

Xây dựng đề cương: Kế hoạch Thực hiện Đề tài Machine Learning

Nhóm 3: Phạm Hải Đăng, Đỗ Mạnh Hùng

Đề tài 17: Phân loại các món ăn Việt Nam qua hình ảnh

I. Giới thiệu đề tài

Ẩm thực Việt Nam rất đa dạng và phong phú với nhiều món ăn đặc trưng. Trong bối cảnh trí tuệ nhân tạo ngày càng phát triển, việc áp dụng các mô hình học sâu để nhận diện và phân loại món ăn có thể mang lại nhiều ứng dụng trong thực tế, như gợi ý món ăn, hỗ trợ du lịch ẩm thực hoặc tự động hóa trong các ứng dụng bán hàng.

II. Mục tiêu

- Xây dựng mô hình có thể nhận biết các món ăn phổ biến của Việt Nam như Phở, Bún chả, Bánh mì...
- Đánh giá hiệu năng của mô hình dựa trên các chỉ số chuẩn: accuracy, precision, recall, F1-score.
- Đề xuất mô hình tối ưu nhất và xây dựng demo thử nghiệm.

III. Thu thập và xử lý dữ liệu

1. Nguồn dữ liệu: Google Image, Kaggle, Flickr, hoặc bộ dataset công khai.
2. Gắn nhãn dữ liệu: phân loại ảnh theo các món ăn Việt Nam phổ biến.
3. Tiền xử lý dữ liệu:
 - Resize ảnh về kích thước chuẩn (224x224 hoặc 256x256).
 - Chuẩn hóa dữ liệu ảnh (giá trị pixel).
 - Data augmentation (xoay, lật, crop, thay đổi độ sáng...).
4. Chia dữ liệu: Train (70%), Validation (15%), Test (15%).

IV. Mô hình gợi ý

- Baseline: CNN cơ bản.
- Mô hình nâng cao: EfficientNet, Vision Transformer (ViT).
- So sánh hiệu năng giữa các mô hình.

V. Huấn luyện và đánh giá

1. Framework: TensorFlow/Keras hoặc PyTorch.
2. Huấn luyện:
 - Epoch: 20–50
 - Optimizer: Adam hoặc SGD
 - Loss function: Cross-Entropy Loss
3. Đánh giá:
 - Confusion matrix
 - Accuracy, Precision, Recall, F1-score
 - Biểu đồ loss/accuracy theo epoch

VI. Kết quả dự kiến

- Độ chính xác mô hình đạt từ 80–90%.
- EfficientNet/ViT có thể vượt trội so với CNN cơ bản.
- Xây dựng demo: nhập ảnh → dự đoán món ăn.

VII. Kết luận và hướng phát triển

Kết luận: Mô hình phân loại món ăn Việt Nam qua hình ảnh có tính khả thi, ứng dụng cao.

Hướng phát triển:

- Thu thập thêm nhiều dữ liệu đa dạng hơn.
- Mở rộng số lượng món ăn.
- Ứng dụng vào hệ thống gợi ý món ăn hoặc ứng dụng di động.

VIII. Mốc thời gian

- Tuần 1–2: Chọn đề tài, tìm hiểu lý thuyết, nghiên cứu mô hình CNN, EfficientNet, ViT, lập kế hoạch thu thập dữ liệu.
- Tuần 3–4: Thu thập dữ liệu, gắn nhãn, tiền xử lý, chia train/validation/test.
- Tuần 5–6: Cài đặt môi trường, xây dựng mô hình CNN cơ bản, huấn luyện và đánh giá ban đầu.
- Tuần 7: Huấn luyện EfficientNet, áp dụng transfer learning, so sánh kết quả với baseline.
- Tuần 8: Triển khai Vision Transformer, huấn luyện và đánh giá, so sánh với EfficientNet và CNN.
- Tuần 9: Phân tích kết quả, viết báo cáo chi tiết, chuẩn bị slide thuyết trình.
- Tuần 10: Nộp báo cáo và thuyết trình, tổng kết đề tài.