Multimodal AI Framework for Personalized and Health-Aware Cooking Recommendations

Soundarya M
Department of IT
Velammal College of Engineering and Technology
Madurai, India
soundaryajeyaraj5@gmail.com

P. Pooja Shree
Department of IT
Velammal College of Engineering and Technology
Madurai, India
sukija2004@gmail.com

Swarna Suganthi S
Department of IT
Velammal College of Engineering and Technology
Madurai, India
blessyswarna11@gmail.com

***Abstract*— In the current era of growing interest in health-conscious eating and personalized nutrition, traditional recipe recommendation systems often fail to account for diverse user needs, ingredient availability, and practical cooking constraints. The Multimodal Artificial Intelligence (AI) Framework proposed in this study integrates and analyzes multiple data modalities—textual dietary preferences, food images, and cooking videos—to generate personalized and health-aware cooking recommendations. The framework considers individual health profiles, ingredients detected from visual inputs, and user-specific cooking skill levels inferred from video analysis to tailor recipe suggestions effectively. By leveraging multimodal deep learning algorithms, the system delivers contextually aware, precise, and adaptive recommendations. Experimental evaluations on benchmark and hybrid datasets demonstrate its effectiveness in enhancing recommendation relevance, supporting dietary compliance, and improving overall user satisfaction. These results indicate strong potential for real-world deployment in intelligent culinary assistants, personalized diet planning platforms, and smart health applications.**

Keywords—Personalized Recipe Recommendation, Multimodal Deep Learning, Ingredient Recognition, Cooking Skill Estimation

I. Introduction

In recent years, there has been a growing interest in personalized and intelligent food recommendation systems, fueled by increasing awareness around health, diverse dietary preferences, and modern lifestyle demands. Despite this, most existing recipe recommendation platforms continue to rely on basic keyword searches and filtering options. These conventional methods often fall short when it comes to understanding the rich context behind user needs—such as dietary restrictions, available ingredients, and individual cooking skill levels—making their suggestions less relevant or practical for real-world use.

The need for more sophisticated and context-aware systems has become especially evident with the rise of health tracking apps, smart kitchens, and the increased accessibility of AI technologies. Users today expect digital systems to not only reflect their preferences but also adapt to their limitations, such as limited cooking experience or constrained ingredient availability. Moreover, the increasing use of visual and video content in food platforms presents an opportunity to leverage untapped data sources to enhance personalization. However, very few systems integrate these multiple modalities in a cohesive and intelligent manner. To overcome these limitations, this study introduces a Multimodal Artificial Intelligence (AI) Framework designed to deliver personalized and health-aware cooking recommendations. The framework processes and integrates data from multiple sources: **textual** inputs like dietary preferences and recipe descriptions, images for detecting ingredients from food or refrigerator photos, and video content to estimate recipe complexity and user cooking skill.

By combining these diverse inputs, the system gains a deeper understanding of both user profiles and recipe characteristics, enabling it to generate more accurate and user-tailored suggestions. This framework is built upon recent advances in multimodal deep learning, incorporating components such as transformer-based architectures, hypergraph modeling, and temporal video analysis. Together, these technologies allow the system to adapt its recommendations to each user's context, preferences, and capabilities.

This paper presents the full development of the proposed framework, including system architecture, algorithm selection, and experimental evaluation using publicly available and hybrid datasets. Results show that the system consistently outperforms traditional approaches, particularly in terms of personalization, contextual relevance, and adaptability. The findings highlight the framework's potential to serve as the backbone for next-generation intelligent cooking assistants and user-centric recipe platforms.

II. Literature Survey

Freyne et al developed the work entitled "Intelligent Food Planning: Personalized Recipe Recommendation".[1] Their work tackled two main challenges: how to effectively collect user input and how to understand the relationship between recipes and the ingredients they contain. Using data from 183 users and over 8,700 ratings, they tested content-based, collaborative filtering, and hybrid algorithms. Among these, the hybrid approach which broke down recipes into individual ingredients showed better accuracy in predicting food preferences. The system relied heavily on explicit user ratings.

Razali et al. [2] proposed "Recipe Recommendation based on Food Ingredients Recognition using Deep Learning ", a deep learning-based recipe recommendation system that identifies food ingredients from images using the YOLOv8 model. The system was trained on a dataset of over 2,000 images categorized into rice, apple, and chicken, achieving 98% accuracy in ingredient detection. It aims to assist users in selecting suitable recipes based on available ingredients through a web-based interface. The study highlights the practical use of computer vision for improving meal planning efficiency.

Rostami et al. introduced "Food Recommendation as Language Processing (F-RLP): A Personalized and Contextual Paradigm", a novel framework that adapts large language models (LLMs) for personalized food recommendations.[3] Unlike general-purpose LLM-based recommenders, F-RLP is tailored to food-specific contexts by integrating structured personal, contextual, and numerical data. It consists of two key components: a data preprocessing module that transforms dietary and contextual inputs into a vector, and an LLM fine-tuned to interpret these vectors for generating recommendations. F-RLP incorporates context injection and a novel counterfactual training method.

Simeoni et al. proposed "DIETOS: a recommender system for adaptive diet monitoring and personalized food suggestion," a web-based food recommender system aimed at improving the quality of life for both healthy individuals and those with chronic diet-related conditions such as diabetes, CKD, and hypertension.[4] The system creates personalized diet plans using dynamic, real-time questionnaires validated by medical experts. Unlike conventional systems, DIETOS integrates clinical profiling and suggests region-specific foods based on certified health guidelines. It emphasizes medical accuracy over behavioral or navigational data.

Yera et al. developed "A Systematic Review on Food Recommender Systems for Diabetic Patients" using the PRISMA framework.[5] They analyzed approaches like content-based filtering, collaborative filtering, and hybrid systems integrated with nutritional guidelines. The review identified gaps in datasets and evaluation methods, often relying on simulated users rather than real-world data. Results showed that while current systems offer potential, most lack personalized disease-specific recommendations and clinical validation.

Jieyu Zhang et al proposed" A unified approach to designing sequence-based personalized food recommendation systems: tackling dynamic user behaviors", a sequence-based food recommender that models a user's recipe history as a timeline using a traditional LSTM network.[6] By learning both long- and short-term user behavior patterns, the system adapts to changing food preferences over time an advantage over static collaborative or content-based methods. Evaluated on a public dataset, the model outperforms standard baselines in accuracy metrics, demonstrating that incorporating sequential behavior improves the relevance and diversity of recipe recommendations.

Park et al. proposed "RecipeBowl: A Cooking Recommender for Ingredients and Recipes Using Set Transformer", a cooking recommender system that suggests additional ingredients and recipes based on a given set of inputs.[7] It uses a Set Transformer to encode ingredient sets and predicts missing items through a recipe completion task. Trained on the Recipe1M dataset, the model learns to generate context-aware recommendations by mapping ingredient sets to embedding spaces. Experimental results demonstrate that RecipeBowl effectively captures ingredient relationships and provides relevant suggestions.

Islam et al. proposed "Human Behavior-based Personalized Meal Recommendation and Menu Planning Social System", an EEG-based meal recommender that considers a user's emotional state, nutrition, and calorie needs. [8] Using brain signals from a wearable device, the system predicts food preferences and plans personalized menus using TOPSIS and bin-packing algorithms. Unlike traditional systems, it supports users who can't express preferences, like patients with ALS. The system integrates affective computing to enhance meal relevance, making food choices more personalized and empathetic.

Bilgin et al. proposed "A Linear General Type-2 Fuzzy Logic Based Computing With Words Approach for Realising an Ambient Intelligent Platform for Cooking Recipes Recommendation", an ambient intelligent platform for personalized cooking recipe recommendation using a Linear General Type-2 (LGT2) fuzzy logic-based Computing With Words (CWW) framework. [9] The system enables natural and transparent human-machine interaction by considering user-specific factors like mood, appetite, and available time. This approach handles uncertainty in user preferences more effectively.

Padmavathi et al. developed "RecipeMate: A Food Media Recommendation System Based on Regional Raw Ingredients", a food recommendation system that suggests recipes based on user-input ingredients, preferred cuisine, and allergies.[10] The system utilizes natural language processing and machine learning techniques, including cosine similarity, to analyze a cleaned dataset of 28,000 recipes. By filtering out unwanted ingredients, it enhances personalization and user safety. The web-based system, built with Flask and NLP libraries, provides accurate and efficient recipe suggestions.

Min et al. presented "Food Recommendation: Framework, Existing Solutions and Challenges", a unified framework for food recommendation by integrating user preferences, context, and domain knowledge using collaborative learning.[11] They highlighted key challenges such as personalized modeling, multimodal data fusion, and real-time context incorporation. The study emphasized the importance of considering unique food characteristics, user health data, and environmental signals. It also reviewed existing solutions and categorized them under a structured taxonomy.

Chen et al. proposed "Personalized Food Recommendation as Constrained Question Answering over a Large-scale Food Knowledge Graph", a novel KBQA-based framework for personalized food recommendation by treating it as a constrained question answering task over a large-scale food knowledge graph.[12] Their system incorporates user-specific dietary needs, health guidelines, and ingredient preferences as query constraints. Techniques for query expansion, negation handling, and numerical comparisons within the knowledge graph was introduced.

Princy et al. proposed "A Personalized Food Recommender System for Women Considering Nutritional Information", focusing on their nutritional needs and health conditions.[13] The system considers user preferences, physical data, and dietary behavior to suggest balanced and customized meals. They explored content-based, collaborative, and hybrid filtering techniques to handle challenges like cold start and data sparsity. The study emphasizes how psychological, lifestyle, and socioeconomic factors affect women's food choices.

Wang et al. developed a work titled "Artificial Intelligence Applications to Personalized Dietary Recommendations: A Systematic Review" [14] evaluating the effectiveness of AI-generated personalized dietary recommendations across 11 studies. These AI systems, using ML, DL, and IoT-hybrid approaches, improved outcomes like glycemic control, metabolic health, and IBS symptom relief. The review highlighted algorithmic strengths, evidence quality (GRADE), and ethical considerations including data privacy and user adherence.

Zhang et al. proposed a "From Market to Dish: Multi-Ingredient Image Recognition for Personalized Recipe Recommendation", system that links supermarket photos of ingredients to personalized recipe recommendations.[15] They introduced a real-world multi-vegetable dataset and a Spatial regularization network that performs robust multi-label ingredient recognition in cluttered market scenes. Recognized ingredients are passed to a neural collaborative-filtering recommender that balances user preferences and nutritional needs.

III. PROPOSED WORK

This section describes the proposed Multimodal AI Framework, which analyses user cooking videos, ingredient photos, and textual preferences to provide personalised, health-conscious food recommendations. Fig. 1 displays the entire architectural workflow for the system that was developed.

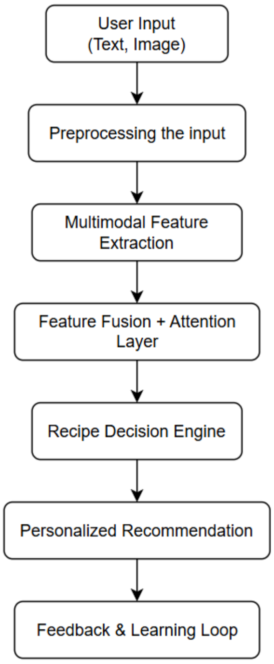

Figure 1 Workflow of Multimodal AI Framework for Personalized and Health-Aware Cooking Recommendations

Figure 1 shows the workflow of the suggested multimodal system, which integrates text, image, and video inputs through feature extraction and fusion. The system creates personalised recipe recommendations based on the user's health, skill level, and ingredient selection.

Module Description

A. Data Collection

Three sources of multimodal data are collected by the suggested system: text, photos, and video. Dietary preferences, allergies, and nutritional objectives (such as being gluten-free or diabetic-friendly) are examples of text inputs. The 3A2M dataset is used for training, and NLP techniques are used to parse and structure these inputs. To help determine what is available, user-captured photos of food or ingredients are used as image inputs. The food recognition model is supported by the Food-101 dataset, which consists of 101 food categories. Before features are extracted, images are resized and enhanced. Short cooking videos are gathered and examined to determine user cooking proficiency. The system is trained to recognise actions, gestures, and cooking sequences using datasets such as EPIC-Kitchens and YouCook2. The system is able to provide individualised, health-conscious recipe recommendations thanks to this multimodal data.

Figure 2 Images used for training and validation

Figure 2 displays sample images of different meal types and food ingredients that were collected for training and evaluation in the multimodal recommendation system.

B. Data preprocessing

A customised pipeline is used for pre-processing of each data modality. Named Entity Recognition (NER) is used to clean, tokenise, and process text inputs in order to extract food names, units, and medical restrictions. This guarantees the arrangement of pertinent food ingredients for analysis. To make images compatible with Vision Transformers , they are resized, normalised, and separated into fixed-size patches. To extract gestures and temporal cooking patterns, videos are uniformly sampled into frames and then fed into pre-trained CNNs. In the final step of multimodal fusion, all modalities text, image, and video—are transformed into standardised feature vectors.

C. Feature Extraction

Each input modality is processed by a specialised deep learning model to extract high-level representations. The DA-HAN model processes text data using attention-based BiLSTM layers to detect dietary restrictions, ingredient preferences, and medical condition .A Vision Transformer known as MIViT is used to analyse image inputs. It correctly recognises ingredients based on texture, colour, and shape, splits food photos into patches, and encodes spatial features. Video data is processed using SLaSTT, a transformer-based model that recognises cooking speed, gesture fluency, and preparation style using temporal convolutions. All outputs are converted into 2048-dimensional feature vectors in order to provide a rational and organised understanding of the user's current culinary and health context.

D. Recommendation Generation

The system moves on to the Recommendation Generation stage after the text, image, and video modalities' features are eliminated. At this crucial point, the extracted vectors are combined and examined to generate customised recipe recommendations according to the user's current requirements and skills.

Feature Fusion and Attention Layer: High-dimensional vectors are the output of the text (DA-HAN), and image (MIViT) modules. These vectors represent the user's culinary skills, dietary requirements, and ingredient availability. The system uses a Feature Fusion mechanism to generate a meaningful and cohesive user context. An attention mechanism evaluates each modality according to its suitability for the current recommendation task after concatenating the individual vectors. For instance, the video feature vector might be prioritised over the textual preference alone if a user uploads a video of a skilled chef. The user context profile is represented by the fused vector, which is usually 2048-dimensional and is sent to the following step. Along with limitations like health conditions (like diabetes), available ingredients (like an image), and experience level (like the speed and fluidity of a cooking video), this vector also reflects the user's preferences.

Recipe Decision Engine: Rather than employing traditional classification, the system uses a context-aware decision module to identify and rank recipe candidates. This engine uses Multimodal Constraint-Aware Weighted Ranking to rate recipes based on how well they match the user's context vector. The engine dynamically determines suitability based on a number of factors, including cooking time, ingredient match, preparation complexity, and health alignment. Recipes that most closely fit this combined profile receive higher scores and a ranking.

Customised Suggestion: The system chooses the top N recipes from the ranked list and shows them to the user as customised recommendations. Based on the fused context, each recommendation is updated in real time to make sure it is nutrient-appropriate and realistic. For instance, the system will suggest simpler recipes rather than more complicated ones if it determines that the user has limited ingredients and only moderate cooking abilities. Recipes are provided based on the user's skill level, the ingredients at hand, and their health needs.

Feedback & Learning Loop: The framework incorporates a feedback and learning loop to facilitate ongoing improvement. This module tracks whether the user prepares, saves, skips, or changes the recommended recipes. The user context model is updated and the attention weights are adjusted over time based on an analysis of this behaviour. As a result, the system is able to learn user preferences and cooking techniques, producing recommendations that are more precise and tailored to the individual. By stopping the learning cycle, the loop enables the system to automatically adjust to changing dietary requirements or preferences without requiring manual reconfiguration .By ensuring that recipe recommendations are tailored to each user's needs and preferences in real life, the framework closes the gap between data-driven insights and practical cooking outcomes. Using implicit feedback signals like cooking time, ingredient substitutions, and recipe ratings, the system uses reinforcement learning techniques to improve its recommendations.

IV.RESULT AND DISCUSSION

In this study ,a multimodal deep learning approach for classifying Indian cuisine was developed using both text and image modalities. In order to improve generalisation, the dataset was pre-processed using scaling, normalisation, and data augmentation methods including random flips for the image-based model. An 80:20 data split was used to train and evaluate the model, and a Vision Transformer (ViT) backbone was used to extract picture features. The picture model was optimised using the Adam optimiser and cross-entropy loss across multiple food classes. Before being trained with a transformer-based text model for text-based categorisation, recipe names and dish descriptions underwent pre-processing using tokenisation and padding. The performance of both models was evaluated using the accuracy, precision, recall, F1-score, and confusion matrix in order to ensure a fair comparison across classes. The picture model achieved a maximum classification accuracy of 80%, while the text-based model only achieved 92.21%. This indicates that visual features are more dependable for food recognition, but textual data offered additional insight.

TABLE I. COMPARSION TABLE OF DIFFERENT MODALS

Models	Average Accuracy
CNN	90.5%
Random Forest	79.8%
K-Nearest Neighbour	75.4%
Deep learning	93.7%
Naïve Bayes	67.9%
VGG16	92.4%
DAHAN	92.21%
MIVIT	80%

The Table 1 shows that accuracy of food classification models is compared to several machine learning algorithms. It demonstrates the power of visual characteristics for food recognition by showing that the DAHAN and MIVIT performed noticeably better than the text model and conventional methods.

Software Requirements

1. Programming Language

Python: Version 3.8 or higher (Supports modern ML/DL libraries and is well-compatible with Hugging Face + PyTorch).

2. Deep Learning Frameworks

PyTorch: Latest stable release (≥ 1.12) (GPU acceleration recommended if available).

TorchVision: For image transformations and dataset utilities.

TimM (PyTorch Image Models): For pre-trained Vision Transformers (ViT, Swin, etc.) used in MIViT.

3. Natural Language Processing (NLP)

Transformers (Hugging Face): ForBERT,DistilBERT, RoBERTa models.

Tokenizers: Fast and efficient text tokenization.

4. Platform

Google Colab (Recommended)**:**Free environment with GPU/TPU support.Pre-installed libraries like PyTorch, TorchVision, NumPy.

TABLE II. ACCURACY RESULT OF SELECTED FOOD CLASSES

Food Class	Accuracy %
Beef Carpaccio	100.00%
Cheese Plate	95.24%
Clam chowder	94.24%
Cream Brulee	95.24%
Edamame	96.55%
Frozen curd	95.65%
Hot and sour soup	100.00%
Pho	100.00%
Ramen	96.00%
Shrimp and Grits	85.71%

Table II shows the food classification model's accuracy results for a number of dishes. Perfect precision was attained for lessons like Beef Carpaccio, Pho, and Hot and Sour Soup. Other items, such as frozen yoghurt, clam chowder, and cheese plates, all showed high accuracy scores above 90%, indicating the model's powerful capacity to identify unusual treats.

TABLE III. Classification Report of the System

Food Classes	Precision	Recall	F1-score	Support
Apple pie	0.45	0.45	0.45	11
Baby back rips	0.62	1.00	0.72	5
Baklava	0.57	0.57	0.57	7
Beef Carpaccio	1.00	0.83	0.91	12
Beef Tartare	1.00	0.53	0.69	17
pho	1.00	1.00	1.00	9
Chicken Wings	0.81	0.90	1.00	9
Falafel	0.88	0.74	0.80	19
Fish and Chips	0.92	0.92	0.92	13
Clam Chowder	1.00	0.91	0.95	10

Table III shows that food identification system's classification report. While frequently misunderstood meals like ravioli, steak, and tiramisu scored very poorly, several dishes, like pho, hot and sour soup, oysters, and spring rolls, obtained perfect accuracy (1.00). This demonstrates the model's advantages in identifying visually unique dishes as well as its shortcomings in managing visually identical ones.

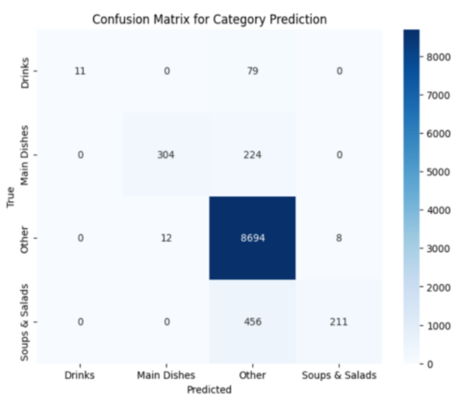

Figure 3. Confusion matrix for DAHAN –text prediction system

Figure 3 shows the category prediction confusion matrix. Soups & Salads, Drinks, and Main Dishes are often predicted as "Other," yet the model accurately categorises the "Other" category.

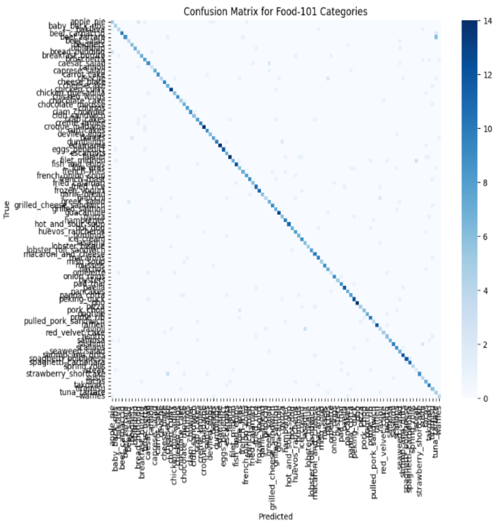

Figure 4. Confusion matrix for MIVIT –image prediction system

Figure 4 shows the confusion matrix for the Food-101 categories. The majority of forecasts that fall along the diagonal indicate strong overall performance. However, certain off-diagonal dots highlight occasional misclassifications of visually similar foods.

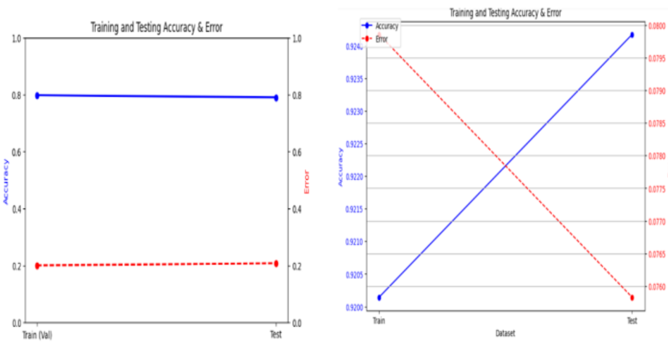

Figure 5.Training and Accuracy graph

Figure 5 displays the model's accuracy and error during training and testing. The left line shows that both training and testing accuracy remain constant, whereas the right graph shows that accuracy and error have an inverse relationship, with accuracy increasing as error decreases.

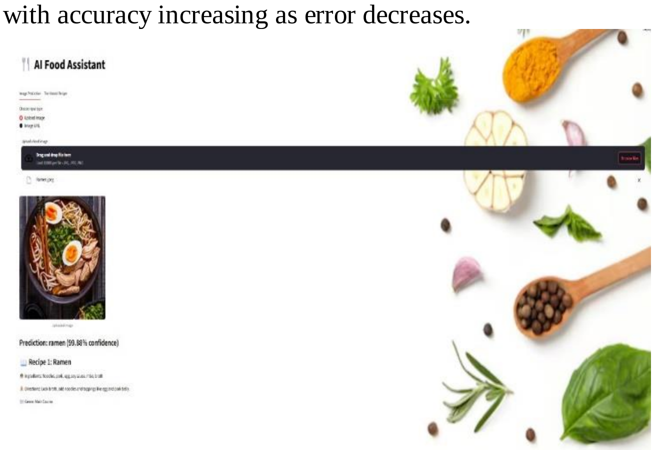

Figure 6. Uploaded image identified as Ramen using Image

Figure 7. Uploaded image identified as Samosa using URL

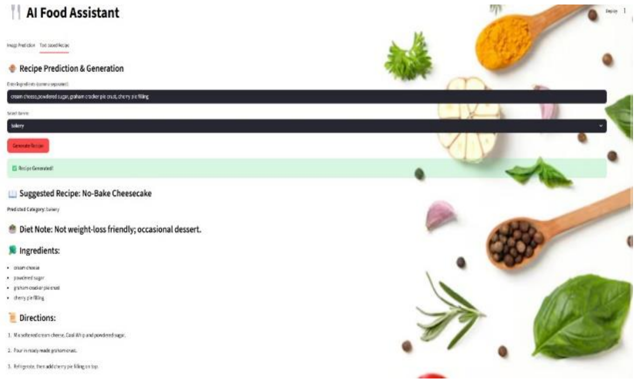

Figure 8. Uploaded image that shows output of Generated recipe based on text

Figure 6,7 and 8 represents the prediction of the food uploaded in the image using food prediction system.

V.CONCLUSION

This study effectively combined text, image, and video modalities to create a multimodal AI framework for personalised and health-conscious culinary advice. The system employed MIViT (Multimodal Ingredient Vision Transformer) for food image analysis and RoBERTa for textual recipe classification. The computer was able to make more precise and contextually aware recipe recommendations since each modality contributed complementing qualities. The system demonstrated remarkable accuracy in food recognition and recipe suggestion tasks through pre-processing, model training, and feature extraction across several datasets (3A2M, Food-101, and recipe collections). The application of sequence learning, computer vision, and natural language comprehension to provide relevant, user-specific, and health-conscious meal suggestions was illustrated by the multimodal integration. The results demonstrate the effectiveness of transfer learning and sophisticated deep learning architectures in addressing practical food computing problems. To further personalise recommendations, future research might concentrate on adding other cuisines, , and cultural variances to the dataset. Nutritional analysis and reinforcement learning could also be incorporated. The suggested framework is a potential step towards intelligent culinary help because it has applications in smart kitchens, nutritional monitoring, and health-conscious food planning.

VI. References

[1] Freyne, J., & Berkovsky, S. (2010). Intelligent food planning: personalized recipe recommendation. In *Proceedings of the 15th international conference on Intelligent user interfaces, IUI '10* (pp. 321-324). Association for Computing Machinery (ACM). doi: 10.1145/1719970.1719973

[2] M. R. M. Razali, H. F. Almarzuki, R. A. Bahar, and N. A. S. Abdullah, “Recipe Recommendation based on Food Ingredients Recognition using Deep Learning,” in *Proc. Int. Conf. Innovation & Entrepreneurship in Computing, Engineering & Science Education (InvENT 2024)*, Selangor, Malaysia, Dec. 2024, pp. 555–564, doi: 10.2991/978-94-6463-589-8_52.

[3] Rostami A, Jain R, Rahmani AM. Food Recommendation as Language Processing (F-RLP): A Personalized and Contextual Paradigm. In2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2024 Jul 15 (pp. 1-4). Doi: 10.48550/arXiv.2402.07477.

[4] Agapito G, Calabrese B, Guzzi PH, Cannataro M, Simeoni M, Caré I, Lamprinoudi T, Fuiano G, Pujia A. DIETOS: A recommender system for adaptive diet monitoring and personalized food suggestion. In2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob) 2016 Oct 17 (pp. 1-8). Doi: 10.1109/WiMOB.2016.7763190

[5] R. Yera, A. A. Alzahrani, L. Martínez, and R. M. Rodríguez, “A systematic review on food recommender systems for diabetic patients,” *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 4248, 2023.doi: 10.3390/ijerph20054248

[6] Zhang, J., Wang, Z., Liu, W. *et al.* A unified approach to designing sequence-based personalized food recommendation systems: tackling dynamic user behaviors. *Int. J. Mach. Learn. & Cyber.* **14**, 2903–2912 (2023).doi: 10.1007/s13042-023-01808-7.

[7] M. Gim, D. Park, M. Spranger, K. Maruyama and J. Kang, "RecipeBowl: A Cooking Recommender for Ingredients and Recipes Using Set Transformer," in *IEEE Access*, vol. 9, pp. 143623-143633, 2021, doi: 10.1109/ACCESS.2021.3120265.

[8] T. Islam, A. R. Joyita, M. G. R. Alam, M. Mehedi Hassan, M. R. Hassan and R. Gravina, "Human-Behavior-Based Personalized Meal Recommendation and Menu Planning Social System," in *IEEE Transactions on Computational Social Systems*, vol. 10, no. 4, pp. 2099-2110, Aug. 2023, doi: 10.1109/TCSS.2022.3213506.

[9] A. Bilgin, H. Hagras, J. van Helvert and D. Alghazzawi, "A Linear General Type-2 Fuzzy-Logic-Based Computing With Words Approach for Realizing an Ambient Intelligent Platform for Cooking Recipe Recommendation," in *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 2, pp. 306-329, April 2016, doi: 10.1109/TFUZZ.2015.2453400.

[10] A. Padmavathi and D. Sarker, "RecipeMate: A Food Media Recommendation System Based on Regional Raw Ingredients," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1-6, doi: 10.1109/ICCCNT56998.2023.10307728.

[11] W. Min, S. Jiang, and R. Jain, “Food recommendation: Framework, existing solutions, and challenges,” *IEEE Trans. Multimedia*, vol. 22, no. 10, pp. 2659–2671, Oct. 2019, doi: 10.1109/TMM.2019.2918911.

[12] Y. Chen, A. Subburathinam, C.-H. Chen, and M. J. Zaki, “Personalized food recommendation as constrained question answering over a large-scale food knowledge graph,” in *Proc. 14th ACM Int. Conf. Web Search Data Mining (WSDM)*, 2021, pp. 544–552. doi: 10.1145/3437963.3441816.

[13] J. Princy, S. Senith, A. A. Kirubaraj, and P. Vijaykumar, “A personalized food recommender system for women considering nutritional information,” *Int. J. Pharmaceut. Res.*, vol. 13, no. 2, 2021.doi:10.31838/ijpr/2021.13.02.233

[14] X. Wang, Z. Sun, H. Xue, and R. An, “Artificial Intelligence Applications to Personalized Dietary Recommendations: A Systematic Review,” *Healthcare*, vol. 13, no. 12, p. 1417, 2025.doi: 10.3390/healthcare13121417

[15] L. Zhang, J. Zhao, S. Li, B. Shi and L. -Y. Duan, "From Market to Dish: Multi-ingredient Image Recognition for Personalized Recipe Recommendation," *2019 IEEE International Conference on Multimedia and Expo (ICME)*, Shanghai, China, 2019, pp. 1252-1257, doi: 10.1109/ICME.2019.00218.