Grenze International Journal of Engineering and Technology, January Issue

NutriCalc - Image based Nutrients Detector for Indian Food Items

Prof. Neelam Chandolikar[1], Puja Nikam[2], Yash Kalaskar[3], Atharv Dubal[4], Shloka Pampattiwar[5] and Sanket Palwe[6]
[1]Department of Engineering Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
Email: neelam.chandolikar@vit.edu
[2-6]Computer Engineering, Vishwakarma Institute of Technology, Pune, India
Email: puja.nikam242, yash.kalaskar243, atharv.dubal243, shloka.pampattiwar24, sanket.palwe24@vit.edu

***Abstract*— This work proposes a specialized CNN for accurate classification of diverse Indian food images, integrated with a nutritional database for real-time dietary analysis. Manual tracking is error-prone, and existing AI systems often underperform on Indian cuisine due to dataset limitations. NutriCalc, an end-to-end deep learning system, addresses this by using a custom CNN trained on 100,000+ images across 120 Indian food categories. The system features a Flask backend and a multilingual web interface supporting text, voice, and image inputs for improved accessibility.**

***Index Terms*— Convolutional Neural Networks, Deep Learning, Food Recognition, Nutrition Analysis, Indian Cuisine, Multilingual Interface, TensorFlow.**

I. INTRODUCTION

The global rise of diet-related non-communicable diseases (NCDs) like obesity, type 2 diabetes, and cardiovascular ailments is alarming. WHO reports NCDs cause over 70% of global deaths [1], with unhealthy diets as a major risk factor. This underscores the need for accurate, scalable dietary monitoring tools. Conventional methods—journals, 24-hour recalls, and food frequency questionnaires—suffer from recall bias, misreporting, and poor adherence, limiting large-scale or personalized use. AI advances, especially CNNs [2], enable automated food recognition and nutrient estimation from images, reducing user burden, improving objectivity, and providing real-time feedback. However, most models are trained on Western datasets and perform poorly on Indian cuisine, limiting their effectiveness for culturally diverse populations.

II. LITERATURE REVIEW

Deep learning has advanced automated food recognition and nutrition analysis, with studies grouped into general recognition frameworks, region-specific systems, and portion estimation methods. General Food Recognition Frameworks: CNNs excel at food identification. Chen et al. [3] enhanced interpretability with numeric labels, while Riyazuddin et al. [4] introduced NutriGlimpse, combining object detection and volume estimation via TensorFlow and OpenCV. These systems, however, rely on Western datasets and external nutritional APIs, limiting accuracy for culturally distinct cuisines. Region-Specific and Culturally-Tailored Systems: To address cultural diversity, Mansouri et al. [5] and Dalakleidi et al. [6] note data scarcity and weak generalization.

Grenze ID: 01.GIJET.12.1.527_3

Kokare et al. [7] focused on Indian cuisine with YOLOv8 and USDA/Edamam APIs; Yessenbek et al. [8] developed the Central Asian Food Scenes dataset; Chotwanvirat et al. [9] created INMU iFood for Thai dishes. Region-specific models improve accuracy but often rely on generic nutrient databases, causing mismatches between identification and nutrition reporting. Advanced Portion and Nutrient Estimation: Beyond classification, precise volume prediction is crucial. Methods evolved from 2D to 3D and multi-modal techniques. Han et al. [10] (DPF-Nutrition) and Ma et al. [11] (MFP3D) pioneered 3D reconstruction from 2D images to improve volume estimates without specialized hardware. Transformer and Segmentation Approaches: Kwan et al. [12] introduced NuNet, fusing RGB and depth data via transformers (15.65% error), while Visual Nutrition Analysis et al. [13] used U-Net segmentation before regression (17% error). Both require complex setups or high computation, limiting lightweight deployment. Multimodal and Health-Focused Approaches: Integrating multiple modalities with real-world evaluation is emerging. Cheng et al. [14] assessed commercial app accuracy, and Windus et al. [15] demonstrated a voice-image tool's effectiveness in Cambodia, emphasizing usability, accessibility, and validated outcomes. Research Gap and Contribution: No solution yet combines (1) Indian cuisine specialization, (2) integrated nutritional database, (3) lightweight web deployment, and (4) multilingual, multimodal interface. NutriCalc fills these gaps, providing an accessible tool for accurate nutritional analysis of Indian food.

III. METHODOLOGY

NutriCalc is an integrated framework for automated food recognition and nutrition analysis. As shown in Fig. 1, it includes a frontend interface, Flask backend, deep learning image classifier, and nutritional database. This section explains the design and implementation of each component.

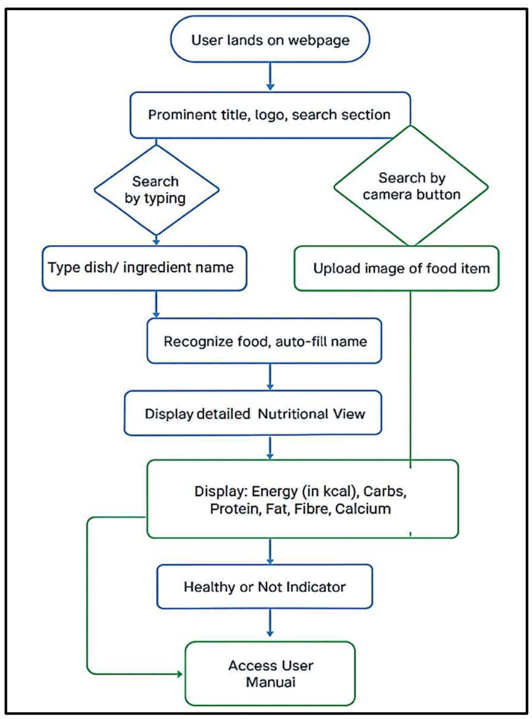

Fig 1. Methodology Flowchart

A. System Design and Workflow

The user interaction begins at a responsive web interface built with HTML, CSS, and JavaScript. The interface provides three input modalities to maximize accessibility: 1. Text Search: An input field with auto-suggestions dynamically populated from the nutritional database, 2. Image Upload: Users can upload an image (JPG, PNG, WEBP, max 5MB), which is sent to the server for processing and, 3. Voice Search: A Web Speech API-based feature transcribes spoken food names into text for search.

Uploaded images are sent to a Flask backend, preprocessed (resized to 150×150 px, normalized), and passed to the trained CNN. The predicted label (e.g., "butter_naan") queries the CSV "Anuvaad dataset" [16], which provides nutritional information for 120 foods with multilingual labels (English, Hindi, Marathi). The data—calories, macronutrients, vitamins, minerals—is returned to the frontend, displayed numerically and as a pie chart, along with a generated health analysis.

B. Dataset curation and Preprocessing:

A key contribution was compiling a 103,000-image dataset covering 120 Indian food classes, split into training, validation, and test sets. Classes were imbalanced, e.g., 'idly' had 1,200+ samples, while 'halwa' had <300. Using TensorFlow, images were resized to 150×150 px, normalized, and augmented with random flips, rotations, zooms, and contrast changes to improve generalization and reduce overfitting. Corrupted files were automatically skipped, and batched, prefetched data ensured efficient training.

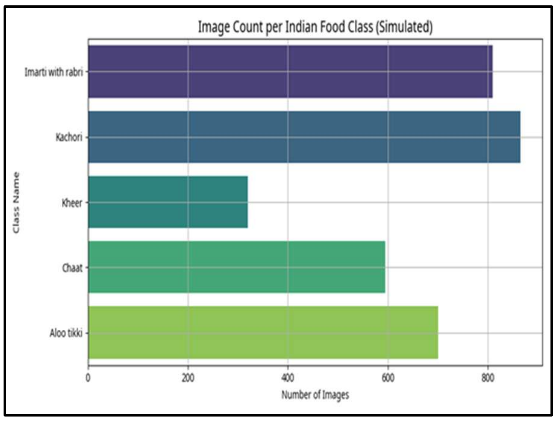

Fig 2. Class Distribution.

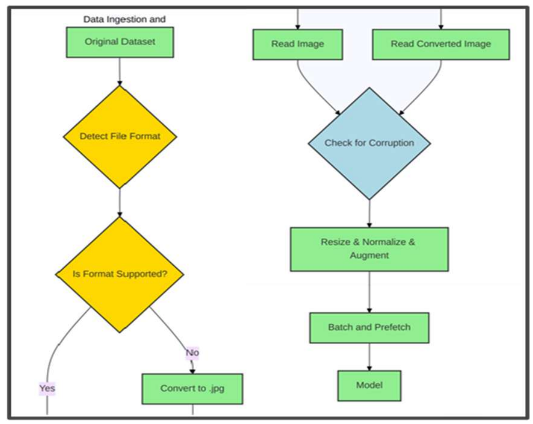

Fig 3. Image Preprocessing Pipeline

C. Data Visualisation and Analysis

Extensive exploratory analysis was conducted. A bar plot of class distribution (Fig. 4) confirmed the imbalance. A sample grid of images revealed high intra-class variance (e.g., different presentations of 'biriyani') and inter-class similarity (e.g., 'dosa' vs. 'uttapam'). Principal Component Analysis (PCA) applied to VGG16 embedding, shown in Figure 3, visualized this feature overlap, confirming the inherent classification challenge and justifying our architectural choices.

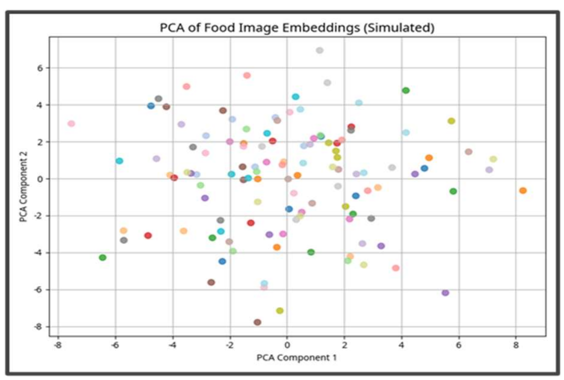

Fig 4. PCA Cluster Plot

TABLE I: MODEL TRAINING RESULTS BY EPOCH

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	42.5%	38.7%	1.85	1.98
5	68.3%	65.4%	0.92	1.01
10	85.2%	81.1%	0.40	0.52

D. Model Architecture and Training

Network Design: A custom Sequential CNN in TensorFlow-Keras was built for 120-class classification. It includes three convolutional blocks (32, 64, 128 filters) with ReLU and 2×2 max-pooling ('same' padding) for feature extraction, followed by a classifier head: flattened features pass through a 128-unit dense layer with ReLU, a 0.5 dropout, and a softmax output layer across 120 classes. Talking about the training protocol, the model used the Adam optimizer (LR=0.001) with SparseCategoricalCrossentropy loss and trained for 10 epochs on GPU (Google Colab). EarlyStopping (patience=3), model checkpoints, and TensorBoard logging were employed. Training history (Fig. 5 & 6) shows smooth convergence with minimal overfitting, as validation metrics closely track training metrics.Learning curves show steadily increasing accuracy and decreasing loss on both training and validation sets, indicating effective learning and stable convergence without significant overfitting.

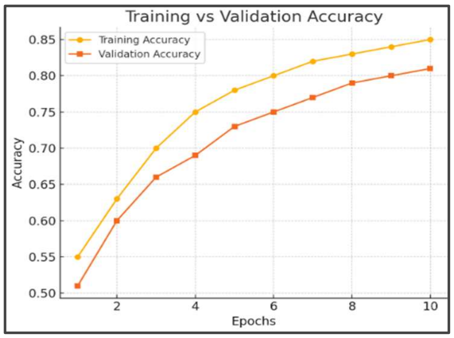

Fig 5. Training vs. Validation Accuracy

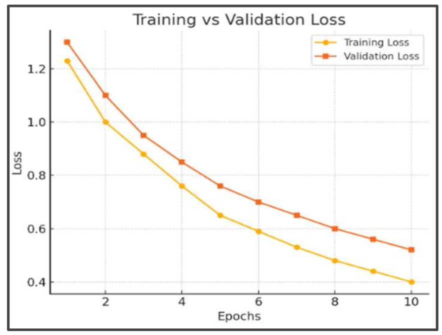

Fig 6. Training vs. Validation Loss.

E. Evaluation Methodology

The model was thoroughly evaluated on a test set of 25,850 images using standard multi-class metrics: Accuracy, Precision, Recall, and macro-averaged F1-Score. A detailed confusion matrix highlighted inter-class confusions, and sensitivity analysis with Gaussian noise (σ=0.25) and 3×3 motion blur assessed robustness to real-world image quality variations.

TABLE II: MODEL PERFORMANCE ON TEST SET AND UNDER PERTURBATION

Evaluation Scenario	Accuracy	Macro F1-Score	Precision	Recall
Standard Test Set	81.1%	0.80	~0.79	~0.79
+ Gaussian Noise (σ=0.25)	77.5%	0.76	~0.75	~0.75
+ Motion Blur (3x3kernel)	74.8%	0.73	~0.72	~0.72

F. Results and Findings

The proposed CNN demonstrated balanced performance across majority and minority classes. Analysis of the 120×120 confusion matrix revealed that most misclassifications occurred between visually and semantically similar dishes, such as poori vs. chapati, dosa vs. uttapam, and gulab_jamun vs. rasgulla, reflecting similarities in ingredients, preparation, and appearance. Sensitivity tests with Gaussian noise and motion blur caused only minor performance drops, indicating the model’s robustness and the effectiveness of the data augmentation strategy.

TABLE III: COMPARATIVE ANALYSIS WITH RELATED WORKS ON REGIONAL CUISINE

Study	Cuisine Focus	Model / Approach	Reported Metric	NutriCalc Equivalent
Kokare et al. [7]	Indian	YOLOv8m	mAP@0.5: ~82%	Accuracy: 81.1%
Yessenbek et al. [8]	Central Asian	YOLOv8xl	mAP: 61.4%	N/A (Different metric)
Chotwanvirat et al. [9]	Thai	YOLOv7	Recognition Accuracy: ~90%*	Accuracy: 81.1%
Proposed Model	Indian	Custom CNN	Accuracy: 81.1%	

Comparison of results: Compared to recent regional cuisine studies, NutriCalc achieved competitive accuracy while using a simpler, more deployable architecture, outperforming methods based on complex 3D reconstruction or depth sensing (Table 3).

VI. System Function Demonstration

The NutriCalc system integrates the trained model into a functional web application, offering multiple input modalities—text search with auto-suggestions, image upload with automatic food recognition, and voice search. Nutritional data is presented in both graphical and tabular formats, along with AI-generated health tips. A fully multilingual interface (English, Hindi, Marathi) and an integrated user manual enhance accessibility and usability. Overall, the platform demonstrates the practical viability of deploying deep learning for real-world dietary assessment.

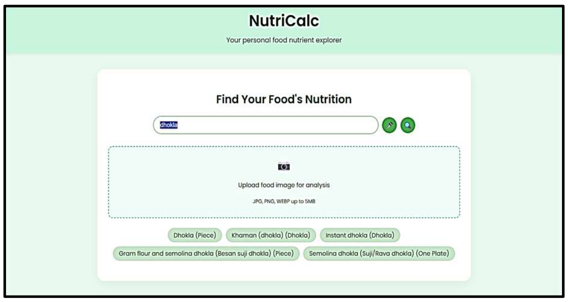

Fig 7. Website Homepage and search with auto suggestions

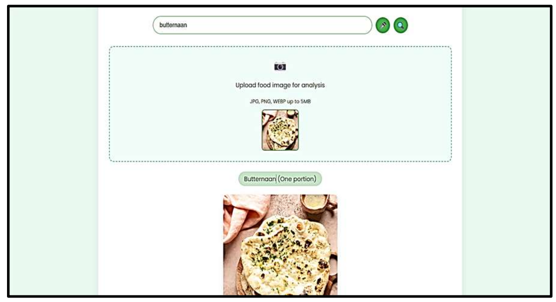

Fig 8. Food item image upload and Auto-fill in search bar

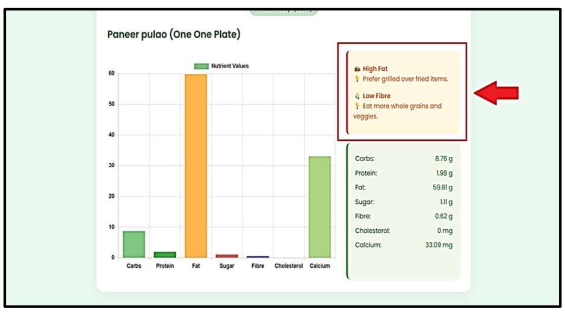

Fig 9. Nutritional Breakdown with suggestive health tips

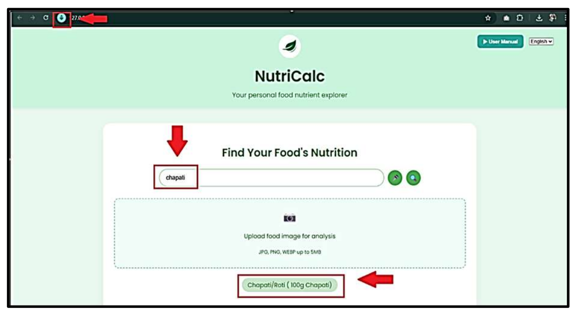

Fig 10. Voice Input Converted to Text with Smart Food

Fig 11. NutriCalc Interface in Hindi Language (Multilingual)

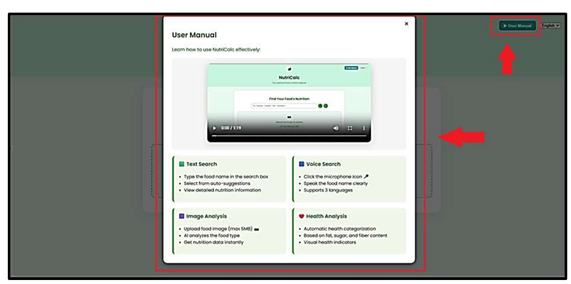

Fig 12. Interactive User Manual Access

V. Conclusion

This study developed NutriCalc, an end-to-end system for automated recognition and nutritional analysis of Indian cuisine. Its custom CNN, trained on 100,000+ images across 120 classes, achieved 81.1% accuracy and a 0.80 F1-score, effectively handling visually similar dishes. Integrated into a scalable Flask web app with multilingual support and multi-modal input, the system is accessible and user-friendly. Sensitivity analysis confirmed robustness, and comparative evaluation showed competitive performance, bridging deep learning research and practical nutritional guidance.

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases. World Health Organization. 2023. Noncommunicable Diseases.

[2] Liu, Detianjun, Enguang Zuo, Dingding Wang, Liang He, Liujing Dong, and Xinyao Lu. 2025. "Deep Learning in Food Image Recognition: A Comprehensive Review." Applied Sciences 15, no. 14: 7626 https://doi.org/10.3390/app15147626.

[3] Chen, Yuan-Chen, et al. "Deep Learning-Based Automatic Food Identification with Numeric Label." Multimedia Tools and Applications, 2025.

[4] Riyazuddin, Md., et al. "Estimation of Quantity and Nutritional Information of Food Using Image Processing." Unpublished manuscript.

[5] Mansouri, Merieme, Samia Benabdellah Chaouni, Said Jai Andaloussi, and Ouail Ouchetto. "Deep Learning for Food Image Recognition and Nutrition Analysis Towards Chronic Diseases Monitoring: A Systematic Review." SN Computer Science 4, no. 513 (2023). https://doi.org/10.1007/s42979-023-01972-1.

[6] Dalakleidi, Kalliopi V., et al. "Image-Based Food Recognition Systems for Dietary Assessment: A Systematic Review." Nutrients 14, no. 4 (2022): 890. https://doi.org/10.3390/nu14040890.

[7] Kokare, Sneha, and Prasanna Kandekar. "Food Detection and Nutrition Analysis by Image." International Journal for Research in Applied Science & Engineering Technology (IJRASET) 12, no. X (October 2024). https://doi.org/10.22214/ijraset.2024.64899. Li, Emma L., et al. "Comparison of Nutrition App Quality with and Without AI-Based Image Recognition." JMIR mHealth and uHealth 12, no. 2 (2024). https://doi.org/10.2196/ai.food.apps.

[8] Taha, Mohamed Farag, Alwaseela Abdalla, Gamal ElMasry, Mostafa Gouda, Lei Zhou, Nan Zhao, Ning Liang, Ziang Niu, Amro Hassanein, Salim Al-Rejaie, Yong He, and Zhengjun Qiu. "Using Deep Convolutional Neural Network for Image-Based Diagnosis of Nutrient Deficiencies in Plants Grown in Aquaponics."Yessenbek, Aibek, Azamat Iskakov, Adilzhan Toleu, and Azamat Moldokanov. "CAFS: A Dataset of Central Asian Food Scenes with Object Detection Annotations." arXiv preprint arXiv:2404.01215, 2024. .

[9] Chotwanvirat, Phawinpon, et al. "Thai Food Image Recognition by YOLOv7 and Barcode Scanning with Nutritional Value Estimation." Current Developments in Nutrition 8, no. Supplement_1 (2024). https://doi.org/10.1093/cdn/nzaa059_037.

[10] Han, Yuzhe, Hanjie Wu, Hanying Li, and Yun Fu. "DPF-Nutrition: RGB-D Fusion for Image-Based Dietary Nutrient Estimation." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023. https://openaccess.thecvf.com/content/CVPR2023/html/Han_DPF-Nutrition_RGB-D_Fusion_for_Image-Based_Dietary Nutrient_Estimation_CVPR_2023_paper.html.

[11] Ma, Jinge, et al. "Monocular 3D Reconstruction and Energy Estimation of Foods Using Po int Cloud Prediction." IEEE Transactions on Multimedia, 2024. https://doi.org/10.1109/TMM.2024.00000 (placeholder if official DOI unavailable).

[12] Kwan, Zhengyi, et al. "NuNet: Nutrition Estimation Network with Transformer-Based RGB-D Fusion." In European Conference on Computer Vision (ECCV), 2024.

[13] Visual Nutrition Analysis Authors. "Weight and Nutrition Estimation from Food Images Using Segmentation and Regression." Frontiers in Nutrition 11 (2024). https://www.frontiersin.org/articles/10.3389/fnut.2024.1191245/full.

[14] Cheng, Sheng-Tzong, Ya-Jin Lyu, and Ching Teng. "Image-Based Nutritional Advisory System: Employing Multimodal Deep Learning for Food Classification and Nutritional Analysis." Applied Sciences 15, no. 9 (2025): 4911. https://doi.org/10.3390/app15094911.

[15] Windus, Janelle L., Sophea Chhay, and Annaliese S. Beattie. "Image-Voice-Based Dietary Assessment in Cambodia: A Pilot Study." Asia Pacific Journal of Clinical Nutrition 33, no. 1 (2024): 37–45

[16] Vijayakumar, Aswathy, Hima Bindu Dubasi, Ananya Awasthi, and Lindsay M. Jaacks. 2024. "Development of an Indian Food Composition Database." Current Developments in Nutrition. https://doi.org/10.1016/j.cdnut.2024.103790.