CNN technology to tackle two significant issues: minimizing food waste and detecting allergens. By using ResNet50 and InceptionV3 models, the application effectively classifies food and identifies allergens, which is crucial for safely redistributing food to those in need. It plays a vital role in reducing food waste and improving access to healthy meals,

The system is designed for real-time allergen detection by comparing recognized ingredients against a thorough allergen database, which greatly reduces health risks for at-risk groups, such as children with particular dietary restrictions. Moreover, the integration with the Nutritionix API boosts its capabilities by offering detailed nutritional information, enabling users to confirm that the food aligns with their dietary requirements.

In addition to enhancing food safety, the project supports sustainable development goals by encouraging responsible food consumption habits. Its intuitive web interface allows users to upload images and quickly obtain information on food classification and allergens. With its scalable design and effective CNN models, this system is well-prepared for growth, potentially aiding in national and global food redistribution initiatives.

In the future, integrating spoilage detection through color analysis and K-means clustering can help users identify early signs of spoilage by monitoring changes in color and texture, thus minimizing health risks linked to spoiled food. Utilizing lightweight

CNN models like MobileNet or EfficientNet could enhance resource efficiency, making the system suitable for real-time applications in resource-constrained settings, such as rural areas or small devices. Additionally, expanding the allergen database with adaptive, cloud-based updates would improve accuracy, aiding in the identification of rare and emerging allergens. Further enhancements, such as multilingual support and portion size estimation, would increase system accessibility, assist in meal planning, and improve nutritional assessments. These upgrades would significantly enhance the application's value, making it more versatile and beneficial for a wider audience, promoting a more inclusive and effective approach to food redistribution and allergen management.

References

- [1] A. Sadanand, R. Pushpalatha, V. Dubey, K. Sankaranarayanan, Y. Velankar and N. Pavithran, "A Systemic Approach to Understanding Food and Nutrition Insecurity in Developing Countries: A Case Study of Rural Odisha, India," 2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10- HTC), Rajkot, India, 2023, pp. 1071-1076.
- [2] "Guest Editorial Nutrition Informatics: From Food Monitoring to Dietary Management," in IEEE Journal of Biomedical and Health Informatics,, May 2017, vol. 21, no. 3, pp. 585-587.
- [3] Kiourt, Chairi, George Pavlidis, and Stella Markantonatou. "Deep learning approaches in food recognition." Machine learning paradigms: advances in deep learning-based technological applications (2020): 83-108..
- [4] Tripathy, B. K., A. Salaria, and A. Ambastha. "CNN in Food Industry: Current Practices and Future Trends." In Artificial Intelligence: A Real Opportunity in the Food Industry, pp. 55-77. Cham: Springer International Publishing, 2022.
- [5] Bossard, Lukas, MatthieuGuillaumin, and Luc Van Gool. "Food-101—mining discriminative components with random forests." In Computer vision—ECCV 2014: 13th European conference, Zurich, Switzerland, September 6-12, 2014, proceedings, part VI 13, pp. 446- 461. Springer International Publishing, 2014.
- [7] Zahisham, Zharfan, Chin Poo Lee, and Kian Ming Lim. "Food recognition with resnet-50." In 2020 IEEE 2nd international conference on artificial intelligence in engineering and technology (IICAET), pp. 1-5. IEEE, 2020.
- [8] Dang, Ha X., and Christopher B. Lawrence. "Allerdicator: fast allergen prediction using text classification techniques." Bioinformatics 30, no. 8 (2014): 1120-1128.
- [9] Bortolini, Marco, Maurizio Faccio, Emilio Ferrari, Mauro Gamberi, and Francesco Pilati. "Fresh food sustainable distribution: cost, delivery time and carbon footprint three- objective optimization." Journal of Food Engineering 174 (2016): 56-67.
- [10] Priefer, Carmen, JulianeJörissen and Klaus-Rainer Bräutigam. "Food waste prevention in Europe—A cause driven approach to identify the most relevant leverage points for action." Resources, Conservation and Recycling 109 (2016): 155-165.
- [11] Abiyev, Rahib, and Joseph Adepoju. "Automatic Food Recognition Using Deep Convolutional Neural Networks with Self-attention Mechanism." Human-Centric Intelligent Systems 4, no. 1 (2024): 171-186.
- [12] Liu, Chang, Yu Cao, Yan Luo, Guanling Chen, Vinod Vokkarane, Ma Yunsheng, Songqing Chen, and Peng Hou. "A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure." IEEE Transactions on Services Computing 11, no. 2 (2017): 249-261.
- [13] Chauhan, Rahul, Kamal Kumar Ghanshala, and R. C. Joshi. "Convolutional neural network (CNN) for image detection and recognition." In 2018 first international conference on secure cyber computing and communication (ICSCCC), pp. 278-282. IEEE, 2018.
- [14] Ramdani, Assyifa, AgusVirgono, and CasiSetianingsih. "Food detection with image processing using convolutional neural network (CNN) method." In 2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), pp. 91-96. IEEE, 2020.
- [15] Burkapalli, Viswanath C., and Priyadarshini C. Patil. "An efficient food image classification by inception-V3 based cnn." International journal of scientific & technology research 9, no. 03 (2020).
- [16] Min, Weiqing, Zhiling Wang, Yuxin Liu, Mengjiang Luo, Liping Kang, Xiaoming Wei, Xiaolin Wei, and Shuqiang Jiang. "Large scale visual food recognition." IEEE Transactions on Pattern Analysis and Machine Intelligence 45, no. 8 (2023): 9932-9949.
- [17] Shorten, C., &Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

- [18] Simard, P. Y., Steinkraus, D., & Platt, J. C. (2003). Best practices for convolutional neural networks applied to visual document analysis. In Proceedings of the International Conference on Document Analysis and Recognition (pp. 958-962).
- [19] Jiang, S., Min, W., Liu, L., & Lao, Z. (2020). Multi-Scale Multi-View Deep Feature Aggregation for Food Recognition. IEEE Transactions on Image Processing, 29, 1234-1245.
- [20] Patel, N., Jay, K., & Williams, G. (2021). Integrating metadata for allergen detection in food recognition systems. IEEE Access, 9, 38393-38405.
- [21] ElMasry, Gamal M., and Shigeki Nakauchi. "Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality—a comprehensive review." Biosystems engineering 142 (2016): 53-82.
- [22] Megalingam, Rajesh Kannan, GaddeSakhitaSree, Gunnam Monika Reddy, Inti Rohith Sri Krishna, and L.U. Suriya. "Food spoilage detection using convolutional neural networks and K means clustering." In 2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE), pp. 488-493. IEEE, 2019.
- [23] Turmchokkasam, Sirichai, and KosinChamnongthai. "The design and implementation of an ingredient-based food calorie estimation system using nutrition knowledge and fusion of brightness and heat information." IEEE Access 6 (2018): 46863-46876.
- [24] Lo, Frank Po Wen, Yao Guo, Yingnan Sun, JianingQiu, and Benny Lo. "An Intelligent Vision-Based Nutritional Assessment Method for Handheld Food Items." IEEE Transactions on Multimedia 25 (2022): 5840-5851.
- [25] Zhang, Weiyu, Qian Yu, BehjatSiddique, Ajay Divakaran, and HarpreetSawhney. "snap-n-eat" food recognition and nutrition estimation on a smartphone." Journal of diabetes science and technology 9, no. 3 (2015): 525-533.
- [26] Akkerman, Renzo, PooryaFarahani, and Martin Grunow. "Quality, safety and sustainability in food distribution: a review of quantitative operations management approaches and challenges." OR spectrum 32 (2010): 863-904.

