HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

LƯƠNG ĐỨC THUẬN

NGHIÊN CỨU ÁP DỤNG HỌC TỰ GIÁM SÁT CẢI THIỆN PHÂN LOẠI BỆNH THỰC VẬT TRONG ĐIỀU KIỆN THIẾU DỮ LIỆU GÁN NHÃN

CHUYÊN NGÀNH : KHOA HỌC MÁY TÍNH

MÃ SỐ : 8.48.01.01

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC

TS. Hoàng Văn Thành

GS.TS. Từ Minh Phương

HÀ NỘI - 2025

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. Hoàng Văn Thành

(Ghi rõ học hàm, học vị)

Phản biện 1: ……..

Phản biện 2: ……...……...…

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ..... giờ ..... ngày ..... tháng ..... năm .....

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

Trong bối cảnh Cách mạng công nghiệp lần thứ tư, trí tuệ nhân tạo và thị giác máy tính ngày càng đóng vai trò trung tâm trong phát triển nông nghiệp thông minh, đặc biệt ở bài toán giám sát và chẩn đoán bệnh cây trồng. Việc phát hiện sớm bệnh hại có ý nghĩa quyết định đối với năng suất, chất lượng nông sản và giảm tác động môi trường, thúc đẩy chuyển dịch từ canh tác dựa trên kinh nghiệm sang canh tác dựa trên dữ liệu hình ảnh.

Mặc dù các mô hình học sâu cho thấy hiệu quả cao trong nhận diện bệnh thực vật, việc triển khai thực tế vẫn bị hạn chế nghiêm trọng bởi sự khan hiếm dữ liệu gán nhãn chất lượng cao. Quá trình gán nhãn đòi hỏi chuyên môn sâu, chi phí lớn, trong khi dữ liệu ngoài thực địa thường nhiễu, không đồng nhất và mất cân bằng lớp, làm suy giảm khả năng tổng quát hóa của mô hình học giám sát truyền thống.

Trước thực trạng đó, luận văn tập trung nghiên cứu học tự giám sát như một hướng tiếp cận tiềm năng nhằm khai thác hiệu quả dữ liệu hình ảnh không nhãn. Bằng cách ứng dụng các phương pháp học tương phản trên các kiến trúc mạng nơ-ron tích chập tối ưu tài nguyên và thực nghiệm trên các bộ dữ liệu như PlantDoc, nghiên cứu hướng tới nâng cao hiệu quả phân loại bệnh trong điều kiện thiếu dữ liệu gán nhãn. Trên cơ sở đó, luận văn thực hiện đề tài "Nghiên cứu áp dụng học tự giám sát cải thiện phân loại bệnh thực vật trong điều kiện thiếu dữ liệu gán nhãn".

Nội dung đề án sẽ bao gồm phần mở đầu, phần kết luận và các chương nội dung được tổ chức như sau:

CHƯƠNG 1: TỔNG QUAN VỀ BÀI TOÁN PHÂN LOẠI BỆNH THỰC VẬT TRONG ĐIỀU KIỆN THIẾU DỮ LIỆU GÁN NHÃN

- Trình bày về tầm quan trọng và các ứng dụng của thị giác máy tính trong nông nghiệp nói chung và bài toán phân loại bệnh thực vật nói riêng.
- Phân tích sâu về sự khan hiếm dữ liệu được gán nhãn. Nêu bật tác động của chúng đến hiệu suất mô hình.
- Giới thiệu tổng quan về các giải pháp tiềm năng, dẫn dắt đến sự cần thiết của áp dụng học Tự Giám Sát.

- Tóm tắt các vấn đề và xác định mục tiêu cần giải quyết trong các chương tiếp theo.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

- Các thành phần chính của mạng nơ-ron tích chập
- Các mô hình CNN nổi tiếng phù hợp với bài toán phân loại bệnh thực vật.
- Học Tự Giám Sát (SSL) nguyên lý và kiến trúc chung và các kĩ thuật liên quan.

CHƯƠNG 3: ĐỀ XUẤT MÔ HÌNH ÁP DỤNG SSL CHO BÀI TOÁN TOÁN PHÂN LOẠI BỆNH THỰC VẬT TRONG ĐIỀU KIỆN THIẾU DỮ LIỆU GÁN NHÃN

- Dựa trên backbone có ít tham số, phù hợp với bài toán như mobilenet v2.
- Thực nghiệm các kĩ thuật học tự giám sát
- Xây dựng mô hình đề xuất
- Luận giải tại sao việc áp dụng phương pháp này lại mang lại lợi ích vượt trội. học tự giám sát cung cấp điểm khởi đầu tốt cho các đặc trưng
- Mô tả chi tiết các thành phần, luồng dữ liệu và hàm mất mát của mô hình.

CHƯƠNG 4: THỰC NGHIỆM VÀ ĐÁNH GIÁ

- Mục tiêu thực nghiệm để kiểm chứng hiệu quả của mô hình.
- Trình bày các kịch bản thực nghiệm
- Phân tích, so sánh, trực quan hoá và bình luận về kết quả, làm rõ ưu thế của mô hình đề xuất.

CHƯƠNG 1: TỔNG QUAN VỀ BÀI TOÁN PHÂN LOẠI BỆNH THỰC VẬT TRONG ĐIỀU KIỆN THIẾU DỮ LIỆU GÁN NHÃN

Sự phát triển của nông nghiệp thông minh đã thúc đẩy ứng dụng thị giác máy tính trong chẩn đoán bệnh cây trồng, nơi các mô hình học sâu cho thấy hiệu quả nổi bật. Tuy nhiên, việc phụ thuộc vào dữ liệu gán nhãn thủ công đang hạn chế khả năng mở rộng và triển khai thực tế của các hệ thống này do chi phí và thời gian thu thập lớn. Tình trạng khan hiếm dữ liệu định danh, đặc biệt với các bệnh hiếm và biến thể phức tạp, làm suy giảm hiệu suất của các phương pháp phân loại truyền thống. Trước bối cảnh đó, Chương 1 tập trung phân tích các thách thức từ thiếu hụt dữ liệu gán nhãn và giới thiệu học tự giám sát như một hướng tiếp cận trọng tâm nhằm khai thác dữ liệu không nhãn, qua đó làm rõ mục tiêu, phạm vi và đối tượng nghiên cứu của luận văn.

1.1. Thị giác máy tính và vai trò trong nông nghiệp chính xác

Các phương pháp giám sát và chẩn đoán bệnh cây trồng truyền thống ngày càng bộc lộ hạn chế trước bối cảnh biến đổi khí hậu và sự gia tăng của các chủng bệnh mới. Việc dựa vào quan sát thủ công không chỉ phụ thuộc nhiều vào kinh nghiệm cá nhân mà còn tốn kém nguồn lực và khó áp dụng hiệu quả trên quy mô lớn, dẫn đến nguy cơ phát hiện bệnh chậm và gia tăng chi phí can thiệp.

Trong bối cảnh đó, các hệ thống giám sát và chẩn đoán bệnh cây trồng tự động được xem là giải pháp phù hợp. Khả năng theo dõi liên tục trên diện rộng giúp phát hiện sớm bệnh, giảm sai lệch do con người và cung cấp dữ liệu đáng tin cậy phục vụ các biện pháp xử lý kịp thời, qua đó nâng cao hiệu quả sản xuất và thúc đẩy phát triển nông nghiệp thông minh, bền vững.

Trong lĩnh vực nhận diện bệnh thực vật, thị giác máy tính đã chuyển từ các phương pháp trích xuất đặc trưng thủ công sang các mô hình học sâu. Mạng nơ-ron tích chập cho phép tự động học biểu diễn đặc trưng từ ảnh thô, giúp hệ thống thích

nghi tốt hơn với điều kiện tự nhiên phức tạp và đạt độ chính xác cao trong phân biệt các bệnh có biểu hiện tương đồng.

Tuy nhiên, các phương pháp học giám sát hiện nay phụ thuộc lớn vào dữ liệu gán nhãn quy mô lớn, gây khó khăn khi mở rộng sang cây trồng mới hoặc các bệnh hiếm gặp. Hạn chế này tạo động lực cho việc nghiên cứu các hướng tiếp cận nhằm khai thác hiệu quả nguồn dữ liệu hình ảnh chưa được gán nhãn.

1.2. Tổng quan bài toán phân loại bệnh thực vật trong điều kiện thiếu dữ liệu gán nhãn

Trong thực tế phân loại bệnh thực vật, việc xây dựng tập dữ liệu huấn luyện chất lượng cao gặp nhiều khó khăn do yêu cầu nhãn chính xác phụ thuộc vào chuyên môn sâu về nông học. Quá trình thu thập và gán nhãn vì vậy tiêu tốn nhiều thời gian và chi phí, dẫn đến tình trạng khan hiếm dữ liệu, đặc biệt đối với cây trồng bản địa và các loại bệnh ít gặp. Bên cạnh đó, dữ liệu thu thập ngoài thực địa thường chịu ảnh hưởng của nhiễu nền, điều kiện ánh sáng không ổn định và sự không đồng nhất về hình thức, làm suy giảm tính đại diện của tập huấn luyện.

Sự thiếu hụt dữ liệu gán nhãn tác động trực tiếp đến hiệu suất của các mô hình học sâu. Khi quy mô dữ liệu không tương xứng với độ phức tạp của mô hình, nguy cơ quá khớp gia tăng và mô hình có xu hướng học các yếu tố ngẫu nhiên thay vì các đặc trưng bệnh học cốt lõi. Đồng thời, biên phân tách giữa các lớp trong không gian đặc trưng trở nên kém ổn định, dễ bị chi phối bởi các yếu tố ngoại cảnh, làm giảm độ chính xác và khả năng thích nghi với dữ liệu mới.

Trong bối cảnh đó, bài toán phân loại bệnh thực vật dưới điều kiện thiếu dữ liệu gán nhãn trở nên đặc biệt thách thức do biến thiên nội lớp lớn và sự tương đồng cao giữa các lớp bệnh. Tình trạng mất cân bằng dữ liệu càng làm gia tăng thiên lệch của các phương pháp học giám sát truyền thống, khiến các bệnh hiếm nhưng nguy hiểm khó được nhận diện chính xác. Những hạn chế này cho thấy sự cần thiết phải nghiên cứu các hướng tiếp cận mới, tận dụng dữ liệu không nhãn để học biểu diễn hiệu quả hơn và giảm phụ thuộc vào giám sát thủ công.

1.3. Các nghiên cứu liên quan gần đây

1.3.1. Các phương pháp nghiên cứu truyền thống trước đây

Trong giai đoạn đầu, nhận diện bệnh thực vật chủ yếu dựa trên các phương pháp học máy truyền thống với đặc trưng thủ công, vốn phụ thuộc nhiều vào tri thức chuyên gia và chỉ hiệu quả trong môi trường kiểm soát, dẫn đến khả năng tổng quát hóa hạn chế trong điều kiện tự nhiên.

Sự phát triển của mạng nơ-ron tích chập đã chuyển hướng sang học biểu diễn đầu cuối, giúp nâng cao độ chính xác đáng kể nhưng đồng thời làm gia tăng sự phụ thuộc vào dữ liệu gán nhãn quy mô lớn, khiến mô hình dễ quá khớp khi dữ liệu huấn luyện hạn chế.

Học chuyển giao và các phương pháp kết hợp góp phần giảm yêu cầu dữ liệu và cải thiện hiệu năng, song vẫn chịu ảnh hưởng bởi sự khác biệt miền dữ liệu và chưa giải quyết triệt để bài toán thiếu nhãn.

Trong bối cảnh đó, học tự giám sát nổi lên như hướng tiếp cận hiệu quả, cho phép khai thác dữ liệu không nhãn để học biểu diễn đặc trưng bền vững, từ đó giảm phụ thuộc vào nhãn thủ công và cải thiện khả năng tổng quát hóa trong môi trường canh tác thực tế.

1.3.2. Phương pháp áp dụng học tự giám sát

Học tự giám sát là hướng tiếp cận hiệu quả để học biểu diễn đặc trưng từ dữ liệu ảnh không gán nhãn thông qua các tác vụ tiền đề, góp phần giảm phụ thuộc vào gán nhãn thủ công và nâng cao khả năng tổng quát hóa trong các bài toán thị giác nông nghiệp.

Các nghiên cứu gần đây cho thấy việc tiền huấn luyện mô hình trên dữ liệu nông nghiệp chưa gán nhãn giúp cải thiện đáng kể hiệu suất phân loại bệnh thực vật, thậm chí vượt trội so với các phương pháp học giám sát truyền thống, đồng thời tăng khả năng tập trung vào vùng bệnh quan trọng.

Tuy nhiên, khi triển khai trên các kiến trúc mạng nhẹ cho thiết bị biên, hiệu quả của học tự giám sát thường suy giảm nếu thiếu kỹ thuật tối ưu phù hợp. Điều này cho thấy nhu cầu nghiên cứu các chiến lược học tự giám sát trực tiếp cho mạng nhẹ nhằm xây dựng hệ thống chẩn đoán bệnh thực vật hiệu quả và tiết kiệm tài nguyên.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1. Các thành phần chính của mạng nơ-ron tích chập

Một mạng nơ-ron tích chập bao gồm lần lượt một lớp đầu vào, nhiều lớp ẩn, và một lớp đầu ra. Lớp đầu vào sẽ khác nhau tùy theo các ứng dụng cụ thể. Các lớp ẩn, được coi là khối lõi của kiến trúc CNN, bao gồm một chuỗi các lớp tích chập, lớp gộp và cuối cùng xuất đầu ra thông qua lớp kết nối đầy đủ.

2.2. Các mô hình CNN nổi tiếng phù hợp với bài toán phân loại bệnh thực vật.

2.2.1. Quá trình huấn luyện mô hình học sâu

Quy trình huấn luyện bắt đầu từ việc thu thập và chuẩn hóa dữ liệu ảnh bệnh thực vật, kết hợp tăng cường dữ liệu nhằm bù đắp sự hạn chế về số lượng mẫu gán nhãn và phản ánh tốt hơn sự đa dạng trong điều kiện thực tế. Tiếp theo, mô hình học sâu được thiết kế và huấn luyện với kiến trúc phù hợp, đi kèm tinh chỉnh siêu tham số và các kỹ thuật điều chuẩn để nâng cao khả năng hội tụ và giảm quá khớp. Sau huấn luyện, mô hình được đánh giá bằng các chỉ số hiệu năng, phân tích lỗi và kiểm chứng trên dữ liệu thực địa, trước khi triển khai và giám sát nhằm đảm bảo hiệu quả chẩn đoán ổn định theo thời gian.

2.2.2. Các mô hình CNN nổi tiếng phù hợp

Trong bối cảnh nông nghiệp thông minh với tài nguyên phần cứng hạn chế, các mô hình chẩn đoán bệnh cây trồng cần đồng thời đảm bảo độ chính xác và hiệu quả tính toán, thúc đẩy việc ưu tiên các kiến trúc mạng tích chập tối ưu tài nguyên cho triển khai trên thiết bị biên.

MobileNet và EfficientNet là những kiến trúc tiêu biểu đáp ứng yêu cầu này nhờ khả năng giảm chi phí tính toán trong khi vẫn duy trì năng lực trích xuất đặc trưng, đặc biệt phù hợp cho các kịch bản dữ liệu gán nhãn hạn chế và học tự giám sát.

Bên cạnh đó, các biến thể gọn nhẹ của ResNet tiếp tục đóng vai trò kiến trúc tham chiếu ổn định, tạo cơ sở đánh giá hiệu quả của học tự giám sát, đồng thời nhấn mạnh rằng cơ chế huấn luyện, đặc biệt là khai thác dữ liệu không nhãn, có vai trò then chốt đối với hiệu năng tổng thể của mô hình.

2.3. Học tự giám sát: nguyên lý và kiến trúc chung và các kĩ thuật liên quan

2.3.1. Định nghĩa

Học tự giám sát là mô thức học máy cho phép mô hình tự sinh tín hiệu huấn luyện từ cấu trúc nội tại của dữ liệu, nhằm học các biểu diễn đặc trưng giàu ngữ nghĩa mà không cần nhãn thủ công.

Quá trình học thường gồm giai đoạn tiền huấn luyện với các tác vụ tiền đề và giai đoạn tinh chỉnh cho nhiệm vụ hạ nguồn, giúp tận dụng hiệu quả dữ liệu không nhãn và giảm phụ thuộc vào dữ liệu gán nhãn hạn chế.

Nhờ khai thác trực tiếp các mối quan hệ nội tại của dữ liệu, học tự giám sát thúc đẩy mô hình học các đặc trưng mức cao, tạo nền tảng quan trọng cho các bài toán thị giác máy tính trong điều kiện thiếu dữ liệu gán nhãn.

2.3.2. Phân loại

Các phương pháp học tự giám sát trong thị giác máy tính được chia thành ba hướng chính gồm tạo sinh, học tương phản và học không tương phản, phản ánh các chiến lược khác nhau trong việc học biểu diễn đặc trưng từ dữ liệu không gán nhãn.

Nhóm phương pháp tạo sinh tập trung vào tái tạo dữ liệu đầu vào hoặc phần bị che khuất, cho phép học phân bố dữ liệu nhưng thường có chi phí tính toán cao và hiệu quả hạn chế với bài toán phân loại.

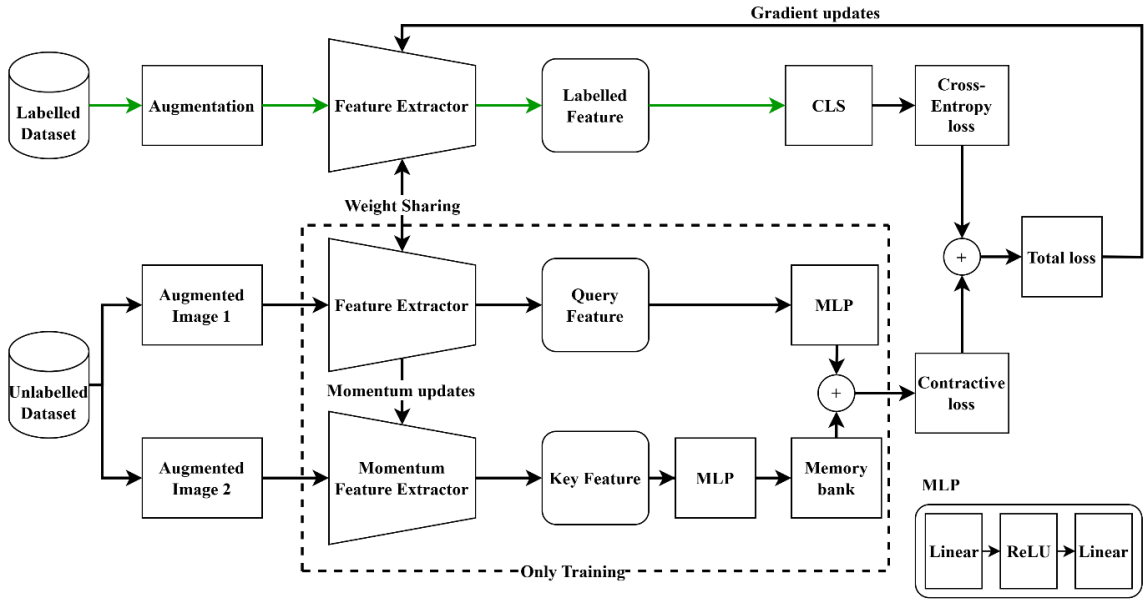
Nhóm học tương phản hướng tới xây dựng không gian biểu diễn giàu ngữ nghĩa bằng cách kéo gần các biến thể của cùng một mẫu và tách xa các mẫu khác, trong đó MoCo v2 thể hiện ưu thế về hiệu quả và tính ổn định trong điều kiện hạn chế dữ liệu và tài nguyên.

Nhóm học không tương phản loại bỏ nhu cầu sử dụng mẫu âm thông qua kiến trúc bất đối xứng, đạt hiệu năng tốt nhưng thường nhạy cảm với thiết lập huấn luyện và kém ổn định trong các bài toán chuyên biệt.

CHƯƠNG 3: ĐỀ XUẤT MÔ HÌNH ÁP DỤNG SSL CHO BÀI TOÁN PHÂN LOẠI BỆNH THỰC VẬT TRONG ĐIỀU KIỆN THIẾU DỮ LIỆU GÁN NHÃN

3.1. Đề xuất kiến trúc mô hình

Trong bài toán phân loại bệnh thực vật, thách thức cốt lõi không nằm ở độ phức tạp của kiến trúc mạng mà ở sự khan hiếm dữ liệu gán nhãn chất lượng cao, khiến các mô hình học giám sát truyền thống dễ rơi vào quá khớp và suy giảm khả năng tổng quát hóa. Trước hạn chế này, luận văn đề xuất một kiến trúc lai ghép kết hợp đồng thời học giám sát và học tự giám sát nhằm khai thác tối đa thông tin từ dữ liệu sẵn có. Cách tiếp cận này hướng mô hình tập trung học các đặc trưng bất biến và giàu ngữ nghĩa thông qua cơ chế học tương phản, từ đó nâng cao hiệu quả phân loại ngay cả khi số lượng mẫu gán nhãn hạn chế.



Hình 2: Mô hình đề xuất

Cụ thể, mô hình tiếp nhận song song dữ liệu có nhãn và dữ liệu phục vụ học tự giám sát, được tăng cường dữ liệu trước khi đưa vào mạng Backbone chung. Từ không gian đặc trưng trích xuất, nhánh học giám sát sử dụng lớp phân loại để tính toán hàm mất mát dựa trên nhãn bệnh, trong khi nhánh học tự giám sát sử dụng mạng chiếu kết hợp cơ chế cập nhật động lượng và bộ nhớ đệm để tối ưu hóa hàm mất mát

tự giám sát. Toàn bộ hệ thống được huấn luyện thông qua hàm mất mát tổng hợp, đảm bảo sự cân bằng giữa độ chính xác phân loại và chất lượng biểu diễn đặc trưng, qua đó cải thiện độ ổn định và khả năng tổng quát hóa của mô hình trong điều kiện thiếu dữ liệu gán nhãn.

3.2. Xây dựng dữ liệu

3.2.1. Tập dữ liệu PlantDoc

Luận văn sử dụng tập dữ liệu PlantDoc để đánh giá mô hình trong bối cảnh triển khai thực tế, với dữ liệu được thu thập ngoài trời và có mức độ phức tạp cao hơn đáng kể so với các tập dữ liệu trong môi trường phòng thí nghiệm.

PlantDoc bao gồm hơn hai nghìn hình ảnh thuộc nhiều loại cây trồng và trạng thái bệnh khác nhau, được thu thập từ nhiều nguồn nhằm đảm bảo sự đa dạng về điều kiện chụp và bối cảnh hình ảnh, phản ánh sát thực tế sản xuất nông nghiệp.

Với đặc trưng phi cấu trúc, nhiễu nền lớn và số lượng mẫu gán nhãn hạn chế theo lớp, PlantDoc đặt ra thách thức rõ rệt cho các phương pháp học giám sát, đồng thời là cơ sở phù hợp để đánh giá hiệu quả của học tự giám sát trong việc cải thiện biểu diễn đặc trưng và khả năng tổng quát hóa của mô hình.

3.2.2. Xây dựng tập dữ liệu SSL

Luận văn xây dựng tập dữ liệu học tự giám sát theo nguyên tắc bảo đảm tính đa dạng hình thái và sự tương đồng ngữ nghĩa với miền dữ liệu mục tiêu, thông qua việc tổng hợp và lọc dữ liệu ngoài thực địa từ nhiều nguồn khác nhau nhằm mở rộng không gian mẫu và hỗ trợ học biểu diễn đặc trưng bền vững.

Các bộ dữ liệu PlantWild, PlantSeg và FieldPlant được khai thác theo vai trò bổ trợ, lần lượt cung cấp dữ liệu quy mô lớn ngoài tự nhiên, thông tin không gian chi tiết của vùng bệnh và bối cảnh thực địa phức tạp, qua đó giúp mô hình học được biểu diễn giàu tính khái quát và khả năng thích nghi khi chuyển miền.

Quy trình lọc dữ liệu không nhãn được thực hiện trong một không gian đặc trưng chung, sử dụng các nguyên mẫu lớp trích xuất từ PlantDoc làm điểm neo ngữ nghĩa, cho phép đánh giá mức độ phù hợp của từng mẫu ứng viên dựa trên độ tương đồng cosin.

Dựa trên các ngưỡng tương đồng khác nhau, luận văn xây dựng nhiều tập dữ liệu học tự giám sát với quy mô tăng dần, tạo cơ sở để phân tích sự đánh đổi giữa chất lượng và số lượng dữ liệu, đồng thời đánh giá tác động của từng lựa chọn đến hiệu quả huấn luyện trong bối cảnh thiếu dữ liệu gán nhãn.

Data size	Image number	Similarity	Max similarity	Min similarity
Small	3088	0.6338	0.8664	0.4422
Medium	3989	0.6068	0.8664	0.3672
Large	5621	0.5851	0.8664	0.3116

Bảng 1: Tập dữ liệu cho quá trình học tự giám sát

3.3. Chiến lược tăng cường dữ liệu

Trong bối cảnh thiếu hụt dữ liệu gán nhãn cho bài toán phân loại bệnh thực vật, luận văn xây dựng một chiến lược tăng cường dữ liệu kép nhằm nâng cao khả năng hội tụ và tính tổng quát hóa của mô hình. Hai luồng tăng cường được thiết kế tách biệt cho học giám sát và học tự giám sát, bảo đảm dữ liệu có nhãn giữ được thông tin bệnh học cốt lõi, trong khi dữ liệu không nhãn được biến đổi mạnh để thúc đẩy việc học các đặc trưng bất biến.

Đối với tập dữ liệu có nhãn Plantdoc, quá trình tăng cường tập trung vào mở rộng tập mẫu huấn luyện và hạn chế quá khớp nhưng không làm sai lệch đặc trưng bệnh lý. Các phép biến đổi hình học và quang học ở mức vừa phải được áp dụng nhằm giúp mô hình thích nghi với sự đa dạng về vị trí, hướng lá và điều kiện chiếu sáng trong thực tế, đồng thời duy trì tính ổn định của đặc trưng phân loại.

Ngược lại, với tập dữ liệu không gán nhãn phục vụ học tự giám sát, tăng cường dữ liệu được thiết kế theo hướng gây nhiễu mạnh và đa dạng. Các phép biến đổi ngẫu nhiên được áp dụng độc lập để tạo ra các cặp mẫu tương phản từ cùng một ảnh, buộc mô hình tập trung vào cấu trúc và ngữ nghĩa tổng thể thay vì chi tiết cục bộ. Cách tiếp cận này giúp cải thiện chất lượng biểu diễn đặc trưng và tạo nền tảng hiệu quả cho giai đoạn tinh chỉnh có giám sát sau đó.

3.4. Thiết kế thành phần mạng nơ-ron

Luận văn tập trung vào bài toán xử lý ảnh trong điều kiện triển khai thực tế hạn chế tài nguyên, đặc biệt trên thiết bị biên trong nông nghiệp. Trước yêu cầu về độ trễ thấp, chi phí triển khai thấp và khả năng hoạt động độc lập với hạ tầng mạng, kiến trúc mạng được lựa chọn theo hướng gọn nhẹ nhưng vẫn bảo đảm năng lực trích xuất đặc trưng hiệu quả.

Trên cơ sở đó, ResNet10T được sử dụng làm mạng xương sống nhờ cấu trúc residual rút gọn, số tham số thấp và cơ chế giảm kích thước hợp lý. Kiến trúc này cho phép mô hình đạt tốc độ suy luận nhanh, duy trì tính ổn định của đặc trưng và phù hợp cho các bài toán thị giác máy tính nông nghiệp trên thiết bị biên.

Để phục vụ học tự giám sát, khối chiếu phi tuyến MLP được thiết kế nhằm ánh xạ đặc trưng từ backbone sang không gian biểu diễn phù hợp cho học tương phản. Cấu trúc mở rộng rồi thu hẹp kết hợp kích hoạt phi tuyến giúp ổn định quá trình huấn luyện, cải thiện tính phân tách của đặc trưng và nâng cao hiệu quả cho giai đoạn tinh chỉnh có giám sát sau đó.

Cuối cùng, bộ phân loại CLS được xây dựng dưới dạng một lớp tuyến tính đơn giản, trực tiếp ánh xạ đặc trưng sang không gian nhãn. Thiết kế tối giản này giúp giảm số tham số, hạn chế overfitting trong điều kiện dữ liệu hạn chế, đồng thời đảm bảo sự tách bạch rõ ràng giữa giai đoạn học biểu diễn và giai đoạn ra quyết định trong bài toán phân loại bệnh thực vật.

3.5. Cơ chế huấn luyện và cập nhập trọng số

Luận văn đề xuất một kiến trúc mạng lai với chiến lược huấn luyện kết hợp giữa học giám sát và học tự giám sát thông qua hai cơ chế cập nhật trọng số khác nhau. Thiết kế này nhằm khai thác hiệu quả dữ liệu không nhãn trong khi vẫn bảo đảm khả năng thích ứng với nhiệm vụ phân loại bệnh thực vật khi số lượng mẫu gán nhãn còn hạn chế.

Nhánh mạng chính được cập nhật trực tiếp bằng lan truyền ngược dựa trên hàm mất mát tổng hợp, kết hợp tín hiệu phân loại từ dữ liệu PlantDoc và tín hiệu học tương phản. Cách tiếp cận này cho phép mô hình đồng thời học các đặc trưng phân biệt lớp và các đặc trưng bất biến, góp phần nâng cao hiệu quả phân loại.

Nhánh mạng phụ trợ được cập nhật theo cơ chế động lượng nhằm duy trì tính ổn định của không gian biểu diễn trong quá trình học tương phản. Việc sử dụng mạng động lượng cùng bộ nhớ đệm hàng đợi giúp mở rộng không gian so sánh đặc trưng, giảm phụ thuộc vào kích thước lô và cải thiện khả năng tổng quát hóa của mô hình trong bối cảnh dữ liệu gán nhãn khan hiếm.

3.6. Xây dựng hàm mất mát

Để giải quyết đồng thời việc phân loại bệnh thực vật chính xác và bài toán học biểu diễn đặc trưng mạnh mẽ trong điều kiện thiếu dữ liệu, luận văn xây dựng một hàm mục tiêu tổng hợp. Hàm mục tiêu này là sự kết hợp tuyến tính giữa hai thành phần: hàm mất mát giám sát và hàm mất mát tự giám sát. Việc thiết kế hàm mất mát đa nhiệm này đóng vai trò như một cơ chế điều hướng, buộc mô hình vừa phải tối ưu hóa độ chính xác phân loại, vừa phải duy trì cấu trúc tô pô của không gian đặc trưng sao cho các biến thể của cùng một loại bệnh luôn nằm gần nhau.

3.6.1. Hàm mất mát phân loại với kỹ thuật làm trơn nhãn

Đối với nhánh giám sát, mục tiêu là giảm thiểu sai số giữa phân phối dự đoán và nhãn thực tế của dữ liệu PlantDoc. Thay vì sử dụng hàm mất mát Entropy chéo tiêu chuẩn, luận văn áp dụng Entropy chéo làm trơn nhãn. Trong dữ liệu bệnh thực vật thường chứa nhiễu, nhãn cứng dễ khiến mô hình quá tự tin vào dự đoán, dẫn đến hiện tượng quá khớp. Do đó, tham số làm trơn được đặt bằng 0.1, phân phối một phần xác suất nhỏ cho các lớp còn lại thay vì tập trung hoàn toàn vào lớp đúng.

Hàm mất mát này khuyến khích mô hình học các đặc trưng tổng quát hơn, thay vì ghi nhớ máy móc các mẫu huấn luyện, từ đó cải thiện khả năng thích nghi trên tập kiểm thử.

3.6.2. Hàm mất mát tương phản InfoNCE cho học tự giám sát

Đối với nhánh tự giám sát, luận văn sử dụng hàm mất mát InfoNCE. Mục tiêu của hàm này là tối đa hóa độ tương đồng giữa véc-tơ truy vấn q và véc-tơ khóa dương $k+$, đồng thời cực tiểu hóa độ tương đồng giữa q và tất cả các véc-tơ khóa âm $k-$ được lưu trữ trong bộ nhớ đệm. Trong không gian đặc trưng, phép đo độ tương đồng được thực hiện thông qua tích vô hướng giữa hai véc-tơ đã chuẩn hóa. Để kiểm soát

độ sắc nhọn của phân phối xác suất, luận văn sử dụng một tham số nhiệt độ $\tau = 0.07$, mô hình phải tìm ra đúng mẫu dương duy nhất trong số hàng nghìn mẫu âm tính lấy từ hàng đợi, buộc mạng Backbone phải trích xuất được các đặc trưng phân biệt cốt lõi của lá bệnh bất chấp các biến đổi về màu sắc hay hình học.

$$\mathcal{L}_{MoCov2} = -\log \frac{\exp\left((q \cdot k^{+})/\tau\right)}{\exp\left((q \cdot k^{+})/\tau\right) + \sum_{j=1}^{K} \exp\left(\left(q \cdot k_j^{-}\right)/\tau\right)}$$

3.6.3. Hàm mất mát tổng hợp và chiến lược trọng số động

Hàm mất mát toàn cục giúp cập nhật trọng số cho mạng là tổng có trọng số của hai thành phần nêu trên, được biểu diễn qua công thức $\mathcal{L}_{total} = \alpha \cdot \mathcal{L}_{SSL} + \beta \cdot \mathcal{L}_{CLS}$. Trong đó, $\mathcal{L}_{CLS}$ là hàm mất mát cho nhánh phân loại, $\mathcal{L}_{SSL}$ là hàm mất mát cho nhánh học tự giám sát. β là trọng số của thành phần phân loại, thường được giữ cố định ở giá trị 1.0 nhằm đảm bảo mục tiêu chính của bài toán luôn được ưu tiên. Ngược lại, hệ số α điều khiển mức độ đóng góp của quá trình tự giám sát để mô hình tận dụng mạnh mẽ các đặc trưng từ học tự giám sát nhằm định hình không gian biểu diễn.

3.7. Các kỹ thuật liên quan

Luận văn sử dụng kỹ thuật tích lũy gradient để khắc phục hạn chế về kích thước lô huấn luyện trong học tương phản, cho phép mô phỏng batch size lớn hơn mà không vượt quá giới hạn bộ nhớ. Cách tiếp cận này giúp duy trì số lượng mẫu âm cần thiết và đảm bảo quá trình huấn luyện học tự giám sát diễn ra ổn định.

Cơ chế Automatic Mixed Precision được áp dụng nhằm tối ưu hiệu năng tính toán và giảm tiêu thụ bộ nhớ GPU. Việc kết hợp các phép toán độ chính xác thấp với duy trì trọng số ở độ chính xác cao giúp tăng tốc huấn luyện trong khi vẫn bảo đảm khả năng hội tụ và chất lượng biểu diễn đặc trưng.

Bên cạnh đó, chiến lược điều chuẩn suy giảm trọng số được tích hợp cùng bộ tối ưu hóa AdamW nhằm hạn chế quá khớp và hướng mô hình tới các nghiệm ổn định hơn. Biện pháp này góp phần cải thiện khả năng tổng quát hóa và nâng cao hiệu năng phân loại bệnh thực vật trong điều kiện dữ liệu gán nhãn hạn chế.

CHƯƠNG 4: THỰC NGHIỆM VÀ ĐÁNH GIÁ

4.1. Cấu hình phần cứng

Giai đoạn đầu tiên bao gồm xây dựng tập dữ liệu tự giám sát và tuyển chọn mẫu dựa trên nguyên mẫu, thực hiện trên máy tính cấu hình cơ bản với Intel Core i5 1135G7 và 8 GB RAM. Việc lọc và chuẩn hóa dữ liệu vận hành trơn tru trên cấu hình này cho thấy chi phí tính toán thấp, cho phép dễ dàng mở rộng hoặc cập nhật dữ liệu đầu vào. Ngược lại, huấn luyện mô hình lai ghép yêu cầu khả năng xử lý song song cao để tính toán ma trận và tích chập, duy trì bộ nhớ và gradient cho cả nhánh giám sát lẫn tự giám sát. Do đó, các thử nghiệm chính thức được triển khai trên nền tảng cloud với GPU P100 16 GB, đảm bảo kích thước lô dữ liệu đủ lớn, hội tụ ổn định hàm mất mát tương phản và rút ngắn thời gian huấn luyện trên dữ liệu phức tạp.

4.2. Các tiêu chí đánh giá

Việc đánh giá hiệu năng là nền tảng để xác định tính khả thi và độ tin cậy của mô hình học sâu trong phân loại bệnh thực vật, đặc biệt khi dữ liệu gán nhãn hạn chế. Các tiêu chí định lượng gồm Accuracy, Recall, Precision và F1-score được sử dụng để thẩm định mô hình. Accuracy phản ánh tỷ lệ phân loại đúng tổng thể, Recall đo khả năng phát hiện đầy đủ các mẫu bệnh, Precision đánh giá độ tin cậy của các cảnh báo dương tính, trong khi F1-score cung cấp thước đo tổng hợp cân bằng giữa Precision và Recall, đặc biệt quan trọng với dữ liệu mất cân bằng.

Accuracy tính tỷ lệ dự đoán đúng trên toàn bộ tập dữ liệu, bao gồm cả True Positive và True Negative, trong khi False Positive và False Negative phản ánh các lỗi cảnh báo nhầm và bỏ sót. Recall tập trung vào khả năng phát hiện các mẫu bệnh, giảm thiểu bỏ sót có thể dẫn đến thiệt hại kinh tế và sinh học. Precision đảm bảo các dự đoán dương tính đáng tin cậy, hạn chế cảnh báo giả, tiết kiệm chi phí và bảo vệ môi trường. F1-score dung hòa giữa Recall và Precision, phản ánh năng lực toàn diện của mô hình, đảm bảo vừa phát hiện hiệu quả bệnh lý vừa tối ưu hóa chi phí xử lý. Tiêu chí này được sử dụng làm cơ sở so sánh giữa các kiến trúc mạng để đánh giá hiệu quả thực tế trong kiểm soát dịch hại nông nghiệp.

4.3. Kết quả thực nghiệm

Trên nền tảng môi trường tính toán và hệ thống tiêu chí đánh giá đã thiết lập, phần này trình bày và phân tích các kết quả định lượng từ quá trình huấn luyện và kiểm thử mô hình. Mục tiêu chính không chỉ là so sánh các chỉ số độ chính xác, mà còn kiểm chứng vai trò của học tự giám sát trong việc nâng cao chất lượng không gian biểu diễn đặc trưng khi dữ liệu gán nhãn hạn chế.

Trong giai đoạn thử nghiệm, các phương pháp học tự giám sát như SimCLR, MoCo v2, MAE và mô hình học giám sát thuần túy được đánh giá trên cùng bộ dữ liệu PlantDoc. Các pretraining này được sử dụng làm đầu vào cho tác vụ hạ nguồn. Với mỗi phương pháp tiền huấn luyện, mô hình được thử nghiệm theo hai giao thức chính. Giao thức thứ nhất giữ cố định backbone và chỉ huấn luyện bộ phân loại trên dữ liệu gán nhãn, nhằm đánh giá chất lượng biểu diễn học được. Giao thức thứ hai fine-tune toàn bộ mô hình end-to-end trên tập dữ liệu downstream, khảo sát khả năng thích nghi và cải thiện biểu diễn khi tham số mô hình được cập nhật đầy đủ.

Số liệu thực nghiệm tại Bảng 2 cho thấy hiệu năng phân loại phụ thuộc chặt chẽ vào sự tương thích giữa phương pháp tiền huấn luyện và chiến lược tinh chỉnh hạ nguồn. Mô hình học giám sát truyền thống đạt độ chính xác tham chiếu 65.25%. Trong các phương pháp học tự giám sát, MoCo v2 thể hiện tính ổn định và hiệu quả vượt trội, đạt 64.87% khi tinh chỉnh toàn bộ, gần bằng mức tham chiếu, nhờ cơ chế cập nhật động lượng duy trì các đặc trưng chất lượng cao ngay cả khi dữ liệu gán nhãn hạn chế. Ngược lại, MAE bộc lộ sự phân hóa rõ rệt: chỉ đạt 28.81% ở chế độ đóng băng tham số do tập trung vào tái tạo điểm ảnh, nhưng tăng lên 64.41% khi tinh chỉnh toàn bộ trọng số. Kết quả này, cùng với nghiên cứu trước, khẳng định việc áp dụng học tự giám sát rập khuôn trên mô hình nhỏ không những không cải thiện mà còn có thể làm suy giảm hiệu suất so với huấn luyện từ đầu hoặc tiền huấn luyện sơ bộ. Điều này củng cố tính thách thức của bài toán và xác nhận MoCo v2 là lựa chọn nền tảng tối ưu cho nghiên cứu.

Pretraining paradigm	**Downtream task**	**Accuracy**	**Precision**	**Recall**	**F1-Score**

Pretrained	Full fine-tuning	**65.25**	**64.98**	**62.4**	**63.66**
SimCLR	Frozen backbone	60.59	58.49	58.77	58.63
	Full fine-tuning	59.75	58.07	59.82	58.93
MoCoV2	Frozen backbone	62.71	59.35	61.53	60.42
	Full fine-tuning	64.87	61.7	65.79	63.68
MAE	Frozen backbone	28.81	39.72	28.62	33.27
	Full fine-tuning	64.41	67.29	64.92	66.08

Bảng 2: Kết quả các phương pháp học giám sát và học tự giám sát trên bộ dữ liệu Plantdoc

Với kiến trúc mô hình đề xuất tích hợp đồng thời nhánh phân loại giám sát và nhánh tương phản tự giám sát, việc xác định tỷ trọng đóng góp tối ưu của từng luồng tín hiệu là yếu tố quyết định sự hội tụ của mạng nơ-ron. Để làm rõ tác động và tương hỗ giữa hai hàm mục tiêu, luận văn thiết kế thực nghiệm định lượng khảo sát hiệu năng khi thay đổi trọng số điều tiết trong hàm mất mát tổng hợp.

Trong thí nghiệm, trọng số beta của hàm mất mát phân loại được giữ cố định ở giá trị một nhằm duy trì ổn định của nhiệm vụ chính, trong khi trọng số alpha của hàm mất mát tương phản được thay đổi từ 0.3 đến 2.0. Mục tiêu là tìm điểm cân bằng lý tưởng, nơi các đặc trưng học được từ nhánh tự giám sát hỗ trợ hiệu quả cho biên quyết định phân loại mà không gây nhiễu loạn tín hiệu giám sát.

Kết quả thực nghiệm cho thấy một xu hướng biến thiên hình parabol rõ rệt của các chỉ số hiệu năng, với đỉnh cao nhất hội tụ tại giá trị alpha bằng 0.5. Tại thiết lập này, mô hình đạt được sự cộng hưởng tối ưu giữa hai nhánh với độ chính xác tổng thể là 69.49% và điểm F1 đạt 70.73%. Ngược lại, khi giá trị alpha được đẩy lên quá cao (từ 0.6 đến 2.0), hiệu năng hệ thống ghi nhận sự suy giảm dần đều, chạm đáy 64.83% tại alpha bằng 2.0. Điều này minh chứng rằng khi tín hiệu tự giám sát chiếm ưu thế quá lớn, nó có xu hướng lấn át tín hiệu định danh, buộc mô hình tập trung quá mức vào việc phân biệt các biến thể ảnh cục bộ thay vì học các đặc trưng ngữ nghĩa tổng quát cần thiết cho việc phân loại bệnh. Do đó, việc lựa chọn alpha bằng 0.5 được xác định là cấu hình tiêu chuẩn cho các thử nghiệm so sánh tiếp theo của luận văn.

alpha	Accuracy	Precision	Recall	F1-Score
alpha=0.3	66.53	69.62	66.2	67.87
alpha=0.4	66.95	69.24	66.83	68.01
alpha=0.5	**69.49**	**72.67**	**68.89**	**70.73**
alpha=0.6	66.53	68.75	66.41	67.56
alpha=0.7	66.53	69.53	66.54	68
alpha=0.8	65.68	67.22	65.19	66.19
alpha=0.9	65.25	66.98	64.68	65.81
alpha=1.0	65.68	68.21	65.3	66.72
alpha=1.5	66.1	71.3	66	68.55
alpha=2.0	64.83	67.97	64.72	66.31

Bảng 3: Thực nghiệm hệ số alpha, beta của mô hình đề xuất trên tập dữ liệu plantdoc

Tiếp nối quy trình xây dựng dữ liệu dựa trên độ tương đồng ngữ nghĩa, luận văn thực nghiệm đánh giá ảnh hưởng của ba tập dữ liệu tự giám sát Small, Medium và Large lên hiệu năng mô hình phân loại. Các tập này được trích xuất từ nguồn dữ liệu thô qua các ngưỡng lọc khác nhau, phản ánh sự đánh đổi giữa độ tinh khiết và đa dạng mẫu.

Mục tiêu chính của thử nghiệm không chỉ là kiểm chứng khả năng tận dụng dữ liệu mở, mà còn làm rõ mối quan hệ phi tuyến giữa số lượng mẫu và hiệu quả huấn luyện trong học tương phản. So sánh kết quả giữa các tập dữ liệu xác nhận rằng trong học tự giám sát, việc kiểm soát nhiễu và đảm bảo độ tương đồng cao với tác vụ đích quan trọng hơn việc tăng số lượng mẫu đầu vào một cách ồ ạt và thiếu chọn lọc.

Dataset	Accuracy	Precision	Recall	F1-Score
PlantDoc	69.49	72.67	68.89	70.73
SSL small (3088)	**72.02**	**73.79**	**71.24**	**72.49**
SSL medium (3989)	70.34	71.94	69.96	70.94
SSL large (5621)	69.49	71.87	69.73	70.78

Bảng 4: Kết quả với các tập dữ liệu không giám sát trên kiến trúc mới

Kết quả tại Bảng 4 cho thấy chất lượng dữ liệu không nhãn có ảnh hưởng quyết định đến hiệu suất phân loại. Khi sử dụng tập SSL Small gồm 3.088 ảnh có độ tương đồng cao với nguyên mẫu bệnh, mô hình đạt độ chính xác 72.02% và F1-Score 72.49%, cải thiện đáng kể so với mức cơ sở 69.49% từ dữ liệu PlantDoc gốc. Ngược lại, khi mở rộng sang tập Medium (3.989 ảnh) và Large (5.621 ảnh) bằng cách hạ ngưỡng lọc, hiệu năng suy giảm trở lại mức ban đầu. Hiện tượng này cho thấy lợi ích từ sự đa dạng mẫu bị triệt tiêu bởi tín hiệu nhiễu và các mẫu ngoại lai. Kết quả khẳng định chiến lược tuyển chọn dựa trên nguyên mẫu là then chốt, đồng thời chỉ ra rằng tập dữ liệu vừa phải nhưng sạch và sát ngữ nghĩa mang lại giá trị huấn luyện cao hơn nhiều so với tập lớn nhưng chứa nhiễu.

Ngoài ra, luận văn cũng tiến hành so sánh phương pháp đề xuất với các nghiên cứu liên quan được công bố trên cùng tập dữ liệu PlantDoc trên các thước đo hiệu năng truyền thống như Độ chính xác (Accuracy), Precision, Recall và F1-Score, luận văn đặc biệt chú trọng phân tích chỉ số về số lượng tham số (Parameters) để đánh giá tính hiệu quả về mặt tài nguyên – một yếu tố then chốt khi triển khai mô hình trên các thiết bị biên trong nông nghiệp thông minh. Các số liệu thống kê chi tiết về hiệu năng của các mô hình được tổng hợp và trình bày tại Bảng 5.

Phân tích số liệu so sánh cho thấy mô hình đề xuất đạt ưu thế vượt trội về cân bằng giữa hiệu suất nhận diện và chi phí tài nguyên. Với độ chính xác 72.02% và F1-score 72.49%, mô hình tích hợp học tự giám sát MoCo v2 trên mạng nhẹ vượt qua cả các kiến trúc CNN sâu như VGG16, ResNet152, cũng như những mô hình phức tạp như InceptionResNet V2 hay Vision Transformer. Điểm nổi bật là sự chênh lệch về số lượng tham số; các mô hình đối chứng cần từ 55 đến gần 140 triệu tham số để đạt độ chính xác 58-70%, trong khi mô hình đề xuất chỉ sử dụng 5.4 triệu tham số mà vẫn đạt hiệu năng cao nhất. Kết quả này chứng minh rằng trong bối cảnh dữ liệu gán nhãn khan hiếm và nhiễu tạp, sử dụng mô hình gọn nhẹ kết hợp với không gian biểu diễn chất lượng cao từ dữ liệu không nhãn hiệu quả hơn việc tăng độ sâu mạng cơ học, đồng thời mở ra khả năng triển khai trên các thiết bị biên hạn chế tài nguyên.

Model	Accuracy	Precision	Recall	F1-Score	Params (M)
DenseNet121 [11]	61.01	-	-	-	8
ResNet152 [11]	60.16	-	-	-	60.2
VGG16 [11]	58.05	-	-	-	**138.4**
Ensemble (CNNs) [12]	68	69	69	68	-
ViT-Base [12]	69	61	59	60	-
InceptionResNet V2 [13]	70.53	-	-	-	55.8
Mô hình đề xuất	**72.02**	**73.79**	**71.24**	**72.49**	5.4

Bảng 5: Hiệu suất các mô hình trên tập dữ liệu plantdoc

4.4. Kết luận chương 4

Chương 4 trình bày kiểm chứng thực nghiệm mô hình lai ghép trên GPU P100 với các chỉ số Accuracy, Precision, Recall và F1-score trên dữ liệu PlantDoc. Kết quả cho thấy MoCo v2 là chiến lược tiền huấn luyện tối ưu cho mạng nhẹ, đạt độ chính xác 64.87% và ổn định hơn MAE.

Tinh chỉnh siêu tham số xác định alpha = 0.5 là điểm cân bằng giữa nhánh giám sát và tự giám sát, đạt độ chính xác cơ sở 69.49%. Chất lượng dữ liệu không nhãn quan trọng hơn số lượng, khi tập Small được lọc nhiễu giúp mô hình đạt 72.02% Accuracy và 72.49% F1-score, trong khi tăng quy mô dữ liệu làm giảm hiệu năng.

So với các mô hình hạng nặng như VGG16, ResNet152 hay Vision Transformer, mô hình đề xuất vượt trội với chỉ 5.4 triệu tham số. Kết quả chứng minh học tương phản tối ưu hóa đặc trưng hiệu quả hơn tăng độ sâu mạng, đồng thời cân bằng Precision và Recall, xử lý tốt mất cân bằng dữ liệu, phù hợp triển khai trên thiết bị biên trong nông nghiệp.

KẾT LUẬN

Luận văn đã hoàn thành các mục tiêu nghiên cứu đặt ra, làm rõ vai trò của nông nghiệp chính xác và chỉ ra hạn chế cốt lõi của các mô hình học sâu hiện nay là sự phụ thuộc lớn vào dữ liệu gán nhãn. Trên cơ sở đó, một kiến trúc lai ghép kết hợp học giám sát và học tự giám sát MoCo v2 với mạng xương sống gọn nhẹ đã được đề xuất và triển khai thành công. Đóng góp nổi bật của nghiên cứu là quy trình tuyển

chọn dữ liệu không nhãn dựa trên độ tương đồng ngữ nghĩa với các nguyên mẫu lớp, cho phép khai thác hiệu quả nguồn dữ liệu mở và cải thiện không gian biểu diễn đặc trưng trong điều kiện dữ liệu hạn chế.

Kết quả thực nghiệm trên tập PlantDoc cho thấy phương pháp đề xuất đạt hiệu năng vượt trội so với các tiếp cận học giám sát truyền thống, đặc biệt khi số lượng mẫu có nhãn bị giới hạn. Tuy nhiên, nghiên cứu cũng chỉ ra một số hạn chế như sự nhạy cảm với nhiễu khi mở rộng quy mô dữ liệu tự giám sát và việc tinh chỉnh siêu tham số chưa được tự động hóa toàn diện. Ngoài ra, giai đoạn tiền huấn luyện tự giám sát vẫn yêu cầu tài nguyên tính toán đáng kể, gây khó khăn cho việc cập nhật mô hình trực tiếp tại biên.

Từ những kết quả đạt được, luận văn khẳng định tính khả thi của việc khai thác cấu trúc dữ liệu không nhãn để giảm phụ thuộc vào giám sát thủ công trong chẩn đoán bệnh thực vật. Hướng nghiên cứu tiếp theo tập trung vào cải tiến chiến lược lọc nhiễu dữ liệu và tối ưu hóa triển khai thông qua các kỹ thuật nén mô hình, nhằm hướng tới xây dựng các hệ thống chẩn đoán tự động hiệu quả, chi phí thấp và phù hợp với môi trường thực tế.