

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



BÁO CÁO GIỮA KỲ

Đề tài: Bài toán nhận diện biển số xe

Giảng viên hướng dẫn: Kim Ngọc Bách
Nhóm lớp 11
Họ và tên sinh viên Đinh Hoàng Vũ
Mã sinh viên B23DCCN941

Hà Nội – 2026

MỤC LỤC

1. Giới thiệu dự án.....	3
1.1. Lý do chọn đề tài	3
1.2. Ý nghĩa của đề tài.....	3
1.2.1. Tính ứng dụng.....	3
1.2.2. Giá trị học thuật	3
2. Cơ sở lý thuyết và công nghệ sử dụng.....	4
2.1. Cơ sở lý thuyết.....	4
2.1.1. Thị giác máy tính	4
2.1.2. Xử lý ảnh số.....	5
2.1.3. Học máy (Machine Learning)	6
2.1.4. Học sâu (Deep Learning)	7
2.1.5. Phát hiện đối tượng.....	9
2.1.6. Nhận dạng ký tự quang học (OCR)	11
2.2. Công nghệ sử dụng.....	12
3. Phân tích yêu cầu của dự án	13
3.1. Chức năng chính	13
3.2. Yêu cầu chức năng	13
3.3. Yêu cầu phi chức năng	13
4. Kế hoạch thực hiện dự án	13
Giai đoạn 1: Nghiên cứu lý thuyết.....	13
Giai đoạn 2: Chuẩn bị dữ liệu.....	13
Giai đoạn 3: Huấn luyện/Fine-tune mô hình.....	13
Giai đoạn 4: Tích hợp hệ thống	14
Giai đoạn 5: Đánh giá và hoàn thiện.....	14

1. Giới thiệu dự án

1.1. Lý do chọn đề tài

Trong bối cảnh ngày nay, trí tuệ nhân tạo (Artitical Intelligence – AI) đang trở thành một trong lĩnh vực trọng tâm của khoa học và công nghệ hiện đại. Đặc biệt, các phương pháp Học sâu (Deep Learning – DL) đã đạt được nhiều thành tựu nổi bật trong các bài toán thị giác máy tính như nhận diện khuôn mặt, phát hiện vật thể và nhận dạng ký tự.

Xuất phát từ mong muốn của bản thân muốn tiếp tục tìm hiểu thêm về lĩnh vực Trí tuệ nhân tạo, em lựa chọn đề tài này nhằm mục tiêu trao đổi thêm kiến thức và kỹ năng liên quan đến Deep Learning for Computer Vison. Việc thực hiện dự án không chỉ giúp củng cố nền tảng lý thuyết mà còn tạo cơ hội tiếp cận các công nghệ và công cụ hiện đại trong lĩnh vực AI.

Bên cạnh đó, trong bối cảnh giao thông ngày càng phức tạp và số lượng phương tiện gia tăng nhanh chóng, nhu cầu ứng dụng các hệ thống tự động trong quản lý và giám sát các phương tiện trở nên cấp thiết. Các giải pháp như nhận diện biển số xe tự động có thể hỗ trợ quản lý bãi đỗ xe, kiểm soát ra vào, giám sát an ninh và xử lý vi phạm giao thông một cách hiệu quả, chính xác và tiêm kiệm nhân lực.

Ngoài ra, phạm vi và độ phức tạp của đề tài em dự tính ở mức phù hợp với học phần Thực tập có số 4 tín chỉ. Dự án tập trung và xây dựng mô hình nhận diện biển số ở quy mô thử nghiệm, sử dụng các công nghệ và công cụ phổ biến, đảm bảo có thể hoàn thành trong khung thời gian của học phần, đồng thời vẫn đáp ứng yêu cầu về tính nghiên cứu và tính ứng dụng.

1.2. Ý nghĩa của đề tài

1.2.1. Tính ứng dụng

Giúp tự động hóa nhiều quy trình trong quản lý và giám sát phương tiện. Cụ thể, hệ thống có thể được triển khai trong các bãi giữ xe thông minh nhằm tự động ghi nhận thông tin phương tiện ra vào, giảm thiểu sai sót do các thao tác thủ công và nâng cao hiệu quả quản lý. Ngoài ra, nhận diện biển số xe đóng vai trò quan trọng trong việc hỗ trợ giám sát an ninh, phát hiện phương tiện vi phạm, thống kê lưu lượng giao thông và phân tích dữ liệu phục vụ công tác quy hoạch đô thị.

1.2.2. Giá trị học thuật

- Giúp vận dụng và trau dồi kiến thức liên quan đến lĩnh vực xử lý ảnh số, thị giác máy tính và học sâu (Deep Learning).
- Kết hợp giữa các thuật toán phát hiện đối tượng (Object Detection) và nhận dạng ký tự quang học (ORC).
- Nghiên cứu và áp dụng các mô hình mạng nơ-ron tích chập (CNN).

- Nâng cao kỹ năng đánh giá mô hình.
- Rèn luyện kỹ năng tự học, nghiên cứu khoa học, tìm kiếm và tổng hợp tài liệu.

2. Cơ sở lý thuyết và công nghệ sử dụng

2.1. Cơ sở lý thuyết

2.1.1. Thị giác máy tính

Computer vision là một lĩnh vực khoa học nghiên cứu các phương pháp cho phép máy tính thu nhận, xử lý, phân tích và hiểu dữ liệu hình ảnh từ thế giới thực nhằm trích xuất thông tin có ý nghĩa dưới dạng số hoặc ký hiệu phục vụ cho quá trình ra quyết định [1], [2]. Dữ liệu hình ảnh có thể tồn tại ở nhiều dạng khác nhau như ảnh tĩnh, chuỗi video, dữ liệu từ camera giám sát hoặc dữ liệu đa chiều từ các thiết bị quét y học [3].

Sự phát triển của thị giác máy tính bắt nguồn từ mục tiêu mở rộng khả năng thị giác của con người thông qua các hệ thống nhân tạo có khả năng nhận diện và diễn giải nội dung hình ảnh số [5]. Theo Marr [5], quá trình nhận thức thị giác có thể được mô hình hóa như một chuỗi các bước xử lý thông tin, từ trích xuất đặc trưng mức thấp đến xây dựng biểu diễn có ý nghĩa ở mức cao hơn.

Nhận diện hình ảnh có thể được xem là quá trình giải quyết bài toán suy luận từ dữ liệu cảm biến sang biểu diễn ký hiệu, trong đó các mô hình toán học và thống kê đóng vai trò trung tâm [6]. Các phương pháp trong thị giác máy tính thường dựa trên sự kết hợp giữa đại số tuyến tính, xác suất thống kê, hình học và tối ưu hóa để xây dựng các thuật toán có khả năng tổng quát hóa tốt trên dữ liệu thực tế [1], [6].

Ngoài vai trò là một lĩnh vực khoa học nghiên cứu nền tảng lý thuyết, thị giác máy tính còn là một ngành kỹ thuật ứng dụng, tập trung vào việc triển khai các mô hình và thuật toán vào các hệ thống thực tế như giám sát giao thông, nhận diện khuôn mặt, xe tự hành và phân tích hình ảnh y tế [7], [8].

Các lĩnh vực quan trọng của thị giác máy tính bao gồm:

- Tái cấu trúc hình cảnh (Scene Reconstruction)
- Phát hiện và theo dõi đối tượng (Object Detection & Tracking)
- Nhận dạng và phân loại (Recognition & Classification)
- Phân đoạn ảnh (Image Segmentation)
- Ước lượng chuyển động (Motion Estimation)
- Phục hồi và nâng cao ảnh (Image Restoration & Enhancement)

Trong bối cảnh đề tài nhận diện biển số xe, thị giác máy tính đóng vai trò nền tảng cho hai nhiệm vụ chính: phát hiện vùng biển số trong ảnh và nhận dạng ký tự trên biển số. Các kỹ thuật hiện đại trong lĩnh vực này, đặc biệt là học sâu (Deep

Learning), đã chứng minh hiệu quả vượt trội so với các phương pháp truyền thống trong môi trường thực tế phức tạp [4].

*Tài liệu tham khảo

- [1] R. Szeliski, Computer Vision: Algorithms and Applications, 2nd ed. Cham, Switzerland: Springer, 2022.
- [2] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach, 2nd ed. Upper Saddle River, NJ, USA: Pearson, 2012.
- [3] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. New York, NY, USA: Pearson, 2018.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2012, pp. 1097–1105.
- [5] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA, USA: W. H. Freeman, 1982.
- [6] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.
- [7] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
- [8] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Hoboken, NJ, USA: Pearson, 2020.

2.1.2. Xử lý ảnh số

Xử lý ảnh (digital image processing) hay xử lý ảnh kỹ thuật số là sử dụng máy tính để xử lý các ảnh kỹ thuật số thông qua một thuật toán. [1] Là một phân ngành tổng thể của xử lý tín hiệu số, xử lý ảnh kỹ thuật số có nhiều lợi thế hơn so với xử lý ảnh analog. Nó cho phép áp dụng nhiều thuật toán hơn cho dữ liệu đầu vào và có thể trách được các vấn đề như sự tích tụ nhiễu và méo hình trong quá trình xử lý. Vì ảnh được định nghĩa trên hai chiều (hoặc nhiều hơn), xử lý ảnh kỹ thuật số có thể được mô hình hóa dưới dạng hệ thống đa chiều. Sự phát triển của xử lý ảnh kỹ thuật số chủ yếu bị ảnh hưởng bởi ba yếu tố: đầu tiên là sự phát triển của máy tính; thứ hai là sự phát triển của toán học; thứ ba là nhu cầu ứng dụng rộng rãi trong môi trường, nông nghiệp, quân sự, công nghiệp và khoa học ý tế ngày càng tăng

*Tài liệu tham khảo

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2018.

2.1.3. Học máy (Machine Learning)

Machine learning là lĩnh vực nghiên cứu các thuật toán cho phép máy tính học từ dữ liệu và cải thiện hiệu suất theo thời gian mà không cần lập trình tường minh cho từng trường hợp cụ thể [1]. Về mặt hình thức, một hệ thống học máy tìm một hàm ánh xạ:

$$f: X \rightarrow Y$$

trong đó X là không gian đặc trưng và Y là không gian đầu ra (nhãn hoặc giá trị dự đoán) [2]

- Biểu diễn dữ liệu và mô hình

Tập dữ liệu huấn luyện được biểu diễn dưới dạng:

$$D = \{(x_i, y_i)\}_{i=1}^n$$

với x_i là vector đặc trưng và y_i là nhãn tương ứng.

Mô hình được tham số hóa bởi θ và mục tiêu là tìm bộ tham số tối ưu:

$$\theta^* = \operatorname{argmin}_\theta L(\theta)$$

trong đó $L(\theta)$ là hàm mất mát đo lường sai số dự đoán [2], [3].

- Hàm mất mát và tối ưu hóa

Hàm mất mát (Loss Function) đóng vai trò trung tâm trong quá trình huấn luyện mô hình. Một số hàm phổ biến gồm:

- Mean Squared Error (MSE) cho bài toán hồi quy
- Cross-Entropy cho bài toán phân loại
- Hinge Loss trong Support Vector Machine (SVM)

Việc tối ưu hóa thường được thực hiện thông qua các thuật toán như Gradient Descent và Stochastic Gradient Descent (SGD), dựa trên đạo hàm của hàm mất mát để cập nhật tham số theo hướng giảm sai số [3].

- Các phương pháp học máy

Học máy được chia thành ba nhóm chính [2]:

- Học có giám sát (Supervised Learning)
- Học không giám sát (Unsupervised Learning)
- Học tăng cường (Reinforcement Learning)

Trong các bài toán thị giác máy tính hiện đại, đặc biệt là nhận diện và phát hiện đối tượng, Deep learning đóng vai trò quan trọng nhờ khả năng tự động học đặc trưng từ dữ liệu lớn [4].

- Nền tảng toán học

Cơ sở lý thuyết của học máy dựa trên:

- Đại số tuyến tính (vector, ma trận)
- Giải tích (đạo hàm, gradient)
- Xác suất thống kê
- Lý thuyết tối ưu

Những nền tảng này cho phép xây dựng các mô hình có khả năng tổng quát hóa và áp dụng hiệu quả trong các hệ thống thực tế [2], [3].

*Tài liệu tham khảo

- [1] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [2] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [3] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

2.1.4. Học sâu (Deep Learning)

Deep learning là một nhánh của Machine learning sử dụng các mạng nơ-ron nhân tạo nhiều tầng (Artificial Neural Networks – ANN) để học biểu diễn dữ liệu ở mức trừu tượng cao [1]. Điểm khác biệt chính của học sâu so với học máy truyền thống là khả năng **tự động trích xuất đặc trưng** từ dữ liệu mà không cần thiết kế thủ công [1], [2].

- Mô hình mạng nơ-ron nhân tạo

Một mạng nơ-ron gồm nhiều lớp (layers), mỗi lớp thực hiện phép biến đổi tuyến tính kết hợp với hàm kích hoạt phi tuyến.

Biểu thức toán học của một lớp:

$$z = Wx + b$$

$$\alpha = \sigma(z)$$

Trong đó:

- x là vector đầu vào
- W là ma trận trọng số
- b là vector bias
- $\sigma(\cdot)$ Là hàm kích hoạt (ReLU, Sigmoid, Tanh)

Việc sử dụng nhiều lớp ẩn giúp mô hình học được các đặc trưng phức tạp hơn [1].

- Lan truyền ngược

Quá trình huấn luyện mạng nơ-ron được thực hiện bằng thuật toán lan truyền ngược (Backpropagation) kết hợp với Gradient Descent nhằm tối thiểu hóa hàm mất mát:

$$\theta = \theta - \eta \nabla L(\theta)$$

Trong đó:

- η là tốc độ học (learning rate)
- $L(\theta)$ là hàm mất mát
- $\nabla L(\theta)$ là gradient

Phương pháp này dựa trên quy tắc đạo hàm chuỗi trong giải tích [2].

- Mạng tích chập (CNN)

Trong lĩnh vực xử lý ảnh, mạng tích chập (Convolutional Neural Network – CNN) được sử dụng phổ biến nhờ khả năng khai thác cấu trúc không gian của ảnh [1].

Phép tích chập được biểu diễn:

$$S(i, j) = (I * K)(i, j)$$

Trong đó:

- I là ảnh đầu vào
- K là kernel (bộ lọc)
- $S(i, j)$ là giá trị sau tích chập

CNN đã đạt hiệu suất vượt trội trong các bài toán phân loại và phát hiện đối tượng [3].

- Vai trò trong nhận diện biển số

Trong hệ thống nhận diện biển số xe:

- CNN được dùng để phát hiện vùng biển số
- Mạng sâu được dùng để nhận dạng ký tự
- Các mô hình hiện đại như YOLO dựa trên kiến trúc CNN

Nhà khả năng học đặc trưng tự động, học sâu cho kết quả chính xác cao trong môi trường thực tế phức tạp

*Tài liệu tham khảo

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

2.1.5. Phát hiện đối tượng

Phát hiện đối tượng (Object Detection) là bài toán trong thị giác máy tính nhằm xác định vị trí và phân loại các đối tượng xuất hiện trong ảnh. Không giống bài toán phân loại ảnh, phát hiện đối tượng yêu cầu đồng thời hai nhiệm vụ: (1) định vị (localization) và (2) phân loại (classification) [1].

- Mô hình toán học

Cho ảnh đầu vào:

$$I \in \mathbb{R}^{H \times W \times C}$$

Hàm phát hiện đối tượng được mô hình hóa như:

$$f(I) = \{(b_i, c_i, s_i)\}_{i=1}^N$$

Trong đó:

- $b_i = (x_i, y_i, w_i, h_i)$ là bounding box
- c_i là nhãn lớp
- s_i là độ tin cậy
- N là một đối tượng phát hiện

Mô hình học tham số θ bằng cách tối thiểu hóa hàm mất mát:

$$\theta^* = \operatorname{argmin}_{\theta} \sum_{i=1}^N L(f(I_i; \theta), y_i)$$

- Hồi qua Bounding Box

Việc định vị đối tượng được thực hiện thông qua hồi quy bounding box. Sai số thường dùng là Smooth L1 Loss [2]:

$$L_{box} = \sum SmoothL1(t_i - \hat{t}_i)$$

Hàm Smooth L1 giúp giảm ảnh hưởng của sai số lớn và tăng tính ổn định khi huấn luyện.

- Độ đo Intersection over Union (IoU)

Chất lượng định vị được đánh giá bằng IoU:

$$IoU = \frac{|B_{pred} \cap B_{gt}|}{|B_{pred} \cup B_{gt}|}$$

Nếu IoU lớn hơn một ngưỡng xác định (ví dụ 0.5), dự đoán được xem là chính xác [3].

- Non-Maximum Suppression (NMS)

Do mô hình có thể dự đoán nhiều bounding box trùng lặp, thuật toán NMS được áp dụng để loại bỏ các hộp dư thừa bằng cách giữ lại hộp có độ tin cậy cao nhất và loại bỏ các hộp có IoU vượt ngưỡng [4].

- Các kiến trúc phát hiện đối tượng

- Tow-Stage Detector
- One-Stage Detector

- Đánh giá mô hình

Hiệu năng mô hình thường được đánh giá bằng mAP (mean Average Precision) [6]:

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c$$

Trong đó C là số lớp và AP_c là Average Precision của lớp c .

*Tài liệu tham khảo

- [1] R. Girshick, “Fast R-CNN,” ICCV, 2015.
- [2] S. Ren et al., “Faster R-CNN,” IEEE TPAMI, 2017.
- [3] J. Redmon et al., “You Only Look Once,” CVPR, 2016.
- [4] J. Redmon and A. Farhadi, “YOLOv3,” arXiv:1804.02767, 2018.

[5] W. Liu et al., “SSD: Single Shot MultiBox Detector,” ECCV, 2016.

[6] T.-Y. Lin et al., “Microsoft COCO,” ECCV, 2014.

2.1.6. Nhận dạng ký tự quang học (OCR)

- Khái niệm

Nhận dạng ký tự quang học (Optical Character Recognition – OCR) là kỹ thuật chuyên đổi hình ảnh chứa văn bản thành dạng ký tự số có thể xử lý bằng máy tính. OCR là một bài toán trong thị giác máy tính kết hợp với học máy [1].

Bài toán OCR được mô hình hóa như một ánh xạ:

$$f: X \rightarrow Y$$

Trong đó:

- X : không gian ảnh đầu vào
- Y : chuỗi ký tự đầu vào

- Mô hình toán học

Cho ảnh văn bản:

$$I \in \mathbb{R}^{H \times W \times C}$$

Mô hình OCR học hàm:

$$\hat{Y} = f(I, \theta)$$

Trong đó:

- θ : tham số mô hình
- $\hat{Y} = (y_1, y_2, \dots, y_T)$: chuỗi ký tự dự đoán

Quá trình huấn luyện tối ưu:

$$\theta^* = \operatorname{argmin}_{\theta} \sum_{i=1}^N L(f(I_i; \theta), y_i)$$

- Các phương pháp OCR

OCR truyền thống

Bao gồm các bước:

1. Tiền xử lý ảnh
2. Phân loại ký tự
3. Trích xuất đặt trưng
4. Phân loại (SVM, KNN, Neural Network)

Hàm phân loại thường dùng Softmax:

$$P(y = k|x) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}}$$

OCR dựa trên Học sâu

Các mô hình hiện đại sử dụng CNN và RNN để nhận dạng chuỗi trực tiếp [2]:

- CNN: trích xuất đặc trưng không gian
- RNN/LSTM: mô hình hóa quan hệ chuỗi

Xác suất chuỗi:

$$P(Y|I) = \prod_{t=1}^T P(y_t|I)$$

Hàm mất mát phô biến:

$$L_{CTC} = -\log P(Y|I)$$

- Đánh giá mô hình OCR

Hàm chỉ số thường dùng:

- Character Error Rate (CER):

$$CER = \frac{S + D + I}{N}$$

- Word Error Rate (WER): Được tính tương tự CER nhưng ở mức từ

*Tài liệu tham khảo

- [1] R. Smith, “An Overview of the Tesseract OCR Engine,” Proc. ICDAR, 2007.
- [2] A. Graves et al., “Connectionist Temporal Classification,” Proc. ICML, 2006.
- [3] B. Shi et al., “CRNN for Image-Based Sequence Recognition,” IEEE TPAMI, 2017.

2.2.Công nghệ sử dụng

- **Python:** được lựa chọn làm ngôn ngữ lập trình chính do hỗ trợ mạnh cho AI, nhiều thư viện học sâu, dễ triển khai và tích hợp
- **OpenCV:** hỗ trợ xử lý ảnh, đọc video, vẽ bounding box, tiền xử lý dữ liệu
- **PyTorch:** framework học sâu được sử dụng để huấn luyện mô hình YOLO, fine-tune mô hình, triển khai inference
- **YOLO:** là mô hình phát hiện đối tượng được sử dụng để phát hiện vị trí biển số xe trong ảnh hoặc video, trích xuất vùng biển số để đưa sang bước nhận dạng ký tự

- **Tesseract OCR hoặc CNNN:** là công cụ nhận dạng ký tự quang học, trong bài toán này Tesseract và CNNN được sử dụng để nhận diện ký tự trên ảnh biển số đã được cắt từ giai đoạn phát hiện

3. Phân tích yêu cầu của dự án

3.1. Chức năng chính

- Phát hiện phương tiện/biển số trong ảnh hoặc video
- Cắt vùng biển số
- Nhận dạng ký tự trên biển số (OCR)
- Hiển thị và lưu kết quả

3.2. Yêu cầu chức năng

- Nhận dữ liệu đầu vào: ảnh, video, ...
- Phát hiện biển số: Sử dụng mô hình YOLO để xác định vị trí biển số, vẽ bounding box quanh biển số
- Nhận dạng ký tự: Cắt vùng biển số, sử dụng OCR để đọc nội dung
- Xuất kết quả: Hiển thị biển số đã nhận dạng và lưu kết quả

3.3. Yêu cầu phi chức năng

- **Độ chính xác:** Tỷ lệ nhận diện đúng $\geq 80\%$
- **Hiệu năng:** Xử lý được tối thiểu 15-20 FPS (video), thời gian xử lý ảnh $< 2s$
- **Tỉnh ổn định**
- **Khả năng mở rộng**

4. Kế hoạch thực hiện dự án

Giai đoạn 1: Nghiên cứu lý thuyết

- Tìm hiểu về: Deep Learning (các lý thuyết liên quan đến bài toán), Object Detection, mô hình YOLOv8
- Tìm hiểu về OCR

Giai đoạn 2: Chuẩn bị dữ liệu

- Thu nhập ảnh biển số
- Gán nhãn bounding box
- Chia train/validation

Giai đoạn 3: Huấn luyện/Fine-tune mô hình

- Fine-tune YOLO với dataset
- Kiểm tra độ chính xác
- Điều chỉnh tham số nếu cần

Giai đoạn 4: Tích hợp hệ thống

Giai đoạn 5: Đánh giá và hoàn thiện

- Đánh giá độ chính xác
- So sánh trước và sau fine-tune
- Hoàn thiện báo cáo
- Chuẩn bị slide bảo vệ