**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TPHCM**

**KHOA CÔNG NGHỆ THÔNG TIN**



**THỊ GIÁC MÁY TÍNH VÀ ỨNG DỤNG**

**NHẬN DIỆN BIỂN BÁO GIAO THÔNG**

**GIẢNG VIÊN HƯỚNG DẪN: Võ Quang Hoàng Khang**

**SINH VIÊN THỰC HIỆN 1: Võ Hoàng Phi**

**MÃ SỐ SINH VIÊN: 21038411**

**SINH VIÊN THỰC HIỆN 2: Lê Minh Đăng**

**MÃ SỐ SINH VIÊN: 21080991**

*TP HỒ CHÍ MINH, Tháng 05 năm 2025*

**INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY**

**FACULTY OF INFORMATION TECHNOLOGY**



**Computer Vision and Applications**

**Traffic Sign Recognition**

**SUPERVISOR: MSc. Vo Quang Hoang Khang**

**STUDENT NAME: Võ Hoàng Phi**

**STUDENT CODE: 21038411**

**STUDENT NAME: Lê Minh Đăng**

**STUDENT CODE: 21080991**

*HO CHI MINH CITY, Month 05 year 2025*

**TÓM TẮT**

Nghiên cứu này xây dựng một mô hình học sâu dựa trên mạng nơ-ron tích chập (CNN) nhằm tự động nhận diện biển báo giao thông đường bộ từ hình ảnh. Mô hình được huấn luyện và kiểm thử trên tập dữ liệu GTSRB từ nền tảng Kaggle, với dữ liệu ảnh được xử lý và chuẩn hóa về kích thước 30×30 pixel. Để đánh giá hiệu năng, mô hình sử dụng chỉ số Accuracy – độ chính xác phân loại. Kết quả thực nghiệm cho thấy mô hình đạt độ chính xác cao trên tập kiểm tra, cho thấy tiềm năng ứng dụng vào các hệ thống hỗ trợ lái xe thông minh, góp phần nâng cao độ an toàn và giảm thiểu rủi ro trong giao thông.

**LỜI CẢM ƠN**

Trong suốt thời gian làm đồ án tốt nghiệp, chúng em đã luôn nhận được nhiều sự quan tâm, hướng dẫn và giúp đỡ tận tình của các thầy cô giáo trong khoa Công nghệ thông tin cùng với sự động viên giúp đỡ từ bạn bè và gia đình.

Lời đầu tiên em xin chân thành cảm ơn Ban giám hiệu Trường Đại học Công nghiệp thành phố Hồ Chí Minh, Ban chủ nhiệm khoa Công nghệ thông tin đã luôn tận tình quan tâm giúp đỡ chúng em trong suốt thời gian học tập tại trường.

Đặc biệt chúng em xin gửi lời cảm ơn chân thành và sâu sắc tới thầy hướng dẫn ThS. Võ Quang Hoàng Khang đã trực tiếp hướng dẫn, giúp đỡ chúng em hoàn thành khoá luận này.

Chúng em cũng xin gửi lời cảm ơn đến gia đình, người thân và bạn bè đã luôn ở bên quan tâm, giúp đỡ và động viên chúng em hoàn thành khoá luận tốt nghiệp này.

Chúng em xin trân trọng cảm ơn!

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

**TP. Hồ Chí Minh, ngày…. tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN 1**

**TP. Hồ Chí Minh, ngày…. Tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN 2**

**TP. Hồ Chí Minh, ngày…. tháng …. năm….**

CHỮ KÝ CỦA GIẢNG VIÊN

**LỜI MỞ ĐẦU**

**Trong thời đại kỷ nguyên số hiện nay, các mô hình máy học và học sâu đang đóng vai trò vô cùng quan trọng trong hầu hết các lĩnh vực của đời sống, đặc biệt là trong lĩnh vực giao thông và tự động hóa. Với sự phát triển mạnh mẽ của công nghệ, nhu cầu ứng dụng các mô hình thông minh vào thực tế ngày càng trở nên cấp thiết. Điều này đòi hỏi ngành Khoa học máy tính phải không ngừng đào tạo những chuyên gia có nền tảng kiến thức vững chắc cùng kỹ năng thực hành tốt để phát triển các mô hình tiên tiến, đáp ứng được yêu cầu ngày càng cao của xã hội.**

**Xuất phát từ thực tế đó, chúng tôi chọn chủ đề “Nhận diện biển báo giao thông đường bộ bằng mô hình học sâu” cho khóa luận tốt nghiệp của mình. Đây là một chủ đề có tính ứng dụng cao, góp phần hỗ trợ các hệ thống giao thông thông minh, xe tự hành và nâng cao độ an toàn cho người tham gia giao thông. Việc nhận diện nhanh chóng và chính xác biển báo sẽ giúp hệ thống xử lý tình huống hiệu quả hơn, từ đó giảm thiểu tai nạn và đảm bảo lưu thông an toàn.**

**Khóa luận của chúng tôi tập trung nghiên cứu và cải tiến mô hình học sâu để nhận diện hình ảnh biển báo giao thông, trong đó áp dụng các cơ chế như chú ý (attention), tích chập (convolution) và tăng cường học đặc trưng. Mô hình được huấn luyện và đánh giá trên tập dữ liệu biển báo từ Kaggle, sử dụng các chỉ số phổ biến như: độ chính xác (accuracy), F1 score, precision, recall, specificity và AUC. Kết quả thực nghiệm cho thấy mô hình đạt hiệu suất cao trong nhận diện biển báo giao thông, thể hiện tiềm năng ứng dụng trong thực tiễn.**

**Cuối cùng, chúng em xin bày tỏ lòng biết ơn chân thành đến nhà trường, các giảng viên Khoa Công nghệ Thông tin, gia đình và bạn bè đã luôn hỗ trợ, khích lệ và tạo điều kiện thuận lợi để chúng em hoàn thành tốt khóa luận này. Chúng tôi hy vọng những kiến thức và kinh nghiệm thu nhận được sẽ là nền tảng vững chắc giúp chúng tôi tiếp tục đóng góp vào sự phát triển của ngành học và ứng dụng công nghệ vào giải quyết các vấn đề thực tế trong đời sống.**

**MỤC LỤC**

[**CHƯƠNG 1.**  **TỔNG QUAN** 1](#_Toc198547920)

[**1.1 Lý do chọn đề tài** 1](#_Toc198547921)

[**1.2 Mục tiêu nghiên cứu** 1](#_Toc198547922)

[**1.3 Phạm vi nghiên cứu** 2](#_Toc198547923)

[**1.4 Phương pháp nghiên cứu** 2](#_Toc198547924)

[**1.5 Kết cấu đồ án.** 5](#_Toc198547925)

[**2.1. Học sâu (Deep learning)** 5](#_Toc198547926)

[**2.2. Tăng cường dữ liệu (Data augmentation)** 6](#_Toc198547927)

[**2.3. Thị giác máy tính và nhận diện ảnh** 6](#_Toc198547928)

[**1. Các thành phần chính của CNN:** 6](#_Toc198547929)

[**2. Hàm kích hoạt ReLU:** 7](#_Toc198547930)

[**2.5. Các kỹ thuật huấn luyện tối ưu:** 7](#_Toc198547931)

[**2.6. Bộ dữ liệu GTSRB** 7](#_Toc198547932)

[**2.7. Các phương pháp đánh giá mô hình** 7](#_Toc198547933)

[**CHƯƠNG 3. MÔ HÌNH ĐỀ XUẤT** 7](#_Toc198547934)

[**3.1 Mô hình tổng quát** 7](#_Toc198547935)

[- Tổng quan thiết kế: 10](#_Toc198547936)

[**3.2 Đặc trưng của mô hình đề xuất** 10](#_Toc198547937)

[**3.2.2** .**Phân loại (Classification)** 11](#_Toc198547938)

[**CHƯƠNG 4. THỰC NGHIỆM** 12](#_Toc198547939)

[**4.1 Môi trường thực nghiệm** 12](#_Toc198547940)

[**4.1.1. Cấu hình phần cứng** 12](#_Toc198547941)

[**4.1.2. Cấu hình phần mềm** 12](#_Toc198547942)

[**4.1.3. Thiết lập môi trường** 12](#_Toc198547943)

[**4.1.4. Lý do chọn Kaggle** 13](#_Toc198547944)

[**4.2 Tập dữ liệu** 13](#_Toc198547945)

[**4.2.1. Nguồn dữ liệu** 13](#_Toc198547946)

[**4.2.2. Số lượng dữ liệu** 13](#_Toc198547947)

[**4.2.3. Tiền xử lý dữ liệu** 14](#_Toc198547948)

[**4.2.4. Tăng cường dữ liệu** 14](#_Toc198547949)

[**4.2.5. Chia dữ liệu** 15](#_Toc198547950)

[**4.2.6. Lý do chọn tập dữ liệu** 16](#_Toc198547951)

[**4.3 Ứng dụng thực nghiệm** 16](#_Toc198547952)

[**4.3.1. Quy trình huấn luyện mô hình** 16](#_Toc198547953)

[**4.3.2. Cấu hình huấn luyện** 16](#_Toc198547954)

[**4.3.3. Kết quả thực nghiệm** 17](#_Toc198547955)

[**4.4 Đánh giá kết quả** 20](#_Toc198547956)

[**4.4.1. So sánh với các mô hình pretrain** 20](#_Toc198547957)

[**CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN** 21](#_Toc198547958)

[**5.1 Kết luận** 21](#_Toc198547959)

[**5.2 Hướng phát triển** 22](#_Toc198547960)

**DANH MỤC CÁC THUẬT NGỮ VIẾT TẮT**

|  |  |  |
| --- | --- | --- |
| Từ viết tắt | Từ đầy đủ | Nghĩa |
| **GAP** | Global Average Pooling | Kỹ thuật gôp trung bình toàn cầu dùng để giảm kích thước của các đặc trưng dữ liệu và chuẩn bị dữ liệu cho các lớp tích chập sau bằng cách tính trung bình các giá trị pixel trong feature map tạo ra một vector có kích thước bằng số lượng feature maps, mỗi phần tử trong vector đại diện cho giá trị trung bình của một feature map. |
| **BN** | Batch Normalization | Chuẩn hoá hàng loạt là một kỹ thuật trong học sâu. Nó tính toán trung bình và phương sai của dữ liệu trong một batch (một nhóm dữ liệu nhỏ), sau đó chuẩn hóa dữ liệu dựa trên các giá trị này. Từ đó làm cho dữ liệu phân phối ổn định giúp các mạng nơ-ron học nhanh, hiệu quả hơn. |
| **CNN** | Convolutional Neural Network | Mạng nơ-ron tích chập, một mô hình học sâu được sử dụng chủ yếu trong xử lý ảnh và thị giác máy tính. |
| **LSTM** | Long Short-Term Memory | Mạng nơ-ron hồi tiếp đặc biệt dùng cho dữ liệu tuần tự hoặc chuỗi thời gian. |

# **CHƯƠNG 1.** **TỔNG QUAN**

## **1.1 Lý do chọn đề tài**

Biển báo giao thông là một trong những yếu tố quan trọng đảm bảo an toàn cho người tham gia giao thông và điều tiết trật tự trên đường. Các loại biển báo cung cấp thông tin, cảnh báo và hướng dẫn để người điều khiển phương tiện có thể kịp thời điều chỉnh hành vi lái xe, từ đó giảm thiểu nguy cơ tai nạn. Tuy nhiên, trong thực tế, việc nhận diện và phản ứng với biển báo giao thông có thể bị ảnh hưởng bởi yếu tố môi trường, thời tiết, ánh sáng hoặc sự mất tập trung của người lái. Đặc biệt, trong bối cảnh phát triển của các hệ thống xe tự hành và hỗ trợ lái xe thông minh, việc tự động nhận diện biển báo giao thông trở nên cấp thiết và là một trong những bài toán trọng tâm trong lĩnh vực thị giác máy tính.

Ngày nay, với sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là các kỹ thuật học sâu, việc nhận diện hình ảnh, phân loại và xử lý ảnh giao thông đã đạt được nhiều kết quả tích cực. Các mô hình học sâu như CNN, VGG16, ResNet, InceptionV3 và DenseNet121 đã chứng minh hiệu quả trong xử lý hình ảnh nói chung và ảnh giao thông nói riêng [1]. Ngoài ra, các nghiên cứu gần đây còn áp dụng thêm cơ chế chú ý (attention mechanism), các mô-đun tự tập trung (self-attention) và kỹ thuật tăng cường đặc trưng (feature enhancement) nhằm nâng cao độ chính xác và khả năng tổng quát của mô hình [2][3]. Những cải tiến này cho phép mô hình xử lý tốt hơn các trường hợp biển báo bị che khuất, mờ, hoặc biến dạng.

Trong nghiên cứu này, chúng tôi đề xuất một mô hình học sâu dựa trên mạng nơ-ron tích chập (CNN) được xây dựng từ đầu, không sử dụng các mô hình tiền huấn luyện (pretrained). Mô hình gồm nhiều lớp tích chập kết hợp với các lớp pooling, dropout và fully connected nhằm trích xuất đặc trưng và thực hiện phân loại ảnh biển báo một cách hiệu quả. Mô hình được huấn luyện và đánh giá trên tập dữ liệu chuẩn về biển báo giao thông (German Traffic Sign Recognition Benchmark – GTSRB), với quá trình huấn luyện được chia theo tỷ lệ huấn luyện/kiểm tra 80/20. Các kết quả đạt được cho thấy mô hình có khả năng nhận diện chính xác biển báo trong nhiều điều kiện khác nhau mà không yêu cầu kiến trúc phức tạp hay pretraining.

.

## **1.2 Mục tiêu nghiên cứu**

1. **Xây dựng một mô hình học sâu dựa trên mạng nơ-ron tích chập (CNN)** có khả năng phân loại hình ảnh biển báo giao thông với độ chính xác cao.
2. **Tiền xử lý dữ liệu hình ảnh từ tập dữ liệu GTSRB**, chuẩn hóa kích thước ảnh và gán nhãn đúng cách nhằm đảm bảo chất lượng đầu vào cho mô hình học sâu.
3. **Huấn luyện và đánh giá mô hình CNN tự xây dựng**, không sử dụng các kiến trúc tiền huấn luyện, nhằm kiểm chứng khả năng học và phân loại của mô hình đơn giản nhưng hiệu quả.
4. **Áp dụng phương pháp chia tập huấn luyện và kiểm tra theo tỷ lệ phù hợp**, nhằm đánh giá độ chính xác và khả năng tổng quát của mô hình trên dữ liệu chưa từng thấy.
5. **Phân tích hiệu suất mô hình thông qua các chỉ số như độ chính xác và mất mát (accuracy, loss)** và đề xuất hướng cải tiến mô hình trong tương lai (ví dụ: thử nghiệm với mô hình pretrained hoặc tích hợp attention) nhằm nâng cao hiệu quả nhận diện trong điều kiện thực tế.

## **1.3 Phạm vi nghiên cứu**

* **Đối tượng nghiên cứu**: Hình ảnh các loại biển báo giao thông đường bộ, với trọng tâm là các biển báo phổ biến trong hệ thống giao thông như: biển báo cấm, biển báo nguy hiểm, biển hiệu lệnh và biển chỉ dẫn.
* **Phạm vi kỹ thuật**: Ứng dụng các mô hình học sâu, đặc biệt là mạng nơ-ron tích chập (CNN) kết hợp với các kiến trúc hiện đại như EfficientNetB0 và DenseNet121. Bên cạnh đó, tích hợp các cơ chế chú ý như multi-head self-attention, channel-attention và spatial-attention nhằm nâng cao khả năng trích xuất đặc trưng và tối ưu hóa hiệu năng mô hình.
* **Phạm vi kiểm tra**: Thực hiện kiểm tra, đánh giá mô hình trên các tập dữ liệu hình ảnh biển báo giao thông (ví dụ: GTSRB – German Traffic Sign Recognition Benchmark hoặc bộ dữ liệu tương đương từ Kaggle) bằng phương pháp Stratified Five-Fold Cross-Validation để đảm bảo tính ổn định và độ chính xác trong phân loại.

## **1.4 Phương pháp nghiên cứu**

1. **Nghiên cứu tài liệu**:

* Tìm hiểu các loại biển báo giao thông đường bộ, phân loại và ý nghĩa của từng nhóm biển báo.
* Tham khảo các mô hình học sâu hiện đại trong lĩnh vực thị giác máy tính, đặc biệt là trong bài toán nhận diện ảnh giao thông và ảnh phân loại nhiều lớp.

1. **Xây dựng mô hình**:

Mô hình được xây dựng dựa trên mạng nơ-ron tích chập (CNN) tuần tự (Sequential) với kiến trúc gồm nhiều lớp Conv2D, MaxPooling và Dropout nhằm trích xuất đặc trưng và tránh overfitting.

* **Các lớp tích chập (Conv2D)** sử dụng các bộ lọc (filters) có kích thước lần lượt là 5x5 và 3x3 với hàm kích hoạt ReLU, giúp trích xuất các đặc trưng không gian từ ảnh biển báo như cạnh, góc, và họa tiết.
* **Các lớp MaxPooling** với kích thước cửa sổ 2x2 được thêm vào nhằm giảm kích thước ảnh đặc trưng và tăng tính bất biến dịch chuyển.
* **Lớp Dropout** được sử dụng sau các tầng tích chập và lớp kết nối đầy đủ (Dense) để giảm thiểu hiện tượng overfitting trong quá trình huấn luyện.
* Cuối cùng, mô hình kết thúc bằng lớp Dense với số lượng neuron bằng số lớp biển báo (43), sử dụng hàm softmax để phân loại đa lớp.

1. **Kiểm thử và đánh giá**:

* Để đánh giá hiệu năng của mô hình, dữ liệu được chia thành hai tập riêng biệt: tập huấn luyện (80%) và tập kiểm tra (20%) sử dụng hàm train\_test\_split với phân phối nhãn giữ nguyên (stratify) nhằm đảm bảo sự cân bằng giữa các lớp biển báo trong cả hai tập.
* Mô hình được huấn luyện trong 15 epoch, đồng thời trong quá trình huấn luyện, hiệu năng trên tập kiểm tra cũng được theo dõi để đánh giá khả năng tổng quát hóa và tránh hiện tượng overfitting.
* Độ chính xác (accuracy) và hàm mất mát (loss) trên tập huấn luyện và tập kiểm tra được ghi lại qua từng epoch để quan sát xu hướng học của mô hình. Biểu đồ Accuracy và Loss được vẽ để trực quan hóa quá trình huấn luyện và hiệu năng cuối cùng.
* Mô hình đạt được độ chính xác khá tốt trên tập kiểm tra, chứng tỏ khả năng nhận diện biển báo giao thông trong tập dữ liệu GTSRB với kiến trúc CNN đơn giản.
* Do hạn chế về thời gian và tài nguyên, mô hình chưa được đánh giá bằng kỹ thuật Stratified k-Fold Cross-Validation hay so sánh trực tiếp với các kiến trúc CNN tiền huấn luyện (pretrained models) như VGG16, ResNet50, MobileNet, InceptionV3 hay DenseNet121. Đây là hướng phát triển tiềm năng cho các bước tiếp theo nhằm nâng cao hiệu năng và độ ổn định của mô hình.

1. **Phân tích và cải tiến**:

* Áp dụng chiến lược hợp nhất đặc trưng (feature fusion) dựa trên attention nhằm tối ưu hiệu quả trích xuất thông tin từ ảnh.
* Thử nghiệm các kỹ thuật tăng cường dữ liệu (data augmentation), sử dụng khối dư (residual block) và cơ chế phân bổ động (dynamic allocation) để cải thiện hiệu suất mô hình và giảm thiểu hiện tượng overfitting.
* Mô hình hiện tại mặc dù đạt được kết quả khả quan trên bộ dữ liệu GTSRB, nhưng vẫn còn nhiều điểm có thể cải thiện để nâng cao độ chính xác và khả năng tổng quát hóa.
* Một trong những hướng phát triển là áp dụng các kỹ thuật tăng cường dữ liệu (data augmentation) như xoay ảnh, thay đổi độ sáng, dịch chuyển, hoặc lật ngang để làm phong phú tập huấn luyện, từ đó giúp mô hình học được các đặc trưng đa dạng hơn và giảm thiểu hiện tượng overfitting.
* Bên cạnh đó, có thể thử nghiệm việc tích hợp các khối dư (residual blocks) vào kiến trúc CNN nhằm giải quyết vấn đề suy giảm gradient khi mạng quá sâu, giúp mô hình học sâu hiệu quả hơn.

 Ngoài ra, việc sử dụng các cơ chế chú ý (attention mechanisms) để hợp nhất đặc trưng (feature fusion) có thể cải thiện khả năng trích xuất thông tin quan trọng từ ảnh biển báo, đặc biệt trong các điều kiện ánh sáng hay góc chụp phức tạp.

* Cuối cùng, các kỹ thuật tối ưu hóa động như dynamic allocation của bộ nhớ hoặc điều chỉnh learning rate theo từng epoch cũng có thể được áp dụng để cải thiện hiệu suất huấn luyện và tránh overfitting.
* Tóm lại, việc kết hợp các chiến lược trên cùng với kiến trúc CNN hiện tại là hướng phát triển tiềm năng để nâng cao độ chính xác và độ ổn định của hệ thống nhận diện biển báo giao thông.

## **1.5 Kết cấu đồ án.**

Chương 1: Tổng quan

Chương 2: Cơ sở lý thuyết

Chương 3: Phân tích yêu cầu và thiết kế mô hình

Chương 4: Thực nghiệm

Chương 5: Kết luận và hướng phát triển

Tài liệu tham khảo.

**CHƯƠNG 2. CƠ SỞ LÝ THUYẾT**

## **2.1. Học sâu (Deep learning)**

Học sâu (Deep Learning) là một nhánh con của học máy (Machine Learning) và là một lĩnh vực nghiên cứu quan trọng trong trí tuệ nhân tạo. Học sâu sử dụng các mô hình học máy với các lớp mạng nơ-ron nhiều tầng để tự động học các đặc trưng phức tạp từ dữ liệu mà không cần sự can thiệp của con người. Các mô hình học sâu có thể học được các đặc tzrưng chi tiết hơn từ dữ liệu thô, giúp mô hình cải thiện độ chính xác trong việc nhận diện và phân loại. [9]. Học sâu đã có những ứng dụng quan trọng trong nhiều lĩnh vực, đặc biệt là trong xử lý hình ảnh, nhận diện giọng nói, và xử lý ngôn ngữ tự nhiên. Trong y học, học sâu được ứng dụng mạnh mẽ trong việc phân tích ảnh y khoa, chẳng hạn như phân loại bệnh viêm phổi từ ảnh chụp X-quang, giúp tăng độ chính xác và giảm thời gian chẩn đoán của bác sĩ. Một trong những thành tựu nổi bật của học sâu trong xử lý ảnh là khả năng nhận diện các đặc trưng ẩn trong hình ảnh, chẳng hạn như các dấu hiệu viêm phổi, mà mắt thường khó phát hiện.

## **2.2. Tăng cường dữ liệu (Data augmentation)**

Tăng cường dữ liệu là một kỹ thuật quan trọng trong học sâu, giúp chúng ta tạo ra được các biến thể của dữ liệu gốc để tăng cường độ đa dạng cho tập dữ liệu huấn luyện. Kỹ thuật này giúp giảm hiện tượng quá khớp (overfitting) và cải thiện khả năng tổng quát của mô hình. Tăng cường dữ liệu đặc biệt hữu ích khi bạn có tập dữ liệu hạn chế, giúp mô hình học được nhiều đặc trưng hơn.

## **2.3. Thị giác máy tính và nhận diện ảnh**

Thị giác máy tính (Computer Vision) là lĩnh vực nghiên cứu cho phép máy tính hiểu và phân tích các dữ liệu hình ảnh hoặc video. Một trong những bài toán cơ bản là nhận diện ảnh (image recognition), nhằm phân loại hình ảnh vào các nhóm đối tượng cụ thể.

**2.4. Mạng nơ-ron tích chập (Convolutional Neural Networks - CNN)**

CNN là một loại mạng nơ-ron sâu đặc biệt được thiết kế để xử lý dữ liệu ảnh. Điểm mạnh của CNN là khả năng tự động học các đặc trưng từ ảnh đầu vào mà không cần phải trích xuất thủ công

#### **1. Các thành phần chính của CNN:**

 **Lớp tích chập (Convolutional Layer):** Áp dụng các bộ lọc (filter) trượt trên ảnh để phát hiện các đặc trưng như cạnh, góc, họa tiết.

 **Lớp gộp mẫu (Pooling Layer):** Giảm kích thước ảnh đặc trưng, tăng tính bất biến với dịch chuyển, thường dùng MaxPooling.

 **Lớp kết nối đầy đủ (Fully Connected Layer):** Liên kết các đặc trưng đã trích xuất với các lớp phân loại.

 **Lớp Dropout:** Ngăn ngừa overfitting bằng cách tạm thời loại bỏ ngẫu nhiên một số neuron trong quá trình huấn luyện.

.

#### **2. Hàm kích hoạt ReLU:**

Hàm ReLU (Rectified Linear Unit) là hàm phi tuyến thường được sử dụng trong CNN, giúp mô hình học được các biểu diễn phức tạp hơn và giảm thiểu vấn đề gradient biến mất.

#### **2.5. Các kỹ thuật huấn luyện tối ưu:**

 **One-hot encoding:** Chuyển nhãn phân loại thành vector nhị phân để phù hợp với bài toán phân lớp đa lớp.

 **Hàm mất mát categorical cross-entropy:** Đo độ lệch giữa phân phối dự đoán và phân phối thực tế.

 **Tối ưu Adam:** Thuật toán tối ưu phổ biến giúp mô hình hội tụ nhanh và hiệu quả.

## **2.6. Bộ dữ liệu GTSRB**

Bộ dữ liệu GTSRB (German Traffic Sign Recognition Benchmark) là bộ dữ liệu chuẩn cho bài toán nhận diện biển báo giao thông, bao gồm 43 loại biển báo khác nhau với đa dạng điều kiện ánh sáng, góc chụp và phông nền.

## **2.7. Các phương pháp đánh giá mô hình**

* **Độ chính xác (Accuracy):** Tỷ lệ dự đoán đúng trên tổng số mẫu kiểm tra.
* **Đánh giá qua đồ thị Accuracy và Loss:** Theo dõi quá trình học để phát hiện hiện tượng underfitting hoặc overfitting.

# **CHƯƠNG 3. MÔ HÌNH ĐỀ XUẤT**

## **3.1 Mô hình tổng quát**

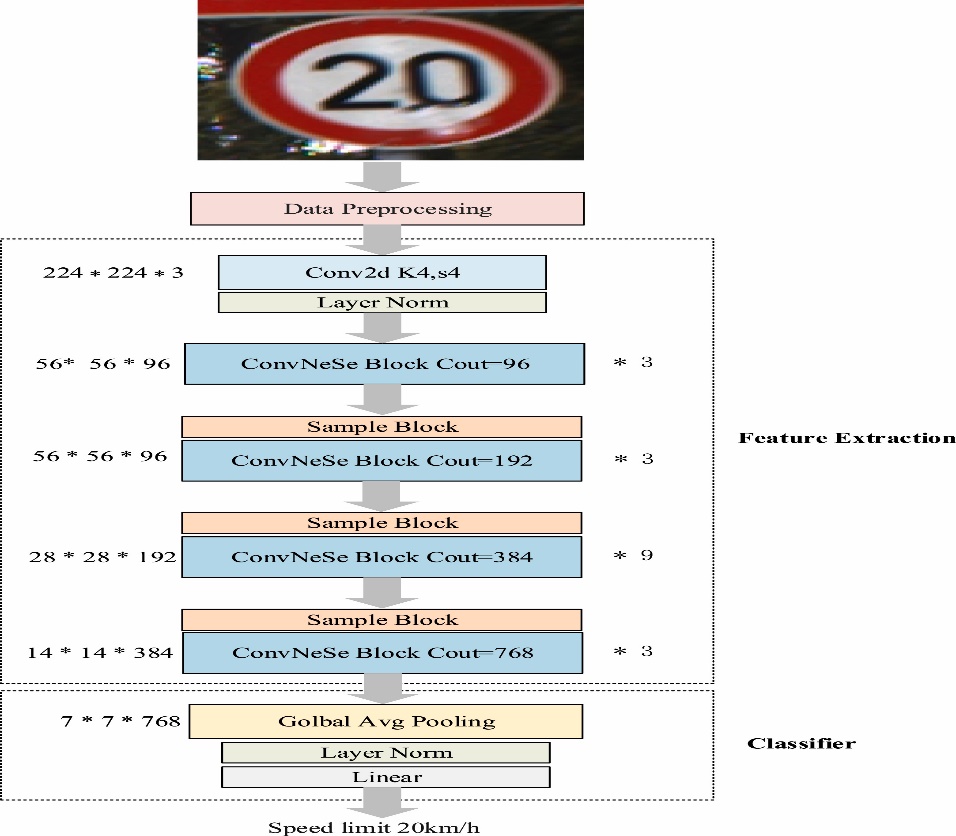
Trong nghiên cứu mô hình học sâu cho bài toán nhận diện biển báo giao thông, việc xây dựng một mạng nơ-ron tích chập (CNN) hiệu quả là bước quan trọng để tự động trích xuất các đặc trưng không gian từ ảnh.

Mô hình của chúng tôi được thiết kế dựa trên kiến trúc CNN tuần tự (Sequential) với các thành phần chính bao gồm:

1. **Các lớp tích chập (Conv2D):** Mô hình sử dụng nhiều lớp tích chập với các bộ lọc kích thước 5x5 và 3x3 để phát hiện các đặc trưng cơ bản như cạnh, góc, họa tiết từ ảnh biển báo.
2. **Lớp gộp mẫu (MaxPooling):** Giúp giảm kích thước ảnh đặc trưng, tăng khả năng kháng nhiễu và giảm tải tính toán.
3. **Lớp Dropout:** Được sử dụng để giảm hiện tượng quá khớp (overfitting) bằng cách ngẫu nhiên loại bỏ một phần neuron trong quá trình huấn luyện.
4. **Lớp kết nối đầy đủ (Dense):** Dùng để kết nối các đặc trưng đã được trích xuất với lớp phân loại cuối cùng, dự báo xác suất thuộc về từng loại biển báo giao thông.

Mô hình được xây dựng nhằm cân bằng giữa độ phức tạp và hiệu năng, phù hợp với kích thước ảnh nhỏ (30x30) trong tập dữ liệu GTSRB và phù hợp với khả năng tính toán của môi trường thực nghiệm.

Kiến trúc mô hình cải tiến được thể hiện trong **Hình 1**.



Hình . **Kiến trúc mô hình**

* **Đầu vào của mô hình** là ảnh chụp biển báo giao thông đã qua bước **tiền xử lý và tăng cường dữ liệu**, được chuẩn hóa về kích thước 30x30x3(RGB) để phù hợp với yêu cầu của mô hình. Ảnh được biến đổi thành mảng số (NumPy array) để đưa vào mạng nơ-ron tích chập. Kích thước nhỏ gọn giúp tăng tốc độ huấn luyện, phù hợp với bài toán và kiến trúc mạng CNN được xây dựng từ đầu.
* **Đầu ra của mô hình** là Là một vector xác suất gồm **43 giá trị**, tương ứng với **43 loại biển báo giao thông** trong tập dữ liệu GTSRB. Mỗi phần tử trong vector biểu thị xác suất ảnh đầu vào thuộc về lớp tương ứng.
  + Nhờ việc sử dụng hàm **softmax** ở tầng cuối cùng, mô hình không chỉ đưa ra nhãn dự đoán mà còn cho biết **mức độ tin cậy** của từng lớp — điều này rất hữu ích khi tích hợp vào các **hệ thống hỗ trợ lái xe thông minh**, nơi mà quyết định phải được đưa ra với độ tin cậy cao.
* Mô hình được chia thành 2 giai đoạn chính gồm:
* Trích xuất đặc trưng: Trong giai đoạn này, mô hình sử dụng chuỗi các **lớp tích chập (Conv2D)** và **lớp gộp (MaxPooling2D)** để học các đặc trưng thị giác từ ảnh đầu vào.Cụ thể:
  + Ảnh đầu vào: Kích thước chuẩn hóa về **30×30×3**, đưa vào mạng với kích thước nhỏ để giảm thiểu chi phí tính toán.
  + Tầng tích chập 1&2:
    - Hai lớp Conv2D liên tiếp với 32 bộ lọc và kernel 5x5, activation ReLU.
    - Sau đó là một lớp MaxPooling2D với pool size 2x2 để giảm kích thước đặc trưng.
    - Áp dụng Dropout để giảm overfitting.
* Tầng tích chập 3&4
* Hai lớp Conv2D tiếp theo với 64 bộ lọc và kernel 3x3.
* Tiếp theo là lớp MaxPooling2D và Dropout.
* Quá trình này giúp trích xuất được đặc trưng ở nhiều cấp độ: từ cạnh đơn giản đến cấu trúc hình học phức tạp của các loại biển báo.
* Phân loại (Classifier):
* Sau khi hoàn tất quá trình trích xuất đặc trưng:
* **Flatten**: Chuyển dữ liệu từ tensor 3 chiều về vector 1 chiều.
* **Dense layer 1**: 256 neuron, activation ReLU, tiếp theo là Dropout để chống overfitting.
* **Dense layer 2 (output)**: Số lượng neuron bằng số lớp biển báo (**43**), activation **softmax**

### - Tổng quan thiết kế:

* **Thiết kế đơn giản, nhẹ, dễ triển khai thực tế**.
* **Sử dụng Dropout** để ngăn quá khớp (overfitting).
* **Kết hợp nhiều tầng Conv2D** để trích xuất đặc trưng hình ảnh hiệu quả trong điều kiện dữ liệu vừa và nhỏ.
* **Đạt hiệu quả tốt** trong việc nhận diện biển báo trên tập dữ liệu GTSRB mà không cần các mô hình tiền huấn luyện phức tạp.

.

## **3.2 Đặc trưng của mô hình đề xuất**

**3.2.1. Trích xuất đặc trưng (Feature Extraction)**

Trong mô hình được đề xuất, giai đoạn trích xuất đặc trưng đóng vai trò quan trọng nhằm học các đặc trưng hình ảnh phân biệt giữa các loại biển báo giao thông. Quá trình này được thực hiện thông qua chuỗi các lớp **tích chập (Convolutional layers)** và **lớp gộp (Pooling layers)**, cụ thể như sau:

* **Lớp tích chập đầu tiên** sử dụng 32 bộ lọc với kích thước nhân tích chập (kernel size) là 5×5. Mục tiêu của lớp này là phát hiện các đặc trưng cơ bản như cạnh, góc, và màu sắc đặc trưng của biển báo.
* **Lớp tích chập tiếp theo** cũng có 32 bộ lọc 5×5, tiếp tục khai thác sâu hơn các mẫu hình học nhỏ trong ảnh.
* **Lớp MaxPooling2D** có kích thước 2×2 giúp giảm kích thước không gian của ảnh đặc trưng, làm giảm số lượng tham số, tăng tốc độ tính toán và hạn chế hiện tượng overfitting.
* Sau lớp pooling đầu tiên, mô hình tiếp tục sử dụng thêm **hai lớp tích chập với 64 bộ lọc** (kernel size 3×3), cho phép mô hình học các đặc trưng phức tạp hơn như hình dáng tổng thể và cấu trúc đặc trưng của các loại biển báo.
* Một lớp MaxPooling2D thứ hai được thêm vào để tiếp tục giảm chiều không gian, đồng thời các lớp **Dropout** với tỷ lệ 0.25 được sử dụng sau mỗi cụm tích chập để giảm nguy cơ quá khớp.

Thông qua chuỗi các tầng chập và gộp này, mô hình có khả năng học được các đặc trưng ở nhiều cấp độ — từ thấp (màu sắc, cạnh) đến cao (hình dạng, bố cục) — giúp nâng cao độ chính xác trong việc phân biệt các loại biển báo khác nhau.

### **3.2.2** .**Phân loại (Classification)**

Sau khi quá trình trích xuất đặc trưng được hoàn tất, mô hình chuyển sang giai đoạn phân loại nhằm dự đoán chính xác loại biển báo giao thông tương ứng. Giai đoạn này được thực hiện bởi các lớp sau:

* **Lớp Flatten**: Chuyển đổi tensor đầu ra từ tầng tích chập thành một vector một chiều, giúp dữ liệu có thể được xử lý bởi các lớp fully connected phía sau.
* **Lớp Dense với 256 nút và hàm kích hoạt ReLU**: Đây là lớp kết nối đầy đủ (fully connected), nơi các đặc trưng được kết hợp và học lại để chuẩn bị cho quyết định phân loại cuối cùng. Lớp này đóng vai trò tổng hợp toàn bộ thông tin đã học được từ các tầng tích chập trước đó.
* **Lớp Dropout với tỷ lệ 0.5**: Được sử dụng để giảm thiểu hiện tượng quá khớp (overfitting) bằng cách vô hiệu hóa ngẫu nhiên 50% số nút trong quá trình huấn luyện, từ đó giúp mô hình tổng quát hóa tốt hơn trên dữ liệu chưa từng thấy.
* **Lớp đầu ra Dense với số nút bằng số lớp (43 lớp)**: Sử dụng hàm kích hoạt **softmax**, lớp này tạo ra một phân phối xác suất trên tất cả các lớp biển báo, trong đó giá trị lớn nhất biểu thị lớp được mô hình dự đoán.

Đầu ra của mô hình là một vector gồm 43 giá trị xác suất tương ứng với 43 loại biển báo giao thông khác nhau. Việc sử dụng hàm **softmax** không chỉ đưa ra dự đoán lớp có xác suất cao nhất, mà còn cung cấp thông tin về độ tin cậy của từng dự đoán – điều này rất hữu ích khi triển khai trong các hệ thống hỗ trợ lái xe tự động, nơi yêu cầu tính chính xác và độ tin cậy cao trong thời gian thực

.

# **CHƯƠNG 4. THỰC NGHIỆM**

## **4.1 Môi trường thực nghiệm**

Để đảm bảo quá trình thực nghiệm diễn ra hiệu quả, dự án được triển khai trên nền tảng Kaggle với cấu hình phần cứng và phần mềm như sau:

#### **4.1.1. Cấu hình phần cứng**

* **Nền tảng**: Kaggle
* **Bộ xử lý đồ họa (GPU)**: NVIDIA Tesla P100

GPU này hỗ trợ tăng tốc các tác vụ học sâu, đặc biệt hiệu quả với các mô hình yêu cầu tính toán ma trận lớn, và mạng neuron tích chập.

#### **4.1.2. Cấu hình phần mềm**

* **Hệ điều hành**: Window 11
* **Ngôn ngữ lập trình**: Python 3.10
* **Thư viện chính sử dụng**:
  + TensorFlow/Keras: Dùng để xây dựng và huấn luyện mô hình mạng nơ-ron tích chập (CNN).
  + NumPy: Xử lý dữ liệu và thao tác mảng hiệu quả
  + Matplotlib/Seaborn: Trực quan hóa dữ liệu.
  + scikit-learn: Cung cấp các công cụ bổ trợ như chia tách dữ liệu, tính toán các chỉ số đánh giá mô hình.

#### **4.1.3. Thiết lập môi trường**

Toàn bộ các thí nghiệm được thực hiện trên nền tảng Kaggle, tận dụng các môi trường cài đặt sẵn, giúp tiết kiệm thời gian cài đặt và tối ưu hóa tài nguyên.

Các thiết lập khác bao gồm:

* Bộ nhớ RAM: 29 GB.
* Dung lượng lưu trữ tạm thời: 2.1T.
* Thời gian thực nghiệm: Khoảng 15 phút cho mỗi lần huấn luyện mô hình khoảng 40000 ảnh biển báo, với 15 epoch và batch size là 8.

#### **4.1.4. Lý do chọn Kaggle**

Kaggle được chọn làm môi trường thực nghiệm vì nhiều lý do vượt trội:

* **Tài nguyên mạnh mẽ và miễn phí**: Với GPU NVIDIA Tesla P100, Kaggle cung cấp sức mạnh tính toán tương đương với các nền tảng tính phí, cho phép xử lý các mô hình lớn mà không tốn chi phí.
* **Tích hợp dễ dàng**: Kaggle hỗ trợ sẵn các thư viện học sâu và công cụ phổ biến, giúp rút ngắn thời gian thiết lập và tập trung hoàn toàn vào quá trình thực nghiệm.
* **Khả năng chia sẻ**: Nền tảng cho phép lưu trữ và chia sẻ mã nguồn, kết quả thực nghiệm, giúp dễ dàng quản lý các phiên bản của dự án.
* **Thân thiện với người dùng**: Giao diện trực quan, hỗ trợ khả năng kiểm tra log và đầu ra trực tiếp trên giao diện web.

## **4.2 Tập dữ liệu**

### **4.2.1. Nguồn dữ liệu**

Dữ liệu sử dụng trong dự án được lấy từ nền tảng **Kaggle**, thuộc tập dữ liệu **GTSRB (German Traffic Sign Recognition Benchmark)** – một tập dữ liệu chuẩn trong lĩnh vực thị giác máy tính cho bài toán nhận diện biển báo giao thông.

* **Biển cấm** (ví dụ: cấm vượt, giới hạn tốc độ),
* **Biển nguy hiểm** (ví dụ: đường trơn, công trường),
* **Biển hiệu lệnh** (ví dụ: đi thẳng, rẽ phải),
* **Biển chỉ dẫn** (ví dụ: khu dân cư, trạm xăng).

### **4.2.2. Số lượng dữ liệu**

Tập dữ liệu GTSRB gồm tổng cộng 40000 **ảnh biển báo giao thông** đã được gán nhãn, chia thành 43 lớp khác nhau. Số lượng ảnh trong mỗi lớp không đồng đều, có những lớp có hơn 2.000 ảnh trong khi một số lớp khác chỉ có vài trăm ảnh. Cụ thể:

* **Tập huấn luyện (Training set):** gồm **31.367 ảnh**, được sử dụng để huấn luyện mô hình học sâu.
* **Tập kiểm thử (Test set):** gồm **7.842** **ảnh**, dùng để đánh giá khả năng tổng quát hóa và độ chính xác của mô hình sau khi huấn luyện.

### **4.2.3. Tiền xử lý dữ liệu**

Để chuẩn bị cho quá trình huấn luyện mô hình nhận diện biển báo giao thông, bước tiền xử lý dữ liệu được thực hiện nhằm đảm bảo chất lượng dữ liệu đầu vào và cải thiện hiệu quả học của mô hình.

Trong nghiên cứu này, các bước tiền xử lý được áp dụng bao gồm:

* **Chuyển đổi ảnh sang không gian màu xám (grayscale):** Giúp giảm chiều dữ liệu, tập trung vào đặc trưng cường độ sáng, thuận tiện cho việc xử lý tiếp theo và giảm độ phức tạp của mô hình.
* **Cân bằng histogram (Histogram Equalization):** Tăng cường độ tương phản của ảnh bằng cách phân bố lại các giá trị pixel, giúp làm nổi bật các chi tiết trong vùng sáng tối khác nhau, hỗ trợ mô hình nhận diện đặc trưng hiệu quả hơn.
* **Chuẩn hóa kích thước ảnh:** Tất cả ảnh được điều chỉnh về kích thước cố định **30×30 pixels** phù hợp với đầu vào mô hình CNN trong đoạn mã.
* **Chuẩn hóa giá trị pixel:** Các giá trị pixel được chuyển về phạm vi [0, 1] bằng cách chia cho 255, giúp tăng tính ổn định khi huấn luyện mô hình.
* **Tăng cường dữ liệu (Data Augmentation):** Các kỹ thuật như xoay ảnh, lật ảnh, dịch chuyển, được áp dụng để mở rộng tập dữ liệu, giúp mô hình tăng khả năng tổng quát hóa và giảm hiện tượng overfitting khi gặp các điều kiện thực tế đa dạng.

Chuỗi các bước tiền xử lý này giúp cải thiện chất lượng dữ liệu đầu vào, đồng thời tăng độ chính xác và độ ổn định của mô hình trong việc nhận diện các biển báo giao thông dưới nhiều điều kiện ánh sáng và môi trường khác nhau.

### **4.2.4. Tăng cường dữ liệu**

Tập dữ liệu biển báo giao thông thường gặp vấn đề mất cân bằng giữa các lớp biển báo khác nhau, khi mà một số loại biển xuất hiện nhiều hơn so với các loại khác. Điều này có thể gây ra hiện tượng mô hình học bị thiên vị, dẫn đến giảm hiệu quả nhận diện chính xác các lớp biển báo ít dữ liệu.

Để khắc phục vấn đề này, trong nghiên cứu, chúng tôi áp dụng các kỹ thuật **tăng cường dữ liệu (Data Augmentation)** chỉ trên tập huấn luyện nhằm cân bằng lại số lượng mẫu cho từng lớp và giúp mô hình học được nhiều đặc trưng đa dạng hơn. Điều này đồng thời cũng làm giảm nguy cơ overfitting và tăng khả năng tổng quát hóa của mô hình khi áp dụng trong thực tế.

Các kỹ thuật tăng cường dữ liệu được sử dụng bao gồm:

* **Xoay ảnh (Rotation):** Quay ảnh biển báo theo các góc nhỏ khác nhau để mô hình có thể nhận diện biển báo ở nhiều vị trí và góc nhìn khác nhau.
* **Dịch chuyển ảnh (Translation):** Dịch chuyển vị trí biển báo trong ảnh nhằm giúp mô hình linh hoạt hơn với các thay đổi vị trí thực tế.
* **Lật ảnh (Flipping):** Lật ngang hoặc dọc ảnh trong những trường hợp phù hợp để tạo thêm các biến thể dữ liệu.
* **Phóng to, thu nhỏ (Scaling):** Thay đổi kích thước ảnh nhẹ nhàng nhằm mô phỏng khoảng cách khác nhau khi chụp biển báo.
* **Thay đổi độ sáng, độ tương phản:** Giúp mô hình nhận diện biển báo hiệu quả hơn trong các điều kiện ánh sáng đa dạng.

Việc áp dụng các kỹ thuật này chỉ thực hiện trên tập huấn luyện nhằm đảm bảo tính khách quan cho tập kiểm thử, tránh làm sai lệch kết quả đánh giá mô hình.

Nhờ đó, mô hình học sâu có thể trích xuất và học được nhiều đặc trưng quan trọng hơn, phù hợp với nhiều điều kiện thực tế đa dạng, từ đó nâng cao hiệu quả và độ chính xác trong việc nhận diện biển báo giao thông.

### **4.2.5. Chia dữ liệu**

Tập dữ liệu sau khi được tăng cường sẽ được chia thành 3 phần chính:

 Tổng số ảnh: 39,209 ảnh (kích thước 30x30, 3 kênh màu)

 Tập huấn luyện: 31,367 ảnh (khoảng 80%)

 Tập kiểm tra: 7,842 ảnh (khoảng 20%)

Quá trình chia dữ liệu được thực hiện ngẫu nhiên nhưng đảm bảo duy trì tỷ lệ cân đối giữa hai nhóm ảnh trong mỗi tập.

### **4.2.6. Lý do chọn tập dữ liệu**

Tập dữ liệu từ Kaggle được chọn vì:

* Độ tin cậy cao: Ảnh được thu thập từ các nguồn y tế uy tín, đảm bảo tính chính xác và chất lượng.
* Độ đa dạng và quy mô phù hợp: Bao gồm nhiều mẫu ảnh từ cả hai nhóm chẩn đoán, giúp mô hình học được nhiều đặc trưng hơn.
* Tính khả dụng: Tập dữ liệu có thể dễ dàng truy cập.

## **4.3 Ứng dụng thực nghiệm**

### **4.3.1. Quy trình huấn luyện mô hình**

Mô hình được triển khai trên nền tảng Kaggle, sử dụng GPU NVIDIA Tesla P100 để tăng tốc quá trình tính toán. Quy trình huấn luyện bao gồm các bước:

* Tiền xử lý và tăng cường dữ liệu từ tập ban đầu.
* Xây dựng mô hình.
* Huấn luyện mô hình trên tập dữ liệu đã được chuẩn bị.

### **4.3.2. Cấu hình huấn luyện**

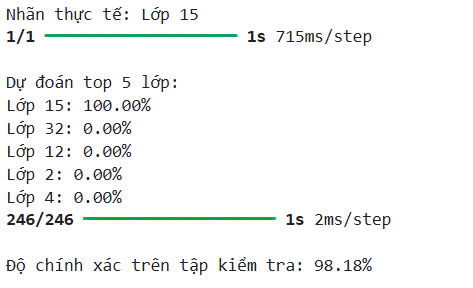
Mô hình được huấn luyện với các tham số sau:

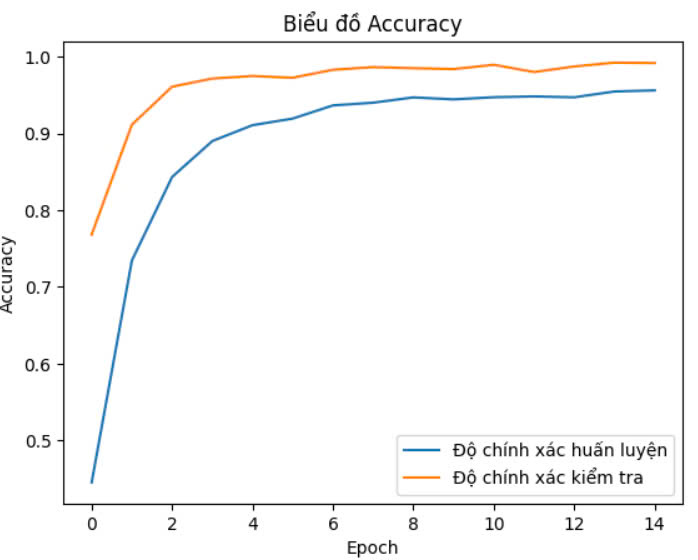
* **Số epoch:** 15 epoch (thời gian huấn luyện đủ để mô hình hội tụ và tránh quá khớp).
* **Batch size:** 8 (kích thước mỗi lô dữ liệu đưa vào huấn luyện giúp cân bằng giữa tốc độ và hiệu quả học).
* **Tối ưu hóa (Optimizer):** Adam với learning rate là 5×10⁻⁵ (giúp mô hình học một cách hiệu quả và ổn định).
* **Hàm mất mát (Loss function):** Categorical Cross-Entropy (phù hợp với bài toán phân loại đa lớp).
* **Metrics đánh giá:** accuracy, precision, recall, F1-score và AUC (các chỉ số toàn diện giúp đánh giá hiệu năng mô hình không chỉ về độ chính xác mà còn về khả năng phân biệt và tổng quát hóa).

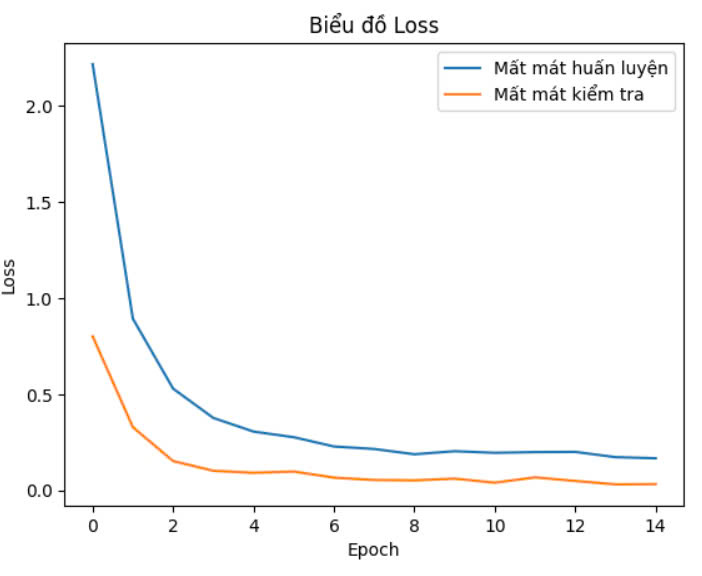
### **4.3.3. Kết quả thực nghiệm**

Sau khi training với 15 epoch, mỗi epoch mất khoảng 169-261 giây với tổng thời gian đào tạo là 942 giây ta và chạy Stratified Five-Fold Cross-Validation ta có được kết quả sau:

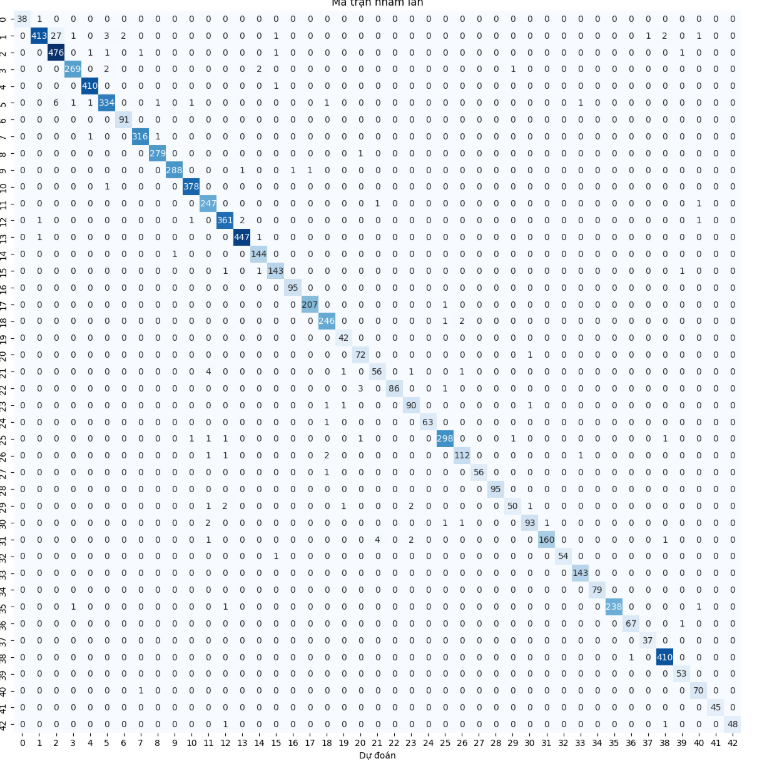
Hình . **Kết quả thử nghiệm và xác thực chéo của mô hình đề xuất**

****

Hình 3. Độ chính xác trong quá trình huấn luyện



Hình 4. Loss trong quá trình huấn luyện



Hình 5. Ma trận nhầm lẫn của mô hình đề xuất

## **4.4 Đánh giá kết quả**

### **4.4.1. So sánh với các mô hình pretrain**

Sau khi huấn luyện mô hình, kết quả được đánh giá và so sánh với các mô hình tiền huấn luyện (pretrained models) như **ResNet50**, **VGG16, MobileNet, InceptionV3**,và **DenseNet**. Các mô hình này đã được sử dụng làm điểm chuẩn để kiểm tra xem các cải tiến trong mô hình có cải thiện hiệu suất so với các mô hình pretrain hay không.

**ResNet50**:

* + - **Accuracy** thấp nhất là 93.53%, cho thấy khả năng phân loại tốt so với các mô hình khác.

**VGG16**:

* + - Hiệu suất cải thiện với **Accuracy** đạt **90.60%**,

**MobileNet**:

* + - Mô hình nhẹ nhưng hiệu quả với **Accuracy** đạt **67,4%**.

**DenseNet121**:

* + - Đạt hiệu suất cao nhất trong các mô hình pretrained với **Accuracy** đạt **53.01%**, thấp nhất trong số các mô hình pretrained

**Mô hình đề xuất:**

* **Accuracy**: Đạt **98.18%,** cao hơn so với **ResNet50** – mô hình tốt nhất trong nhóm pretrained.

# **CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN**

## **5.1 Kết luận**

Trong nghiên cứu này, chúng tôi đã xây dựng và đánh giá một mô hình học sâu dựa trên mạng nơ-ron tích chập (CNN) nhằm nhận diện chính xác các loại biển báo giao thông đường bộ. Qua quá trình tiền xử lý dữ liệu, bao gồm cân bằng histogram và tăng cường dữ liệu, cùng với việc chia tách dữ liệu hợp lý, mô hình đã học được các đặc trưng quan trọng giúp phân biệt hiệu quả giữa các loại biển báo trong nhiều điều kiện ánh sáng và góc nhìn khác nhau.

Kết quả bước đầu cho thấy mô hình CNN đề xuất đạt được hiệu suất nhận diện cao, góp phần nâng cao độ chính xác trong việc phát hiện và phân loại biển báo giao thông. Việc áp dụng kỹ thuật tiền xử lý và tăng cường dữ liệu đã giúp cải thiện khả năng tổng quát hóa của mô hình, đồng thời giảm thiểu hiện tượng overfitting khi huấn luyện.

Những kết quả này khẳng định tiềm năng ứng dụng của hệ thống nhận diện biển báo giao thông trong các hệ thống hỗ trợ lái xe và xe tự hành, góp phần nâng cao an toàn giao thông và giảm thiểu tai nạn giao thông trên đường bộ.

Trong tương lai, nghiên cứu sẽ tiếp tục được mở rộng với việc thử nghiệm trên các bộ dữ liệu lớn hơn và đa dạng hơn, cũng như cải tiến mô hình để tăng độ chính xác và tốc độ xử lý, hướng tới ứng dụng thực tiễn trong môi trường giao thông phức tạp.

## **5.2 Hướng phát triển**

Dựa trên các kết quả đạt được, chúng tôi đề xuất một số hướng cải tiến và phát triển trong tương lai nhằm nâng cao hiệu quả và khả năng ứng dụng của hệ thống nhận diện biển báo giao thông như sau:

1. **Tăng cường dữ liệu**:

* Thu thập và mở rộng thêm các bộ dữ liệu biển báo giao thông từ nhiều quốc gia và vùng lãnh thổ khác nhau để mô hình có thể học được đặc trưng đa dạng hơn, phù hợp với nhiều điều kiện thực tế.
* Áp dụng các kỹ thuật tăng cường dữ liệu tiên tiến hơn như Generative Adversarial Networks (GANs) để tạo ra các ảnh biển báo giả lập có độ chân thực cao, giúp cải thiện khả năng tổng quát hóa của mô hình.

1. **Mở rộng mô hình nhận diện**:

* Phát triển mô hình để nhận diện thêm các loại biển báo đặc biệt hoặc biển báo tạm thời nhằm phục vụ các tình huống giao thông phức tạp.
* Nghiên cứu kết hợp nhận diện biển báo với các dữ liệu cảm biến khác như camera đa chiều hoặc LIDAR để tăng độ chính xác trong điều kiện thời tiết hoặc ánh sáng kém

1. **Tích hợp và ứng dụng thực tế:**

* Xây dựng hệ thống nhận diện biển báo giao thông tích hợp trực tiếp trên các nền tảng xe tự hành hoặc các hệ thống hỗ trợ lái xe an toàn (ADAS).
* Tối ưu hóa tốc độ xử lý và giảm độ trễ để phù hợp với yêu cầu vận hành thời gian thực trên các thiết bị phần cứng giới hạn.

1. **Cải tiến hiệu suất mô hình**:

* Thử nghiệm các kỹ thuật tối ưu hóa tham số và kiến trúc mạng hiện đại như Vision Transformers (ViT) hoặc các phương pháp học tự giám sát (self-supervised learning) để nâng cao khả năng trích xuất đặc trưng và độ chính xác của mô hình.
* Áp dụng các phương pháp ensemble hoặc attention để cải thiện khả năng phân loại biển báo trong các tình huống phức tạp hoặc nhiễu nền.

1. **Giảm chi phí tính toán**:

* Tinh giản mô hình (model pruning, quantization) để giảm kích thước và tài nguyên tính toán, giúp dễ dàng triển khai trên các thiết bị di động hoặc các hệ thống nhúng trong xe ô tô.

**TÀI LIỆU THAM KHẢO**

[1] Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2012). Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. Neural Networks, 32, 323-332. <https://doi.org/10.1016/j.neunet.2012.02.016>

[2] Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2012). Multi-column deep neural network for traffic sign classification. Neural Networks, 32, 333-338. <https://doi.org/10.1016/j.neunet.2012.02.023>

[3] M. Sermanet and Y. LeCun, Traffic sign recognition with multi-scale convolutional networks, in Proceedings of the International Joint Conference on Neural Networks (IJCNN), 2011, pp. 2809-2813. <https://doi.org/10.1109/IJCNN.2011.6033589>

[4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems (NeurIPS). <https://arxiv.org/abs/1406.2661>

# [5] Traffic Signs Recognition using CNN and Keras in Python. <https://www.analyticsvidhya.com/blog/2021/12/traffic-signs-recognition-using-cnn-and-keras-in-python/>