Multi-Ingredient Food Image Classification and Context-Aware Recipe Recommendation System

Team Name : K-Nearest Party

- Ram Bagaria
- Dhanush Biligiri
- Fnu Sowrabh
- Akhilesh Dandavati

Aim of the Project

Develop a two-stage system that:

- Accurately detects multiple ingredients from food images using deep learning.
- Recommends contextually relevant recipes by leveraging detected ingredients and recipe databases.

Problem Statement

Current food image classifiers focus on dish-level recognition rather than granular ingredient detection. Recipe recommendation systems lack integration with visual input, requiring manual ingredient entry. This creates friction in meal planning and exploration.

Goal

- Train a multi-label classifier to identify ingredients from images.
- Design a recipe retrieval engine that maps detected ingredients to recipes with minimal additional components.

Problem Formulation

- **Ingredient Detection**
  - Input: RGB image $I \in \mathbb{R}^{H \times W \times 3}$
  - Output: Multi-label vector $y \in \{0, 1\}^N$, where N = number of ingredient classes
  - Model: Fine-tune a pre-trained CNN (ResNet-152, EfficientNet-B7) or Vision Transformer (ViT) for multi-label classification.
- **Recipe Recommendation**
  - Input: Detected ingredients y, optional filters (cuisine, diet).
  - Output: Ranked recipes $R = r1, r2, ..., rk$, prioritized by ingredient overlap and user constraints.
  - Method: Graph-based ingredient-recipe matching or embedding similarity (e.g., sentence-BERT)

Related works

- **Food Recognition**: CVPR, ICCV, ACM Multimedia.
- **Recipe Retrieval**: ACM RecSys, ACL, NAACL.
- **Multi-Label Classification**: IEEE TPAMI, NeurIPS.

Relevant Background Papers

- **Recipe1M (Salvador et al., 2017)**: Aligns food images with recipe text.
- **Im2Recipe (Carvalho et al., 2018)**: Cross-modal embedding for recipe retrieval.
- **Multi-Ingredient Detection (Wang et al., 2020)**: Attention-driven CNN for ingredient localization.
- **FoodBERT (Wang et al., 2021)**: BERT-based recipe embeddings.

Data Resources

- **Images**
  - Recipe1M+ (1M+ images with ingredient lists)
  - UEC-Food256 (14k images, 256 ingredient classes, bounding boxes).
- **Recipes**
  - Recipe1M (1M recipes with structured ingredients/instructions).
  - Food.com Dataset (500k+ recipes via Kaggle).
- **APIs**
  - Spoonacular (recipe search) and Edamam (nutritional data).

Methodology

- **Data Preprocessing**
  - Image augmentation (RandAugment, MixUp).
  - Recipe text normalization (lemmatization, unit standardization).
- **Ingredient Detection Model**
  - Train ResNet-152/ViT with sigmoid outputs for multi-label classification.
  - Loss: Focal Loss to handle class imbalance.
- **Recipe Engine**
  - Build ingredient-recipe graph; compute Jaccard similarity for ranking.
  - Use FAISS for efficient nearest-neighbor search in embedding space.

Evaluation

- **Ingredient Detection**
  - Metrics: Mean Average Precision (mAP), F1-score, per-class precision/recall.
  - Baselines: Pretrained Food-101 models, Im2Recipe's ingredient module.
- **Recipe Recommendation**
  - Metrics: Precision@10, Recall@10, Mean Reciprocal Rank (MRR)
  - User Study: 50 participants rate relevance of top-5 recommendations.
- **End-to-End Testing:**
  - Measure latency (image-to-recipe time ¡ 2s on Tesla T4 GPU)