

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
KHOA CÔNG NGHỆ THÔNG TIN 1



BÁO CÁO ĐỀ CƯƠNG
MÔN THỰC TẬP CƠ SỞ

**Đề tài: Nghiên cứu, đánh giá và triển khai
mô hình Deep Learning tối ưu cho hệ thống phân loại rác thải**

Giảng viên hướng dẫn:	TS. Kim Ngọc Bách
Sinh viên thực hiện:	Nguyễn Văn Gia Bảo
Mã sinh viên:	B23DCCN079

Hà Nội - 2026

MỤC LỤC

I. GIỚI THIỆU DỰ ÁN.....	3
1. Lý do chọn đề tài.....	3
2. Ý nghĩa và tính ứng dụng.....	3
3. Giá trị học thuật.....	3
II. Cơ sở lý thuyết và công nghệ sử dụng.....	4
1. Deep Learning và Mạng nơ-ron tích chập (CNN).....	4
2. Các mô hình Deep Learning được nghiên cứu	4
3. Transfer Learning	4
4. Công nghệ triển khai	4
III. Phân tích yêu cầu của dự án.....	5
1. Yêu cầu chức năng	5
2. Yêu cầu phi chức năng	5
3. Dữ liệu sử dụng	5
IV. Kế hoạch thực hiện dự án.....	6
V. TÀI LIỆU THAM KHẢO	6

I. GIỚI THIỆU DỰ ÁN.

1. Lý do chọn đề tài

- Ô nhiễm rác thải đang là một trong những vấn đề môi trường nghiêm trọng nhất hiện nay. Việc phân loại rác thải đúng cách là bước đầu tiên và quan trọng nhất trong quy trình xử lý, tái chế rác - nhưng phần lớn vẫn được thực hiện thủ công, phụ thuộc vào ý thức con người và dễ xảy ra sai sót.

- Sự phát triển của Deep Learning, đặc biệt là các mô hình học sâu ứng dụng trong thị giác máy tính, mở ra khả năng tự động hóa quá trình phân loại rác thải với độ chính xác cao. Tuy nhiên, mỗi mô hình Deep Learning có ưu và nhược điểm khác nhau về độ chính xác, tốc độ xử lý và khả năng triển khai thực tế. Việc nghiên cứu, đánh giá và lựa chọn mô hình phù hợp nhất là bài toán có giá trị cả về học thuật lẫn thực tiễn.

2. Ý nghĩa và tính ứng dụng

- Đề tài hướng đến xây dựng một ứng dụng web thực tế cho phép người dùng upload ảnh rác thải và nhận kết quả phân loại tự động kèm hướng dẫn xử lý. Hệ thống có thể ứng dụng trong các hộ gia đình, trường học, khu dân cư nhằm hỗ trợ phân loại rác đúng cách và nâng cao ý thức bảo vệ môi trường.

3. Giá trị học thuật

- Đề tài nghiên cứu và so sánh có hệ thống ba mô hình Deep Learning tiêu biểu là VGG16, ResNet50 và EfficientNet trên cùng một bộ dữ liệu rác thải, đánh giá theo các tiêu chí độ chính xác, tốc độ và hiệu quả triển khai. Kết quả nghiên cứu cung cấp cơ sở thực nghiệm cho việc lựa chọn mô hình phù hợp trong các bài toán phân loại ảnh tương tự.

II. Cơ sở lý thuyết và công nghệ sử dụng.

1. Deep Learning và Mạng nơ-ron tích chập (CNN)

- Deep Learning là một nhánh của Machine Learning, sử dụng các mạng nơ-ron nhân tạo nhiều lớp để học đặc trưng từ dữ liệu [1]. Trong lĩnh vực xử lý ảnh, Mạng nơ-ron tích chập (Convolutional Neural Network - CNN) là kiến trúc phổ biến nhất, có khả năng tự động trích xuất đặc trưng không gian từ ảnh thông qua các lớp tích chập (convolution), pooling và fully connected [2].

2. Các mô hình Deep Learning được nghiên cứu

- VGG16 là mô hình được đề xuất bởi Simonyan và Zisserman (2014), sử dụng kiến trúc đơn giản với 16 lớp, các bộ lọc tích chập kích thước 3×3 xếp chồng liên tiếp. VGG16 được đánh giá cao về độ chính xác nhưng có số lượng tham số lớn, đòi hỏi nhiều bộ nhớ [3].

- ResNet50 được giới thiệu bởi He et al. (2015), giải quyết vấn đề vanishing gradient trong mạng sâu thông qua kỹ thuật skip connection (residual connection). ResNet50 với 50 lớp cho phép huấn luyện mạng sâu hơn mà không bị suy giảm độ chính xác [4].

- EfficientNet được đề xuất bởi Tan và Le (2019), sử dụng phương pháp compound scaling để đồng thời tối ưu độ sâu, độ rộng và độ phân giải của mạng. EfficientNet đạt độ chính xác cao hơn các mô hình trước với số lượng tham số ít hơn đáng kể [5].

3. Transfer Learning

- Transfer Learning là kỹ thuật tái sử dụng trọng số của mô hình đã được huấn luyện trên tập dữ liệu lớn (thường là ImageNet) cho bài toán mới [6]. Kỹ thuật này giúp giảm đáng kể thời gian huấn luyện và lượng dữ liệu cần thiết, đặc biệt hiệu quả khi dữ liệu huấn luyện còn hạn chế.

4. Công nghệ triển khai

- PyTorch là framework Deep Learning mã nguồn mở được phát triển bởi Meta AI, cung cấp API linh hoạt để xây dựng và huấn luyện mô hình học sâu [7].

- Streamlit là thư viện Python cho phép xây dựng ứng dụng web tương tác một cách nhanh chóng mà không cần kiến thức HTML/CSS/JavaScript [8].

- OpenCV là thư viện mã nguồn mở chuyên dụng cho xử lý ảnh và thị giác máy tính [9].

III. Phân tích yêu cầu của dự án.

1. Yêu cầu chức năng

- Hệ thống cần đáp ứng các chức năng chính sau. Về phần nghiên cứu, hệ thống cần huấn luyện và đánh giá 3 mô hình VGG16, ResNet50 và EfficientNet trên cùng một bộ dữ liệu rác thải, so sánh kết quả theo các tiêu chí độ chính xác (accuracy), precision, recall và F1-score. Về phần ứng dụng web, người dùng có thể upload ảnh rác thải, hệ thống tự động phân loại và hiển thị kết quả phân loại kèm hướng dẫn xử lý tương ứng.

2. Yêu cầu phi chức năng

- Giao diện web đơn giản, dễ sử dụng, không yêu cầu người dùng có kiến thức kỹ thuật. Thời gian xử lý mỗi ảnh không quá 5 giây. Hệ thống hoạt động ổn định trên các trình duyệt phổ biến.

3. Dữ liệu sử dụng

- Đề tài sử dụng bộ dữ liệu TrashNet — bộ dữ liệu công khai phổ biến cho bài toán phân loại rác thải, gồm 6 danh mục: cardboard (bìa các tông), glass (thủy tinh), metal (kim loại), paper (giấy), plastic (nhựa) và trash (rác hỗn hợp) với tổng cộng hơn 2.500 ảnh [10].

IV. Kế hoạch thực hiện dự án

Tuần	Nội dung công việc	Kết quả dự kiến
1-2	Tìm hiểu lý thuyết Deep Learning, CNN, Transfer Learning. Nghiên cứu các mô hình VGG16, ResNet50, EfficientNet	Nắm vững lý thuyết, tóm tắt tài liệu
3-4	Chuẩn bị dữ liệu: tải dataset TrashNet, tiền xử lý ảnh, chia tập train/validation/test	Dataset sẵn sàng cho huấn luyện
5-6	Huấn luyện và đánh giá mô hình VGG16 và ResNet50	Kết quả accuracy, F1-score của 2 mô hình
7	Huấn luyện và đánh giá mô hình EfficientNet	Kết quả accuracy, F1-score của EfficientNet
8	So sánh tổng hợp 3 mô hình, lựa chọn mô hình tối ưu	Báo cáo so sánh, chọn được mô hình tốt nhất
9-10	Xây dựng ứng dụng web Streamlit tích hợp mô hình tối ưu	Ứng dụng web hoàn chỉnh, có thể demo
11	Hoàn thiện, kiểm thử, viết báo cáo cuối kỳ	Sản phẩm hoàn thiện, báo cáo hoàn chỉnh

V. TÀI LIỆU THAM KHẢO

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [2] Y. LeCun et al., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE CVPR*, pp. 770–778, 2016.

- [5] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of ICML*, pp. 6105–6114, 2019.
- [6] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [7] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [8] Streamlit Inc., "Streamlit — The fastest way to build data apps," <https://streamlit.io>, 2024.
- [9] G. Bradski, "The OpenCV library," *Dr. Dobb's Journal of Software Tools*, 2000.
- [10] G. Thung and C. Yang, "TrashNet: A dataset for trash classification," Stanford University, <https://github.com/garythung/trashnet>, 2016.