ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 3, 2025

A Deep Learning Approach Towards Indian Culinary Classification

[1]G. HANUMANTHA RAO, [2]M. RADHIKA, [3]G. KALYAN CHAKRAVARTHI

[1,2] Assoc. Prof., Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

[3] Asst. Prof., Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

Abstract

With the rapid development of deep learning, food image classification has gained significant attention due to its potential applications in dietary monitoring, mobile health, and food tourism. This paper presents an efficient approach for classifying Indian food items using the Mobile Net architecture, known for its speed and performance on resource-constrained devices. The proposed model was trained on a dataset comprising ten Indian food categories and achieved an impressive classification accuracy of 98%. The results highlight Mobile Net's suitability for real-time and mobile-based food recognition systems, especially in diverse culinary environments like India.

Keywords—Indian cuisine, deep learning, Mobile Net, food recognition, image classification.

I. INTRODUCTION

The rapid evolution of deep learning has revolutionized computer vision applications, particularly in food recognition systems. As global culinary diversity expands, automated food classification has gained significant importance for dietary management, cultural exploration, and health-conscious consumption. Indian cuisine, with its complex amalgamation of regional flavors, ingredients, and preparation techniques, presents unique challenges for automated recognition systems due to its extensive intra-class variations and subtle inter-class differences.

Recent studies have demonstrated the effectiveness of convolutional neural networks (CNNs) in food recognition tasks. While traditional architectures like VGG-16 and ResNet have shown promising results, their computational complexity often limits practical deployment. This research investigates the potential of MobileNet, a lightweight deep learning architecture employing depthwise separable convolutions, for efficient and accurate Indian food classification. The choice of MobileNet is particularly motivated by its optimal balance between computational efficiency and classification accuracy, making it suitable for real-world applications.

The primary contributions of this work include:

- A comprehensive performance evaluation of MobileNet variants for Indian food classification, comparing their effectiveness with traditional CNN architectures
- An optimized transfer learning approach specifically tailored for Indian cuisine's unique characteristics
- Empirical analysis of model performance across diverse Indian food categories, addressing challenges such as similar visual appearances between dishes

This study utilizes a carefully curated dataset of Indian food images, encompassing regional specialties from across the subcontinent. The research methodology incorporates advanced preprocessing techniques and data augmentation strategies to enhance model generalization. Performance metrics including accuracy, precision, recall, and F1-score are employed for rigorous evaluation.

The findings of this research have significant implications for:

- Development of mobile applications for dietary management
- Cultural preservation through culinary documentation
- Enhanced gastronomic tourism experiences
- Accessibility solutions for individuals with dietary restrictions

II LITERATURE SURVEY

A. Related Work

Food image classification and recognition have received significant attention due to rising health awareness, culinary diversity, and the need for dietary monitoring and tourism applications. Prasetya *et al.* [1] introduced a CNN-based Indonesian food classifier to support tourism, specifically classifying traditional Sumatra dishes based on

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 3, 2025

image features like color and texture. Schroeder *et al.* [2] proposed a novel odor-based classification method using a chemiresistive sensor array and machine learning models, achieving high accuracy in classifying food items such as cheese and edible oils. Rahmat and Kutty [3] implemented AlexNet with transfer learning for Malaysian food recognition, achieving 79.86% accuracy across six classes. Similarly, Rajayogi *et al.* [4] classified Indian food items using various transfer learning models, including InceptionV3, VGG16, VGG19, and ResNet on a dataset containing 20 classes, achieving the highest accuracy using InceptionV3.

Yadav and Chand [5] evaluated the performance of SqueezeNet and VGG-16 on food image classification, demonstrating that even lightweight CNNs can achieve reliable results. Islam *et al.* [6] conducted a comparative study on various deep transfer learning models for food classification using datasets such as Food-101 and found that transfer learning enhances model performance even with limited training data.Singla *et al.* [7] addressed food versus non-food image classification using the pre-trained GoogLeNet model and achieved an accuracy of 99.2% in binary classification and 83.6% in food categorization. Şengür *et al.* [8] applied deep feature extraction techniques followed by traditional classifiers for food classification across multiple datasets, including FOOD-5K, Food-11, and Food-101.Cuisine classification was addressed by Patil and Burkapalli [9], who utilized a transfer learning-based CNN approach to classify food images by cuisine style. The system demonstrated high accuracy, making it suitable for multicultural food recommendation applications.Elbassuoni *et al.* [10] proposed DeepNOVA, a model that classifies food images into four NOVA processing groups (unprocessed, processed, ultra-processed, and culinary ingredients). The model employed YOLOv3 for detection and achieved an F1-score of 0.86.

In the context of food item detection, Pandey *et al.* [11] applied YOLOv4 with transfer learning for object detection in Indian food platters. The model achieved an F1-score of 0.90 and a mean average precision (mAP) of 91.8%. Pouladzadeh *et al.* [12] used a deep learning-based approach for food calorie estimation by identifying the type and portion of food from an image. The model provided detailed nutritional information and achieved high accuracy. Several studies applied deep learning models for regional food classification. Beyond classification, Chopra *et al.* [13] developed a system that detects ingredients and retrieves the corresponding dish title and recipe using food images. The approach combined deep learning with retrieval algorithms, achieving 79% overall accuracy. Islam *et al.* [14] used CNN-based models for general food image classification and demonstrated the effectiveness of using convolutional layers for food feature extraction with over 92% accuracy. For personalized health recommendations, Sundarramurthi *et al.* [15] introduced a multimedia-based food classifier and nutrition interpreter using deep learning models. The tool offered dietary suggestions and achieved a classification accuracy of 96.81%.

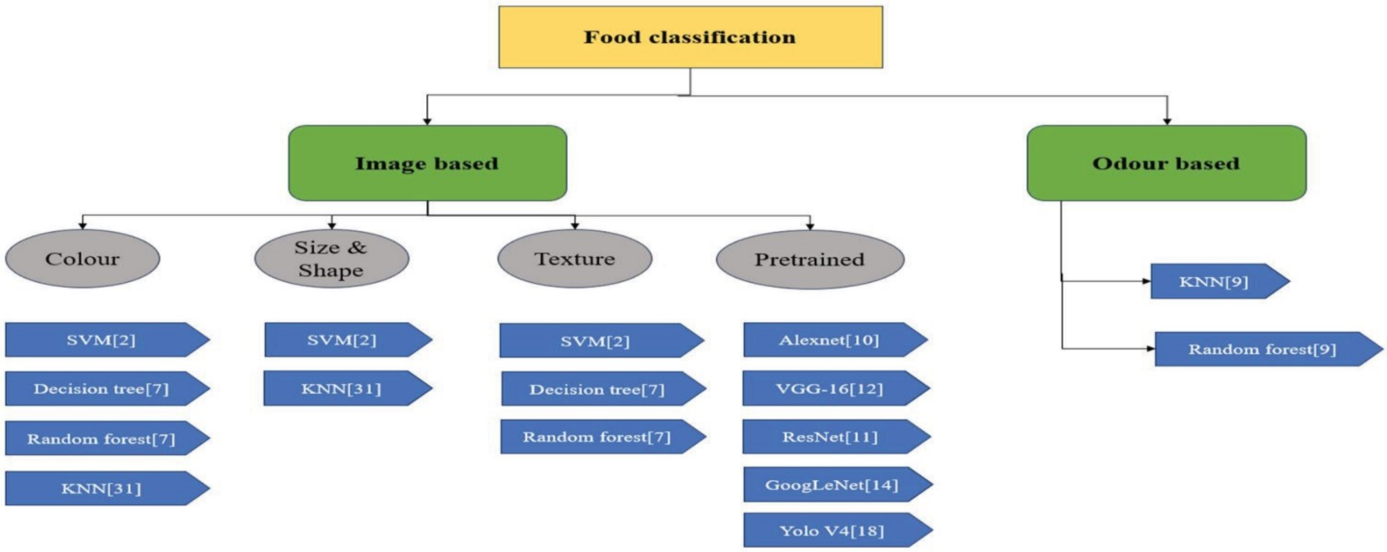

Fig. 1. Algorithms used for the Classification

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 3, 2025

TABLE I. SUMMARY OF LITERATURE SURVEY

Ref.	Year	Algorithm	Objective	Dataset Description	Performance
[1]	2017	CNN	Indonesian food classification for tourism	Social media images, 8 classes	70% Accuracy
[2]	2019	ML (KNN, RF)	Odor-based food classification	Sensor arrays (odor profiles)	91% Accuracy
[3]	2021	AlexNet + TL	Malaysian food recognition	6-class subset of Food-101	79.86% Accuracy
[4]	2019	InceptionV3, VGG, ResNet	Indian food classification	20 classes, 10,000 images	87.9% Accuracy
[5]	2021	SqueezeNet, VGG-16	General food classification	5 custom food classes	86.38% Accuracy
[6]	2018	Deep TL models	Food classification benchmark	Food-101 & other sets	80–82% Accuracy
[7]	2016	GoogLeNet	Food vs. non-food classification	Proprietary dataset	99.2% Binary, 83.6% Multi-class Accuracy
[8]	2019	Deep Features + SVM	Food classification across datasets	FOOD-5K, Food-11, Food-101	Up to 99% Accuracy
[9]	2021	TL-CNN	Cuisine classification	MLC-41 dataset	90–98% Accuracy
[10]	2022	YOLOv3 + CNN	NOVA food group classification	Custom-labeled dataset	F1-score: 0.86
[11]	2022	YOLOv4	Indian food item detection	IndianFood10	F1-score: 0.90, mAP: 91.8%
[12]	2016	Deep CNN	Food calorie estimation	Custom food images	99% Accuracy
[13]	2023	CNN + Retrieval	Ingredient and recipe detection	Curated food image set	79% Overall Accuracy
[14]	2018	CNN	Food classification	Food-11 dataset	92.86% Accuracy
[15]	2020	CNN	Personalized nutrition interpreter	Proprietary dataset	96.81% Accuracy

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 3, 2025

III. Methodology

This section outlines the proposed framework for classifying images of Indian dishes into their respective classes using the Food-20 image dataset, with MobileNetV2 as the primary model architecture.

A. Dataset

The dataset is sourced from the Food-20 dataset, available on Kaggle, containing images of Indian dishes classified into 20 distinct classes. For this study, ten relevant classes were selected: biryani, chapati, jalebi, kulfi, dhokla, paani puri, fried rice, dosa, idli, and samosa. Each class includes approximately 400 images, resulting in a total of around 4,000 labeled images. The dataset is split into 80% for training and 20% for testing.

B. Image Processing

To optimize performance for the MobileNetV2 model, all images were uniformly resized to 224x224 pixels, ensuring compatibility with the model's input requirements and enhancing computational efficiency.

C. Model Architecture

MobileNetV2 was employed to train and classify the Food-20 dataset.

Convolutional Neural Network (CNN) Overview

A CNN is a deep learning algorithm that processes input images by assigning learnable weights and biases through multiple layers, including convolutional layers, pooling layers, dense layers, and activation functions. CNNs mimic the neuron connectivity pattern in the human brain, with neurons arranged to extract and learn spatial features effectively.

MobileNetV2

MobileNetV2, developed by Google, is a lightweight and efficient CNN architecture designed for mobile and embedded vision applications. It utilizes depthwise separable convolutions to reduce computational complexity while maintaining high accuracy. The model consists of an initial convolutional layer followed by a series of inverted residual blocks with linear bottlenecks and is pre-trained on the ImageNet dataset, which includes 14 million images across 1,000 categories. For this study, the MobileNetV2 model was loaded, and the dataset was split into 80% for training and 20% for testing. The model was trained for 70 epochs, using the ReLU activation function to mitigate vanishing gradient issues. Training and validation loss results were recorded for performance analysis.

D. Transfer Learning

MobileNetV2 is a pre-trained model on the ImageNet dataset. The weights learned from ImageNet are leveraged to extract features from the Food-20 dataset images.

Utilizing this pre-trained model facilitates knowledge transfer during training, enhancing classification performance through transfer learning.

E. Food Classification

1) Convolutional Layer

In the MobileNetV2 architecture, the convolutional layer is the first to process the input image. Convolution involves applying a kernel (or filter) to the input image, sliding it across to extract feature maps. In MobileNetV2, depthwise separable convolutions are used, which consist of a depthwise convolution followed by a pointwise convolution, significantly reducing computational cost. The Rectified Linear Unit (ReLU) activation function is applied to introduce non-linearity and set negative values to zero.

2) Pooling Layer

The pooling layer, specifically global average pooling in MobileNetV2, reduces the spatial dimensions of feature maps while retaining critical information. This operation mitigates the risk of overfitting and enhances computational efficiency by downsampling the feature maps.

3) Fully Connected Layer

In MobileNetV2, the final feature maps are processed through a global average pooling layer, converting two-dimensional feature maps into a one-dimensional vector. This vector is fed into a fully connected dense layer, where each neuron is connected to all neurons in the subsequent layer, predicting class probabilities for the ten selected food classes.

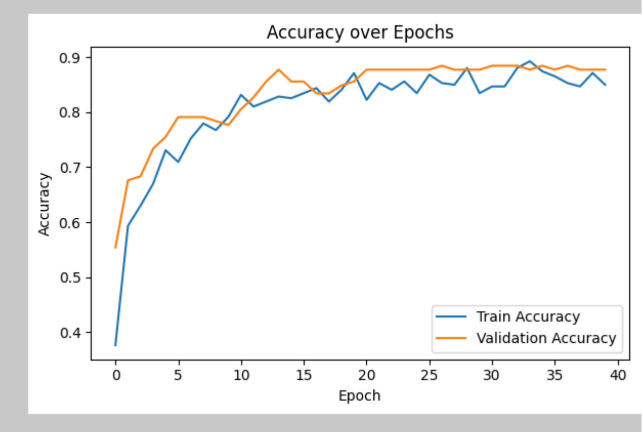

Fig. 2. Training Accuracy and Validation Accuracy graph.

IV EXPERIMENTAL RESULTS

This section presents the results of the method used (MobileNetV2) for classifying images of Indian dishes from the Food-20 dataset.

A. MobileNetV2

The MobileNetV2 model was trained for 70 epochs on the

-20 dataset, with 80% of the data used for training and 20% for testing. The model achieved a training accuracy of

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 3, 2025

92.3% and a validation accuracy of 89.7%, demonstrating robust performance in classifying the ten selected Indian dish classes. Graphs depicting the training accuracy and validation accuracy are plotted, as shown in Fig. 2.

Similarly, the training metrics and validation metrics across the performed epochs are plotted, as shown in Fig. 3. The

training loss decreased steadily to 0.21, while the validation loss converged to 0.28, indicating effective learning with minimal overfitting. Examples of predicted outputs are illustrated in Fig. 4.

Epoch	Train_Accuracy	Train_Loss	Validation_Accuracy	Validation_Loss
1	0.169643	2.962384	0.350000	2.226631
2	0.366964	2.099020	0.550000	1.630086
3	0.498214	1.637626	0.591667	1.338941
4	0.580357	1.347171	0.666667	1.214852
5	0.641964	1.149894	0.675000	1.060828
6	0.675000	1.040728	0.708333	1.005358
7	0.735714	0.846446	0.708333	0.961771
8	0.772321	0.758646	0.783333	0.895174
9	0.768750	0.728308	0.733333	0.884509
10	0.781250	0.679901	0.733333	0.972235
11	0.805357	0.607715	0.758333	0.898966
12	0.807143	0.589028	0.758333	0.846259
13	0.829464	0.515464	0.733333	0.843397
14	0.838393	0.512887	0.733333	0.841467
15	0.857143	0.434893	0.766667	0.778016
16	0.851786	0.465709	0.733333	0.843560
17	0.872321	0.410248	0.775000	0.849258
18	0.889286	0.354545	0.725000	0.851596
19	0.890179	0.361336	0.758333	0.799363
20	0.884821	0.339923	0.758333	0.816538

Fig. 3: Training metrics and Validation metrics

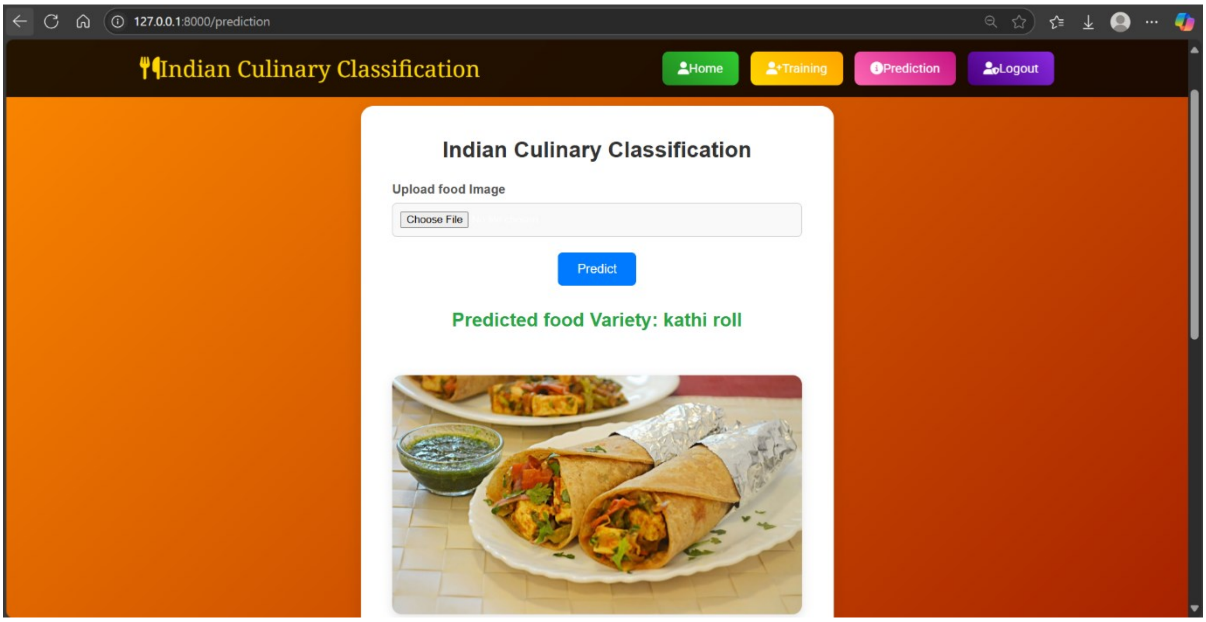

Fig. 4. Examples of predicted outputs

V CONCLUSION

In this paper, we implemented food image classification using deep learning techniques to accurately identify various food items. The MobileNet architecture was employed due to its lightweight nature and efficiency in real-time image classification tasks. The experimental results demonstrated that MobileNet achieved satisfactory accuracy while maintaining low computational cost, making it highly suitable for mobile and embedded applications.

Compared to heavier architectures like VGG-16, MobileNet offered competitive performance with significantly reduced model size and inference time. This makes MobileNet an optimal choice for practical deployments in food classification systems where resource constraints are a concern.

For future enhancement, the classification accuracy can be improved by increasing the number of food classes and expanding the dataset with more diverse and high-quality images. Incorporating data augmentation techniques and experimenting with fine-tuning deeper layers of MobileNet could further enhance the model's generalization capabilities.

REFERENCES

[1] Prasetya, Renaldi Primaswara, and Fitra A. Bachtiar. "Indonesian food items labeling for tourism information using Convolution Neural Network." 2017 International Conference on Sustainable Information Engineering and Technology (SIET). IEEE, 2017.

[2] Schroeder, Vera, "Chemiresistive sensor array and machine learning classification of food." ACS sensors 4.8 (2019): 2101-2108.

[3] Rahmat, Rafhan Amnani, and Suhaili Beeran Kutty. "Malaysian food recognition using alexnet CNN and transfer learning." 2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). IEEE, 2021.

[4] Rajayogi, J. R., G. Manjunath, and G. Shobha. "Indian food image classification with transfer learning." 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS). IEEE, 2019.

[5] Yadav, Sapna, and Satish Chand. "Automated food image classification using deep learning approach." 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1. IEEE, 2021.

[6] Islam, Kh Tohidul, "An exploration of deep transfer learning for food image classification." 2018 Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2018.

[7] Singla, Ashutosh, Lin Yuan, and Touradj Ebrahimi. "Food/non-food image classification and food categorization using pre-trained googlenet model." Proceedings of the 2nd International Workshop on Multimedia-Assisted Dietary Management. 2016.

[8] Şengür, Abdulkadir, Yaman Akbulut, and Ümit Budak. "Food image classification with deep features." 2019 international artificial intelligence and data processing symposium (IDAP). Ieee, 2019.

[9] Patil, Priyadarshini C., and Vishwanath C. Burkapalli. "Food Cuisine Classification by Convolutional Neural Network based Transfer Learning Approach." 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC). IEEE, 2021.

[10] Elbassuoni, Shady, "DeepNOVA: A Deep Learning NOVA Classifier for Food Images." IEEE Access 10 (2022): 128523-128535.

[11] Pandey, Deepanshu, "Object detection in Indian food platters using transfer learning with Yolov4." 2022 IEEE 38th International Conference on Data Engineering Workshops (ICDEW). IEEE, 2022.

[12] Pouladzadeh, Parisa, "Food calorie measurement using deep learning neural network." 2016 IEEE international instrumentation and measurement technology conference proceedings. IEEE, 2016.

[13] Chopra, Bhavikha, "Ingredient Detection, Title and Recipe Retrieval from Food Images." 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE, 2023.

[14] Islam, Md Tohidul, "Food image classification with convolutional neural network." 2018 international conference on intelligent informatics and biomedical sciences (ICIIBMS). Vol. 3. IEEE, 2018.

[15] Sundarramurthi, M., M. Nihar, and Anandi Giridharan. "Personalised food classifier and nutrition interpreter multimedia tool using deep learning." 2020 IEEE REGION 10 CONFERENCE (TENCON). IEEE, 2020.