International Research Journal of Education and Technology

Peer Reviewed Journal, **ISSN: 2581-7795**

Food classification using Deep learning

Dr.V.R.sugumaran , Research Guide

Gokulnath.R, Abishek.A, Aswath.B

Final Year Student

Department of Computer Science and Engineering,

E.G.S. PILLAY ENGINEERING COLLEGE, Nagapattinam, India

ABSTRACT

Food classification using deep learning has emerged as a significant research area within computer vision due to its wide applications in dietary monitoring, healthcare management, smart restaurants, agriculture, and automated food supply chains. Accurate recognition and categorization of food items from digital images is a challenging task because of high intra-class variability, inter-class similarity, complex backgrounds, occlusion, varying illumination conditions, and differences in presentation styles. Traditional machine learning approaches rely heavily on handcrafted features such as color histograms, texture descriptors, and shape-based representations, which often fail to generalize across diverse food datasets. To overcome these limitations, this study proposes a robust deep learning-based framework for automatic food image classification.The proposed system leverages Convolutional Neural Networks (CNNs) for hierarchical feature extraction and classification. Transfer learning techniques are employed using pre-trained architectures such as Google's InceptionV3, Microsoft's ResNet, and Oxford University's VGG16 to enhance performance while reducing computational cost and training time. The models are fine-tuned on benchmark food image datasets including Food-101 and custom-curated datasets containing multi-class Indian and international cuisine categories. Extensive data preprocessing techniques such as image resizing, normalization, augmentation (rotation, flipping, zooming, brightness variation), and noise reduction are applied to improve model robustness and prevent overfitting.The architecture integrates multiple convolutional layers followed by pooling layers, batch normalization, dropout regularization, and fully connected dense layers with Softmax activation for multi-class classification. Cross-entropy loss is used as the optimization objective, and adaptive optimizers such as Adam and stochastic gradient descent (SGD) are evaluated to determine optimal convergence behavior. Performance metrics including accuracy, precision, recall, F1-score, confusion matrix analysis, and top-5 accuracy are used to evaluate model effectiveness.Experimental results demonstrate that deep learning-based models significantly outperform traditional machine learning classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests when applied to high-dimensional image data. Among the tested architectures, fine-tuned ResNet-based models achieved superior accuracy due to residual learning mechanisms that mitigate vanishing gradient problems and enable deeper network training. Data augmentation and transfer learning contributed substantially to generalization performance, particularly when training data was limited.Furthermore, the study explores practical deployment aspects including mobile-based food recognition systems, cloud-integrated calorie estimation platforms, and real-time classification using edge computing devices. The proposed framework can be extended to incorporate nutritional analysis, ingredient detection, and portion size

estimation using segmentation models and attention-based mechanisms. The integration of multimodal learning, combining visual features with textual menu descriptions, is also discussed as a future enhancement to improve classification reliability.The results indicate that deep learning provides a scalable, accurate, and efficient solution for automated food recognition. This research contributes to the advancement of intelligent dietary assessment systems and supports applications in personalized nutrition, obesity prevention, hospital diet monitoring, and smart cafeteria management. Future work focuses on improving dataset diversity, reducing model bias, optimizing lightweight architectures for real-time deployment, and incorporating explainable AI techniques to enhance transparency and user trust in food classification systems.

1. INTRODUCTION:

The rapid advancement of artificial intelligence and computer vision technologies has significantly transformed the way visual data is processed and interpreted. Among various computer vision applications, food image classification has emerged as a highly relevant and impactful research domain due to its direct implications in healthcare, nutrition monitoring, lifestyle management, smart agriculture, and the food service industry. The increasing use of smartphones, social media platforms, and food delivery applications has resulted in an enormous growth of food-related image data, thereby creating opportunities for automated food recognition systems. Accurate classification of food images can assist in calorie estimation, dietary assessment, diabetic monitoring, obesity control, and personalized nutrition planning.Food classification, however, presents unique challenges compared to general object recognition tasks. Unlike rigid objects such as vehicles or furniture, food items often lack fixed shapes and exhibit significant intra-class variations due to differences in cooking styles, ingredients, garnishing, lighting conditions, camera angles, and background clutter. Simultaneously, many food categories demonstrate high inter-class similarity, making discrimination between visually similar dishes complex even for humans. Traditional image processing and machine learning techniques rely heavily on handcrafted feature extraction methods such as color histograms, Scale-Invariant Feature Transform (SIFT), Local Binary Patterns (LBP), and texture descriptors. Although these approaches provide moderate performance in controlled environments, they struggle to generalize across diverse real-world scenarios due to limited feature representation capacity.The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized image classification tasks by enabling automatic hierarchical feature extraction directly from raw pixel data. CNN-based architectures have demonstrated remarkable performance improvements in large-scale visual recognition benchmarks such as ImageNet. Inspired by this success, researchers have increasingly adopted deep neural networks for food image classification tasks. Pre-trained deep architectures such as AlexNet, VGG16, ResNet, InceptionV3, and MobileNet have been widely utilized through transfer learning techniques to address limited dataset availability and computational constraints.Transfer learning enables models pre-trained on large datasets to adapt to domain-specific tasks such as food recognition with relatively smaller labeled datasets. Publicly available food datasets such as Food-101 have played a crucial role in accelerating research in this field by providing standardized benchmarks for performance evaluation. However, challenges persist when models are applied to regional cuisines, mixed dishes, and real-time mobile deployment scenarios. Cultural diversity in food presentation further complicates classification tasks, especially in countries with rich culinary traditions.Beyond classification, modern research has extended toward multi-task learning approaches that integrate food detection, segmentation, ingredient recognition, and nutritional value estimation. Attention mechanisms and transformer-based models are increasingly being explored to enhance contextual understanding and fine-grained classification accuracy. Additionally, lightweight deep learning architectures are being developed to support edge computing devices and smartphone-based applications, enabling real-time food recognition without reliance on high-end computational infrastructure.Despite significant progress, several research gaps remain. Many existing models suffer from dataset bias, limited generalization across unseen categories, and high computational complexity. Moreover, explainability and interpretability of deep learning models are critical in healthcare-related applications where decision transparency is essential. Addressing these challenges requires comprehensive dataset augmentation strategies, domain adaptation techniques, and the integration of explainable artificial intelligence frameworks.This study aims to contribute to the ongoing research in food classification using deep learning by exploring robust CNN-based architectures, optimizing hyperparameters, and evaluating performance across diverse datasets. The proposed framework emphasizes scalability, accuracy, and practical deployment feasibility. By leveraging advanced deep learning methodologies, this research seeks to enhance automated dietary monitoring systems and support intelligent healthcare applications.In summary, food classification using deep learning represents a multidisciplinary intersection of computer vision, nutrition science, healthcare technology, and artificial intelligence. Continued advancements in this domain hold the potential to revolutionize dietary management systems and improve global health outcomes through data-driven, automated food recognition solutions.

II. EXISTING SYSTEM:

Existing systems for food classification primarily rely on deep learning–based image recognition frameworks built upon Convolutional Neural Networks (CNNs). Early approaches adopted transfer learning from large-scale visual datasets such as ImageNet, enabling models to leverage pretrained weights for feature extraction and fine-tuning on food-specific datasets. Benchmark datasets like Food-101 have been widely used to evaluate classification accuracy across 101 food categories under diverse presentation conditions.Popular pretrained architectures employed in existing systems include VGG16, ResNet, InceptionV3, and MobileNet. These models extract hierarchical spatial features from food images and use fully connected layers with Softmax activation for multi-class classification. Among them, residual learning in ResNet-based systems has shown improved convergence and deeper network capability, while MobileNet-based systems are preferred for lightweight mobile deployment.Most existing systems apply data preprocessing techniques such as resizing, normalization, and augmentation (rotation, scaling, and flipping) to enhance robustness and reduce overfitting. Optimization algorithms like Adam and stochastic gradient descent (SGD) are commonly used to minimize categorical cross-entropy loss during training. Performance is typically measured using accuracy, precision, recall, F1-score, and confusion matrix analysis.However, current systems face several limitations. Many models struggle with high intra-class variation and inter-class similarity among visually similar dishes. Complex backgrounds, occlusion, and varying lighting conditions reduce classification reliability in real-world environments. Additionally, existing frameworks often require large labeled datasets and high computational resources, limiting scalability in resource-constrained devices. Although some systems integrate calorie estimation and ingredient recognition modules, these components are not fully standardized and may lack cross-cultural generalization.

III.PROPOSED SYSTEM:

Hybrid Deep Learning Architecture with Transfer Learning

The proposed system utilizes a hybrid Convolutional Neural Network (CNN) framework combining feature extraction from pretrained models such as ResNet and MobileNet. Transfer learning is applied to fine-tune the higher layers for domain-specific food categories, reducing training time while improving classification accuracy and generalization.

Comprehensive Dataset Augmentation and Domain Expansion

To address dataset bias, the system integrates advanced data augmentation techniques including rotation, scaling, color jittering, contrast adjustment, and random cropping. In addition to benchmark datasets like Food-101, a custom dataset representing regional and mixed cuisine categories is incorporated to improve cross-domain adaptability.

Attention-Based Feature Enhancement

An attention mechanism is embedded within the CNN architecture to focus on discriminative food regions while suppressing background noise. This improves recognition performance in complex real-world environments with cluttered backgrounds and varying illumination.

Lightweight and Edge-Optimized Deployment

The proposed system integrates model compression techniques such as pruning and quantization to reduce computational complexity. This enables real-time deployment on mobile devices and edge computing platforms without significant loss in accuracy.

Multi-Task Learning for Nutritional Insight

Beyond classification, the system is extended to include optional modules for ingredient detection and calorie estimation, enabling practical applications in dietary monitoring and healthcare support systems

Explainable AI Integration

To enhance transparency, visualization techniques such as Grad-CAM are incorporated to highlight important image regions influencing classification decisions. This improves interpretability and user trust, particularly in medical and nutrition-related applications

IV.SYSTEM FLOWCHART DESCRIPTION

The system flow for food classification using deep learning begins with the image acquisition stage, where food images are captured using smartphones, digital cameras, or collected from benchmark datasets such as Food-101. These images serve as the primary input to the system.The next step is data preprocessing, which includes image resizing, normalization, and noise reduction to ensure uniform input dimensions compatible with deep learning architectures. Data augmentation techniques such as rotation, flipping, scaling, and brightness adjustment are applied to increase dataset diversity and reduce overfitting.Following preprocessing, the images are passed to the feature extraction module based on Convolutional Neural Networks (CNNs). In the proposed system, transfer learning is applied using pretrained architectures such as ResNet or MobileNet. These models extract hierarchical spatial features from input images through multiple convolutional and pooling layers.The extracted feature maps are then forwarded to the feature optimization stage, where batch normalization and dropout layers are applied to improve generalization and prevent overfitting. If incorporated, an attention mechanism highlights the most relevant food regions while suppressing background interference.Next, the processed features enter the classification layer, typically consisting of fully connected dense layers followed by a Softmax activation function to predict probability scores for each food category. The class with the highest probability is selected as the final output.During training, the system computes categorical cross-entropy loss and updates model weights using optimization algorithms such as Adam or stochastic gradient descent (SGD). Model performance is evaluated using metrics including accuracy, precision, recall, and F1-score.Finally, the classified output is displayed to the user through a graphical interface or mobile application. Optionally, the system may provide additional modules such as calorie estimation or ingredient identification to enhance practical usability.

V. SYSTEM ARCHITECTURE AND WORKING MODEL

The proposed system architecture for food classification using deep learning is designed to ensure high accuracy, scalability, and real-time deployment capability. The architecture consists of multiple interconnected modules, including image acquisition, preprocessing, feature extraction, classification, and output generation.The system begins with the input layer, where food images are collected from user devices or standardized datasets such as Food-101. These images may vary in size, resolution, lighting conditions, and background complexity. To ensure consistency, all images undergo a preprocessing stage that includes resizing to fixed dimensions, pixel normalization, and optional noise filtering.Following preprocessing, the images are fed into a deep Convolutional Neural Network (CNN) for hierarchical feature extraction. The architecture leverages transfer learning from pretrained models such as ResNet and MobileNet. These networks consist of multiple convolutional layers that automatically learn low-level features (edges, textures), mid-level features (shapes, patterns), and high-level semantic representations (dish-specific characteristics). Residual connections in ResNet help mitigate vanishing gradient problems, while MobileNet enables lightweight and efficient computation for edge devices.The extracted feature maps pass through pooling layers to reduce spatial dimensions and computational complexity. Batch normalization is applied to stabilize learning, and dropout regularization is incorporated to prevent overfitting. If required, an attention mechanism is integrated into the architecture to focus on relevant food regions while suppressing background noise, thereby improving classification accuracy in complex real-world environments.The optimized feature representations are then flattened and passed into fully connected dense layers. The final classification layer employs a Softmax activation function to compute probability distributions across multiple food categories. The predicted class corresponds to the highest probability score.During the training phase, the model minimizes categorical cross-entropy loss using optimization algorithms such as Adam or stochastic gradient descent (SGD). Backpropagation is used to update network weights iteratively. Performance evaluation is conducted using metrics including accuracy, precision, recall, F1-score, and confusion matrix analysis.In the deployment phase, the trained model is integrated into a web or mobile-based application for real-time food recognition. Optional modules for calorie estimation and ingredient analysis can be incorporated to enhance functionality in healthcare and dietary monitoring applications.Overall, the proposed system architecture provides a structured pipeline that integrates advanced deep learning models, optimization techniques, and deployment strategies to achieve reliable and efficient food classification performance

VI. SYSTEM DESIGN AND DATA FLOW MODEL

The system design for food classification using deep learning is structured as a modular and scalable framework that ensures efficient data processing, accurate classification, and real-time deployment capability. The architecture is divided into functional components that manage data acquisition, preprocessing, feature extraction, model training, inference, and result visualization.

1. Overall System Design

The proposed system follows a layered architecture consisting of:

Presentation Layer – User interface for uploading or capturing food images.

Application Layer – Handles preprocessing, model inference, and output generation.

Model Layer – Deep learning module responsible for feature extraction and classification.

Data Layer – Stores training datasets, trained model weights, and prediction logs.

The design ensures modularity, allowing independent optimization of each component without affecting the overall system.

2. Data Acquisition Module

The process begins with the collection of food images from cameras, mobile devices, or benchmark datasets such as Food-101. Images may vary in resolution, background complexity, lighting, and orientation, which introduces variability in input data.

3. Data Preprocessing Module

Before feeding the images into the deep learning model, preprocessing operations are performed:

Image resizing to standardized dimensions (e.g., 224×224 pixels).

Pixel value normalization to scale intensities between 0 and 1.

Data augmentation (rotation, flipping, zooming, brightness adjustment).

Noise reduction and background filtering (optional).

This stage enhances data consistency and reduces overfitting during model training.

4. Feature Extraction Module

The preprocessed images are forwarded to a Convolutional Neural Network (CNN) for automatic feature extraction. Transfer learning is implemented using pretrained architectures such as ResNet or MobileNet.

The CNN extracts hierarchical representations:

Low-level features (edges and textures)

Mid-level features (patterns and shapes)

High-level semantic features (dish identity)

Pooling layers reduce dimensionality, while batch normalization improves training stability.

5. Classification Module

The extracted features are flattened and passed to fully connected dense layers. A Softmax activation function generates probability scores for each food category. The class with the highest probability is selected as the predicted output.During training, categorical cross-entropy loss is computed, and model parameters are updated using optimization algorithms such as Adam or stochastic gradient descent (SGD).

6. Data Flow Model

The data flow in the system follows a sequential pipeline:

Input Image → Preprocessing → Augmentation → CNN Feature Extraction → Feature Optimization → Dense Layers → Softmax Classification → Output Prediction

During inference mode, only forward propagation is executed, ensuring real-time prediction capability. In training mode, both forward and backward propagation are performed for weight updates.

7. Output and Integration

The final prediction is displayed through a graphical interface or mobile application. Optional modules for calorie estimation and ingredient detection can be integrated to extend the system’s functionality for healthcare and dietary monitoring applications.

CONCLUSION:

Food classification using deep learning demonstrates significant improvements in accuracy and robustness compared to traditional image processing methods. By leveraging advanced Convolutional Neural Network architectures such as ResNet and pretrained datasets like Food-101, the proposed framework effectively handles complex visual variations in food images. The integration of transfer learning, data augmentation, and optimization techniques enhances generalization performance across diverse cuisines. Furthermore, the system supports scalable deployment for real-time applications in healthcare, dietary monitoring, and smart food services. Overall, deep learning-based food classification provides a reliable and efficient solution for automated food recognition systems..

ACKNOWLEDGMENT:

We would like to express our sincere gratitude to all individuals and organizations who contributed to the successful completion of this research. We extend our heartfelt thanks to our institution and faculty members for their guidance, encouragement, and support throughout the study. We also acknowledge the developers and maintainers of publicly available datasets such as Food-101, and the creators of deep learning frameworks and pretrained models including ResNet and MobileNet, whose resources made this research possible. Finally, we appreciate the constructive feedback from peers and reviewers that helped refine the methodology and presentation of this work

REFERANCE:

*Bossard, L., Guillaumin, M., & Van Gool, L. (2014). Food-101 – Mining Discriminative Components with Random Forests. In European Conference on Computer Vision (ECCV), 446–461.

*Kagaya, H., Aizawa, K., & Ogawa, M. (2014). Food Detection and Recognition Using Convolutional Neural Network. In IEEE International Conference on Multimedia & Expo Workshops, 1–6.

*He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.

*Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. In International Conference on Learning Representations (ICLR).

*Howard, A. G., Zhu, M., Chen, B., et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861.

*Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818–2826.