Thai Food Detection and Classification

Napat Jirataranon
Faculty of ICT, Mahidol University, Nakhon Pathom, Thailand 73170
napat.jir@student.mahiol.ac.th

Natthaphat Pintip
Faculty of ICT, Mahidol University, Nakhon Pathom, Thailand 73170
natthaphat.pin@student.mahidol.ac.th

Sittikorn Maneewong
Faculty of ICT, Mahidol University, Nakhon Pathom, Thailand 73170
sittikorn.man@student.mahidol.ac.th

***Abstract*— This project introduces an intelligent Thai food detection and classification system aimed at assisting health-conscious individuals in estimating their calorie intake from images. By integrating YOLOv11 for object detection and ResNet34 for image classification, the system can identify multiple Thai dishes from an uploaded image, classify them, and provide corresponding calorie estimates. This solution is designed for real-time application and enhances user convenience by automating food tracking**

Keywords— Thai food detection, YOLOv11, ResNet34, calorie estimation, object detection, image classification, deep learning

I. Introduction

In recent years, personal health monitoring has gained increasing attention, especially in the context of diet and calorie tracking. However, manual logging of meals can be time-consuming and inaccurate. This project proposes a visual-based solution that allows users to simply upload an image of a meal and automatically receive food item classifications and estimated calorie counts.

The key contributions of this work are:

- An end-to-end pipeline that detects and classifies Thai food from images.
- Integration of YOLOv11 for object detection and ResNet34 for dish classification.
- Calorie estimation based on THFOOD-50 dataset labels.

II. Methodology

A. Dataset

Two primary datasets were used in this project:

- ***Platefood Dataset*** *(via Roboflow): Contains annotated images of Thai food plates used for training the object detection model (YOLOv11).*
- ***THFOOD-50 Dataset****: Contains labeled images of 50 Thai dish categories, used to train the ResNet34 classifier.*

Images were cleaned using a custom `datasetcleaner.py` *script to eliminate corrupted files.*

B. Model Development

- ***Object Detection***
  - *Model: YOLOv11 (s and m variants)*
  - *Fine-tuned on Platefood dataset with frozen backbone layers*
  - *Optimizer: AdamW with custom learning rate*
  - *Loss functions: Included box loss, classification loss, and DFL loss*
  - *Confidence threshold adjustments were explored to improve detection of underrepresented classes (e.g., plates)*
- ***Classification***
  - *Model: ResNet34*
  - *Fully connected layer replaced with Dropout + Linear classifier head*
  - *Data augmentation applied to reduce overfitting:*
  - *Horizontal flips*
  - *Random color jitter*
  - *Random resized crop*
  - *Gaussian blur*
  - *Normalization followed ImageNet mean and std*

III. Results

A. Experimental Setup

All models were trained for 100 epochs with early stopping based on validation loss. Experiments were run on a GPU-enabled environment with evaluation metrics logged every epoch. Models were compared using precision-recall and F1-Confidence curves.

B. Performance Metrics

The following metrics were used to evaluate both object detection and classification:

- *mAP@0.5 and mAP@0.5:0.95: For measuring localization accuracy.*
- *Precision, Recall, F1 Score: For measuring classification quality.*
- *Confusion Matrix (Normalized): To visualize per-class performance.*

C. Model Comparison

Three versions of YOLOv11 were compared

- *Model A (First): Best for detecting plates (Plate PR AUC = 0.364)*
- *Model B (train14): Highest F1 Score (0.68)*
- *Model C (train22): Highest mAP@0.5 (0.720) and balanced dish detection*

Conclusion: Model A is the best choice for plate and dish detection balance. Model C excels in dish recognition, while Model B offers best overall classification balance but poor plate detection.

IV. Conclusions

This project successfully integrates YOLOv11 and ResNet34 for real-time Thai food recognition and calorie estimation. While dish classification performs well, plate detection remains a challenge due to dataset imbalance. Future improvements could include focal loss, threshold tuning, and data augmentation targeting underrepresented classes like plates.

References

[1] Chakkrit Termritthikun. (2021). THFOOD-50: A dataset for Thai food image classification.Available online: https://github.com/chakkritte/THFOOD-50

[2] Pathirana, S. (2023). Plate Food Dataset. Roboflow Universe Dataset Repository. Available online: https://universe.roboflow.com/subhash-pathirana-cdryb/plate_food

[3] Ultralytics. (2024). YOLOv11: You Only Look Once Object Detection and Classification. Ultralytics YOLO. Available online: https://docs.ultralytics.com

[4] PyTorch. (2024). PyTorch: An open-source machine learning framework. Meta AI. Available online: https://pytorch.org

[5] Gradio. (2024). Gradio: Build & Share Delightful Machine Learning Apps. Hugging Face. Available online: https://www.gradio.app