**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ GIAO THÔNG VẬN TẢI**

**KHOA CÔNG NGHỆ THÔNG TIN**

**---🙢🏵🙠---**

A picture containing text, clipart

Description automatically generated

**BÁO CÁO MÔN HỌC**

**TRÍ TUỆ NHÂN TẠO**

**Tên đề tài:**

**ỨNG DỤNG MẠNG NEURAL** **TRONG**

**PHÂN LOẠI VÀ NHẬN DIỆN BIỂN SỐ XE**

|  |  |
| --- | --- |
| **Sinh viên thực hiện:** | **PHAN MINH KHOA**  **ĐINH XUÂN HIẾU**  **NGUYỄN MINH VŨ**  **NUYỄN VĂN KHÁNH**  **VŨ DOÃN HÀ** |
| **Lớp:** | **72DCHT21** |
| **Giáo viên hướng dẫn:** | **ĐOÀN THỊ THANH HẰNG** |

**Hà Nội – 2023**

# **BẢNG PHÂN CÔNG**

Thành viên nhóm 10:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STT** | **Mã SV** | **Thành viên** | **Công việc** | **Ghi chú** |
| 1 | 72DCHT20022 | Nguyễn Minh Vũ | Thiết kế giao diện  người dùng | Nhóm trưởng |
| 2 | 72DCHT20021 | Đinh Xuân Hiếu | Thu thập dữ liệu |  |
| 3 | 72DCHT20039 | Vũ Doãn Hà | Làm báo cáo, thu thập hình ảnh |  |
| 4 | 72DCHT20027 | Phan Minh Khoa | Huấn luận hệ thống |  |
| 5 | 72DCHT20020 | Nguyễn Văn Khánh | Kiểu tra hệ thống |  |

**MỤC LỤC**

**[LỜI NÓI ĐẦU 4](#_Toc7089)**

**[CHƯƠNG 1: GIỚI THIỆU TỔNG QUAN 5](#_Toc9531)**

[1.1. Giới thiệu về trí tuệ nhân tạo. 5](#_Toc12796)

[1.1.1. Trí tuệ nhân tạo là gì ? 5](#_Toc32155)

[1.1.2. Lịch sử AI. 6](#_Toc2315)

[1.1.3. Lĩnh Vực AI. 7](#_Toc4094)

[1.2 Giới thiệu đề tài 8](#_Toc25296)

[1.3 Lí do chọn đề tài 9](#_Toc18701)

[1.4 Nội dung đề tài 9](#_Toc31019)

[1.5 Mục tiêu đề tài 11](#_Toc893)

[1.6 Giới hạn đề tài 12](#_Toc2806)

[1.7 Phương pháp nghiên cứu 12](#_Toc24045)

[1.8 Đối tượng và phạm vi nghiên cứu 12](#_Toc21204)

**[CHƯƠNG 2: CƠ SỞ LÝ THUYẾT 13](#_Toc27886)**

[2.1 Giới thiệu mạng Neural tích chập CNN 13](#_Toc11682)

[2.2 Mô hình mạng neural convolutional (CNN) 16](#_Toc4657)

[1. Kiến Trúc: 16](#_Toc13678)

[2. Lớp Tích Chập và Lớp Gộp: 16](#_Toc18984)

[3. Hàm Kích Hoạt ReLU: 16](#_Toc32025)

[4. Dropout: 17](#_Toc2969)

[5. Normalization: 17](#_Toc14862)

[6. Fully Connected Layers: 17](#_Toc28925)

[7. Hàm Loss: 17](#_Toc21688)

**[CHƯƠNG 3: THIẾT KẾ HỆ THỐNG 18](#_Toc24990)**

[3.1. Mô hình hệ thống . 18](#_Toc23962)

[3.2 Thiết kế hệ thống bằng ngôn ngữ Matlab. 18](#_Toc24779)

**[CHƯƠNG 4: KẾT QUẢ 25](#_Toc5176)**

[4.1 Hệ thông nhận diện loại biển số xe 25](#_Toc7078)

[4.2 Hệ thống nhận diện đọc biển số xe 26](#_Toc22063)

**[KẾT LUẬN 27](#_Toc21070)**

**[TÀI LIỆU THAM KHẢO 28](#_Toc16432)**

**LỜI NÓI ĐẦU**

Ngày nay, chúng ta đã bước vào thời đại 4.0, thời kì của Công nghệ thông

tin, đặc biệt là trí tuệ nhân tạo là yếu tố quan trọng nhất quyết định sự thành công của mỗi ngành hay mỗi quốc gia. Trí tuệ nhân tạo đã và đang làm thay đổi

cuộc sống của chúng ta, với sự phát triển mạnh mẽ của việc áp dụng các nghiên

cứu về trí tuệ nhân tạo áp dụng cho cuộc sống. Tất cả các ngành như: Quân đội,

y tế, giáo dục,kinh tế thương mại, tài chính,… Đều có thể áp dụng trí tuệ nhân

một cách rộng rãi. Việc áp dụng trí tuệ nhân tạo để giải quyết các vấn đề trong

xã hội và việc phát triển kinh tế đang được nhà nước khuyến khích và đầu tư rất

lớn.Trên thế giới cũng như Việt Nam, công nghệ thông tin có ảnh hưởng rất

mạnh mẽ đến sự phát triển của đất nước và thế giới đặc biệt là trí tuệ nhân tạo.

Nó chở thành một yếu tố không thể thiếu và có tính quyết định đến sự thành

công hay thất bại của nhiều ngành ở nước ta.

Công nghệ thông tin đang phát triển với tốc độ khá mạnh mẽ và được ứng

dụng rộng rãi trong tất cả các lĩnh vực, đặc biệt là trong công tác ứng dụng công

nghệ vào trong cuộc sống. Như chúng ta đã biết, sức mạnh của một nền kinh tế

phụ thuộc rất lớn vào các hoạt động trong nước của các doanh nghiệp, vì vậy sự

thành công trong kinh doanh của doanh nghiệp không những là mục tiêu của

riêng doanh nghiệp, mà nó còn là nhân tố quyết định vị thế của đất nước trên

trường quốc tế. Nhà nước đang đi vào phát triển dịch vụ, và đầu tư mạnh mẽ vào

trí tuệ nhân tạo hay còn gọi là AI. Vì thế, việc đó đầu nó và phát triển nó đang là

một xu thế rất hot và rất được ưu chuộng hiện này.Chính vì vậy thông qua việc

học môn trí tuệ nhân tạo (AI) nhốm em đã nghĩ ra ý tưởng là ứng dụng mạng neural trong kiểm soát quy trình và chất lượng trong sản phục vụ cho việc chứng minh áp dụng trí tuệ nhân tạo mang lại lợi ích tối ưu về trí tuệ và thời gian cho con người.

Với sự hướng dẫn, động viên tận tình của cô Đoàn Thị Thanh Hằng , em đã hiểu được cách thức hoạt động của mô hình mạng neural và hoàn thành bài báo cáo này. Do chưa có nhiều kinh nghiệm nghiên cứu, thực hành nên em cũng không tránh khỏi những thiếu sót. Em rất mong nhận được sự thông cảm và góp ý của thầy cô để đề tài của em được hoàn thiện hơn.

***Chúng em xin chân thành cảm ơn!***

# **CHƯƠNG 1: GIỚI THIỆU TỔNG QUAN**

## **Giới thiệu về trí tuệ nhân tạo.**

### Trí tuệ nhân tạo là gì ?

Để hiểu về trí tuệ nhân tạo (Artificial Intelligence - AI), chúng ta có thể bắt đầu với khái niệm về sự bay nhân tạo, tức là máy bay. Loài người đã lâu đã khao khát tạo ra một thiết bị có khả năng di chuyển trên không mà không phụ thuộc vào địa hình dưới mặt đất, nói cách khác là một máy có thể bay. Không ngạc nhiên khi những ý tưởng đầu tiên về máy bay xuất phát từ nghiên cứu về cách con chim bay. Những máy bay đầu tiên được thiết kế dựa trên nguyên lý "vỗ cánh" như con chim chỉ có thể bay được quãng đường ngắn. Lịch sử hàng không thực sự bước sang một giai đoạn mới khi anh em nhà Wright thiết kế máy bay dựa trên nguyên lý của khí động học (aerodynamics).

Các máy bay hiện đại ngày nay không nhất thiết phải tuân theo nguyên lý bay của con chim, nhưng vẫn có khả năng bay giống như chim về dáng vẻ và thậm chí tốt hơn. Quay lại câu hỏi về Trí tuệ nhân tạo, nó là sự thông minh của máy mà con người tạo ra. Ngay từ khi chiếc máy tính điện tử đầu tiên xuất hiện, những nhà khoa học máy tính đã hướng đến việc phát triển hệ thống máy tính (bao gồm cả phần cứng và phần mềm) sao cho nó có khả năng thông minh như loài người. Mặc dù ước mơ này vẫn còn xa vời, nhưng đã có những thành tựu đáng kể. Các hệ thống như phần mềm chơi cờ trên siêu máy tính GeneBlue đã có thể đánh bại vua cờ thế giới và phần mềm có khả năng chứng minh các bài toán hình học.

Trí tuệ nhân tạo không chỉ là việc mô phỏng cách con người thực hiện một nhiệm vụ mà còn bao gồm khả năng làm tốt hơn hoặc ít nhất là tương đương với con người trong một số lĩnh vực. Có nhiều cách tiếp cận để tạo ra trí thông minh của máy, bao gồm việc nghiên cứu cách bộ não người tạo ra trí thông minh và mô phỏng nguyên lý đó. Tuy nhiên, có những cách tiếp cận hoàn toàn khác, không tuân theo nguyên lý sản sinh trí thông minh của con người mà vẫn tạo ra máy có khả năng thông minh, giống như máy bay hiện đại có cơ chế bay khác biệt với con chim.

Tóm lại, trí tuệ nhân tạo là khả năng của máy khi thực hiện các nhiệm vụ mà con người thường xuyên phải đối mặt. Đánh giá sự thông minh của máy không chỉ dựa trên cách nó thực hiện một nhiệm vụ mà còn dựa trên kết quả hoặc dáng vẻ bên ngoài mà nó thể hiện. Môn học Trí tuệ nhân tạo nhằm cung cấp các phương pháp để máy tính có khả năng giải quyết bài toán, lập luận, học, và thực hiện nhiều công việc khác nhau.

### Lịch sử AI.

Vào năm 1943, Warren McCulloch và Walter Pitts bắt đầu thực hiện nghiên cứu về ba cơ sở lý thuyết cơ bản: triết học cơ bản và chức năng của các neuron thần kinh; phân tích các mệnh đề logic; và lý thuyết dự đoán của Turing. Các tác giả nghiên cứu đề xuất mô hình neuron nhân tạo, mỗi neuron đặc trưng bởi hai trạng thái "bật" và "tắt" và phát hiện mạng neuron có khả năng học.

Thuật ngữ "Trí tuệ nhân tạo" (Artificial Intelligence - AI) được đặt ra bởi John McCarthy tại Hội thảo đầu tiên về chủ đề này vào mùa hè năm 1956. McCarthy cũng đề xuất ngôn ngữ lập trình Lisp, một trong những ngôn ngữ lập trình hàm tiêu biểu được sử dụng trong lĩnh vực AI. Alan Turing đưa ra "Turing test" như một phương pháp kiểm chứng hành vi thông minh.

Trong thập kỷ 60 và 70, Joel Moses phát triển chương trình Macsyma - chương trình toán học sử dụng cơ sở tri thức đầu tiên thành công. Marvin Minsky và Seymour Papert đưa ra các chứng minh về giới hạn của các mạng neuron đơn giản. Ngôn ngữ lập trình logic Prolog xuất hiện và được phát triển bởi Alain Colmerauer. Ted Shortliffe xây dựng thành công một số hệ chuyên gia đầu tiên hỗ trợ chẩn đoán y học, sử dụng ngôn ngữ luật để biểu diễn tri thức và suy diễn.

Vào đầu những năm 1980, nghiên cứu về AI liên quan đến các hệ chuyên gia (expert systems) - một dạng của chương trình AI mô phỏng tri thức và kỹ năng phân tích của chuyên gia con người - đạt được sự thành công.

Trong những năm 1990 và đầu thế kỷ 21, AI đạt được những thành tựu lớn, được áp dụng trong logic, khám phá dữ liệu, chẩn đoán y học và nhiều lĩnh vực ứng dụng khác trong công nghiệp. Sự thành công này phản ánh sự tăng khả năng tính toán của máy tính, sự tập trung vào giải quyết các bài toán cụ thể, xây dựng mối quan hệ giữa AI và các lĩnh vực khác để giải quyết các bài toán tương tự, và sự chuyển giao của các nhà nghiên cứu sang các phương pháp toán học chặt chẽ và chuẩn xác.

### Lĩnh Vực AI.

Lập luận và suy diễn tự động là hai khái niệm chủ chốt trong lĩnh vực Trí Tuệ Nhân Tạo (AI), đóng vai trò quan trọng trong quá trình xử lý tri thức và ra quyết định. Trong AI, lập luận được định nghĩa là quá trình suy luận logic, sử dụng để rút ra các kết luận mới từ những giả định đã được đưa ra và được biểu diễn dưới dạng cơ sở tri thức. Điều này đòi hỏi có các phương pháp lưu trữ cơ sở tri thức và thủ tục lập luận trên nền cơ sở tri thức đó.

Để máy tính có khả năng lưu trữ và xử lý tri thức, cần sự hiện diện của các phương pháp biểu diễn tri thức. Ngôn ngữ biểu diễn tri thức được đánh giá là "tốt" nếu nó có khả năng biểu đạt cao và hiệu quả của thuật toán lập luận trên nó. Tính biểu đạt thể hiện khả năng biểu diễn rộng lớn trong một miền ứng dụng, trong khi tính hiệu quả của thuật toán lập luận thể hiện chi phí về thời gian và không gian. Thách thức là tìm ra sự cân bằng giữa tính biểu đạt và hiệu quả.

Lập kế hoạch là khả năng suy ra các mục đích cần đạt được và xác định dãy hành động cần thực hiện để đạt được mục đích đó. Trong AI, quá trình lập kế hoạch được sử dụng để tự động hóa quyết định và giải quyết vấn đề.

Học máy, một lĩnh vực nổi bật của AI, tập trung vào khả năng máy tính học từ dữ liệu và cải thiện hiệu suất theo thời gian. Xử lý ngôn ngữ tự nhiên là một nhánh khác của AI, chuyên tâm vào ứng dụng trên ngôn ngữ của con người như nhận dạng tiếng nói, dịch tự động và tìm kiếm thông tin.

Cuối cùng, hệ chuyên gia là một ứng dụng quan trọng của lập luận và suy diễn trong AI. Cung cấp khả năng suy luận, các hệ chuyên gia như MYCIN và DENDRAL trong lĩnh vực y học đã đóng góp quan trọng trong việc đưa ra quyết định dựa trên thông tin phức tạp.

Tóm lại, lập luận và suy diễn tự động không chỉ là khái niệm mà còn là cốt lõi của sự phát triển trong lĩnh vực Trí Tuệ Nhân Tạo. Chúng tạo nên cơ sở hạ tầng cho máy tính có khả năng học và ra quyết định thông minh, đóng vai trò quyết định trong hệ thống AI ngày nay.

## **1.2 Giới thiệu đề tài**

Mạng neural tích chập CNN là một trong những mô hình mạng Học sâu phổ biến nhất hiện nay, có khả năng nhận dạng và phân loại hình ảnh với độ chính xác rất cao, thậm chí còn tốt hơn con người trong nhiều trường hợp. Mô hình này đã và đang được phát triển, ứng dụng vào các hệ thống xử lý ảnh lớn của Facebook, Google hay Amazon… cho các mục đích khác nhau, như các thuật toán gắn thẻ tự động, tìm kiếm ảnh hoặc gợi ý sản phẩm cho người tiêu dùng.

Mạng CNN với kiến trúc thay đổi, có khả năng xây dựng liên kết chỉ sử dụng một phần cục bộ trong ảnh kết nối đến node trong lớp tiếp theo thay vì toàn bộ ảnh như trong mạng nơ ron truyền thẳng.Nhờ sự phát triển mạng CNN, chúng em có cơ sở thực hiện đề tài “Ứng dụng mạng Neural trong nhận diện và phân loại biển số xe”.

Việc ứng dụng học máy (machine learning) trong nhận diện biển số xe là một trong những ứng dụng phổ biến và hữu ích trong lĩnh vực thị giác máy tính. Quá trình này giúp tự động hóa và cải thiện khả năng nhận diện biển số xe từ hình ảnh, giảm thiểu sự phụ thuộc vào các phương pháp truyền thống.

Mô hình học máy được huấn luyện trên một lượng lớn dữ liệu biển số xe, trong đó mỗi biển số xe được gắn với nhãn tương ứng. Các thuật toán như Convolutional Neural Networks (CNNs) thường được sử dụng để học đặc trưng từ ảnh và tự động trích xuất các đặc điểm quan trọng.

Thông qua việc sử dụng dữ liệu huấn luyện, mô hình có khả năng học được các mẫu phổ biến trong biển số xe, bao gồm cả sự biến đổi về kích thước, màu sắc, và vị trí trên hình ảnh. Khi đã được huấn luyện đủ, mô hình có thể nhận diện biển số xe trên các hình ảnh mới một cách chính xác và hiệu quả.

Ứng dụng này không chỉ giúp tăng cường hiệu suất của các hệ thống giám sát an ninh và giao thông mà còn có ứng dụng rộng rãi trong các lĩnh vực như quản lý đỗ xe, thu phí tự động, và theo dõi phương tiện giao thông công cộng. Việc tích hợp học máy trong nhận diện biển số xe không chỉ mang lại sự thuận tiện mà còn đóng góp vào việc xây dựng các hệ thống thông minh và an toàn trên đường.

**1.3 Lí do chọn đề tài**

Cùng với sự phát triển khoa học kỹ thuật, nhu cầu đi lại của con người ngày càng tăng, lưu lượng giao thông ngày càng lớn. Với số lượng phương tiện giao thông ngày càng lớn và còn không ngừng tăng thì việc quản lý các phương tiện giao thông gặp rất nhiều khó khăn do đó cần có một hệ thống tự động. Một trong những hệ thống đó là hệ thống nhận dạng biển số xe. Đó là một hệ thống có khả năng “đọc” và “hiểu”các biển số xe một cách tự động. Trong luận văn này, người thực hiện đã xây dựng hệ thống để quản lý bãi giữ xe bởi vì các bãi giữ xe hiện nay còn có vấn đề bất cập như: Tốn nhiều nhân công. An toàn chưa cao, vẫn còn xảy ra hiện tượng mất xe. Vé xe bằng giấy, dễ bị mất hay nhàu nát.

**1.4 Nội dung đề tài**

Cũng như mọi hệ thống khác, hệ thống này yêu cầu có phần cứng và phần mềm. Phần cứng có phần chính là WebCam để thu nhận hình ảnh và phần mềm sẽ phân tích hình ảnh đó để lấy ra các ký tự trên biển số xe.

Đầu tiên, hình ảnh trong biến imgfildata được chuyển đổi sang định dạng grayscale nếu nó không phải là ảnh grayscale ban đầu. Sau đó, một ngưỡng (threshold) xám được áp dụng để chuyển đổi hình ảnh sang hình ảnh nhị phân, làm nổi bật các ký tự trên nền xám. Nhị phân hóa là quá trình biến một ảnh xám thành ảnh nhị phân. Ảnh nhị phân chỉ chứa hai giá trị 0 hoặc 1 (hoặc 0 và 255 tùy theo quy định của cấu trúc ảnh). Theo đó, giá trị 0 sẽ là giá trị ứng với những điểm đen trên ảnh và giá trị 1 (hoặc 255) sẽ là giá trị ứng với những điểm trắng. Công thức để nhị phân ảnh là đặt tất cả các giá trị của ảnh xám về 0 nếu chúng bé hơn một ngưỡng T cho trước và  về 255 nếu chúng lớn hơn ngưỡng T đó. Vấn để ở đây là với ngưỡng T như thế nào thì cho ta được tấm ảnh nhị phân như ý muốn, nghĩa là ảnh làm nổi các vùng đối tượng và nền. Việc xác định giá trị T là rất khó vì chúng phụ thuộc vào từng điều kiện chiếu sáng khác nhau của môi trường. Với môi trường này T nhận một giá trị, môi trường khác lại một giá trị khác. Các kỹ thuật nhị phân hóa có thể được sử dụng, ví dụ như phân ngưỡng cố định (fixed thresholding) hoặc phân ngưỡng tự động (adaptive thresholding), tùy thuộc vào đặc điểm của hình ảnh.

Ngưỡng hóa ảnh thành ảnh nhị phân: quá trình ngưỡng hóa ảnh thành ảnh nhị phân là một bước quan trọng trong việc chuẩn bị dữ liệu hình ảnh cho quá trình nhận diện biển số xe. Bước này giúp chuyển đổi hình ảnh từ dạng liên tục thành dạng nhị phân, trong đó các vùng sáng và tối được phân biệt nhau dựa trên một ngưỡng cường độ pixel.Đầu tiên, sử dụng hàm `graythresh`, ngưỡng sáng được tự động xác định dựa trên phân phối cường độ pixel trong ảnh gốc. Kết quả ngưỡng này được sử dụng để chuyển đổi ảnh thành ảnh nhị phân bằng hàm `im2bw`. Các pixel có cường độ lớn hơn ngưỡng sẽ trở thành 1 (trắng), trong khi các pixel có cường độ thấp hơn ngưỡng sẽ trở thành 0 (đen).

Loại bỏ nhiễu và tinh chỉnh ảnh nhị phân: sử dụng hàm `bwareaopen` để loại bỏ các vùng nhỏ có diện tích ít hơn một ngưỡng cụ thể. Điều này giúp làm sạch và tạo ra ảnh nhị phân cuối cùng, trong đó các vùng quan trọng và đối tượng sáng trên nền tối được làm nổi bật, chuẩn bị cho các bước xử lý và nhận diện tiếp theo trong quá trình nhận diện biển số xe.

Phát hiện ký tự: Hình ảnh nhị phân được phân đoạn để tách riêng các ký tự. Các thuật toán phân đoạn như Connected Component Labeling (CCL) có thể được sử dụng để gắn nhãn cho các ký tự riêng biệt. Các thuộc tính của các vùng như hình bao quanh (bounding box) được trích xuất để xác định vị trí và kích thước của từng ký tự.

Nhận dạng ký tự: Mỗi ký tự được tách ra từ hình ảnh và đưa vào quá trình nhận dạng. Các phương pháp nhận dạng ký tự có thể sử dụng, bao gồm các mô hình Machine Learning như Convolutional Neural Networks (CNN), kỹ thuật Template Matching, hoặc các thuật toán nhận dạng ký tự tùy chỉnh. Đối với mỗi ký tự, quá trình nhận dạng sẽ trả về kết quả tương ứng là một ký tự hoặc một danh sách các ký tự có thể tương đồng.

## **1.5 Mục tiêu đề tài**

Đề tài "Ứng dụng mạng Neural trong kiểm soát quy trình và chất lượng trong sản xuất" có thể áp dụng trong nhiều lĩnh, là một phần quan trọng trong quản lý bãi giữ xe. Nó bao gồm các hoạt động và hệ thống được thiết lập để đảm bảo rằng quy trình sản xuất diễn ra hiệu quả, đồng nhất, và đáp ứng được các tiêu chí chất lượng cụ thể. Kiểm soát quy trình bãi giữ xe giúp đảm bảo rằng các phương tiện xe cộ được quản lí nghiêm ngặt, đảm bảo an ninh, bên cạnh đó còn cõ thể ứng dụng việc nhận dạng phân loại biển số xe vào việc hộ trợ điều tra đường đi của phương tiện. Điều này bao gồm việc theo dõi các thông số kí tự, nhận diện kí tự, hình ảnh và đảm bảo rằng đọc và phân loại được các biển số xe của phương tiện di chuyển. Từ đó, giúp hạn chế những vấn đề bất cập của bài giữ xe truyền thống hiện nay, xây dựng một mô hình hệ thống quản lí bãi giữ xe tự động.

**1.6 Giới hạn đề tài**

Việc có nhiều biển số xe với định dạng và độ sáng khác nhau gây khó khăn cho việc nhận dạng. Do thời gian thực hiện đề tài không cho phép nên người thực hiện giới hạn các biển số và điều kiện như sau: Biển số có chữ đen, còn nguyên vẹn, không bị tróc sơn hay rỉ sét, không bị che khuất, không nghiêng, biển số xe chiếm ưu thế trong ảnh. Hình chụp biển số không bị mờ, ký tự biển số còn phân biệt, nhận dạng được bằng trực quan. Không bị nhiễu bởi ánh sáng làm ảnh chụp bị chói.

**1.7 Phương pháp nghiên cứu**

Trong phần này, chúng em thiết kế mô hình nhận dạng chẩn đoán dựa trên mạng CNN và ứng dụng học máy (machine learning) trong lĩnh vực thị giác máy tính. Dựa trên các phương pháp nghiên cứu chính như phương pháp phân tích, tham khảo tài liệu, phương pháp tổng hợp tài liệu lý thuyết…

## **1.8 Đối tượng và phạm vi nghiên cứu**

Các đối tượng cần nghiên cứu và phạm vi nghiên cứu có thể giải quyết được đề tài:

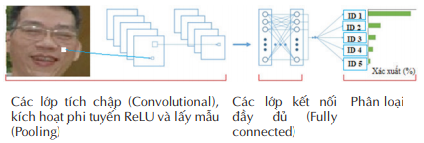
* Đối tượng nghiên cứu*:* Trí tuệ nhân tạo, lập trình matlab.
* Phạm vi nghiên cứu:Mạng CNN, thuật toán tối ưu hóa SGD, sử dụng Toolbox Deep Learning của MATLAB và các hàm liên quan để xây dựng và huấn luyện một mô hình phân loại ảnh dựa trên AlexNet.

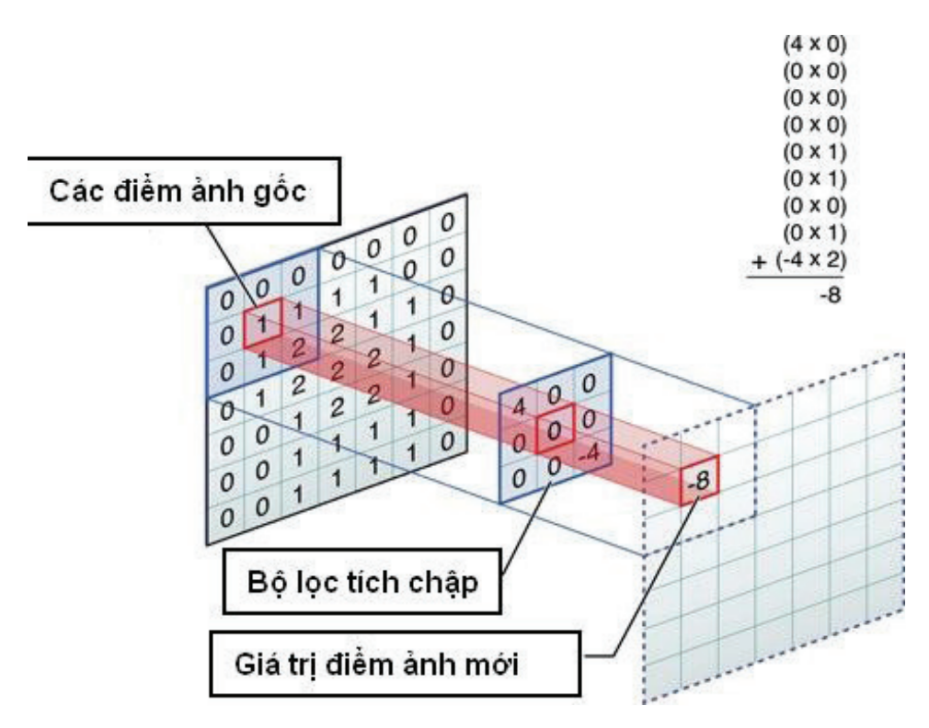
# **CHƯƠNG 2: CƠ SỞ LÝ THUYẾT**

## **2.1 Giới thiệu mạng Neural tích chập CNN**

Hình 2.1 trình bày một kiến trúc mạng CNN, các lớp cơ bản trong một mạng CNN bao gồm: lớp tích chập (Convolutional); lớp kích hoạt phi tuyến ReLU (Rectified Linear Unit); lớp lấy mẫu (Pooling); lớp kết nối đầy đủ (Fully connected) được thay đổi về số lượng và cách sắp xếp để tạo ra các mô hình huấn luyện phù hợp cho từng bài toán khác nhau. Các lớp tích chập (Convolutional), kích hoạt phi tuyến ReLU và lấy mẫu (Pooling) Các lớp kết nối đầy đủ (Fully connected) Phân loại Hình 1. Kiến trúc cơ bản của một mạng CNN.

Lớp tích chập: đây là thành phần quan trọng nhất trong mạng CNN, thể hiện sự liên kết cục bộ thay vì kết nối toàn bộ các điểm ảnh. Các liên kết cục bộ được tính toán bằng phép tích chập giữa các giá trị điểm ảnh trong một vùng ảnh cục bộ với các bộ lọc filters có kích thước nhỏ.

Kiến trúc cơ bản của mạng CNN



Bộ lọc tích chập được sử dụng trên ma trận điểm ảnh

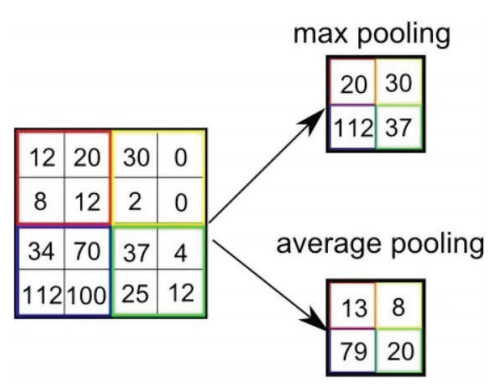
Lớp tích chập: đây là thành phần quan trọng nhất trong mạng CNN, thể hiện sự liên kết cục bộ thay vì kết nối toàn bộ các điểm ảnh. Các liên kết cục bộ được tính toán bằng phép tích chập giữa các giá trị điểm ảnh trong một vùng ảnh cục bộ với các bộ lọc filters có kích thước nhỏ.

Bộ lọc tích chập được sử dụng trên ma trận điểm ảnh. Trong hình 2.2, bộ lọc được sử dụng là một ma trận có kích thước 3x3, bộ lọc này dịch chuyển lần lượt qua từng vùng ảnh đến khi hoàn thành quét toàn bộ bức ảnh, tạo ra một bức ảnh mới có kích thước nhỏ hơn hoặc bằng với kích thước ảnh đầu vào. Kích thước này được quyết định tùy theo kích thước các khoảng trắng được thêm ở viền bức ảnh gốc và được tính theo công thức sau:

Trong đó: O: kích thước ảnh đầu ra; i: kích thước ảnh đầu vào; p: kích thước khoảng trắng phía ngoài viền của ảnh gốc; k: kích thước bộ lọc; s: bước trượt của bộ lọc. Như vậy, sau khi đưa một bức ảnh đầu vào cho lớp tích chập nhận được kết quả đầu ra là một loạt ảnh tương ứng với các bộ lọc đã được sử dụng để thực hiện phép tích chập. Các trọng số của các bộ lọc này được khởi tạo ngẫu nhiên trong lần đầu tiên và sẽ được cập nhật trong quá trình huấn luyện. - Lớp kích hoạt phi tuyến ReLU: được xây dựng để đảm bảo tính phi tuyến của mô hình huấn luyện sau khi đã thực hiện một loạt các phép tính toán tuyến tính qua các lớp tích chập. Lớp kích hoạt phi tuyến sử dụng các hàm kích hoạt phi tuyến như ReLU hoặc sigmoid, tanh… để giới hạn phạm vi biên độ cho phép của giá trị đầu ra. Trong số các hàm kích hoạt này, hàm ReLU được chọn do cài đặt đơn giản, tốc độ xử lý nhanh mà vẫn đảm bảo được tính toán hiệu quả. Phép tính toán của hàm ReLU chỉ đơn giản là chuyển tất cả các giá trị âm thành giá trị 0. Lớp ReLU được áp dụng ngay phía sau lớp tích chập, với đầu ra là một ảnh mới có kích thước giống với ảnh đầu vào, các giá trị điểm ảnh cũng hoàn toàn tương tự, trừ các giá trị âm đã bị loại bỏ.

𝑓(𝑥) = 𝑚𝑎𝑥(0, 𝑥) (2.2)

Lớp lấy mẫu: được đặt sau lớp tích chập và lớp ReLU để làm giảm kích thước ảnh đầu ra trong khi vẫn giữ được các thông tin quan trọng của ảnh đầu vào. Việc giảm kích thước dữ liệu có tác dụng làm giảm được số lượng tham số cũng như tăng hiệu quả tính toán. Lớp lấy mẫu cũng sử dụng một cửa sổ trượt để quét toàn bộ các vùng trong ảnh như lớp tích chập, và thực hiện phép lấy mẫu thay vì phép tích chập, sẽ chọn lưu lại một giá trị duy nhất đại diện cho toàn bộ thông tin của vùng ảnh đó.



Phương thức Avarage Pooling và Max Pooling

Hình ảnh thể hiện các phương thức lấy mẫu thường được sử dụng nhất hiện nay, đó là Max Pooling (lấy giá trị điểm ảnh lớn nhất) và Avarage Pooling (lấy giá trị trung bình của các điểm ảnh trong vùng ảnh cục bộ).

Như vậy, với mỗi ảnh đầu vào được đưa qua lấy mẫu sẽ thu được một ảnh đầu ra tương ứng, có kích thước giảm xuống đáng kể nhưng vẫn giữ được các đặc trưng cần thiết cho quá trình tính toán và nhận dạng.

Lớp kết nối đầy đủ: được thiết kế tương tự như trong mạng nơ ron truyền thống, tất cả các điểm ảnh được kết nối đầy đủ với node trong lớp tiếp theo.

So với mạng nơ ron truyền thống, các ảnh đầu vào của lớp này đã có kích thước được giảm bớt rất nhiều, đồng thời vẫn đảm bảo các thông tin quan trọng của ảnh cho việc nhận dạng. Do vậy, việc tính toán nhận dạng sử dụng mô hình truyền thẳng đã không còn phức tạp và tốn nhiều thời gian như trong mạng nơ ron truyền thống.

## **2.2 Mô hình mạng neural convolutional (CNN)**

AlexNet là một mô hình mạng neural convolutional (CNN) tiên tiến, được thiết kế để giải quyết bài toán nhận diện hình ảnh. Mô hình này đã giành chiến thắng trong cuộc thi ImageNet Large Scale Visual Recognition Challenge (ILSVRC) năm 2012 và đã mở đường cho sự phổ biến của các mô hình CNN mạnh mẽ trong thị giác máy tính.

Dưới đây là một số đặc điểm chính của AlexNet:

1. Kiến Trúc:

AlexNet có kiến trúc sâu với 8 lớp, bao gồm 5 lớp tích chập và 3 lớp fully connected (hoặc dense). Là một trong những mô hình đầu tiên với kiến trúc sâu, nó đã đặt nền tảng cho các kiến trúc mạng neural sâu sau này.

2. Lớp Tích Chập và Lớp Gộp:

Các lớp tích chập giúp trích xuất các đặc trưng cấp cao từ ảnh, trong khi các lớp gộp giảm kích thước không gian của đặc trưng.

3. Hàm Kích Hoạt ReLU:

AlexNet sử dụng hàm kích hoạt ReLU (Rectified Linear Unit) thay vì các hàm kích hoạt truyền thống như sigmoid hay tanh. ReLU giúp ngăn chặn hiện tượng biến mất gradient và làm tăng tốc độ học của mô hình.

4. Dropout:

Để giảm overfitting, AlexNet sử dụng lớp dropout, tắt ngẫu nhiên một số nơ-ron trong quá trình huấn luyện.

5. Normalization:

AlexNet sử dụng local response normalization (LRN) để tăng cường tính tổng quát và tăng cường khả năng tự kháng cự của mô hình.

6. Fully Connected Layers:

Sau các lớp tích chập và gộp, AlexNet kết thúc bằng ba lớp fully connected để thực hiện phân loại. Các lớp này có số nơ-ron lớn, tạo ra một không gian đặc trưng cấp cao để thực hiện phân loại.

7. Hàm Loss:

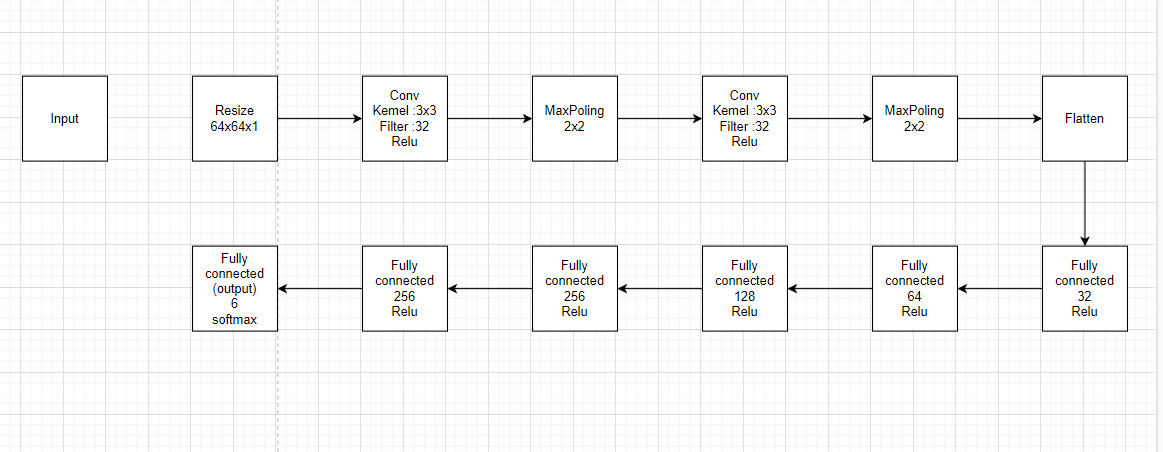
AlexNet sử dụng softmax activation và cross-entropy loss để thực hiện phân loại.Mặc dù AlexNet có kiến trúc mạnh mẽ, nhưng do kích thước của nó so với các mô hình hiện đại, nó có thể không hiệu quả như các mô hình mới được phát triển. Tuy nhiên, nó vẫn là một phần quan trọng trong lịch sử phát triển của thị giác máy tính và deep learning.

# **CHƯƠNG 3: THIẾT KẾ HỆ THỐNG**

## **3.1. Mô hình hệ thống .**

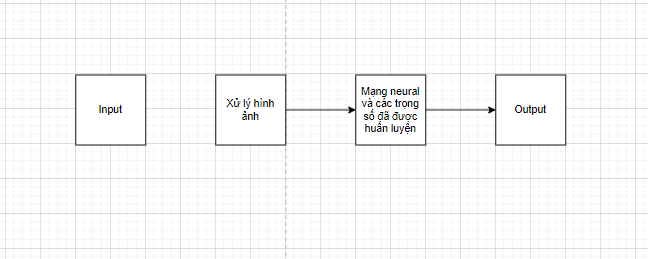
### 3.1.1 Mô hình mạng neural netword.

Ảnh input sẽ được resize và cho qua 2 lớp convolution, giữa 2 lớp sẽ là lớp max pooling trước khi qua max pooling 1 lần nữa và qua 1 lớp flatten , cuối cùng là lớp fully connected.



### 3.1.2 Mô hình sản phẩm.

Dựa vào mô hình mạng CNN đã huấn luyện , chúng em tạo thành 1 chương trình có chức năng nhận diện và phân loại biển số xe. Mô hình sẽ nhận ảnh từ các file ảnh đã được thu thập và tiến hành nhận diện thời gian thực . Sơ đồ khối của sản phẩm đc thể hiện như sau :



## **3.2 Thiết kế hệ thống bằng ngôn ngữ Matlab.**

3.2.1 Thu thập dữ liệu.

3.2.1.1 Dữ liệu huấn luyện nhận dạng, phân loại biển số xe.

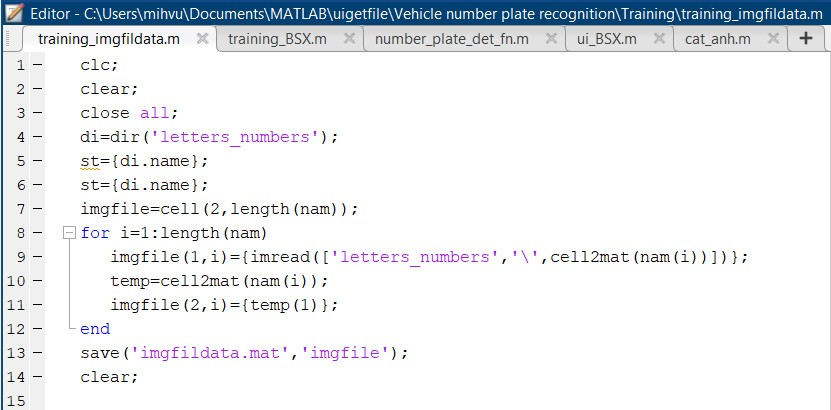


3.2.2.2 Dữ liệu huấn luyện nhận dạng kí tự biển số xe

### 

3.2.2 Huấn luyện hệ thống

3.2.2.1 Huấn luyện nhận dạng kí tự biển số xe

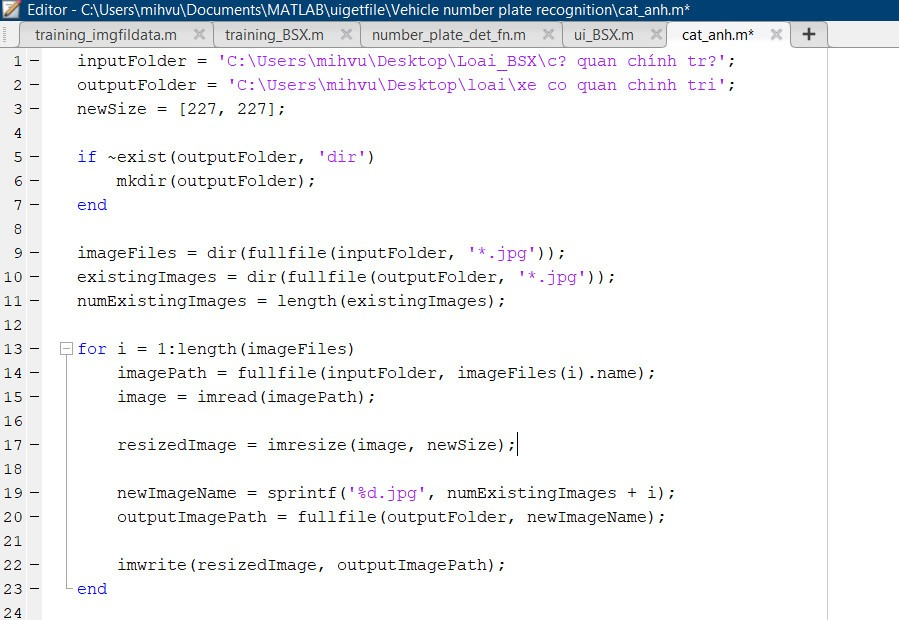


Đoạn mã MATLAB trên được thiết kế để xử lý hình ảnh trong thư mục 'letters\_numbers'. Trước hết, môi trường làm việc được làm sạch để đảm bảo một bắt đầu mới. Sau đó, thông tin về các file và thư mục trong 'letters\_numbers' được thu thập thông qua hàm dir, và các tên file được lưu vào một cell array có tên là nam.

Tiếp theo, một biến cell array có tên là imgfile được tạo để lưu trữ thông tin về hình ảnh và ký tự đầu tiên của tên file tương ứng. Một vòng lặp được sử dụng để đọc từng file hình ảnh trong thư mục 'letters\_numbers'. Hàm imread được sử dụng để đọc hình ảnh, và hàm fullfile được sử dụng để tạo đường dẫn đầy đủ đến file. Thông tin này sau đó được lưu vào imgfile.

Cuối cùng, thông tin được lưu vào một tệp có tên 'imgfiledata.mat' thông qua hàm save. Điều này giúp bảo quản dữ liệu để có thể sử dụng lại trong tương lai. Cuối cùng, một lệnh clear được sử dụng để làm sạch không gian làm việc và giải phóng bộ nhớ.

3.2.2.2. Trình cắt ảnh theo kích thước [227,227]



Trong đoạn mã MATLAB trên, một quy trình chuyển đổi và thay đổi kích thước ảnh được thực hiện từ thư mục nguồn đến thư mục đích. Đầu tiên, đường dẫn của thư mục đầu vào và đầu ra được đặt cho biến `inputFolder` và `outputFolder` tương ứng. Kích thước mới mong muốn cho ảnh được đặt là [227, 227] trong biến `newSize`.

Để đảm bảo rằng thư mục đầu ra tồn tại, một điều kiện kiểm tra được thực hiện. Nếu thư mục chưa tồn tại, nó sẽ được tạo mới thông qua hàm `mkdir`.

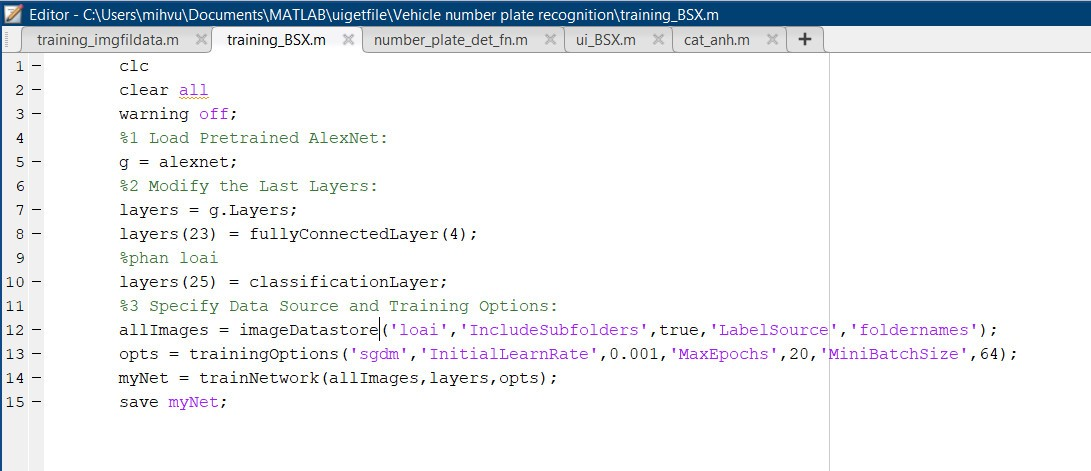
Danh sách tất cả các tệp tin hình ảnh trong thư mục đầu vào được lấy thông qua hàm `dir`, và danh sách các tệp tin hình ảnh đã tồn tại trong thư mục đầu ra cũng được thu thập. Số lượng ảnh đã tồn tại được đếm để tạo tên mới cho ảnh thêm vào.

Tiếp theo, vòng lặp `for` được sử dụng để duyệt qua từng tệp tin hình ảnh trong thư mục đầu vào. Mỗi ảnh được đọc thông qua hàm `imread`. Sau đó, kích thước của ảnh được thay đổi bằng hàm `imresize` để đảm bảo rằng nó có kích thước mới như đã định sẵn.

Tên mới cho ảnh được tạo thông qua hàm `sprintf` bằng cách sử dụng số thứ tự mới (tăng dần) và định dạng '.jpg'. Cuối cùng, ảnh mới được lưu vào thư mục đầu ra bằng hàm `imwrite`.

Tóm lại, đoạn mã thực hiện một quy trình hiệu quả để chuyển đổi và thay đổi kích thước ảnh từ thư mục nguồn sang thư mục đích, tránh việc ghi đè lên các ảnh đã tồn tại trong thư mục đích.

3.2.2.3 Huấn luyện nhận dạng, phân loại biển số xe



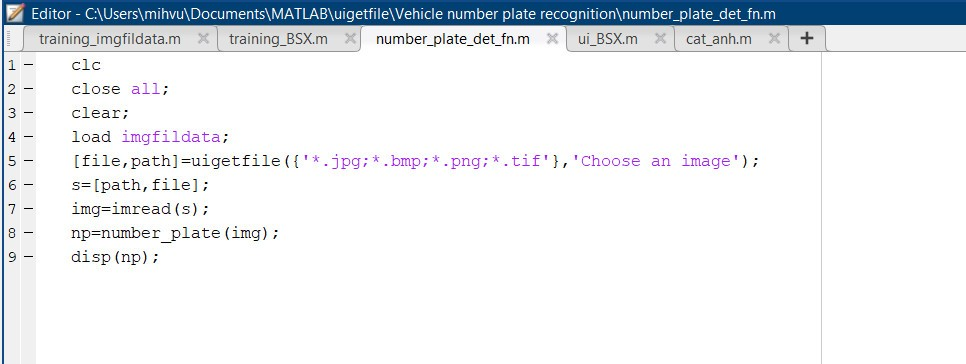
Đoạn mã MATLAB trên thực hiện quá trình huấn luyện một mô hình mạng nơ-ron sâu dựa trên kiến trúc AlexNet, đã được tiền huấn luyện trên một lượng lớn dữ liệu. Trước tiên, mọi thông báo và biến được xoá để làm sạch môi trường làm việc. Bước tiếp theo là tải kiến trúc AlexNet sử dụng hàm `alexnet`, và sau đó sửa đổi lớp cuối cùng của mô hình để phù hợp với số lượng lớp đầu ra cần phân loại.

Sử dụng `imageDatastore`, một nguồn dữ liệu hình ảnh được tạo từ thư mục 'loai', trong đó các hình ảnh được tổ chức vào các thư mục con tương ứng với nhãn của chúng. Tùy chọn huấn luyện được xác định bao gồm việc sử dụng thuật toán tối ưu SGD, tốc độ học ban đầu là 0.001, số epoch là 20, và kích thước batch là 64.

Sau đó, mô hình được huấn luyện thông qua hàm `trainNetwork` sử dụng dữ liệu từ nguồn đã xác định và cấu trúc mô hình mới. Kết quả của quá trình huấn luyện được lưu xuống tệp tin 'myNet.mat' để có thể tái sử dụng mô hình đã được huấn luyện mà không cần thực hiện lại quá trình này từ đầu.

Tổng cộng, đoạn mã thực hiện quá trình chuyển đổi kiến trúc AlexNet tiền huấn luyện thành một mô hình phân loại tùy chỉnh trên dữ liệu cụ thể và lưu kết quả để sử dụng trong các ứng dụng phân loại ảnh.

3.2.2.4 Kiểm tra mô hình hệ thống imgfildata



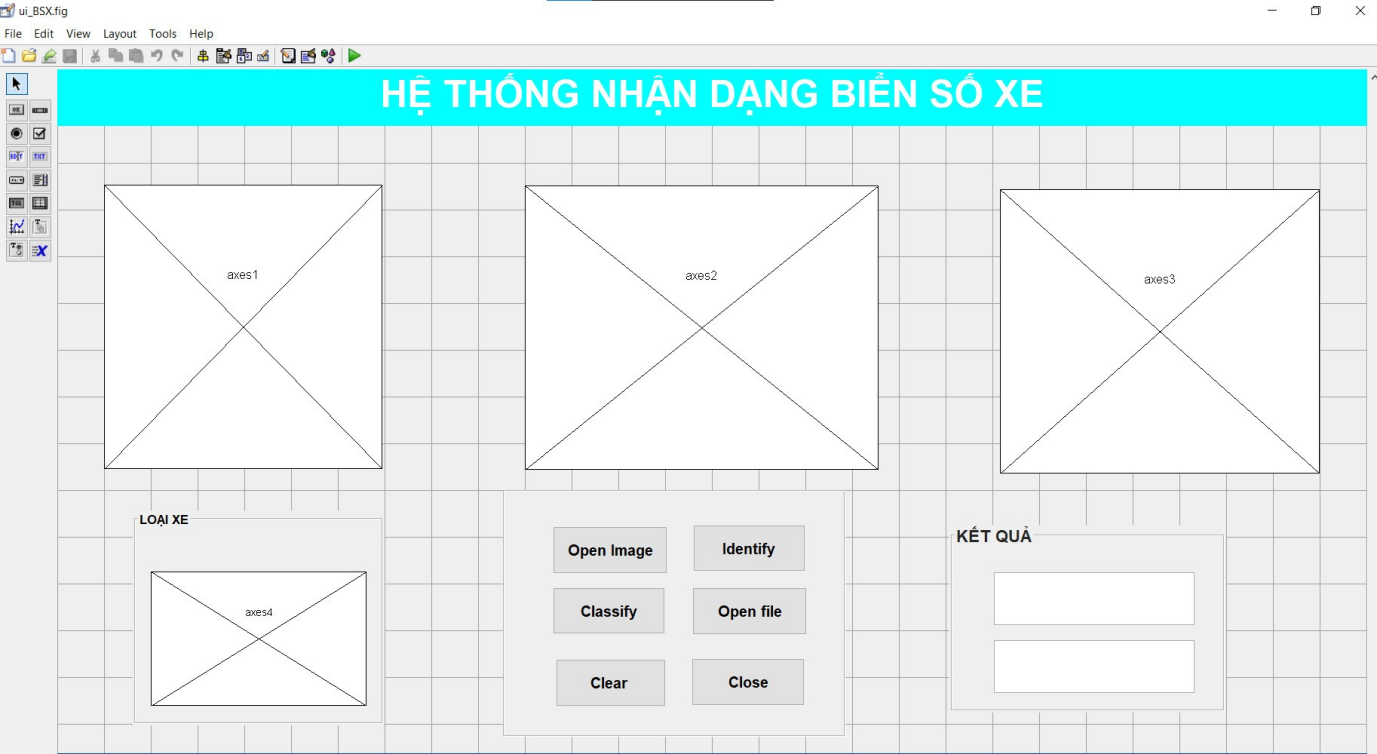
Đoạn mã MATLAB trên được thiết kế để thực hiện nhận diện biển số xe trên một hình ảnh đã cho. Trước hết, môi trường làm việc được làm sạch bằng cách xóa mọi thông báo từ cửa sổ lệnh (`clc`), đóng tất cả các cửa sổ hình ảnh đang mở (`close all`), và xóa tất cả các biến từ bộ nhớ (`clear`).

Dữ liệu cần thiết cho quá trình nhận diện biển số xe được tải từ tệp 'imgfildata.mat' sử dụng hàm `load imgfildata`. Sau đó, người dùng được yêu cầu chọn một tệp hình ảnh thông qua hộp thoại chọn tệp (`uigetfile`). Đường dẫn đầy đủ của tệp hình ảnh được lưu vào biến `s`.

Hình ảnh được đọc từ tệp đã chọn bằng cách sử dụng `imread` và lưu vào biến `img`. Tiếp theo, hàm `number\_plate` được gọi để nhận diện biển số xe trên hình ảnh, và kết quả được lưu vào biến `np`.

Cuối cùng, thông tin về biển số xe được hiển thị lên cửa sổ lệnh thông qua hàm `disp(np)`. Tổng cộng, đoạn mã này tạo ra một quy trình tương tác, cho phép người dùng chọn một hình ảnh và hiển thị thông tin về biển số xe từ ảnh đó lên cửa sổ lệnh MATLAB.

3.2.3 Thiết kế giao diện người dùng



Chương trình MATLAB trên đại diện cho một ứng dụng GUI được tạo bằng công cụ GUIDE của MATLAB, giúp người dùng tương tác với hình ảnh thông qua một giao diện đồ họa thân thiện. Cửa sổ GUI bao gồm nút "Open Image" để người dùng có thể chọn một tập tin hình ảnh từ máy tính. Sau khi chọn hình ảnh, người dùng có thể nhấn các nút chức năng như "Identify " để nhận diện biển số xe, "Classify" để phân loại biển số xe, "Open file" để hiển thị thông tin kết quả của biển số xe đã được nhận diện. Nút “Clear” để xóa trình hiện tại và nút "Đóng" có thể được sử dụng để đóng cửa sổ hiện tại.

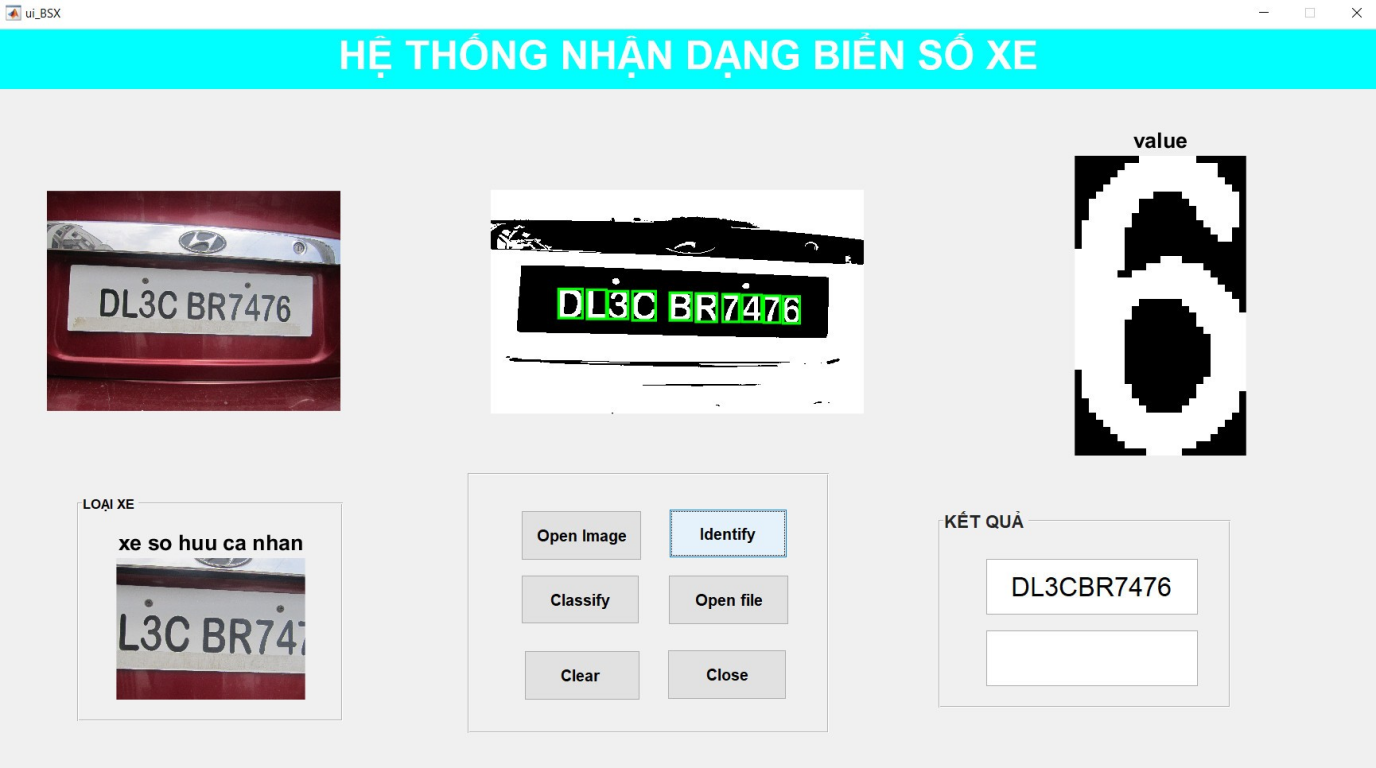
# **CHƯƠNG 4: KẾT QUẢ**

**4.1 Hệ thông nhận diện loại biển số xe**





**4.2 Hệ thống nhận diện đọc biển số xe**



**KẾT LUẬN**

Trí tuệ nhân tạo (AI) đã mang đến một cuộc cách mạng trong lĩnh vực công nghệ và ứng dụng. Mạng nơ ron nhân tạo, một phần quan trọng của AI, đã chứng minh khả năng ấn tượng của nó trong nhiều ứng dụng thực tế.

Mạng nơ ron nhân tạo đóng vai trò quan trọng trong việc xử lý thông tin và học từ dữ liệu. Với cấu trúc mô phỏng theo cách làm việc của hệ thống thần kinh của con người, mạng nơ ron nhân tạo có khả năng tự học và tự điều chỉnh thông qua việc huấn luyện trên tập dữ liệu lớn.

Ứng dụng của mạng nơ ron nhân tạo rất đa dạng. Trong lĩnh vực thị giác máy tính, chúng được sử dụng để nhận diện và phân loại hình ảnh. Trong xử lý ngôn ngữ tự nhiên, chúng giúp dịch thuật tự động, tạo ra bài viết tự động và phân loại văn bản. Trong lĩnh vực xe tự lái, mạng nơ ron nhân tạo hỗ trợ trong việc nhận dạng và dự đoán hành vi giao thông. Điều này chỉ là một số ví dụ, và ứng dụng của mạng nơ ron nhân tạo ngày càng mở rộng và đa dạng.Tuy nhiên, việc huấn luyện mạng nơ ron nhân tạo không phải là một nhiệm vụ dễ dàng. Nó đòi hỏi tập dữ liệu lớn, tính toán mạnh mẽ và kiến thức chuyên sâu về các thuật toán và kiến trúc mạng. Người ta cần chú ý đến việc xử lý dữ liệu, lựa chọn và điều chỉnh các siêu tham số, và đánh giá hiệu suất của mạng.

Mặc dù còn nhiều thách thức và hạn chế, mạng nơ ron nhân tạo đang phát triển mạnh mẽ và trở thành một công cụ quan trọng trong nhiều lĩnh vực. Sự tiến bộ trong phần cứng và phần mềm đã giúp nâng cao hiệu suất và khả năng của mạng nơ ron nhân tạo, và dự kiến rằng chúng sẽ tiếp tục đóng vai trò quan trọng trong tương lai.

Trí tuệ nhân tạo và mạng nơ ron nhân tạo đang mở ra những cánh cửa mới cho sự tiến bộ và sự phát triển trong xã hội và kinh tế. Sự ứng dụng của chúng trong các lĩnh vực như y tế, tự động hóa công nghiệp, tài chính, và nhiều lĩnh vực khác đang tạo ra những tiềm năng và triển vọng vô cùng hứa hẹn cho tương lai.

Tóm lại, mạng nơ ron nhân tạo là một công cụ hứa hẹn trong lĩnh vực trí tuệ nhân tạo. Với khả năng học và xử lý thông tin mạnh mẽ, chúng đang thúc đẩy sự phát triển và cách mạng hóa nhiều lĩnh vực khác

# **TÀI LIỆU THAM KHẢO**

1. "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

Link: http://www.deeplearningbook.org/

1. "Neural Networks and Deep Learning" by Michael Nielsen.

Link: http://neuralnetworksanddeeplearning.com/

1. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron.
2. "Pattern Recognition and Machine Learning" by Christopher M. Bishop.
3. PGS. TS Trương Ngọc Sơn, *Trí tuệ nhân tạo cơ sở và ứng dụng,* Đại học quốc gia TPHCM, 2020.