

BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC ĐIỆN LỰC



BÙI THANH TÙNG

**NHẬN DẠNG VÀ TÌM KIẾM MẪU VẬT
TRONG KHO ẢNH TANG VẬT BẰNG
CÔNG NGHỆ THỊ GIÁC MÁY TÍNH**

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

HÀ NỘI, 2024

BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC ĐIỆN LỰC

BÙI THANH TÙNG

**NHẬN DẠNG VÀ TÌM KIẾM MẪU VẬT
TRONG KHO ẢNH TANG VẬT BẰNG
CÔNG NGHỆ THỊ GIÁC MÁY TÍNH**

Chuyên ngành : Công nghệ thông tin
Mã số : 8480201

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

Người hướng dẫn khoa học: TS. NGÔ MINH DŨNG

LỜI CẢM ƠN

Đầu tiên, em xin bày tỏ lòng biết ơn sâu sắc tới Tiến sĩ Ngô Minh Dũng, người đã tận tâm hướng dẫn và hỗ trợ em xuyên suốt quá trình nghiên cứu và hoàn thành luận văn thạc sĩ. Bên cạnh đó, em xin chân thành cảm ơn các thầy cô tại Đại học Điện lực, những người đã nhiệt tình chỉ dẫn và hỗ trợ em trong suốt quá trình học tập. Em cũng muốn gửi lời biết ơn sâu sắc đến gia đình và bạn bè, những người đã tạo điều kiện và động viên em để hoàn thành tốt chương trình thạc sĩ.

Dù đã cố gắng hoàn thiện đề tài, em hiểu rằng vẫn còn những thiếu sót. Em rất mong nhận được sự chỉ dẫn và góp ý từ thầy cô và các bạn. Em xin chân thành cảm ơn!

Hà Nội, ngày 08 tháng 06 năm 2024.

Học viên cao học

Bùi Thanh Tùng

LỜI CAM ĐOAN

Em xin cam đoan rằng đây là nghiên cứu khoa học do chính em thực hiện dưới sự hướng dẫn của Tiến sĩ Ngô Minh Dũng. Các kết quả và nội dung nghiên cứu đều trung thực và chưa được công bố trước đây. Dữ liệu và số liệu thống kê sử dụng trong phân tích được lấy từ các nguồn tài liệu khác nhau, đã được trích dẫn đầy đủ trong phần tài liệu tham khảo.

Hà Nội, ngày 08 tháng 06 năm 2024.

Học viên cao học

Bùi Thanh Tùng

MỤC LỤC

MỤC LỤC	I
DANH MỤC TỪ VIẾT TẮT.....	III
DANH MỤC BẢNG BIỂU	IV
DANH MỤC HÌNH VẼ	V
MỞ ĐẦU	1
1. Lý do chọn đề tài.....	1
2. Mục đích nghiên cứu.....	1
3. Cấu trúc của luận văn.....	2
CHƯƠNG 1: TỔNG QUAN NHẬN DẠNG VÀ TÌM KIẾM.....	3
1.1 Lý thuyết nhận dạng.....	3
1.2 Tổng quan về xử lý ảnh.....	3
1.2.2 Một số khái niệm cơ bản trong xử lý ảnh.	4
1.3 Phương pháp tìm kiếm hình ảnh.	8
1.4 Bài toán nhận dạng và tìm kiếm từ đặc trưng mẫu vật.	9
1.5 Phương pháp tiếp cận để giải quyết bài toán.	11
1.5.1 Phương pháp Học sâu.	11
1.6 Kết luận chương 1.....	16
CHƯƠNG 2: MẠNG NƠ-RON TÍCH CHẬP.....	17
2.1 Kiến trúc mạng nơ-ron tích chập (CNN)	17
2.2 Học chuyển giao và tinh chỉnh mô hình huấn luyện.....	20
2.2.2 Lợi ích của việc học chuyển giao.	23
2.2.3 Hạn chế của việc học chuyển giao.....	23
2.3 Kiến trúc của VGG16.	24
2.4 Ứng dụng mạng VGG16 vào bài toán phân loại và tìm kiếm mẫu vật.....	25
2.5 Kết luận chương 2.....	26
CHƯƠNG 3: Hệ thống nhận dạng và tìm kiếm mẫu vật.	27
3.1 Kiến trúc hệ thống.....	27

II

3.2 Mô-đun quản lý và lưu trữ dữ liệu.	31
3.3 Mô đun nhận dạng và tìm kiếm mẫu vật.....	31
3.4 Mô-đun huấn luyện mô hình.	32
3.5 Môi trường huấn luyện.....	36
3.6 Cài đặt cấu trúc mạng huấn luyện VGG16.	38
3.7 Ví dụ về các đặc trưng được phân tích từ mạng VGG16.....	37
3.8 Các thành phần phía máy chủ.	38
3.9 Ứng dụng máy khách.	43
3.10 Thực nghiệm và kết quả.....	46
3.10.1 Dữ liệu và môi trường thực nghiệm.....	46
3.10.2 Kết quả và đánh giá.	47
3.11 Kết luận chương 3	49
KẾT LUẬN	51
I. TÓM TẮT VỀ CÁC KẾT QUẢ ĐẠT ĐƯỢC.	51
II. NHỮNG VẤN ĐỀ CHƯA ĐƯỢC GIẢI QUYẾT BỞI LUẬN VĂN.....	51
III. HƯỚNG PHÁT TRIỂN	51
MỤC LỤC THAM KHẢO	53

III

DANH MỤC TỪ VIẾT TẮT

STT	Từ viết tắt	Ý nghĩa
1	CSDL	Cơ sở dữ liệu
2	CNN	Convolutional Neural Network – Mạng nơ ron tích chập
3	ReLU	Rectified Linear Unit – Tính chỉnh đơn vị tuyến tính
4	GPU	Graphics Processing Unit – Bộ vi xử lý đồ họa

IV

DANH MỤC BẢNG BIỂU

<i>Bảng 3-1: Thống kê về tập dữ liệu thực nghiệm.</i>	<i>46</i>
<i>Bảng 3-2: Giá trị tham số tối ưu</i>	<i>48</i>
<i>Bảng 3-3: Kết quả phân loại</i>	<i>48</i>

DANH MỤC HÌNH VẼ

<i>Hình 1-1: Mô tả hệ thống nhận dạng.....</i>	<i>3</i>
<i>Hình 1-2: Các bước xử lý trong hệ thống xử lý ảnh.</i>	<i>4</i>
<i>Hình 1-3: Hình ảnh thu được và hình ảnh khi qua xử lý.</i>	<i>4</i>
<i>Hình 1-4: Các thách thức, trở ngại trong nhận dạng.</i>	<i>10</i>
<i>Hình 1-5: Các loại mẫu vật ma túy phổ biến.</i>	<i>10</i>
<i>Hình 1-6: Mô tả liên kết học sâu, học máy và trí tuệ nhân tạo.....</i>	<i>12</i>
<i>Hình 1-7: Cấu trúc về mạng lưới thần kinh.</i>	<i>12</i>
<i>Hình 2-1: Phép tính tích chập trên ảnh màu.....</i>	<i>17</i>
<i>Hình 2-2: Mô hình cơ bản của mạng nơ-ron tích chập.</i>	<i>18</i>
<i>Hình 2-3: Một ma trận điểm ảnh sử dụng ma trận tích chập</i>	<i>18</i>
<i>Hình 2-4: Ảnh khi tích chập khi thêm viên trắng.</i>	<i>19</i>
<i>Hình 2-5: Ví dụ về lớp lấy mẫu</i>	<i>20</i>
<i>Hình 2-6: Mô hình trí tuệ nhân tạo</i>	<i>21</i>
<i>Hình 2-7: Hiệu quả dùng khi transfer learning.</i>	<i>23</i>
<i>Hình 2-8: Kiến trúc của VGG16.</i>	<i>24</i>
<i>Hình 3-1: Mô tả hoạt động máy khách và máy chủ.</i>	<i>28</i>
<i>Hình 3-2: Sơ đồ hoạt động của hệ thống</i>	<i>30</i>
<i>Hình 3-3: Biểu đồ mô tả các thành phần mô-đun huấn luyện mô hình.</i>	<i>33</i>
<i>Hình 3-4: Các công cụ phát triển cho phát triển học sâu.....</i>	<i>37</i>
<i>Hình 3-5: Cách Keras triển khai định nghĩa lớp trong mạng CNN.....</i>	<i>39</i>
<i>Hình 3-6: Hình ảnh thể hiện các đặc trưng do mạng VGG16 phân tích.</i>	<i>38</i>
<i>Hình 3-7: Biểu đồ mô tả các thành phần phía máy chủ.....</i>	<i>40</i>
<i>Hình 3-8: Sơ đồ ca sử dụng phía máy khách.</i>	<i>44</i>
<i>Hình 3-9: Một số hình ảnh trong tập dữ liệu.</i>	<i>47</i>

VI

<i>Hình 3-10: Giao diện tìm kiếm và trả kết quả.....</i>	<i>49</i>
--	-----------

MỞ ĐẦU

1. Lý do chọn đề tài

Hiện nay, việc quản lý và tìm kiếm mẫu vật trong kho tàng vật ở nước ta chủ yếu dựa vào các phương pháp thủ công. Quy trình này không chỉ tiêu tốn nhiều công sức và thời gian mà còn dễ dẫn đến sai sót trong việc phân loại và truy xuất thông tin. Do đó, việc tự động hóa quy trình nhận dạng và tìm kiếm mẫu vật dựa trên các đặc trưng hình ảnh là một nhu cầu cấp thiết, đặc biệt trong bối cảnh số lượng mẫu vật ngày càng tăng và tính chất phức tạp ngày càng đa dạng.

Các nghiên cứu hiện có đã tập trung vào việc ứng dụng xử lý ảnh cơ bản để giải quyết vấn đề này. Tuy nhiên, các giải pháp hiện tại vẫn chưa đáp ứng được yêu cầu thực tiễn do khả năng nhận dạng và tìm kiếm chính xác còn hạn chế, đặc biệt khi đối mặt với những thách thức như:

- Sự đa dạng về chủng loại mẫu vật.
- Chi tiết phức tạp và cấu trúc hình dạng khó phân tích.
- Sự biến đổi về màu sắc và đặc điểm bề mặt.

Đồng thời, việc xây dựng một hệ thống tự động không chỉ đòi hỏi khả năng phân tích đặc trưng chính xác mà còn cần đảm bảo hiệu quả trong các tình huống dữ liệu lớn. Đây là một thách thức quan trọng trong nghiên cứu và ứng dụng thực tiễn.

Trong những năm gần đây, sự phát triển của công nghệ học sâu (deep learning) và các mô hình thị giác máy tính (computer vision) đã mở ra nhiều cơ hội mới để giải quyết các vấn đề phức tạp trong nhận dạng và tìm kiếm. Những tiến bộ này, kết hợp với khả năng xử lý mạnh mẽ của các thế hệ siêu máy tính hiện đại, đã chứng minh hiệu quả vượt trội trong việc phân tích hình ảnh và trích xuất đặc trưng chính xác.

Nhận thức được tầm quan trọng và tiềm năng ứng dụng thực tiễn của công nghệ này, em đã chọn đề tài **“Nhận dạng và tìm kiếm mẫu vật trong kho tàng vật bằng công nghệ thị giác máy tính”**. Mục tiêu của nghiên cứu là phát triển một hệ thống tự động, ứng dụng các mô hình học sâu tiên tiến để nhận dạng và tìm kiếm hiệu quả các mẫu vật phổ biến tại nước ta. Qua đó, đề tài không chỉ đóng góp vào việc cải thiện hiệu suất quản lý mà còn góp phần thúc đẩy ứng dụng công nghệ hiện đại trong các lĩnh vực thực tiễn.

2. Mục đích nghiên cứu

Vì thời gian nghiên cứu hạn chế, luận văn chủ yếu tập trung vào việc tìm hiểu những khó khăn và thách thức trong lĩnh vực nhận dạng và mô tả đặc trưng mẫu vật. Từ đó nghiên cứu và cài đặt mô hình học sâu được huấn luyện trên một số lượng mẫu nhỏ, nhằm ứng dụng vào việc nhận dạng và tìm kiếm từ đặc trưng hình ảnh trong kho ảnh mẫu vật.

Những nỗ lực này không chỉ giúp làm rõ các vấn đề lý thuyết và thực tiễn liên quan đến nhận dạng mẫu vật mà còn đặt nền móng cho việc phát triển các hệ thống tự động hóa hiệu quả hơn trong tương lai.

3. Cấu trúc của luận văn

Căn cứ vào mục tiêu cụ thể đã nêu ở phần trước, cấu trúc luận văn gồm ba chương với nội dung như sau:

Chương 1: Đề cập đến các khó khăn, thách thức trong bài toán nhận dạng và tìm kiếm mẫu vật từ các đặc trưng được nhận dạng và việc triển khai phương pháp học sâu để giải quyết các vấn đề.

Chương 2: Phân tích về kiến trúc mạng nơ-ron VGG16 một phần cốt lõi trong việc xác định các đặc trưng từ mẫu vật để có thể nhận dạng và tìm kiếm mẫu vật, chương này sẽ tìm hiểu và áp dụng để xử lý các thách thức trong việc xác định các đặc trưng mẫu vật.

Chương 3: Mô tả những bộ phận cốt lõi trong hệ thống nhận dạng và tìm kiếm mẫu vật từ các đặc trưng như cấu hình máy chủ, máy trạm, luồng nghiệp vụ, phương pháp xử lý dữ liệu thô, cách cài đặt để huấn luyện mô hình và ứng dụng mô hình học sâu được lựa chọn trong chương 2 vào hệ thống. Tiếp theo sẽ trình bày các kết quả thực nghiệm với mô hình huấn luyện từ tập dữ liệu được cung cấp.

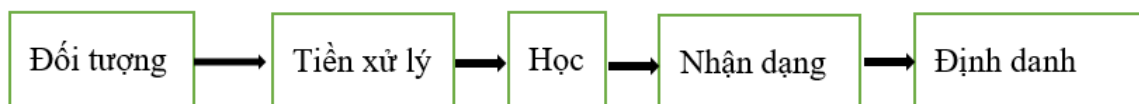
CHƯƠNG 1: TỔNG QUAN NHẬN DẠNG VÀ TÌM KIẾM

1.1 Lý thuyết nhận dạng.

Nhận dạng là thành phần của xử lý ảnh, đối tượng cần nhận dạng phân biệt theo một mô hình được huấn luyện từ các mẫu và gán nhãn chính xác khi được biết trước. Trong đó nhận dạng dựa trên mẫu trong quá trình huấn luyện mô hình được gọi là học có giám sát, ngược lại là học không giám sát.

Học có giám sát: Liên quan đến việc phân loại dựa trên tri thức biết trước. Phương pháp này dựa trên các mẫu tiêu chuẩn để so sánh với mẫu đang được phân tích. Ví dụ, trong viễn thám, người ta có thể nhằm mục đích phân biệt giữa một cánh đồng lúa, đất cằn cỗi hoặc một khu rừng có các đặc điểm đã biết. Khó khăn chính nằm ở việc phát triển một hệ thống có thể nhận dạng các vật thể trong hình ảnh bằng cách khớp chúng với các mẫu được xác định trước và gán chúng cùng một lớp cụ thể. Quá trình đối sánh này được xây dựng gọi là phân lớp hay cây quyết định.

Học không giám sát: Phương pháp này liên quan đến nhiệm vụ tự phân biệt các lớp học khác nhau và xác định các đặc điểm cơ bản cho từng lớp. Nó liên quan đến khó khăn của việc học mà không có hướng dẫn, sự khó khăn do không thể định rõ số lượng lớp và các đặc điểm khác biệt của những thực thể. Mục đích chính của phương pháp là phân loại các nhóm theo mọi cách có thể và chọn phương pháp tối ưu. Bắt đầu từ tập dữ liệu, các kỹ thuật xử lý ảnh được sử dụng để phân loại và tinh chỉnh dần dần để đạt được kết quả phân loại chuẩn xác nhất. Bất kể kỹ thuật nhận dạng hoặc mô hình nào được sử dụng, thì một hệ thống nhận dạng có thể tuân theo một mô hình cụ thể sau:



Hình 1-1: Mô tả hệ thống nhận dạng.

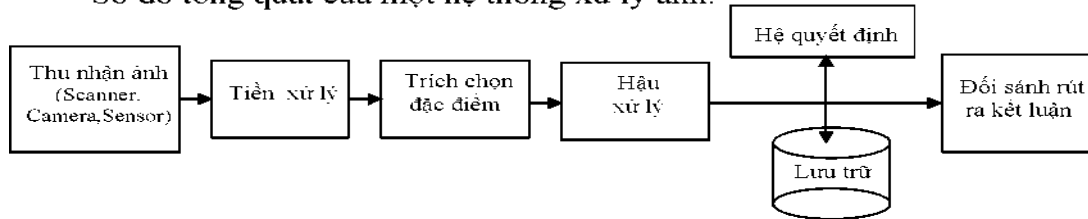
1.2 Tổng quan về xử lý ảnh.

Một yếu tố quan trọng trong tương tác giữa con người và máy tính chính là xử lý hình ảnh. Xử lý hình ảnh là một ngành thuộc khoa học máy tính và trí tuệ nhân tạo, chuyên về phân tích và xử lý ảnh số. Gần đây, xử lý hình ảnh đã phát triển vượt bậc nhờ sự tiến bộ nhanh chóng của phần cứng máy tính, dẫn đến nhiều ứng dụng thực tiễn trong đời sống.

Xử lý ảnh là quy trình liên quan đến việc xử lý ảnh nguyên bản để thu được kết quả cụ thể. Kết quả của quá trình là một hình ảnh được tối ưu hoặc đưa ra một quyết định. Một bức ảnh có thể được coi như một nhóm các điểm ảnh, trong đó mỗi điểm

ảnh thể hiện một mức độ sáng nhất định để xác định sự xuất hiện của một thực thể trong không gian. Trong lĩnh vực xử lý ảnh, mối quan hệ này thường được mô tả bởi hàm số với biến trong không gian hai chiều, các thực thể có thể được xác định dựa trên các đặc điểm cụ thể như tọa độ không gian hoặc các pixel riêng lẻ.

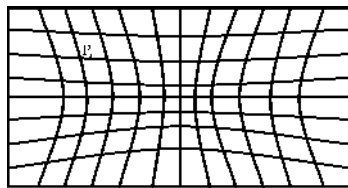
Sơ đồ tổng quát của một hệ thống xử lý ảnh:



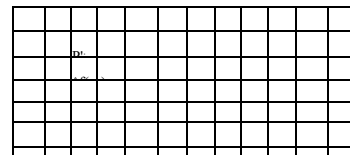
Hình 1-2: Các bước xử lý trong hệ thống xử lý ảnh.

1.2.1. Một số khái niệm cơ bản trong xử lý ảnh.

- * *Hình ảnh và điểm ảnh*: Hình ảnh là một nhóm các điểm ảnh, mỗi điểm ảnh mang thông tin về màu sắc và độ sáng tại một vị trí trong không gian của vật thể.
- * *Mức xám, màu*: Chi tiết về cường độ sáng và màu sắc tại một vị trí cụ thể.
- * *Điều chỉnh khi bị biến dạng*: Hình ảnh được tạo thành thường xuyên bị biến đổi do ảnh hưởng của các công cụ quang học và điện tử.



Ảnh thu nhận



Ảnh mong muốn

Hình 1-3: Hình ảnh thu được và hình ảnh khi qua xử lý.

Để điều chỉnh hình dạng hoặc biến dạng của một hình ảnh các phép chiếu thường được sử dụng. Các phép chiếu này thường dựa vào một nhóm các điểm được kiểm soát gọi là điểm điều chỉnh hoặc mốc cố định trên cả hình ảnh gốc và hình ảnh mục tiêu.

Giả sử (P_i, P'_i) $i = 1, n$ có n các tập điều khiển

Tìm hàm $f: P_i \rightarrow f(P_i)$ sao cho: $\sum_{i=1}^n ||f(P_i) - P'_i||^2 \rightarrow \min [1]$

* *Khử ảnh nhiễu*: có hai loại nhiễu.

- Nhiễu hệ thống: có đặc điểm là tuân theo các quy định hoặc trình cụ thể, để khử nhiễu có thể áp dụng phép biến đổi hình ảnh.
- Nhiễu ngẫu nhiên: đặc điểm là không tuân theo đặc trưng bởi các quy định hoặc trình tự cụ thể nào, thường được gây ra bởi các yếu tố không xác định như nhiễu điện từ, thay đổi trong hoàn cảnh môi trường, hoặc sai số trong khi thu thập dữ liệu. Trong trường, hoặc sai số trong quá trình thu thập dữ liệu. Trong những tình huống này, các phương pháp lọc truyền thống được

thực hiện để giải quyết nhiều.

* *Chỉnh mức xám*: Nhằm giải quyết sự không đồng đều trong quá trình hiệu chỉnh mức xám có hai hướng tiếp:

- Tăng số lượng các mức màu xám thông qua kỹ thuật nội suy, nhằm mục đích tạo thành mức xám trung bình để tăng cường độ mịn của hình ảnh.
- Giảm số lượng các mức màu xám thông qua việc gộp các mức xám gần đó lại với nhau.

Trong tình huống chỉ có hai mức màu xám, ảnh có thể được chuyển đổi thành định dạng đen trắng. Kỹ thuật này thường được dùng khi chuyển đổi ảnh màu sang đen trắng để in trên máy in thang độ xám nhằm tái tạo hình ảnh.

* *Lựa chọn đặc trưng*: đối với một đối tượng thì tùy vào nhu cầu cụ thể trong việc xử lý hình ảnh sẽ lựa chọn ra những đặc trưng nhất định của hình ảnh:

Đặc trưng không gian: liên quan đến vị trí và cấu trúc không gian của các điểm cụ thể bao gồm các khía cạnh như đặc điểm phân bố, sự phân tán mức xám, biên độ điểm cong trong hình ảnh.

Đặc trưng biến động: đặc trưng của hình ảnh tạo ra thông qua việc áp dụng lọc vùng bằng cách sử dụng mặt nạ đặc trưng bao gồm các cấu trúc hình dạng khác nhau như hình tam giác, cung tròn và hình chữ nhật.

Đặc trưng biên và đường biên: đóng một vai trò chính trong việc xác định biên vật thể, do đó tạo điều kiện thuận lợi cho việc trích xuất các đặc trưng bất biến cần thiết cho nhận dạng đối tượng. Những đặc điểm này có thể được xác định bằng toán tử đạo hàm bậc nhất, đạo hàm bậc hai, toán tử phân tích hướng, toán tử điểm giao cắt.

Chọn lọc đặc trưng chính xác là một yếu tố cần thiết cho việc nhận diện đối tượng chính xác trong khi vẫn duy trì hiệu quả xử lý nhanh và tối ưu hóa bộ nhớ.

1.2.2. Các kỹ thuật chính trong xử lý ảnh

Xử lý không gian ảnh (Spatial Domain Processing)

Đây là phương pháp thao tác trực tiếp trên các điểm ảnh (pixel) của hình ảnh trong không gian tọa độ.

- Các phép toán điểm ảnh (Point Operations): Thay đổi giá trị độ sáng của từng pixel, chẳng hạn như tăng/giảm độ sáng hoặc cải thiện độ tương phản. Ví dụ, ánh xạ giá trị pixel thông qua các hàm toán học như tuyến tính hoặc phi tuyến.
- Lọc ảnh không gian (Spatial Filtering): Sử dụng các mặt nạ (kernel) hoặc bộ lọc áp dụng lên từng điểm ảnh và vùng lân cận để:
 - + Làm mờ ảnh (Blurring): Giảm chi tiết, làm mịn ảnh.
 - + Làm sắc nét (Sharpening): Tăng cường biên và các chi tiết.

- + Lọc nhiễu (Noise Reduction): Loại bỏ các yếu tố nhiễu trong ảnh.

Xử lý miền tần số (Frequency Domain Processing)

Phương pháp này chuyển hình ảnh từ không gian pixel sang miền tần số thông qua các phép biến đổi toán học, chẳng hạn như biến đổi Fourier.

- Biến đổi Fourier (Fourier Transform): Chuyển đổi hình ảnh thành các thành phần tần số, giúp phân tích các thông tin như họa tiết và biên.
- Lọc miền tần số:
 - + Lọc thông thấp (Low-pass filter): Loại bỏ tần số cao để làm mờ ảnh.
 - + Lọc thông cao (High-pass filter): Loại bỏ tần số thấp để tăng cường chi tiết.
- Nén ảnh (Image Compression): Giảm dung lượng ảnh bằng cách loại bỏ các thông tin không cần thiết dựa trên tần số, ví dụ như phương pháp JPEG.

Phân đoạn ảnh (Image Segmentation)

Phân đoạn ảnh là quá trình chia ảnh thành các vùng hoặc đối tượng cụ thể dựa trên các đặc tính như màu sắc, cường độ sáng, hoặc kết cấu.

- Ngưỡng hóa (Thresholding): Áp dụng ngưỡng để phân chia ảnh thành các vùng foreground (đối tượng quan tâm) và background.
- Phân đoạn dựa trên biên (Edge-based Segmentation): Xác định biên của các đối tượng dựa trên sự thay đổi cường độ sáng.
- Phân đoạn theo vùng (Region-based Segmentation): Chia ảnh thành các vùng đồng nhất dựa trên đặc điểm như độ sáng hoặc màu sắc.

Phát hiện và trích xuất đặc trưng (Feature Detection and Extraction)

Nhóm kỹ thuật này tập trung vào việc xác định và trích xuất các đặc điểm quan trọng từ ảnh, phục vụ cho các ứng dụng như nhận diện, theo dõi đối tượng, hoặc phân loại.

- Phát hiện cạnh (Edge Detection): Sử dụng các thuật toán như Sobel, Canny để xác định biên của các đối tượng.
- Phát hiện góc (Corner Detection): Tìm các điểm góc trong ảnh, chẳng hạn như thuật toán Harris Corner.
- Trích xuất đặc trưng (Feature Extraction): Xác định các đặc trưng độc lập, chẳng hạn như các điểm đặc trưng SIFT (Scale-Invariant Feature Transform) hoặc SURF (Speeded-Up Robust Features).

Khôi phục ảnh (Image Restoration)

Tập trung vào việc khôi phục ảnh về trạng thái ban đầu hoặc chất lượng tốt hơn từ dữ liệu bị nhiễu hoặc không hoàn chỉnh.

- Lọc nhiễu (Denoising): Sử dụng các kỹ thuật lọc không gian hoặc tần số để loại bỏ nhiễu mà không làm mất thông tin quan trọng.

- Khử mờ (Deblurring): Xử lý để phục hồi các ảnh bị mờ do chuyển động hoặc các vấn đề trong quá trình chụp ảnh.

1.2.3. Ứng dụng thực tế của xử lý ảnh

Xử lý ảnh đóng vai trò quan trọng trong lĩnh vực y tế, hỗ trợ bác sĩ chẩn đoán và điều trị bệnh hiệu quả hơn. Các hệ thống phân tích hình ảnh y tế như X-quang, MRI và CT scan giúp phát hiện các bệnh lý như ung thư, bệnh tim mạch, và tổn thương não một cách chính xác. Ngoài ra, xử lý ảnh còn hỗ trợ phẫu thuật bằng cách cung cấp hình ảnh 3D chi tiết, giúp bác sĩ thực hiện các ca phẫu thuật phức tạp với độ chính xác cao. Công nghệ này cũng được áp dụng trong phân tích tế bào và mô học, mang lại nhiều tiến bộ trong nghiên cứu y học.

Ứng dụng trong an ninh và giám sát

Trong lĩnh vực an ninh, xử lý ảnh được sử dụng để xây dựng các hệ thống nhận diện khuôn mặt, kiểm tra danh tính tại sân bay, ngân hàng, hoặc các cơ sở quan trọng. Hệ thống camera thông minh sử dụng công nghệ này để phát hiện hành vi đáng ngờ, giám sát an ninh tại các khu vực công cộng. Trong giao thông, xử lý ảnh được dùng để nhận diện biển số xe, phát hiện vi phạm giao thông, và tối ưu hóa quản lý giao thông đô thị.

Ứng dụng trong công nghiệp

Trong công nghiệp, xử lý ảnh là công cụ mạnh mẽ để kiểm tra chất lượng sản phẩm trên dây chuyền sản xuất. Các hệ thống tự động hóa sử dụng công nghệ này để phát hiện lỗi trên sản phẩm, chẳng hạn như vết nứt hoặc lỗi màu sắc. Ngoài ra, robot công nghiệp sử dụng thị giác máy tính để nhận diện và xử lý các vật thể, hỗ trợ quá trình lắp ráp, đóng gói và kiểm soát quy trình sản xuất hiệu quả hơn.

Ứng dụng trong giao thông thông minh

Hệ thống giao thông thông minh sử dụng xử lý ảnh để giám sát và quản lý lưu lượng giao thông. Các camera giao thông phân tích hình ảnh để phát hiện hành vi vi phạm, nhận diện người đi bộ, và tối ưu hóa tín hiệu đèn giao thông. Đặc biệt, trong lĩnh vực xe tự hành, xử lý ảnh giúp xe nhận diện làn đường, tín hiệu giao thông, và các vật cản để đảm bảo di chuyển an toàn.

Ứng dụng trong thương mại và giải trí

Trong thương mại điện tử, xử lý ảnh hỗ trợ phân tích hành vi khách hàng, tối ưu hóa trải nghiệm mua sắm và xây dựng các hệ thống tìm kiếm sản phẩm dựa trên hình ảnh. Trong lĩnh vực giải trí, công nghệ này được sử dụng để chỉnh sửa ảnh, thêm hiệu ứng đặc biệt, và tạo ra các mô phỏng 3D trong phim hoặc trò chơi điện tử. Các ứng dụng thực tế tăng cường (AR) cũng dựa trên xử lý ảnh để tạo ra trải nghiệm tương tác hấp dẫn cho người dùng.

Ứng dụng trong địa lý và viễn thám

Hình ảnh vệ tinh kết hợp với xử lý ảnh giúp phân tích môi trường, giám sát tài nguyên thiên nhiên, và dự báo thời tiết. Trong nông nghiệp, công nghệ này hỗ trợ giám sát cây trồng, phân tích đất đai, và phát hiện sâu bệnh để tối ưu hóa năng suất.

Ngoài ra, xử lý ảnh còn được sử dụng trong việc giám sát thiên tai, theo dõi lũ lụt, sạt lở đất, và cháy rừng, giúp giảm thiểu tác động của các thảm họa thiên nhiên.

Ứng dụng trong giáo dục và nghiên cứu

Xử lý ảnh hỗ trợ số hóa tài liệu, giúp bảo tồn các tư liệu quan trọng trong thư viện số. Công nghệ này cũng hỗ trợ phân tích và trực quan hóa dữ liệu, giúp các nhà nghiên cứu xử lý thông tin nhanh hơn. Trong giáo dục, xử lý ảnh giúp tạo ra các bài giảng trực quan, mô phỏng 3D, và các công cụ học tập sinh động, tăng cường hiệu quả giảng dạy và học tập.

Ứng dụng trong trí tuệ nhân tạo

Xử lý ảnh là nền tảng của các ứng dụng trí tuệ nhân tạo và học sâu, đặc biệt là trong thị giác máy tính. Các hệ thống nhận diện khuôn mặt, phân loại đối tượng, và đọc văn bản từ hình ảnh (OCR) được phát triển dựa trên công nghệ này. Ngoài ra, xử lý ảnh còn hỗ trợ các hệ thống tương tác người-máy như phân tích biểu cảm khuôn mặt và chuyển động cơ thể, mở ra nhiều cơ hội mới trong lĩnh vực thực tế ảo và thực tế tăng cường.

1.2.4. Thách thức và triển vọng

Mặc dù xử lý ảnh đã đạt được nhiều thành tựu đáng kể, lĩnh vực này vẫn đối mặt với các thách thức như:

- Xử lý ảnh chất lượng thấp hoặc bị nhiễu.
- Tăng cường hiệu suất thuật toán trong các ứng dụng thời gian thực.
- Phát triển các phương pháp đảm bảo tính minh bạch và an toàn khi sử dụng trí tuệ nhân tạo trong xử lý ảnh.

Với sự phát triển không ngừng của công nghệ, xử lý ảnh sẽ tiếp tục đóng vai trò then chốt trong việc nâng cao hiệu quả của các hệ thống thông minh và mở ra những cơ hội mới trong nhiều lĩnh vực khác nhau.

1.3 Phương pháp tìm kiếm hình ảnh.

Ban đầu, các phương pháp tìm kiếm hình ảnh thường tìm kiếm sự tương đồng theo nhãn. Hai bức ảnh được đánh giá là tương đồng nhau nếu như chúng có cùng nhãn, ví dụ như bức ảnh về hoa, bức ảnh về con người. Tuy nhiên, mức độ tương đồng này vẫn chưa đủ đối với các ứng dụng cấp cao hơn. Các bài toán sau này cần tìm kiếm sự tương đồng ở mức độ lớn, chi tiết hơn, kể cả khi hai bức ảnh có cùng nhãn, nhằm nhận diện sự khác nhau dù nhỏ nhất.

Một phương pháp xây dựng sự giống nhau về hình ảnh đó là sử dụng các thuật toán trích xuất đặc trưng (feature) của ảnh như bộ lọc SIFT (Scale-Invariant Feature Transform) [2], HOG (Histogram of Oriented Gradient) [3], rồi sử dụng các đặc trưng đó để tính toán sự tương đồng giữa hai bức ảnh. Phương pháp này đã được sử dụng trong các nghiên cứu [4] [5]. Tuy nhiên, những mô hình này bị giới hạn bởi khối lượng tính toán quá lớn. Trong những năm gần đây, mô hình mạng nơ ron tích chập

(CNN) được sử dụng phổ biến cho việc nhận dạng và phân loại hình ảnh đã đem lại một cách tiếp cận mới cho việc tính toán độ tương đồng hình ảnh. Các mô hình học sâu, đặc biệt là CNN có khả năng tìm các đặc trưng từ bậc thấp bậc cao với độ chính xác ổn định, điều này giúp rất nhiều trong việc trích xuất các đặc trưng chính của bức ảnh để phục vụ quá trình tìm kiếm thông qua độ tương đồng giữa đặc trưng hình ảnh. Các lớp CNN kế tiếp nhau sẽ biểu thị hình ảnh theo các mức độ trừu tượng khác nhau. Lớp cuối cùng làm một véc tơ đại diện cho ảnh, có thể dùng để làm đặc trưng tính toán độ tương đồng hình ảnh. Để so sánh độ tương đồng hình ảnh từ các véc tơ đặc trưng thu được, có thể sử dụng các phép đo khoảng cách. Một số phép đo thông dụng bao gồm L1-norm (khoảng cách Manhattan), L2-norm (khoảng cách Euclid), hay độ tương đồng Cosine (cosine similarity). Cần chú ý rằng, khoảng cách và độ tương đồng tỉ lệ nghịch với nhau, khoảng cách càng lớn thì độ tương đồng càng nhỏ và ngược lại. Để làm tăng độ chính xác của việc tìm kiếm ảnh thông qua tính toán độ tương đồng như trên, có thể áp dụng bước phân loại ảnh trước, sau đó tiến hành tính toán độ tương đồng trên các ảnh có cùng nhãn (loại) với ảnh được tìm kiếm. Các ảnh không cùng nhãn cũng có thể được tính toán độ tương đồng, nhưng nằm ở phần sau của kết quả tìm kiếm.

1.4 Bài toán nhận dạng và tìm kiếm từ đặc trưng mẫu vật.

Nhận dạng vật thể trong ảnh là một nhiệm vụ cơ bản trong thị giác máy tính, đóng vai trò quan trọng trong nhiều ứng dụng công nghệ hiện đại. Từ đó bài toán phát triển thành các bài toán khác như phân loại, xác định vị trí và phân đoạn vật thể đều phụ thuộc vào bài toán này. Tuy nhiên, mặc dù nó đã tồn tại hàng thế kỷ nhưng để máy tính có thể phân tích các đặc trưng trong hình ảnh vẫn gặp phải nhiều trở ngại và vấn đề. Trong đó, những trở ngại và khó khăn lớn [6] phải kể đến:

- Khó khăn trong việc nhận dạng vật thể trong ảnh là sự biến thiên trong điểm nhìn. Đối tượng có thể xuất hiện với các biến thể từ các điểm nhìn khác nhau trong bức ảnh, dẫn đến các biến thể hình ảnh không giống nhau. Điều này làm cho việc huấn luyện một mô hình máy học để hiểu và nhận dạng sẽ đối mặt với sự khó khăn trong việc xử lý tất cả các biến thể.

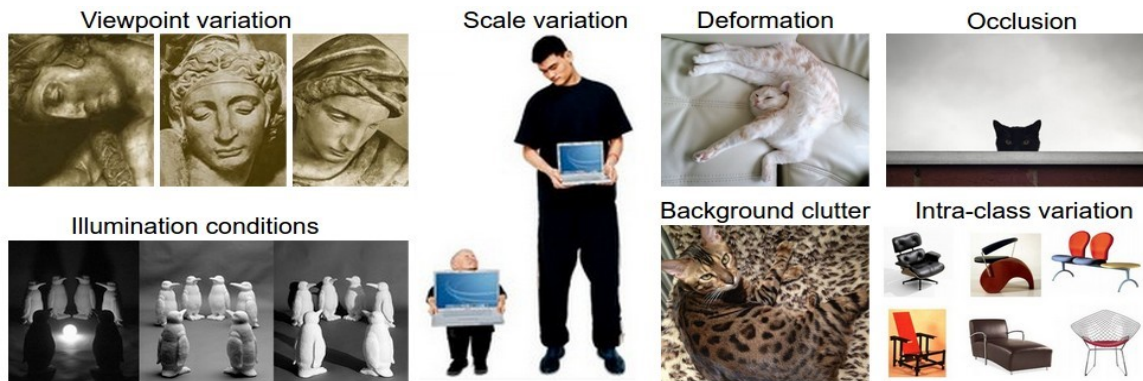
- Nhận dạng vật thể trong ảnh gặp thách thức trong việc phân tích đối tượng có kích thước đa dạng. Trong thực tế, các vật thể có thể xuất hiện trong các bức ảnh ở nhiều kích thước đa dạng điều này làm cho việc nhận dạng trở nên khó khăn.

- Một trở ngại lớn đối với việc nhận dạng vật thể là khả năng nhận diện đối tượng bị ảnh hưởng bởi sự thay đổi ánh sáng. Sự thay đổi ánh sáng có thể tác động mạnh đến chi tiết hình ảnh, đặc biệt là ở các điểm ảnh với cường độ sáng yếu.

- Một khó khăn lớn để nhận dạng vật thể là khi một phần của nó được giấu sau các đối tượng khác trong ảnh. Đôi khi, vật thể không hiển thị trong toàn bộ hình dạng hoặc không hiển thị hoàn chỉnh do các đối tượng khác che lấp chúng trong bức ảnh. Điều này làm giảm chi tiết và đặc trưng của vật thể dẫn đến tăng độ phức tạp cho việc nhận dạng.

- Khó khăn tiếp nữa là nền không đồng nhất và lộn xộn có thể làm giảm hiệu quả nhận dạng đối tượng. Trong nhiều trường hợp, các đối tượng cần nhận dạng thường khó phân biệt vì chúng bị lẫn vào nền hình ảnh, sự đồng nhất về màu sắc và họa tiết giữa nền và vật thể làm cho việc nhận dạng trở nên khó khăn hơn, ngay cả với những người có thể nhìn thấy bằng mắt con người.

- Khó khăn lớn trong việc nhận dạng vật thể là sự đa dạng về loại vật thể làm cho việc nhận dạng trở nên khó khăn hơn do sự biến đổi về hình dạng, màu sắc, và kết cấu. Dẫn đến việc xây dựng các mô hình nhận dạng đối tượng trở nên phức tạp hơn, đặc biệt là trong việc nhận dạng các biến thể của cùng một loại vật thể, chẳng hạn như cần nhận diện chính xác các loại ghế và phân tách chúng khỏi các vật thể khác như bàn.



Hình 1-4: Các thách thức, trở ngại trong nhận dạng.

Bài toán nhận dạng và tìm kiếm mẫu vật dựa trên các đặc trưng có thể được coi là một ứng dụng của bài toán phân loại và tìm kiếm. Với một hình ảnh đầu vào của một mẫu vật nó sẽ được phân loại thành loại mẫu vật gì. Sau đó các hình ảnh mẫu vật cùng loại giống nó nhất sẽ được tính toán và trả về kết quả tìm kiếm.



Hình 1-5: Các loại mẫu vật ma túy phổ biến.

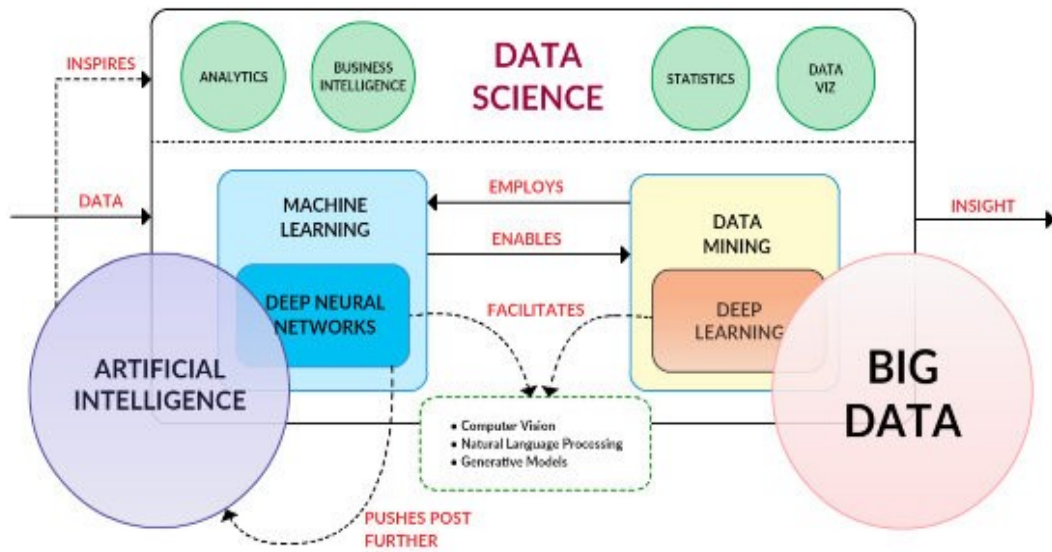
1.5 Phương pháp tiếp cận để giải quyết bài toán.

Hiện nay để cải tiến thuật toán nhận dạng đã có rất nhiều nghiên cứu và bài viết khoa học được công bố, một trong những phương pháp đầu tiên đó là phương pháp xử lý ảnh. Phương pháp này chú trọng vào việc nghiên cứu các thuật toán để trích chọn đặc trưng từ bức ảnh đầu vào ví dụ các đặc trưng về sắc thái màu, kích thước, hình thức và kết cấu từ ảnh đầu vào để nhận dạng mẫu vật [7]. Các phương pháp này không khả thi đối với nhiều loại mẫu vật vì chúng chỉ phù hợp với một tập hợp ảnh đầu vào cụ thể trong khi sự biến thiên về hình thức, kết cấu, kích thước, màu sắc và các yếu tố khác của mẫu vật là quá phức tạp do đó kết quả trên phương pháp này không cao.

Các thách thức chính trong việc xác định mẫu vật cụ thể là ma túy là sự đa dạng và sự tương đồng về hình thái của các mẫu vật, có thể trông giống hệt nhau khi nhìn bằng mắt thường. Nói chung, mẫu vật bề ngoài có thể được xác định bởi hoa văn, màu sắc và hình dạng. Hình dạng mẫu vật có thể được chia thành nhiều loại: tròn, viên nang, hình bầu dục [8]. Các đặc điểm khác biệt cũng có thể được quan sát đối với màu sắc mẫu vật và hoa văn. Trong nhận dạng mẫu vật thông thường, người dùng cố gắng xác định một loại mẫu vật không xác định bằng cách nhập thủ công các tính năng của mẫu vật trên trang trên hệ thống, việc này tốn nhiều thời gian quá trình. Trong các nghiên cứu trước đây, một số vấn đề đã không được giải quyết hiệu quả trong việc nhận dạng mẫu vật [8], chẳng hạn như ngẫu nhiên vị trí của mẫu vật và sự hiện diện của nhiều loại mẫu vật trong một hình ảnh. Góc quay của mẫu vật cũng khó xác định và bình thường hóa cho từng loại mẫu vật. Hơn nữa, rất khó để thực hiện phân loại tự động các mẫu vật bằng hình ảnh truyền thống. Bây giờ, chúng ta sẽ phân tích chi tiết về kỹ thuật trích chọn đặc trưng và phương pháp hiện tại phổ biến nhất để giải quyết bài toán nhận dạng và tìm kiếm từ đặc trưng mẫu vật trong ảnh là phương pháp Học sâu.

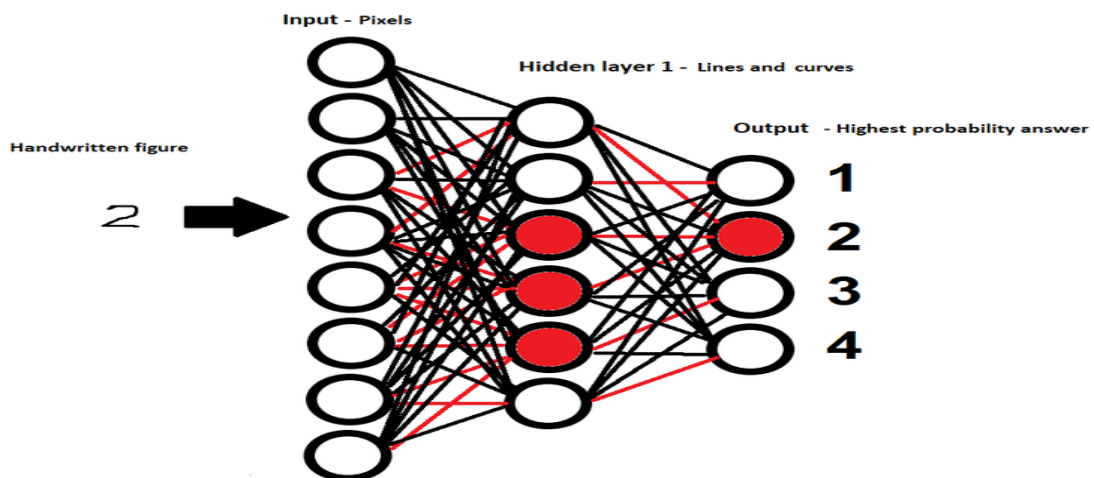
1.5.1 Phương pháp Học sâu.

Một lĩnh vực đặc trưng của học máy là học sâu, trong phương pháp này các mô hình học máy được thiết kế để mô phỏng cách não người hoạt động thông qua các mạng nơ-ron nhân tạo sâu. Nó trở nên phổ biến trong thời gian gần đây do việc các nhà khoa học có khả năng áp dụng các công nghệ máy tính tiên tiến với khả năng tính toán hiệu quả trong việc xử lý khối lượng lớn dữ liệu âm thanh, văn bản, hình ảnh. Hình 1.7 [9] mô tả học sâu là một nhánh của học máy một lĩnh vực con của trí tuệ nhân tạo:



Hình 1-6: Mô tả liên kết học sâu, học máy và trí tuệ nhân tạo.

Vì cách thức hoạt động của chúng nên huấn luyện mạng sử dụng phương pháp học sâu được biết đến như là mạng nơ-ron sâu (Deep Neural Network). Mỗi nơ-ron trong mạng nơ-ron đóng vai là một hàm toán học, đầu vào tập hợp dữ liệu và biến đổi chúng thành định dạng có thể dễ dàng điều chỉnh và xử lý và sau đó cung cấp đầu ra cho kết quả. Một mạng lưới thần kinh bao gồm các nơron được sắp xếp theo lớp gồm ba thành phần chính là lớp đầu vào, lớp ẩn và lớp đầu ra. Trong hình 1.7, nhiều lớp ẩn có thể xuất hiện trong một mạng nơ-ron, lớp ẩn trong mạng nơ-ron nhận các tín hiệu đầu vào từ lớp nơ-ron trước đó và xử lý chúng để truyền đạt cho các lớp tiếp theo.



Hình 1-7: Cấu trúc về mạng lưới thần kinh.

Trong mạng học sâu dành cho nhận dạng ảnh thường là các lớp tích chập (convolutional layers). Sau đó, các bộ lọc (filters) được sử dụng để xác định các đặc trưng quan trọng như đốm màu, cạnh hoặc góc trong ảnh đầu vào. Các lớp tiếp theo sẽ áp dụng thông tin này để nhận dạng các thành phần đặc trưng của vật thể trong ảnh từ các đường và cạnh của chúng. Cuối cùng, lớp cao nhất của mạng huấn luyện sẽ xác định và phân loại vật thể trong ảnh.

Việc phân tích để hiểu xử lý dữ liệu phức tạp bằng cách phân chia thông tin thành nhiều lớp đơn giản phương pháp này được gọi là "Học sâu" (Deep Learning). Cách tổng hợp và biểu diễn thông tin qua nhiều tầng lớp, mạng học sâu có khả năng nhận dạng và học các đặc trưng phức tạp trong dữ liệu, từ các đặc điểm cơ bản như cạnh, góc, cho đến các đặc trưng phức tạp hơn như thành phần của đối tượng, mối liên quan giữa các đối tượng, và các quan hệ phức tạp khác.

Mặc dù học sâu có nhiều ưu điểm đặc biệt đối với các bài toán có sự phức tạp, nhưng nó vẫn đối mặt với một số hạn chế đáng kể. Học sâu không phải là giải pháp cho tất cả các loại vấn đề vì nó có nhiều hạn chế, một hạn chế lớn nhất là để huấn luyện mô hình học sâu một cách hiệu quả, cần phải cung cấp một khối lượng dữ liệu đầu vào rất lớn. Học sâu cũng yêu cầu sức mạnh tính toán lớn, đặc biệt là trong quá trình huấn luyện các mô hình lớn với hàng triệu hoặc thậm chí hàng tỷ tham số. Điều này đòi hỏi sử dụng các máy tính cỡ lớn hoặc hệ thống tính toán phân tán giúp có thể phân phối công việc để huấn luyện các mô hình này. Đây là một vấn đề của học tập sâu liên quan đến chi phí huấn luyện. Huấn luyện một mô hình có thể tốn thời gian, tài nguyên, từ việc chuẩn bị dữ liệu đầu vào, xác định kiến trúc mô hình, đến việc điều chỉnh tham số và đánh giá hiệu suất điều này đòi hỏi sự đầu tư nhiều vì chúng tạo ra chi phí lớn. Tuy nhiên, phương pháp học tập chuyển giao (Transfer Learning) đã giúp giảm bớt đi một số hạn chế của học sâu, đặc biệt là liên quan đến số lượng dữ liệu huấn luyện, những vấn đề này sẽ được phân tích chi tiết trong các chương sau. Ngoài những hạn chế liên quan đến kích thước dữ liệu đầu vào, học sâu cũng không đủ thông minh để nhận dạng và hiểu các logic phức tạp như con người, và không đảm bảo rằng trí tuệ nhân tạo có thể đưa ra các kết luận logic chính xác.

Mục đích của luận văn là phân tích và ứng dụng một mô hình để xử lý vấn đề được đề cập trong phần mở đầu. Với khả năng mạnh mẽ vượt trội của học sâu trong việc xác định các đặc trưng phù hợp đối với bài toán nhận dạng và sử dụng các đặc trưng được nhận dạng để tìm kiếm vật thể là cụ thể là trong trường hợp này là các loại mẫu vật ma túy nên em đã chọn sử dụng giải pháp sử dụng phương pháp học sâu. Để giải quyết bài toán, luận văn đã triển khai huấn luyện sử dụng bộ dữ liệu đầu vào để đánh giá và xác minh nhận định này. Kết quả chi tiết sẽ được nêu chi tiết trong Chương 3 – Kết quả từ đánh giá thực nghiệm.

1.5.2 . Phương pháp dựa trên trích xuất đặc trưng truyền thống

Phương pháp này sử dụng các kỹ thuật xử lý ảnh và thị giác máy tính truyền thống để phân tích và trích xuất các đặc trưng quan trọng từ hình ảnh, sau đó sử dụng các thuật toán so khớp và phân loại để nhận dạng hoặc tìm kiếm mẫu vật. Các bước chính bao gồm trích xuất đặc trưng, đối sánh đặc trưng, và phân loại mẫu vật.

1. Trích xuất đặc trưng

Trích xuất đặc trưng là bước quan trọng nhằm biểu diễn hình ảnh dưới dạng các thông tin đặc trưng có thể phân biệt được. Trong phương pháp truyền thống, đặc trưng có thể được phân thành hai loại chính:

a. Đặc trưng cục bộ (Local Features)

Đặc trưng cục bộ biểu diễn các điểm nổi bật (keypoints) hoặc các vùng nhỏ trong hình ảnh. Những đặc trưng này có khả năng ổn định trước các biến đổi về góc nhìn, ánh sáng, và tỉ lệ.

– SIFT (Scale-Invariant Feature Transform):

SIFT là một thuật toán phổ biến để phát hiện và mô tả đặc trưng cục bộ. Nó tạo ra các điểm đặc trưng không thay đổi khi hình ảnh bị biến đổi kích thước hoặc góc nhìn. Quy trình của SIFT gồm:

- + Phát hiện các điểm đặc trưng (keypoints) bằng cách sử dụng Gaussian Difference.
- + Mô tả keypoints bằng các vector đặc trưng dựa trên gradient định hướng của từng vùng lân cận.

– SURF (Speeded-Up Robust Features):

Là phiên bản nhanh hơn của SIFT, sử dụng tích phân Haar wavelet để phát hiện và mô tả đặc trưng. SURF thích hợp cho các ứng dụng thời gian thực nhưng đôi khi kém chính xác hơn SIFT trong các hình ảnh phức tạp.

– ORB (Oriented FAST and Rotated BRIEF):

ORB là thuật toán hiệu quả về tính toán, kết hợp hai kỹ thuật:

- + FAST để phát hiện các điểm đặc trưng.
- + BRIEF để tạo các mô tả đặc trưng nhỏ gọn. ORB thường được sử dụng trong các ứng dụng yêu cầu xử lý nhanh, như thị giác máy tính trên thiết bị nhúng.

b. Đặc trưng toàn cục (Global Features)

Đặc trưng toàn cục biểu diễn các thông tin tổng thể của hình ảnh, thường được sử dụng để nhận diện các đối tượng lớn hoặc toàn cảnh.

- **Histogram of Oriented Gradients (HOG):**

HOG mô tả sự phân bố hướng gradient trong hình ảnh, rất hiệu quả để nhận diện đối tượng như người hoặc xe. HOG chia hình ảnh thành các ô (cells), tính toán histogram gradient trong mỗi ô, sau đó chuẩn hóa để tạo vector đặc trưng.

- **Đặc trưng màu sắc và kết cấu:**

- + Histogram màu: Thống kê tần suất xuất hiện của các giá trị màu trong hình ảnh, giúp phân biệt đối tượng theo màu sắc.
- + Local Binary Patterns (LBP): Mô tả kết cấu của hình ảnh bằng cách phân tích các mẫu pixel lân cận. LBP đặc biệt hiệu quả trong nhận diện khuôn mặt và phân loại hình ảnh.

2. Đối sánh đặc trưng

Sau khi trích xuất đặc trưng, quá trình đối sánh được thực hiện để xác định sự tương đồng giữa các đặc trưng từ hình ảnh đầu vào và hình ảnh mẫu.

- **Đo lường khoảng cách:**

Khoảng cách giữa các vector đặc trưng được sử dụng để đánh giá độ tương đồng. Các phương pháp đo lường phổ biến bao gồm:

- + Khoảng cách Euclid: Tính độ dài tuyến tính giữa hai vector.
- + Khoảng cách Manhattan: Tổng giá trị tuyệt đối của sự khác biệt giữa các thành phần vector.
- + Cosine Similarity: Đo lường góc giữa hai vector, đặc biệt hiệu quả với đặc trưng đã được chuẩn hóa.

- **Phương pháp đối sánh nhanh:**

Trong các ứng dụng với dữ liệu lớn, các phương pháp như KD-Tree hoặc FLANN (Fast Library for Approximate Nearest Neighbors) được sử dụng để tăng tốc độ tìm kiếm.

3. Phân loại mẫu vật

Sau khi đối sánh đặc trưng, quá trình phân loại hoặc nhận dạng mẫu vật được thực hiện để xác định đối tượng trong hình ảnh.

a. Phương pháp phân loại đơn giản

2. **K-Nearest Neighbors (KNN):** Phân loại đối tượng dựa trên k mẫu gần nhất trong không gian đặc trưng. KNN đơn giản nhưng hiệu quả với dữ liệu nhỏ.

b. Mô hình máy học

- **Support Vector Machines (SVM):** Tìm một siêu phẳng phân tách giữa các lớp trong không gian đặc trưng. SVM thường được sử dụng cho các bài toán có dữ liệu phức tạp.
- **Random Forest:** Áp dụng các cây quyết định để phân loại dựa trên đặc trưng. Random Forest có khả năng xử lý dữ liệu không tuyến tính tốt hơn so với SVM.

c. Phân cụm (Clustering)

Nếu dữ liệu không có nhãn, các thuật toán phân cụm như K-Means hoặc DBSCAN được sử dụng để nhóm các đối tượng tương tự vào một cụm.

1.6 Kết luận chương 1.

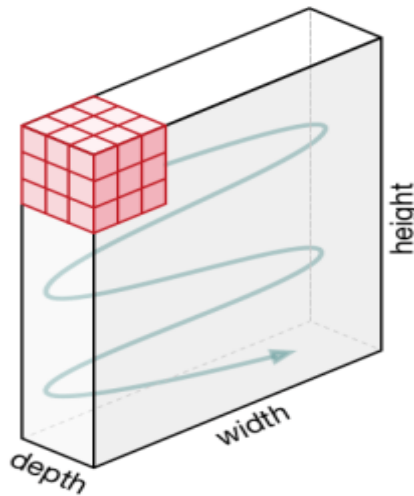
Chương này đã đưa ra tổng quát về hệ thống nhận dạng và tìm kiếm các cách tiếp cận trong việc xây dựng hệ thống nhận dạng và tìm kiếm mẫu vật từ đặc trưng được nhận dạng. Ngoài ra, luận văn tập trung vào việc xem xét các hướng tiếp cận để giải quyết vấn đề nhằm mục đích phân tích các ưu điểm và các nhược điểm, thách thức và hạn chế trong việc giải quyết bài toán nhận dạng nói chung và bài toán nhận dạng và tìm kiếm mẫu vật nói riêng.

CHƯƠNG 2: MẠNG NƠ-RON TÍCH CHẬP

Trong các mô hình học sâu, mạng nơ-ron tích chập (CNN) là một trong những mô hình được sử dụng phổ biến nhất. Nó hỗ trợ việc phát triển các hệ thống thông minh với độ chính xác tối ưu mô hình này đã và đang được áp dụng để đưa ra các chức năng như tìm kiếm ảnh, gợi ý các sản phẩm phù hợp cho khách hàng hoặc thuật toán gắn thẻ tự động ở Google, Amazon và Facebook.

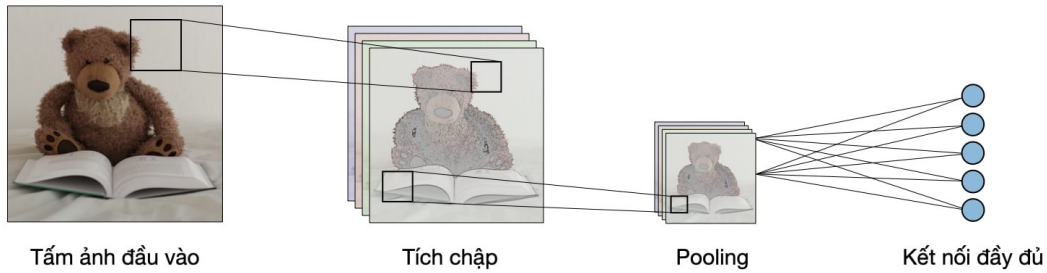
2.1 Kiến trúc mạng nơ-ron tích chập (CNN)

Một loại mô hình mạng nơ-ron nhân tạo được triển khai phổ biến trong lĩnh vực xử lý ảnh và thị giác máy tính là mạng nơ-ron tích chập, một kiểu mạng nơ-ron có cấu trúc truyền thẳng vì dữ liệu không có lượt lặp hoặc chu trình mà di chuyển theo một đường thẳng từ đầu vào đến đầu ra. Do khả năng xử lý nhanh và độ chính xác ưu việt của nó nên thường được áp dụng cho nhận dạng vật thể cũng như nhiều ứng dụng khác trong các lĩnh vực trí tuệ nhân tạo và thị giác máy tính. Trong mạng nơ-ron tích chập, các lớp có ba chiều: chiều rộng, chiều cao và chiều sâu, khác với một chiều trong mạng nơ-ron truyền thống (Hình 2.1). Bên cạnh đó chia sẻ tham số và kết nối cục bộ là hai khái niệm chính trong mạng nơ-ron tích chập. Điều này giúp tăng tốc độ tính toán bằng cách tối ưu số lượng trọng số cần thiết để huấn luyện.



Hình 2-1: Phép tính tích chập trên ảnh màu.

Một mạng CNN cơ bản có các lớp được sắp xếp khác nhau để thiết kế các mô hình huấn luyện phù hợp với từng bài toán riêng biệt. Các lớp này bao gồm lớp tích chập (convolutional), lớp kích hoạt phi tuyến ReLU (Rectified Linear Unit), Lớp lấy mẫu (Pooling Layer).



Hình 2-2: Mô hình cơ bản của mạng nơ-ron tích chập.

- Lớp tích chập

Lớp tích chập là lớp đầu tiên trong việc trích xuất đặc điểm từ hình ảnh đầu vào. Tích chập bảo đảm mối quan hệ giữa các điểm ảnh bằng cách sử dụng các bộ lọc nhỏ để thực hiện tích chập giữa các giá trị điểm ảnh trong một vùng ảnh cục bộ.

1	0	0	1	0		1	0	0	5	1	3		
0	1	1	0	1	×	0	1	1	2				
1	0	1	0	1		1	0	1					
1	0	0	1	0									
0	1	1	0	1									
1	0	0	1	0		1	0	0	5	1	3		
0	1	1	0	1	×	0	1	1	2	3	2		
1	0	1	0	1		1	0	1	2	2	4		
1	0	0	1	0									
0	1	1	0	1									

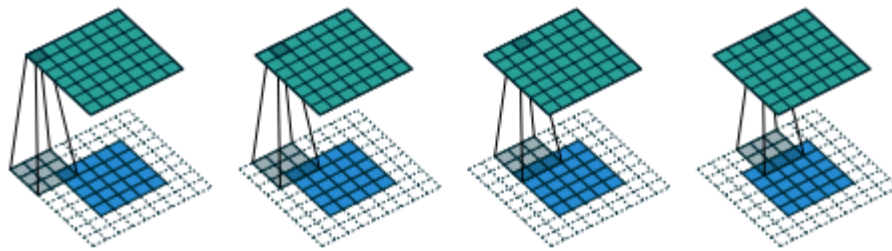
Hình 2-3: Một ma trận điểm ảnh sử dụng ma trận tích chập

Trên Hình 2.3 [10], ta có thể nhìn thấy rằng bộ lọc là một ma trận có kích thước cụ thể 3 x 3 và được áp dụng lần lượt lên từng khu vực của ảnh đầu vào trong mỗi lần áp dụng, giá trị của vùng ảnh và bộ lọc tích chập được nhân với nhau, các kết quả sau đó được kết hợp để tạo ra một giá trị duy nhất. Chúng tạo ra một bức ảnh mới có kích thước bằng hoặc nhỏ hơn bức ảnh đầu vào chứa các đặc trưng được trích chọn từ hình ảnh đầu vào bằng cách di chuyển bộ lọc trên toàn bộ bức ảnh. Kích thước này được xác định dựa trên kích thước các khoảng trống thêm vào viền bức ảnh gốc, theo công thức tính toán [11]:

$$m = \frac{n+2*i-j}{k} + 1$$

Trong đó:

- m: kích thước ảnh đầu ra
- n: kích thước ảnh đầu vào
- i: là kích thước của khoảng trắng thêm vào xung quanh viền của ảnh gốc
- j: kích thước bộ lọc
- k: bước trượt của bộ lọc



Hình 2-4: Ảnh khi tích chập khi thêm viền trắng.

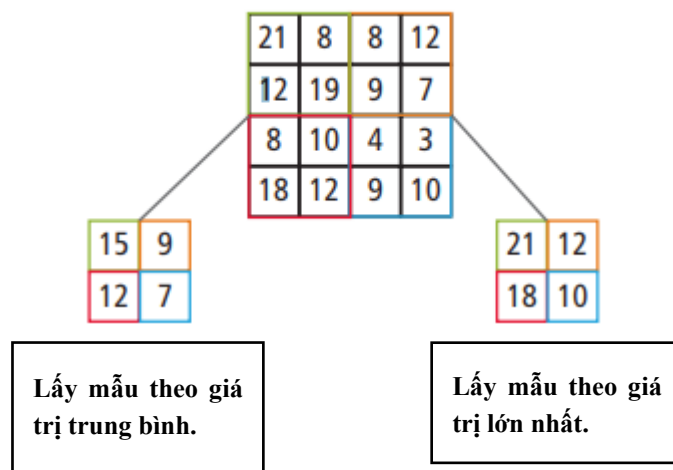
Sau khi qua lớp tích chập, ảnh đầu vào sẽ tạo ra một loạt các ảnh tương ứng với mỗi một bộ lọc được tính toán trên ảnh đầu vào. Các ảnh tương ứng chứa các đặc điểm được trích xuất từ ảnh đầu vào. Ban đầu trọng số trong bộ lọc được thiết lập ngẫu nhiên sau đó chúng sẽ được điều chỉnh và cải thiện qua quá trình huấn luyện.

- Lớp kích hoạt phi tuyến ReLU:

Để giữ tính phi tuyến cho mạng, đầu ra ở các lớp tuyến tính sẽ bị giới hạn bởi lớp kích hoạt phi tuyến tính. Các hàm kích hoạt như ReLU, tanh hoặc sigmoid thường được sử dụng trong lớp kích hoạt phi tuyến để kiểm soát phạm vi giá trị đầu ra. Ở đây, hàm ReLU được chọn trong số các hàm kích hoạt này vì nó có thể được cài đặt dễ dàng và đạt được tốc độ xử lý nhanh mà vẫn bảo đảm hiệu quả tính toán do chỉ đơn giản là chuyển đổi các giá trị âm thành 0. Ngay sau các lớp tích chập, lớp ReLU thường được sử dụng để làm hàm kích hoạt cho từng đầu ra các bộ lọc tích chập. Điều này đảm bảo hiệu suất tối ưu của mô hình và tăng tốc độ tính toán. Các giá trị điểm ảnh của bức ảnh đầu ra sau lớp ReLU sẽ tương tự như ảnh đầu vào sau khi xóa bỏ các giá trị âm.

- Lớp lấy mẫu:

Sử dụng sau khi các lớp tích chập và hàm kích hoạt ReLU là một phần thiết yếu của mạng CNN. Mục tiêu của lớp lấy mẫu là giữ đặc trưng quan trọng trong khi thu nhỏ kích thước dữ liệu đầu vào khi đó hiệu quả tính toán tăng lên đồng thời giảm số lượng tham số. Một cửa sổ trượt được sử dụng bởi lớp lấy mẫu để duyệt qua toàn bộ khu vực của hình ảnh giống như lớp tích chập. Tuy nhiên, lớp này sử dụng phép lấy mẫu nhằm xác định một giá trị đại diện từ khu vực đó để mô tả toàn bộ dữ liệu của khu vực đó. Nhờ đó, kích thước của các vùng được giảm đi so với sử dụng phép tích chập [12].



Hình 2-5: Ví dụ về lớp lấy mẫu

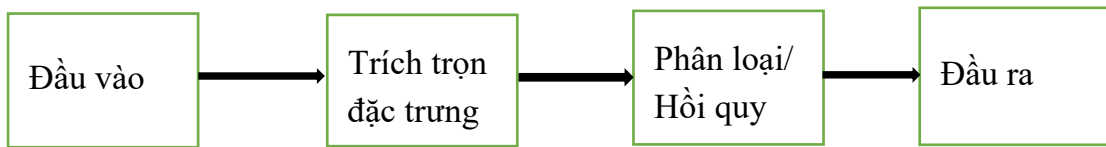
Các kỹ thuật lấy mẫu phổ biến hiện nay là lấy giá trị điểm ảnh tối đa và tính giá trị trung bình của các điểm ảnh trong khu vực cục bộ của ảnh [12], được mô tả tại Hình 2.5.

- Lớp kết nối đầy đủ:

Trong lớp kết nối đầy đủ có cấu trúc thiết kế giống như mạng nơ-ron truyền thống, nơi tất cả các điểm ảnh đều kết nối hoàn toàn với các nút trong lớp tiếp theo. Các ảnh đầu vào được cung cấp cho lớp này vẫn chứa các thông tin cần thiết cho việc nhận dạng, nhưng kích thước của chúng đã được giảm so với mạng nơ-ron truyền thống. Từ đó, tính toán nhận dạng được thực hiện bằng cách sử dụng mạng nơ-ron tích chập đơn giản hơn và tiết kiệm thời gian hơn so với mạng nơ-ron truyền thống.

2.2 Học chuyển giao và tinh chỉnh mô hình huấn luyện.

Trước năm 2012, phần lớn các mô hình học sâu đều được tạo thành từ hai bước riêng biệt:



Hình 2-6: Mô hình trí tuệ nhân tạo

Kỹ thuật trích chọn đặc trưng là kỹ thuật rút trích đặc trưng trong tập dữ liệu từ kiến thức của con người trong các vấn đề cần giải quyết. Do đó, các đặc trưng này được làm thủ công, kỹ thuật này là một thành phần của mô hình, từ đó các đặc trưng được học và đưa ra dự đoán bằng cách sử dụng các phương pháp học máy như phân lớp hoặc hồi quy.

Trong mô hình học sâu đã thiết lập hai bước này với nhau, các lớp đầu tiên là bước trích chọn đặc trưng (Feature Extractor), trong khi bước tiếp theo áp dụng các phương pháp máy học như phân lớp hoặc hồi quy để tạo ra kết quả. Khác với kỹ thuật trích chọn đặc trưng truyền thống, trong mô hình học sâu đặc trưng có thể lấy tự động từ dữ liệu thông qua quá trình học mà không cần định nghĩa trước.

Nhiều yếu tố góp phần vào tạo nên sự khác biệt về hiệu suất giữa các mô hình học sâu hiện đại và các phương pháp cũ như việc sử dụng các bộ dữ liệu mở rộng, khả năng xử lý song song và bản chất động của các lớp tính năng trong mô hình. Các mô hình này có khả năng thay đổi và điều chỉnh các tính năng của chúng để nắm bắt chính xác các đặc điểm cần thiết để tạo ra đầu ra mong muốn thông qua truyền ngược, không giống như các phương pháp cũ dựa vào các tính năng được thiết kế trước của con người. Một điểm quan trọng là khả năng mã hóa: Trong giai đoạn ban đầu việc chuyển đổi việc dữ liệu thành một loạt các biến ẩn do bước trích chọn đặc trưng đã lọc ra các thuộc tính thiết yếu và biến chúng thành một tập dữ liệu nhỏ gọn có thể giải quyết được bởi phần giải mã của mạng giống hệt nhau. Liên quan đến các thuộc tính khác chúng ta có thể linh hoạt điều chỉnh chuyển đổi đầu vào thành định dạng phù hợp ở giai đoạn tiếp theo để đảm bảo kết quả tối ưu.

Mỗi loại mạng học sâu đều sở hữu một khả năng độc đáo để rút trích các đặc trưng cụ thể từ các dạng dữ liệu khác nhau. Điều này tạo thành nguyên tắc cốt lõi của việc học chuyển giao, cho phép tận dụng các tính năng thu được từ việc học dữ liệu lớn để xác định các đặc trưng cho mô hình mà không cần huấn luyện lại toàn bộ từ đầu.

Trong giai đoạn đầu khi các kỹ thuật học sâu mới đang phát triển, một khái niệm phổ biến nhưng sai lầm đã chiếm ưu thế trong cộng đồng học sâu: nếu không có một số lượng dữ liệu huấn luyện lớn, việc tạo ra một mô hình học sâu mạnh mẽ

được coi là không thể. Nó đã từng là một khái niệm đúng đắn và hợp lý, vì các mô hình học máy này bao gồm nhiều lớp với hàng nghìn nơ-ron và hàng triệu tham số, yêu cầu sức mạnh tính toán mạnh mẽ cho việc huấn luyện. Đồng thời, một mô hình học sâu được huấn luyện trên một tập dữ liệu riêng biệt ví dụ như dữ liệu từ nhiệm vụ phân loại hoa quả có thể không chuyển giao hiệu quả khi sang sang tập dữ liệu khác. Điều này chứng minh rằng việc dùng mô hình học sâu đã được huấn luyện trên một tập dữ liệu khác có thể ảnh hưởng đến độ chính xác của mô hình do sự thay đổi trong cấu trúc dữ liệu, điểm và sự đa dạng của các thành phần trong hình ảnh. Do đó, một phương pháp mới đã ra đời để giải quyết các vấn đề hiện có, được biết đến với tên gọi học chuyển giao [13].

Học chuyển giao học trong lĩnh vực học máy thường đòi hỏi phải sử dụng một mô hình đã được huấn luyện trước với một tập dữ liệu phong phú và đa dạng và ứng dụng vào một tác vụ hoặc lĩnh vực mới. Cách tiếp cận này cho phép áp dụng kiến thức từ việc huấn luyện trên tập dữ liệu đa dạng và phong phú mà không cần bắt đầu huấn luyện một mô hình ngay từ đầu, do đó tiết kiệm tài nguyên tính toán và thời gian. Hơn nữa, học chuyển giao có thể nâng cao chất lượng mô hình trên các tập dữ liệu mới nhờ vào việc sử dụng các đặc trưng thu được từ dữ liệu lớn. Hiện tại, một kỹ thuật phổ biến khi sử dụng một tập dữ liệu có quy mô nhỏ để huấn luyện mô hình sử dụng phương pháp học chuyển giao đó là sử dụng mạng nơ-ron tích chập được huấn luyện trên tập dữ liệu phong phú và đa dạng như ImageNet có 1,2 triệu hình ảnh với 1 nghìn nhãn lớp. Phương pháp này có thể sử dụng theo hai cách được nêu dưới đây: cách đầu tiên chỉ sử dụng mạng nơ-ron tích chập như một bộ trích xuất đặc trưng cho tập dữ liệu huấn luyện mới, thay thế các lớp được kết nối đầy đủ ở cuối mạng trong khi duy trì các tham số cố định cho tất cả các lớp còn lại. Cách hai đồng thời tham gia vào việc tinh chỉnh, điều chỉnh và tinh chỉnh tất cả hoặc một số lớp nhất định trong mạng. Quan điểm tái sử dụng mạng nơ-ron tích chập xuất phát từ việc các đặc trưng học được trong các lớp mạng đầu tiên là chung, tổng quát nhất và có thể áp dụng cho nhiều tác vụ nhận dạng hình ảnh. Các lớp ban đầu của mạng nơ-ron tích chập thường xác định các đặc điểm cơ bản như các cạnh, hình dạng cơ bản hoặc thuộc tính màu sắc. Ngược lại, các lớp sau trong mạng nơ-ron tích chập thường bao gồm các tính năng phức tạp hơn để củng cố tính cụ thể và sự khác biệt của tính năng, phục vụ cho các nhiệm vụ nhận dạng cụ thể.

Hai phương pháp chính trong học chuyển giao là:

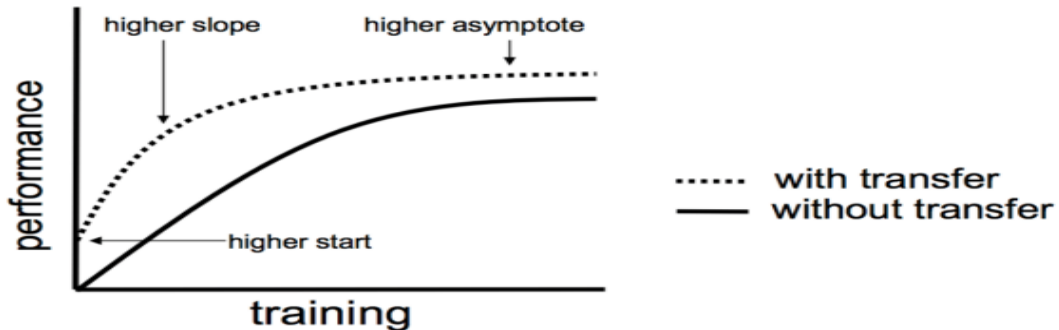
Một phương pháp là sử dụng một mô hình đã được huấn luyện trước để trích xuất các đặc trưng từ dữ liệu đầu vào, sau đó sử dụng các đặc trưng này như đầu vào cho một mô hình học máy khác. Thay vì tinh chỉnh mô hình, giữ nguyên các trọng

số của mô hình đã huấn luyện trước và chỉ sử dụng đầu ra của nó để huấn luyện mô hình mới.

Phương pháp khác là bắt đầu với một mô hình đã được huấn luyện trước trên một tập dữ liệu lớn và chung, sau đó tinh chỉnh các trọng số của mô hình bằng cách tiếp tục huấn luyện nó trên một tập dữ liệu mới và cụ thể hơn. Tinh chỉnh thường chỉ thay đổi một phần của mô hình, chẳng hạn như các lớp cuối cùng hoặc các lớp được thêm vào, trong khi các lớp đầu vẫn giữ nguyên.

2.2.2 Lợi ích của việc học chuyển giao.

- Bằng cách sử dụng các đặc trưng thu được từ mô hình sẵn có, giúp cắt giảm đáng kể thời gian dành cho việc huấn luyện mô hình mới. Điều này là do mô hình có thể tận dụng các trọng số đã có sẵn, dẫn đến quá trình huấn luyện hiệu quả hơn.
- Hiệu quả: Mô hình có sẵn đã cung cấp một sự chính xác cao vì vậy học trên mô hình có sẵn thì mô hình chuyển giao (transfer model) sẽ nâng cao chính xác này thay vì phải bắt đầu trạng thái có sự chính xác thấp. Như trong Hình 2.7 cho chúng ta thấy khi sử dụng học chuyển giao mô hình sẽ tiếp tục tăng độ chính xác hiệu quả mô hình (performance).



Hình 2-7: Hiệu quả dùng khi transfer learning.

2.2.3 Hạn chế của việc học chuyển giao.

- Học chuyển giao là phương pháp không dễ sử dụng nếu bạn sai lầm trong quá trình học chuyển giao từ mô hình có sẵn như thêm bớt các lớp thì khi học sẽ dẫn đến sự chính xác sẽ bị giảm xuống thấp.
- Học chuyển giao chỉ có thể dùng mô hình có sẵn có nếu có sự liên quan đến mô hình đầu ra mong muốn.

2.4 Ứng dụng mạng VGG16 vào bài toán phân loại và tìm kiếm mẫu vật.

Thực tế hiện nay các mẫu vật ma túy dạng viên nén được phân biệt với nhau dựa trên hình dáng, màu sắc và logo in trên viên ma túy với số lượng chủng loại xuất hiện ngày càng nhiều phụ thuộc từng nơi sản xuất. Do vậy việc nhận dạng để phân loại và tìm kiếm thông tin về mẫu vật ma túy rất quan trọng trong quá trình điều tra truy nguyên nguồn gốc. Trong luận văn, em xin tập trung vào nghiên cứu thử nghiệm nhận dạng và tìm kiếm thông tin của 7 loại ma túy dạng viên nén xuất hiện nhiều nhất tại thời điểm hiện tại trong kho ảnh tang vật bằng phương pháp ứng dụng các mô hình học sâu.

Dựa trên những kết luận rút ra từ hiệu quả của học chuyển giao được nêu trong phần 2.2, đặc biệt là với kiến trúc mạng tích chập trong việc giải quyết các khó khăn bởi kích thước cơ sở dữ liệu hạn chế, luận văn này xin đưa ra phương pháp để giải quyết các vấn đề khó khăn:

1) Ứng dụng kiến trúc mạng VGG16 bằng cách sử dụng mô hình được huấn luyện trước có nguồn gốc từ tập dữ liệu ImageNet. Việc lựa chọn kiến trúc VGG16 dựa trên các đặc điểm vượt trội của nó, bao gồm kích thước và độ sâu mạng tăng cường so với các kiến trúc thay thế như AlexNet, việc triển khai các bộ lọc giảm dần kích thước giữa các lớp để phù hợp với cả các đặc trưng chung và riêng, và áp dụng các kỹ thuật chuẩn hóa cục bộ tạo điều kiện cho việc chuẩn hóa các lớp, do đó đẩy nhanh quá trình hội tụ. Mạng VGG16 được huấn luyện với kho lưu trữ hình ảnh ImageNet là một cơ sở dữ liệu hình ảnh lớn để thúc đẩy nghiên cứu trong các hệ thống nhận dạng đối tượng. ImageNet bao gồm hơn 14 triệu hình ảnh đã được gán nhãn cụ thể để biểu thị các đối tượng hiện diện trong mỗi hình ảnh. Ngoài ra, ImageNet bao gồm hơn 1000 danh mục, với một danh mục đại diện bao gồm các mẫu vật như trái cây và động vật, mỗi danh mục chứa hàng trăm hình ảnh.

2) Thiết lập một bộ dữ liệu hình ảnh phục vụ cho mục đích huấn luyện, bao gồm 7 loại mẫu vật ma túy ở dạng viên nén, với hình ảnh được lựa chọn theo các tiêu chí về kích thước, thuộc tính màu sắc và độ sắc nét của hình ảnh và được chú thích tỉ mỉ

3) Tinh chỉnh lại mô hình để tối ưu độ chính xác phân loại với 7 loại ma túy dạng viên nén và tìm kiếm thông tin về loại ma túy tương đồng nhất với hình ảnh ma túy cần tìm. Như đã được chứng minh trong các nghiên cứu khoa học nhờ khả năng trích chọn đặc trưng tự động của mạng VGG16, việc áp dụng học chuyển giao với các mô hình mạng nơ-ron tích chập trong trường hợp kích thước của CSDL ảnh không quá lớn có thể mang lại nhiều lợi ích và hiệu quả. Tinh chỉnh mô hình sẽ tính

toán phân phối xác suất ở đầu ra, tận dụng lại đặc trưng từ mô hình có sẵn và giảm chiều dữ liệu.

2.5 Kết luận chương 2.

Chương này tập trung vào phương pháp học sâu trong giải quyết bài toán. Có thể nói rằng bài toán nhận dạng và tìm kiếm mẫu vật từ đặc trưng đã đáp ứng được nhu cầu của người dùng bằng cách sử dụng các kỹ thuật trích chọn đặc trưng trong phương pháp học sâu. Nhưng thực tế là chất lượng ảnh đầu vào là một trong những yếu tố ảnh hưởng đến chất lượng nhận dạng, tìm kiếm. Vì lý do đó, trong hướng tiếp cận thu thập tập mẫu dựa trên kết quả nhận dạng, tìm kiếm ngoài việc lựa chọn mô hình nhận dạng và phương pháp tìm kiếm có thể giải quyết tốt trên tập ảnh đầu vào được kiểm thử chúng ta còn cần phải chú ý tới tiền xử lý, tăng cường chất lượng ảnh đầu vào trước khi nhận dạng.

CHƯƠNG 3: HỆ THỐNG NHẬN DẠNG VÀ TÌM KIẾM MẪU VẬT

3.1 Kiến trúc hệ thống.

Kiến trúc hệ thống nhận dạng và tìm kiếm mẫu vật được thiết kế theo mô hình máy khách và máy chủ gồm năm lớp:

- Tầng thứ nhất là tầng giao diện nhận dạng và tìm kiếm trên môi trường web thông minh, ghi nhận các sự kiện người dùng như việc chụp và gửi ảnh lên máy chủ, và hiển thị kết quả từ máy chủ ra giao diện.
- Tầng thứ hai của máy chủ là tầng điều phối dữ liệu chịu trách nhiệm điều chỉnh các thiết lập truyền tải thông tin dữ liệu với máy khách. Hệ thống áp dụng giao thức HTTP cho việc trao đổi truyền tải thông tin dữ liệu giữa máy chủ và máy khách.
- Tầng thứ ba của máy chủ quản lý hoạt động nghiệp vụ như yêu cầu từ máy khách và đảm bảo rằng hiệu suất nhận dạng, tìm kiếm vẫn được duy trì ổn định dù có nhiều máy khách kết nối cùng lúc.
- Tầng thứ tư phụ trách huấn luyện và cập nhật và duy trì các phiên bản của mô hình nhận dạng sử dụng dữ liệu huấn luyện được cung cấp bởi tầng quản lý dữ liệu phía dưới.
- Tầng máy chủ thứ 5 là tầng quản lý, lưu trữ đặc trưng hình ảnh được nhận dạng để tìm kiếm thông tin mẫu vật. Lưu trữ dữ liệu từ người dùng với mục tiêu đa dạng hóa và nâng cao độ chính xác của mô hình. Việc lưu trữ các bộ ảnh theo cấu trúc để dễ dàng quản lý và đánh giá các phiên bản mô hình huấn luyện.

Tầng 1: Giao diện nhận dạng tìm kiếm mẫu vật.

- Ứng dụng trên môi trường web.

Tầng 2: Máy chủ thiết lập quản lý cấu hình.

- HTTP Server: Thiết lập và duy trì các kết nối với người dùng, bao gồm việc xử lý nhiều yêu cầu đồng thời.

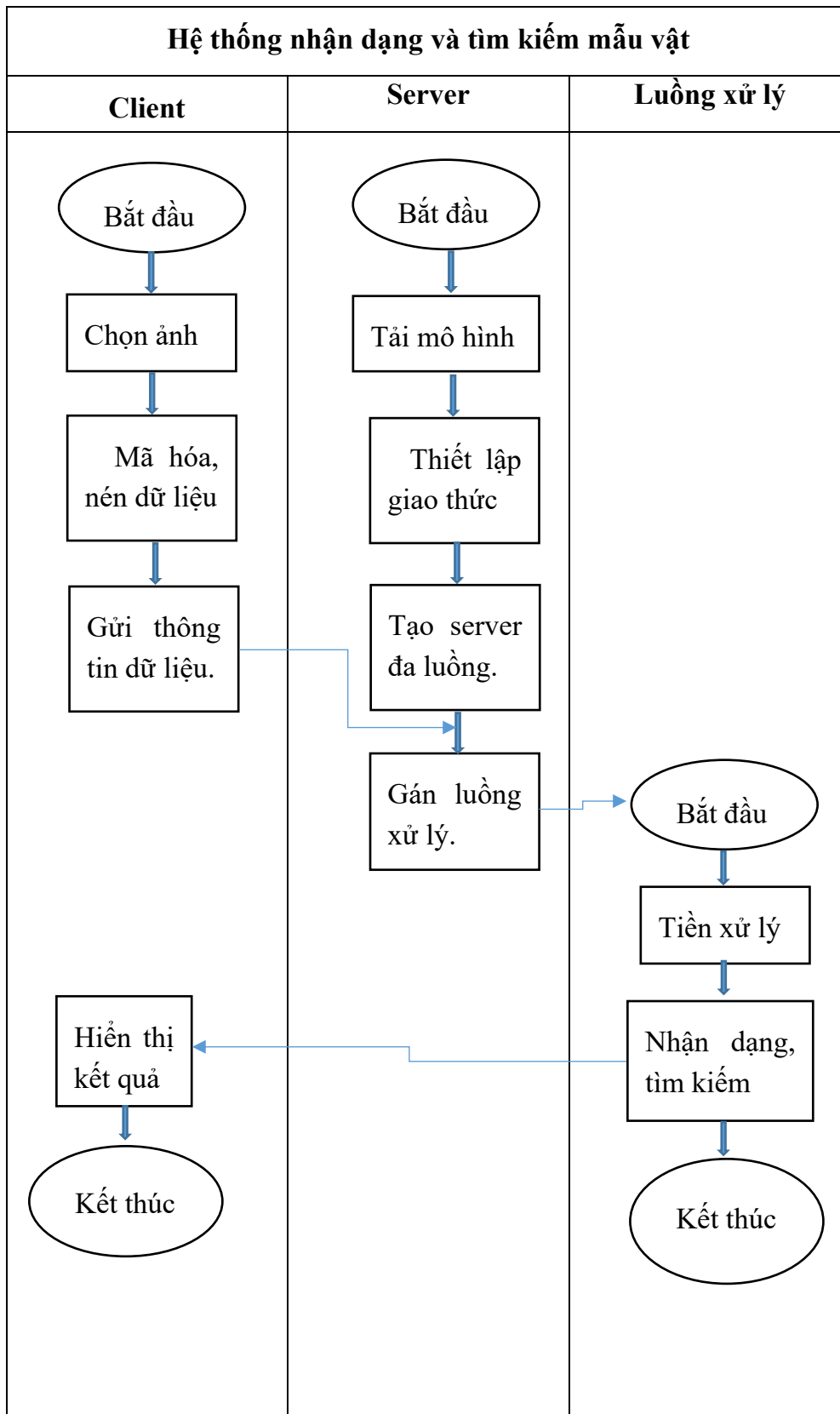
Tầng 3: Máy chủ xử lý nghiệp vụ hệ thống • Cập nhật mô hình nhận dạng đã được huấn luyện lên máy chủ. • Quản lý các luồng xử lý độc lập: ▪ Máy chủ kiểm tra trạng thái của các luồng xử lý và phân phối yêu cầu đến luồng đang không có nhiệm vụ khi có yêu cầu từ máy khách. ▪ Khi luồng xử lý kết thúc công việc nhận dạng và tìm kiếm, máy chủ sẽ giải phóng luồng và cập nhật thông tin trạng thái của luồng.
Tầng 4: Bộ huấn luyện mô hình nhận dạng. • Huấn luyện mô hình nhận dạng: ▪ Truy xuất dữ liệu từ tầng CSDL. ▪ Tính toán tham số để huấn luyện mô hình. ▪ Thực hiện huấn luyện mô hình. ▪ Lưu trữ mô hình đã huấn luyện thành công. • Quản lý và theo dõi các mô hình đã huấn luyện thành công.
Tầng 5: Cơ sở dữ liệu. • Dữ liệu hình ảnh gốc: CSDL ban đầu để huấn luyện mô hình gồm cả ảnh nhiễu và ảnh gốc. • Dữ liệu hình ảnh thực tế: CSDL thu thập được qua các lần xử lý yêu cầu từ máy khách, được huấn luyện bổ sung và cải thiện chất lượng nhận dạng cho mô hình huấn luyện sau. • Lưu trữ đặc trưng được trích xuất từ mô hình nhận dạng.

Hình 3-1: Mô tả hoạt động máy khách và máy chủ.

Sơ đồ mô tả các hoạt động máy chủ và máy khách thực hiện từ khi khởi động cho đến khi kết thúc hệ thống được minh họa trong Hình 3.1:

- Máy khách ứng dụng trên môi trường web:
 - Người dùng mở ứng dụng web.
 - Người dùng chọn ảnh từ thư mục.

- Mã hóa và nén trước khi gửi.
- Nhận và hiển thị kết quả.
- Chương trình máy chủ:
 - Chương trình bắt đầu tải thư viện khởi động.
 - Cập nhật mô hình huấn luyện thành công lên máy chủ.
 - Thiết lập phương thức truyền thông tin dữ liệu giữa máy khách và máy chủ.
 - Khởi tạo các luồng xử lý, bắt đầu ở trạng thái chờ.
 - Phân phối luồng xử lý để nhận và xử lý dữ liệu từ máy khách.
 - Trong luồng xử lý:
 - ✓ Khởi động với trạng thái đang thực hiện.
 - ✓ Giải nén dữ liệu từ máy khách.
 - ✓ Nhận dạng qua mô hình, tìm kiếm thông tin trong cơ sở dữ liệu.
 - ✓ Gửi kết quả nhận dạng kèm thông tin sau khi tìm kiếm về cho máy khách.
 - ✓ Hoàn tất quá trình nhận dạng và tìm kiếm.
 - Khi quá trình nhận dạng và tìm kiếm hoàn tất, trạng thái luồng được cập nhật về chờ xử lý để giải phóng luồng.



Hình 3-2: Sơ đồ hoạt động của hệ thống

3.2 Mô-đun quản lý và lưu trữ dữ liệu.

Dữ liệu hình là thành phần chính trong việc huấn luyện và tối ưu hóa các mô hình nhận dạng, nó ảnh hưởng đến hiệu suất hoạt động và tính chính xác của mô hình. Quản lý và lưu trữ cơ sở dữ liệu hình ảnh hiệu quả là rất quan trọng để đảm bảo rằng quy trình huấn luyện mô hình nhận dạng, tìm kiếm thông tin hình ảnh chính xác và tiết kiệm thời gian, đồng thời tối ưu hóa chất lượng và hiệu suất của mô hình. Hệ thống lưu trữ hình ảnh gồm:

- **Lưu trữ đặc trưng:** lưu giữ các đặc trưng đã được trích xuất từ hình ảnh vào một cơ sở dữ liệu hoặc một cấu trúc dữ liệu, để có thể truy xuất và sử dụng lại sau này. Lưu trữ đặc trưng có thể giúp giảm thiểu việc tính toán lại các đặc trưng, tiết kiệm thời gian và tài nguyên khi xử lý dữ liệu lớn hoặc khi triển khai hệ thống.
- **Thư mục dữ liệu hình ảnh gốc:** bao gồm các thư mục con được tổ chức với các nhiệm vụ khác nhau nhằm hỗ trợ việc xây dựng và huấn luyện mô hình. Mỗi thư mục bao gồm hình ảnh gốc, hình ảnh nghiêng, hình ảnh nhiễu, hình ảnh một nền và hình ảnh nhiều nền.
- **Thư mục hình ảnh thực tế chưa kiểm tra:** gồm ảnh được thu thập từ người dùng tổ chức thành các thư mục con dựa trên 7 loại mẫu ma túy đã được huấn luyện. Thư mục này chưa được sử dụng để huấn luyện do chưa được kiểm duyệt.
- **Thư mục ảnh thực tế đã được kiểm tra:** Các ảnh đã được kiểm tra và gán nhãn chính xác được sắp xếp vào các thư mục con để chuẩn bị cho việc huấn luyện mô hình. Các thư mục chính được đánh số để phân biệt các phiên bản mô hình huấn luyện bổ sung. Việc này giúp tránh nhầm lẫn giữa các phiên bản và bảo đảm sự thống nhất trong quá trình xây dựng mô hình.

3.3 Mô-đun nhận dạng và tìm kiếm mẫu vật.

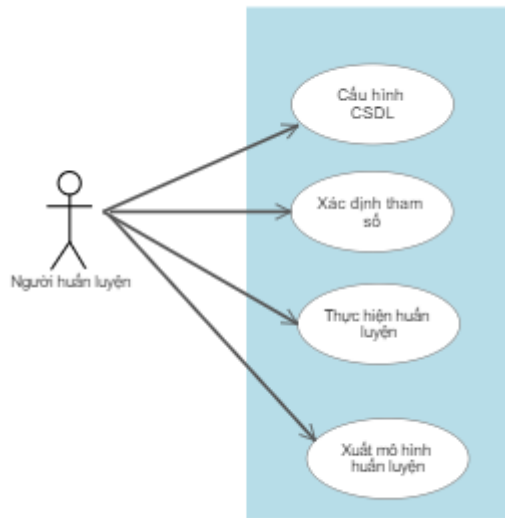
Đây là một thành phần thiết yếu ở tầng thứ ba trong cấu trúc của hệ thống có nhiệm vụ nhận dạng và tìm kiếm mẫu vật. Quá trình này bao gồm hai giai đoạn là nhận dạng mẫu vật từ mô hình học sâu và tìm kiếm thông tin liên quan đến mẫu vật từ đặc trưng được nhận dạng trong hệ thống.

- Mục tiêu: Người dùng có thể nhận dạng mẫu vật từ mô hình học sâu đã được huấn luyện, và sau đó tìm kiếm thông tin liên quan về mẫu vật trong hệ thống.
- Tác nhân, Mô tả tổng quát:

- Tác nhân: Người dùng muốn nhận dạng và tìm kiếm thông tin về mẫu vật.
- Mô tả tổng quát: Người dùng tải lên hình ảnh muốn nhận dạng và tìm kiếm, hệ thống tiến hành nhận dạng thông qua mô hình học sâu được huấn luyện, hệ thống tiếp tục tìm kiếm thông tin liên quan đến mẫu vật và trả kết quả về cho người dùng.
- Kịch bản chính: Hệ thống sử dụng mô hình nhận dạng để phân tích và nhận dạng mẫu vật. Sau khi nhận dạng xong hệ thống sẽ so sánh hình ảnh với cơ sở dữ liệu để tìm ra mẫu vật tương tự.
- Kịch bản phụ: Không.
- Các yêu cầu tiêu chuẩn: Hệ thống cần hỗ trợ các định dạng hình ảnh và dữ liệu khác nhau.
- Điều kiện trước: Cơ sở dữ liệu đặc trưng hình ảnh phải có sẵn trong hệ thống lưu trữ và phải đáp ứng cấu trúc cũng như định dạng được quy định.
- Điều kiện sau: Người dùng nhận được kết quả nhận diện và thông tin liên quan về mẫu vật.

3.4 Mô-đun huấn luyện mô hình.

Đây là một phần thiết yếu ở tầng thứ tư trong cấu trúc của hệ thống có nhiệm vụ huấn luyện từ dữ liệu. Quá trình này bao gồm cả giai đoạn khởi tạo mô hình và tinh chỉnh mô hình để nâng cao độ chính xác và hiệu suất. Việc triển khai, cấu hình và nâng cấp mô-đun huấn luyện được thực hiện mà không gây cản trở hoặc ảnh hưởng đến hoạt động của các phần mềm và dịch vụ khác trên cùng một máy chủ. Các hoạt động chính của mô-đun bao gồm: chuẩn bị và cấu hình dữ liệu huấn luyện, xác định các tham số quan trọng cho mô hình, thực hiện quá trình huấn luyện mô hình, và xuất ra mô hình đã huấn luyện cùng với thông tin liên quan, dựa trên phiên bản của bộ cơ sở dữ liệu hình ảnh đã được áp dụng mô tả trong Hình 3.3



Hình 3-3: Biểu đồ mô tả các thành phần mô-đun huấn luyện mô hình.

Mô tả các ca sử dụng trong mô-đun huấn luyện mô hình

Cấu hình cơ sở dữ liệu hình ảnh:

- Mục tiêu: Để thiết lập các thông tin cơ bản về cơ sở dữ liệu hình ảnh cho việc huấn luyện mô hình bao gồm đường dẫn đến thư mục chứa các ảnh được sử dụng cho huấn luyện mô hình, phiên bản hiện tại của quá trình huấn luyện, số lượng ảnh dùng cho huấn luyện và kiểm thử.
- Tác nhân, Mô tả tổng quát:
 - Tác nhân: Quản lý hệ thống hoặc người quản lý mô-đun huấn luyện mô hình nhận dạng.
 - Mô tả tổng quát: Cần cấu hình thông tin bộ CSDL ảnh phục vụ cho huấn luyện trước khi muốn bắt đầu huấn luyện mới, hoặc huấn luyện bổ sung cho mô hình nhận dạng mẫu.
- Kịch bản chính: Cập nhật thông tin về bộ dữ liệu hình ảnh trong tập tin cấu hình huấn luyện, tạo ra các tập tin mới để lưu trữ đường dẫn đến ảnh huấn luyện và kiểm thử, cũng như các nhãn phân loại tương ứng.
- Kịch bản phụ: Không.
- Các yêu cầu tiêu chuẩn: Đánh giá và xác minh rằng các đường dẫn đến ảnh huấn luyện và kiểm thử không gặp lỗi và đúng định dạng.
- Điều kiện trước: Cơ sở dữ liệu hình ảnh huấn luyện phải có sẵn trong hệ thống lưu trữ và phải đáp ứng cấu trúc cũng như định dạng được quy định.
- Điều kiện sau: Không có điều kiện sau cụ thể phải thực hiện.

Xác định tham số huấn luyện:

- Mục tiêu: Phân tích các thông số cần thiết từ bộ cơ sở dữ liệu hình ảnh và xác định các tham số để thiết lập mô hình và quy trình huấn luyện.
- Tác nhân, Mô tả tổng quát:
 - Tác nhân: Quản lý hệ thống hoặc người quản lý mô-đun huấn luyện mô hình nhận dạng.
 - Mô tả tổng quát: Người quản lý cần xác định các thông số thiết yếu và phân tích các giá trị từ bộ dữ liệu hình ảnh để chuẩn bị cho quy trình huấn luyện mô hình nhận dạng...
- Kịch bản chính: Người quản lý sử dụng tệp thực thi đã được phát triển để tính toán các thông số cần thiết từ bộ dữ liệu hình ảnh đầu vào. Tệp này có thể chứa các hàm xử lý dữ liệu, tính toán, trích chọn đặc trưng, hoặc các tác vụ khác để chuẩn bị dữ liệu cho quá trình huấn luyện. Sau khi hoàn tất tính toán, người quản lý sẽ chỉnh sửa và cập nhật các tham số trong các tệp cấu hình huấn luyện mô hình. Các tham số này bao gồm kiến trúc mạng neuron, tốc độ học (learning rate), số lượng epoch, kích thước batch (batch size), và các tham số khác cần thiết cho việc huấn luyện mô hình.
- Kịch bản phụ: Khi tệp thực thi không thể hoàn tất việc tính toán trên bộ dữ liệu hình ảnh đã cấu hình, tệp cần phải xử lý lỗi đúng cách và thông báo cho người quản lý hoặc người dùng về sự cố.
- Các yêu cầu tiêu chuẩn: Đảm bảo rằng đầu ra của tệp thực thi là các tệp dữ liệu theo định dạng tiêu chuẩn đã được xác định, và các tham số cấu hình mô hình phải được thiết lập và cập nhật phù hợp với mục tiêu huấn luyện.
- Điều kiện trước: Các chi tiết liên quan đến bộ dữ liệu hình ảnh đã được thiết lập chính xác trước khi thực hiện.
- Điều kiện sau: Xác nhận rằng quá trình tính toán các giá trị cần thiết đã diễn ra thành công.

Thực hiện huấn luyện:

- Mục tiêu: Huấn luyện và điều chỉnh mô hình nhận dạng cho hệ thống bằng cách sử dụng bộ dữ liệu hình ảnh đã được chuẩn bị trước bao gồm ảnh gốc, ảnh tự động sinh ra và ảnh thu thập trong quá trình hoạt động.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Quản lý hệ thống hoặc người quản lý mô-đun huấn luyện mô hình nhận dạng.

- Mô tả tổng quát: Sau khi người quản trị đã thiết lập thông tin về CSDL ảnh và các tham số cần thiết được tính toán, có thể thực hiện quá trình huấn luyện mô hình nhận dạng cho hệ thống.
- Kịch bản chính: Người quản lý sử dụng tập tin thực thi để thực hiện các lệnh nhằm bắt đầu hoặc tiếp tục phiên huấn luyện mô hình. Các lệnh này bao gồm: lệnh khởi tạo một phiên huấn luyện mới và lệnh tiếp tục phiên huấn luyện đã bị dừng trước đó.
- Kịch bản phụ: Tập tin thực thi sẽ thông báo lỗi nếu không thể thực hiện huấn luyện với các tham số đầu vào đã cấu hình. Các lỗi có thể gặp phải bao gồm: lỗi trong tập tin định nghĩa mô hình, mô hình huấn luyện trước không hợp lệ, sự cố khi lựa chọn card đồ họa GPU, và lỗi với tập tin huấn luyện khi tạm dừng.
- Các yêu cầu chi tiết: Kết quả của quá trình huấn luyện bao gồm một mô hình nhận dạng mẫu cùng với các tập tin ghi lại nhật ký huấn luyện gồm thông tin cảnh báo hoặc lỗi xảy ra trong quá trình huấn luyện, giúp người quản lý có thể theo dõi và kiểm tra khi cần. Thêm vào đó, thông tin về phiên bản của mô hình nhận dạng cũng được lưu trữ.
- Điều kiện trước: Các thông tin cấu hình về bộ dữ liệu ảnh và tham số cho việc huấn luyện mô hình cần phải được thiết lập một cách chính xác. Mô hình huấn luyện trước và tập tin trạng thái huấn luyện tại thời điểm tạm dừng cũng phải được kiểm tra tính hợp lệ.
- Điều kiện sau: Cung cấp thông báo huấn luyện mô hình đã thành công, cùng với thông tin về phiên bản hiện tại của mô hình và độ chính xác đạt được.

Xuất mô hình đã huấn luyện:

- Mục tiêu: Lưu trữ mô hình đã được huấn luyện hoàn tất để phục vụ cho quá trình nhận dạng mẫu phía máy chủ.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Người quản lý hệ thống hoặc người quản lý mô-đun huấn luyện mô hình nhận dạng.
 - Mô tả chung: Sau khi mô hình đã qua huấn luyện thành công, tiến hành xuất mô hình mới và thay thế mô hình hiện tại.
- Kịch bản chính: Người điều hành hệ thống gọi tập thực thi để thực hiện việc xuất mô hình đã huấn luyện vào thư mục lưu trữ theo cấu hình trong tệp cấu hình chung. Mô hình cũ hiện đang được dùng sẽ được di chuyển vào thư mục lưu trữ các phiên bản không còn hoạt động.

- Kịch bản phụ: Tập thực thi hiển thị thông báo lỗi trong quá trình lưu mô hình đã hoàn thành và chuyển mô hình cũ vào thư mục lưu trữ các phiên bản không còn sử dụng.
- Các yêu cầu chi tiết: Quá trình xuất mô hình cần phải đảm bảo tính chính xác và sự hoàn chỉnh của mô hình đã được huấn luyện.
- Điều kiện trước: Mô hình sau khi huấn luyện phải đầy đủ và hoàn chỉnh, có độ chính xác và hiệu suất đã được đánh giá. Thông tin cấu hình liên quan đến thư mục mô hình đang sử dụng và thư mục lưu các phiên bản cũ phải hợp lệ để quản lý các phiên bản mô hình hiệu quả.
- Điều kiện sau: Xác nhận rằng việc xuất mô hình đã thành công với mô hình đã được lưu và sẵn sàng cho các mục đích sử dụng.

Mô-đun huấn luyện mô hình bao gồm hai phần quan trọng: chương trình huấn luyện sử dụng mạng nơ-ron tích chập VGG16 và công cụ phát triển Keras được cài đặt trên hệ điều hành Ubuntu và thành phần quản lý phiên bản cho mô hình nhận dạng.

3.5 Môi trường huấn luyện.

Môi trường sử dụng để huấn luyện mô hình nhận dạng mẫu vật bao gồm công cụ phát triển Keras được cài đặt trên hệ điều hành Ubuntu 18.04 LTS, Python 3.7.10. Keras là một nền tảng mã nguồn mở dành cho học sâu, được Berkeley Vision and Learning Center (BVLC) phát triển viết bằng C++, CUDA C++, và bao gồm gói hỗ trợ cho các ngôn ngữ lập trình khác như Python và Matlab. Keras là một nền tảng mã nguồn mở cho học sâu, được viết bằng Python, dễ sử dụng và thiết kế đơn giản với cộng đồng lớn và sự phổ biến. Đặc biệt, Keras cung cấp tùy chọn cho người dùng để huấn luyện mô hình trên CPU hoặc GPU, giúp tăng tốc quá trình huấn luyện một cách đáng kể khi GPU được sử dụng thay vì CPU. Keras cũng nổi bật với nhiều ưu điểm như dễ dàng sử dụng, đơn giản hơn Tensor, khả năng thiết kế và xây dựng mô hình nhanh chóng và hỗ trợ cả CPU và GPU. Nó hỗ trợ thiết kế mạng nơ-ron tích chập, mạng nơ-ron hồi quy, hoặc cả hai giúp việc xây dựng mô hình học trở nên đơn giản và hiệu quả.



Hình 3-4: Các công cụ phát triển cho phát triển học sâu.

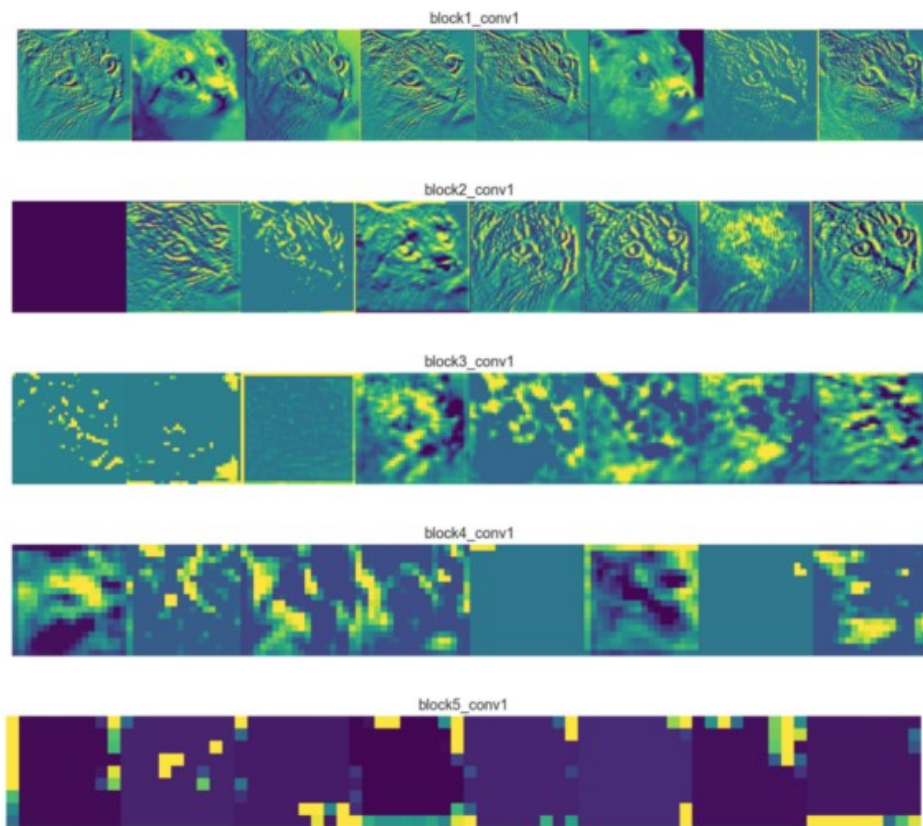
Ngoài Keras, có nhiều công cụ học sâu nổi tiếng khác mà người dùng có thể chọn bao gồm Theano, Torch7, và TensorFlow (xem Hình 3.4). Mỗi công cụ trong lĩnh vực học sâu đều có các ưu điểm riêng, phù hợp với các nhu cầu sử dụng khác nhau. Theano là một công cụ phát triển học sâu truyền thống, được thiết kế để tối ưu hóa tính toán số học trên các cấu trúc dữ liệu đa chiều. Nó phù hợp cho những người muốn can thiệp sâu vào thuật toán và tối ưu hóa mã. Tuy nhiên, việc lập trình từ đầu và tìm kiếm tài liệu hướng dẫn có thể gặp khó khăn. Torch7 là một công cụ phát triển học sâu truyền thống, có nhiều điểm tương đồng với Theano, và được sử dụng phổ biến trong các dự án của Facebook AI. Torch7 dùng ngôn ngữ Lua và mặc dù có nhiều dự án mã nguồn mở sử dụng nó tuy nhiên tài liệu hỗ trợ không phong phú. Công cụ học sâu nổi tiếng TensorFlow do Google phát triển, không chỉ hỗ trợ học sâu mà còn bao gồm các công cụ cho học tăng cường và nhiều thuật toán khác. Mặc dù không hoàn toàn mã nguồn mở, TensorFlow có sự ổn định và hỗ trợ liên tục từ Google. Nó cung cấp nhiều công cụ hỗ trợ quan trọng như công cụ gỡ lỗi và thường đạt hiệu quả tốt trong các bài kiểm tra.

Sau khi đánh giá và so sánh các công cụ phát triển, em đã quyết định sử dụng Keras cho việc triển khai ứng dụng vì những điểm mạnh của nó trong bài toán nhận dạng mẫu vật: Keras là một công cụ phát triển nổi bật trong việc xử lý ảnh, giúp người dùng tinh chỉnh các mô hình một cách dễ dàng để phù hợp với nhu cầu của bài toán mà không cần phải bắt đầu từ đầu.

3.6 Ví dụ về các đặc trưng được phân tích từ mạng VGG16.

Trong mạng CNN đặc biệt là các mạng như VGG16, các lớp tích chập đầu tiên thường được xem như công cụ phân tích các đặc trưng cơ bản từ ảnh đầu vào. Điều này giúp mạng có khả năng học các đặc trưng chung và cấp thấp như các cạnh, góc, hoặc các đốm ảnh màu, trước khi diễn giải các đặc trưng phức tạp hơn tại các lớp cuối cùng để thực hiện các công việc phân loại ảnh. Điều này có thể được quan sát

trong Hình 3.6, với các hình ảnh minh họa các đặc trưng phổ biến của các lớp tích chập trong mạng VGG16 [13].



Hình 3-5: Hình ảnh thể hiện các đặc trưng do mạng VGG16 phân tích.

3.7 Cài đặt cấu trúc mạng huấn luyện VGG16.

Các mô hình và điều chỉnh cho mạng VGG16 cùng các mạng CNN khác được Keras thiết kế để đảm bảo định nghĩa mô hình học sâu rõ ràng và dễ tiếp cận qua cấu trúc văn bản thuần. Chúng ta có thể xem các ví dụ về cách định nghĩa lớp trong mạng CNN, bao gồm lớp dữ liệu đầu vào và lớp tích chập, như được trình bày trong Hình 3.5:

```

traingen = train_generator.flow_from_directory(train_data_dir,
                                              target_size=(224, 224),
                                              class_mode='categorical',
                                              classes=class_subset,
                                              subset='training',
                                              batch_size=BATCH_SIZE,
                                              shuffle=True,
                                              seed=42)

validgen = train_generator.flow_from_directory(train_data_dir,
                                              target_size=(224, 224),
                                              class_mode='categorical',
                                              classes=class_subset,
                                              subset='validation',
                                              batch_size=BATCH_SIZE,
                                              shuffle=True,
                                              seed=42)

```

Hình 3-6: Cách Keras triển khai định nghĩa lớp trong mạng CNN.

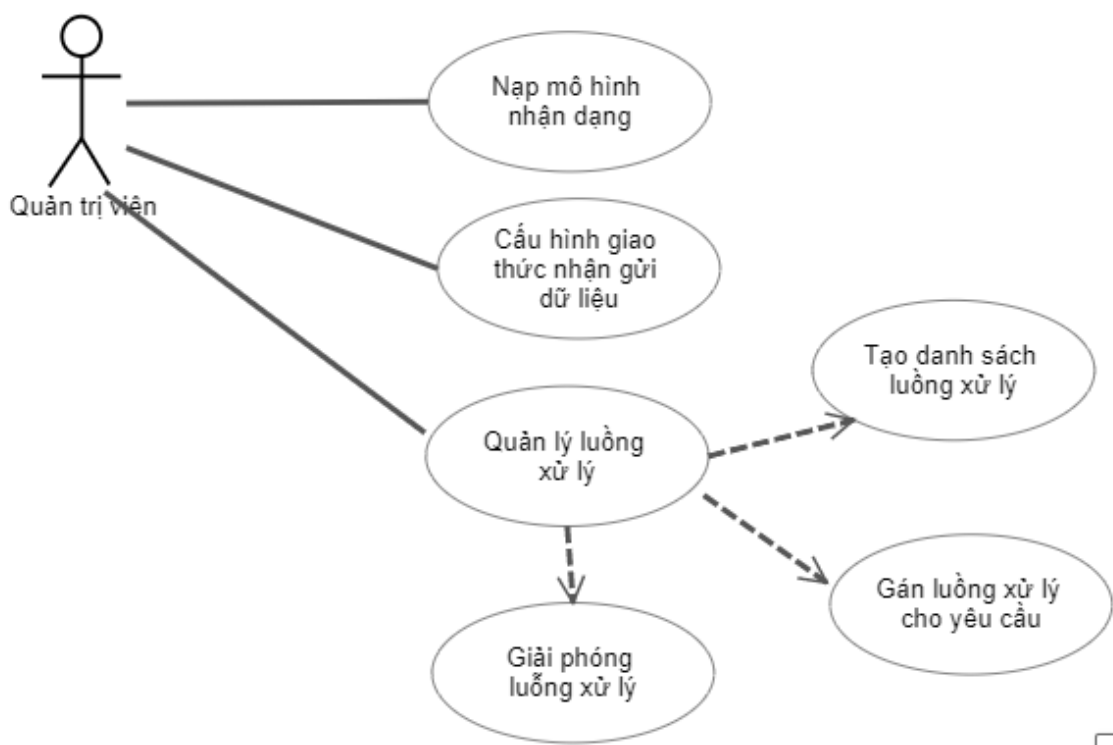
Các tham số của từng lớp trong mạng huấn luyện đã được cài đặt sẵn trong mô hình, và có thể điều chỉnh để phù hợp với yêu cầu của bài toán mà không cần thay đổi cấu trúc mạng cơ bản. Đối với việc điều chỉnh mô hình đã được huấn luyện trước, các tham số quan trọng cần chú ý là:

- Dữ liệu hình ảnh đầu vào: là một phần quan trọng để xác nhận rằng dữ liệu được chuẩn bị một cách chính xác và phù hợp cho việc huấn luyện. Trong công cụ Keras thông tin ảnh đầu vào được định nghĩa thông qua việc cấu hình lớp “train” giúp công cụ có thể thực hiện điều chỉnh kích thước đầu vào để tương thích với mạng tích chập (CNN), tiền xử lý như thay đổi tỉ lệ, cắt ảnh và đối xứng gương là các kỹ thuật cơ bản để tăng cường bộ dữ liệu hình ảnh trong huấn luyện.
- Thông số tỷ lệ học: Tỷ lệ học được thiết lập cho từng lớp và trong các trường hợp huấn luyện từ đầu, tỷ lệ học thường giống nhau ở các lớp với giá trị hợp lý để đảm bảo các đặc trưng được tính toán chính xác. Trong trường hợp điều chỉnh mô hình đã huấn luyện trước, cần giảm tỷ lệ học ở các lớp tích chập để đảm bảo rằng mô hình không bị mất đi các thông tin đặc trưng đã được học. Điều này đảm bảo mô hình học tập hiệu quả và có khả năng tổng quát tốt trên dữ liệu mới, đồng thời tăng cường tỷ lệ học tại các lớp kết nối đầy đủ cuối mạng, giúp mô hình nhanh chóng hội tụ và đạt kết quả tốt hơn.

- Số lượng lớp đầu ra để nhận dạng: được xác định trong lớp cuối cùng của mạng là lớp kết nối đầy đủ, và phản ánh số lớp cần nhận dạng cụ thể. Đối với bài toán nhận dạng mẫu vật ma túy, thông số này sẽ là 7.

3.8 Các thành phần phía máy chủ.

Chương trình phía máy chủ được tổ chức thành nhiều thành phần với các nhiệm vụ khác nhau bao gồm khởi tạo mô hình nhận dạng, quản lý mô hình, thiết lập giao thức giao tiếp máy khách và máy chủ, và xử lý đa luồng để đáp ứng nhiều yêu cầu đồng thời được mô tả chi tiết tại Hình 3.7.



Hình 3-7: Biểu đồ mô tả các thành phần phía máy chủ.

Mô tả các ca sử dụng trong thành phần phía máy chủ:

Tải mô hình nhận dạng:

- Mục tiêu: Tải vào hệ thống mô hình nhận dạng phiên bản mới nhất được xuất ra bởi mô đun bộ huấn luyện mô hình, phục vụ cho việc nhận dạng trong ứng dụng máy khách gửi lên.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Người quản trị hệ thống.
 - Mô tả tổng quát: Để hệ thống thực hiện phân tích và nhận dạng mẫu vật ma túy trong ảnh từ ứng dụng máy khách, mô hình cần phải được tải vào hệ thống. Việc tải mô hình này chỉ diễn ra một lần tại

khởi đầu phiên chạy của máy chủ nhằm tiết kiệm tài nguyên và đảm bảo hệ thống hoạt động ổn định.

- Kịch bản chính: Khi chương trình máy chủ được khởi động, hệ thống sẽ bắt đầu quá trình khởi tạo. Hàm khởi tạo tự động gọi lệnh để tải mô hình nhận dạng vào bộ nhớ hoặc hệ thống lưu trữ. Mô hình nhận dạng được tải từ mô đun huấn luyện mô hình, nơi mà mô hình đã được huấn luyện và xuất ra phiên bản hoàn chỉnh nhất.
- Kịch bản phụ: Nếu quá trình tải mô hình nhận dạng không thành công, chương trình máy chủ sẽ cần thông tin về lỗi xảy ra. Các lỗi có thể xảy ra bao gồm: Tập tin định dạng triển khai của mô hình không hợp lệ, mô hình nhận dạng không tương thích với phiên bản hiện tại của chương trình máy chủ, hoặc các sự cố khác liên quan đến việc tải dữ liệu mô hình.
- Các yêu cầu chi tiết: Phiên bản mô hình nhận dạng là phiên bản cập nhật hoàn chỉnh nhất. Cấu hình hệ thống phải chính xác và hợp lệ.
- Điều kiện trước: Các thư viện yêu cầu cho chương trình máy chủ đã được cài đặt đầy đủ và với các phiên bản chính xác.
- Điều kiện sau: Hệ thống đã hoàn tất các bước chuẩn bị để xử lý và nhận các yêu cầu nhận dạng từ ứng dụng máy khách.

Thiết lập giao thức trao đổi dữ liệu:

- Mục tiêu: Phân phối các yêu cầu từ phía máy khách đến các luồng có sẵn trong danh sách luồng. Khi có yêu cầu đến, nó sẽ được giao cho một luồng đang không xử lý trong danh sách, giúp rút ngắn thời gian chờ đợi và tối ưu hóa tài nguyên.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Thành phần quản lý logic xử lý luồng trong chương trình máy chủ.
 - Mô tả tổng quát: Hệ thống sẽ khởi tạo một số luồng xử lý ngay từ đầu để sẵn sàng tiếp nhận các yêu cầu từ máy khách. Mỗi khi nhận yêu cầu mới từ máy khách, máy chủ sẽ rà soát trạng thái của các luồng trong danh sách đã được khởi tạo sẵn. Nếu có một luồng nào đó đang rảnh, máy chủ sẽ phân phối yêu cầu đó cho luồng này. Luồng sẽ nhận yêu cầu, thực hiện các nghiệp vụ xử lý nhận dạng cần thiết, rồi gửi kết quả về cho máy khách.
- Kịch bản chính: Đây là thành phần chủ yếu trong chương trình máy chủ, quản lý việc phân phối yêu cầu từ phía máy khách tới các luồng đang hoạt động.

- Kịch bản phụ: Nếu tất cả các luồng đang được sử dụng, yêu cầu mới sẽ được xếp vào hàng đợi. Yêu cầu sẽ chờ trong hàng đợi cho đến khi có một luồng sẵn sàng để xử lý.
- Các yêu cầu chi tiết: Có ít nhất một luồng trong danh sách xử lý của máy chủ đang ở chế độ rảnh.
- Điều kiện trước: Các luồng xử lý đã được thiết lập thành công kể từ khi chương trình máy chủ được khởi động.
- Điều kiện sau: Hệ thống xác nhận thành công khi yêu cầu được giao cho một luồng sẵn có.

Kết thúc công việc của luồng xử lý:

- Mục tiêu: Sau khi luồng hoàn tất công việc nhận dạng, cần cập nhật trạng thái của nó, thông báo rằng công việc đã xong và chờ đợi các yêu cầu tiếp theo từ máy chủ.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Thành phần điều phối các hoạt động liên quan đến luồng trong chương trình máy chủ.
 - Mô tả tổng quát: Thành phần này phụ trách điều phối các luồng để xử lý yêu cầu từ máy khách trên máy chủ. Sau khi một luồng đã hoàn thành công việc nhận dạng theo yêu cầu từ máy khách cần thực hiện các bước để cập nhật lại trạng thái của luồng và chuẩn bị sẵn sàng để tiếp nhận và thực hiện các yêu cầu mới tiếp theo từ máy chủ.
- Kịch bản chính: Khi một luồng đã hoàn thành việc xử lý yêu cầu từ máy chủ, cần cập nhật trạng thái của luồng đó từ đang làm việc thành đang rảnh rỗi. Điều này đảm bảo rằng luồng có thể tiếp nhận và giải quyết các yêu cầu tiếp theo từ máy khách một cách hiệu quả.
- Kịch bản phụ: Không có luồng thay thế cụ thể nào được sử dụng.
- Các yêu cầu chi tiết: Đảm bảo rằng khi một luồng hoàn thành việc nhận dạng yêu cầu từ máy khách, nó có thể thông báo về cho thành phần điều phối các hoạt động luồng để cập nhật tình trạng và xử lý các yêu cầu tiếp theo.
- Điều kiện trước: Đảm bảo rằng luồng đã được tạo sẵn và đang trong trạng thái chờ để tiếp nhận và xử lý các yêu cầu từ máy khách.
- Điều kiện sau: Khi luồng xử lý hoàn thành một yêu cầu, thông báo cho thành phần điều phối các hoạt động luồng rằng nhiệm vụ đã được hoàn thành. Thành phần điều phối các hoạt động luồng cập nhật trạng thái của luồng từ trạng thái đang bận sang trạng thái rảnh rỗi, cho phép

luồng này sẵn sàng để nhận yêu cầu mới. Sau đó, thành phần điều phối các hoạt động luồng kiểm tra hàng đợi các yêu cầu từ máy khách. Khi hàng đợi không trống, thành phần điều phối luồng sẽ chỉ định yêu cầu mới cho một luồng đang ở trạng thái rỗi.

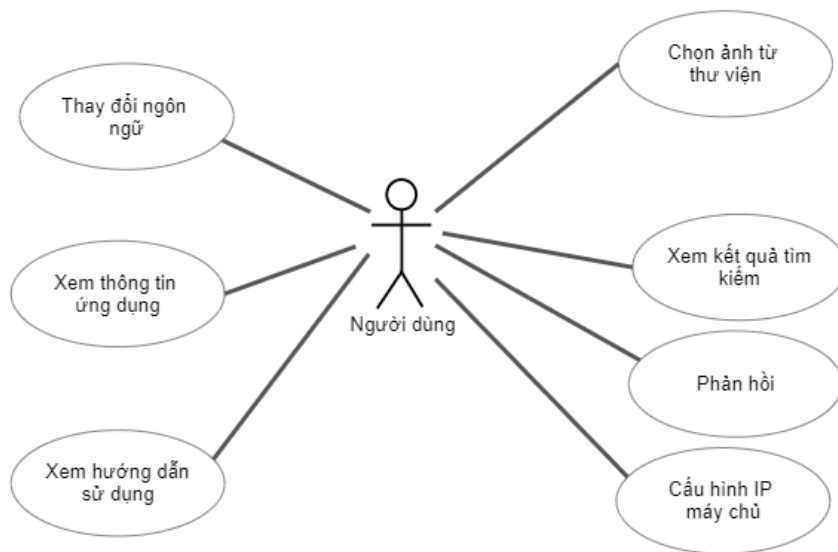
Phân phối yêu cầu từ máy khách đến luồng xử lý:

- Mục tiêu: Tạo sẵn một tập hợp các luồng xử lý mục đích chính là chuẩn bị sẵn các luồng xử lý để đáp ứng việc xử lý các yêu cầu từ phía máy khách một cách hiệu quả, tách riêng nghiệp vụ nhận dạng các luồng xử lý giúp tăng tính hiệu quả và đáp ứng của hệ thống.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Người quản lý hệ thống.
 - Mô tả tổng quát: Trong suốt quá trình vận hành, nhằm tránh tình trạng chậm trễ khi có nhiều yêu cầu đồng thời từ các khách hàng khác nhau, chương trình máy chủ cần áp dụng phương pháp xử lý đa luồng. Mỗi yêu cầu từ phía máy khách được xử lý trên một luồng riêng biệt, không bị ảnh hưởng lẫn nhau và không làm gián đoạn các luồng xử lý khác. Để thuận tiện trong việc quản lý và ngăn ngừa hiện tượng tràn bộ nhớ do việc quản lý luồng không hiệu quả, chương trình chuẩn bị sẵn một tập hợp các luồng xử lý và sử dụng cờ trạng thái để phân phối nhiệm vụ cũng như giải phóng các luồng khi không còn cần thiết.
- Kịch bản chính: Người quản lý khởi động chương trình máy chủ, và hàm khởi tạo của chương trình sẽ tự động thực hiện các lệnh để thiết lập danh sách các luồng xử lý. Số lượng luồng xử lý được xác định dưới dạng hằng số trong tệp cấu hình chung của hệ thống.
- Kịch bản phụ: Chương trình máy chủ thông báo không thể thiết lập các luồng xử lý một cách thành công.
- Các yêu cầu chi tiết: Không.
- Điều kiện trước: Các thư viện cần thiết cho chương trình máy chủ đã được cài đặt đầy đủ và đúng phiên bản được khuyến cáo. Chương trình máy chủ đã hoàn thành các bước tải mô hình được huấn luyện và thiết lập giao thức gửi, nhận dữ liệu.
- Điều kiện sau: Thông báo tạo thành công danh sách các luồng xử lý.

3.9 Ứng dụng máy khách.

Ứng dụng máy khách trên nền website là một thành phần trong hệ thống có nhiệm vụ thu thập dữ liệu hình ảnh đầu vào để phân loại, bao gồm các chức năng

chính dưới đây: Người dùng chọn ảnh từ thư mục của thiết bị hoặc từ các nền tảng khác như mạng xã hội để đưa vào nhận dạng. Sau khi gửi ảnh lên máy chủ để xử lý nhận dạng, ứng dụng sẽ hiển thị thông tin tìm kiếm hoặc thông tin liên quan được trả về từ máy chủ. Bên cạnh đó, ứng dụng còn bao gồm các chức năng bổ sung như: người dùng có thể phản hồi về kết quả nhận dạng, đánh giá tính chính xác của kết quả hoặc cung cấp ý kiến phản hồi khác, cung cấp thông tin chi tiết về ứng dụng và hướng dẫn sử dụng nhằm giúp người dùng tìm hiểu và áp dụng hiệu quả, cùng với việc cấu hình địa chỉ IP của máy chủ mà ứng dụng sẽ gửi yêu cầu tới.



Hình 3-8: Sơ đồ ca sử dụng phía máy khách.

Hình 3.9 mô tả ca sử dụng phía máy khách, tiếp theo chúng ta sẽ đi vào phần đặc tả chi tiết của hai ca sử dụng:

Chọn ảnh trong thư mục:

- Mục tiêu: Chọn một ảnh đã được lưu trong thư viện trên máy để gửi đến máy chủ nhằm yêu cầu nhận dạng và tìm kiếm mẫu vật.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Người dùng ứng dụng.
 - Mô tả tổng quát: Khi người dùng cần nhận dạng và tìm kiếm một loại mẫu vật cụ thể, chẳng hạn như ma túy, nhưng không có kết nối mạng vào thời điểm đó, họ có thể chụp và lưu ảnh mẫu vật vào thư mục ảnh của thiết bị và thực hiện nhận dạng sau khi có kết nối mạng.
- Kịch bản chính: Sau khi người dùng lựa chọn một bức ảnh từ thư mục ảnh trên thiết bị để tiến hành nhận dạng, ứng dụng sẽ mã hóa và nén dữ

liệu ảnh. Người dùng sẽ nhận được thông báo yêu cầu chờ đợi kết quả tìm kiếm nhận dạng.

- Kịch bản phụ: Khi quyền truy cập của ứng dụng bị hạn chế hoặc không thể đọc được thư mục thư viện ảnh của thiết bị, nó cần thông báo cho người dùng về thiếu quyền truy cập và yêu cầu cấp quyền hoặc kiểm tra lại cài đặt.
- Các yêu cầu chi tiết: Không.
- Điều kiện trước: Sau khi gửi dữ liệu lên máy chủ, ứng dụng cần thông báo cho người dùng cần chờ đợi trong khi máy chủ đang xử lý để trả về kết quả tìm kiếm nhận dạng.

Xem kết quả tìm kiếm:

- Mục tiêu: Hiển thị kết quả nhận dạng và tìm kiếm nhận được từ chương trình máy chủ. Kết quả hiển thị cho người dùng bao gồm 10 kết quả tìm kiếm được trả về được sắp xếp theo độ tương đồng từ cao đến thấp.
- Người thực hiện, Mô tả tổng quát:
 - Người thực hiện: Ứng dụng phía máy khách.
 - Mô tả tổng quát: Sau khi nhận được yêu cầu từ máy khách, máy chủ sẽ thực hiện tính toán trên luồng xử lý và gửi kết quả về ứng dụng máy khách, và ứng dụng sẽ hiển thị kết quả này cho người dùng.
- Kịch bản chính: Hiển thị cho người dùng bao gồm 10 kết quả tìm kiếm được trả về được sắp xếp theo độ tương đồng từ cao đến thấp.
- Kịch bản phụ: Xử lý ngoại lệ trong hệ thống như việc nhận dữ liệu không đúng định dạng hoặc bị thiếu thông tin.
- Các yêu cầu chi tiết: Không.
- Điều kiện trước: Sau khi gửi ảnh lên máy chủ, ứng dụng máy khách cần duy trì trạng thái đợi kết quả từ máy chủ.
- Điều kiện sau: Sau khi nhận được kết quả từ máy chủ, ứng dụng cần hiển thị thông tin nhận dạng và tìm kiếm cho người dùng.

Để đảm bảo độ chính xác của kết quả nhận dạng từ mô hình, các ảnh đầu vào phải đáp ứng các yêu cầu sau, nhằm duy trì mức độ tương đồng đủ cao với bộ dữ liệu huấn luyện:

1. Ảnh đầu vào chỉ nên chứa một loại mẫu vật ma túy duy nhất để đảm bảo rằng mô hình nhận dạng không bị phân tán giữa nhiều lớp đối tượng khác nhau.

2. Mẫu vật trong ảnh cần chiếm một tỉ lệ nhất định so với khung ảnh. Nếu mẫu vật quá nhỏ, khả năng tính toán đặc trưng của mô hình có thể bị giảm sút, dẫn đến nhầm lẫn giữa mẫu vật và nền..
3. Mẫu vật trong ảnh không được che khuất quá nhiều bởi các vật thể khác. Nếu bị che lấp, khả năng tính toán đặc trưng không chính xác dẫn đến kết quả nhận dạng của mô hình có thể bị giảm.
4. Ảnh đầu vào phải có độ nét cao và điều kiện ánh sáng đầy đủ để đảm bảo rằng mô hình có thể xử lý và nhận dạng mẫu vật chính xác. Ảnh bị mờ hoặc thiếu sáng có thể làm giảm độ chính xác của nhận dạng.

3.10 Thực nghiệm và kết quả.

3.10.1 Dữ liệu và môi trường thực nghiệm.

Trong luận văn này, em sử dụng tập dữ liệu mẫu vật ma túy được lấy từ Viện khoa học hình sự. Các ảnh được thu thập đa dạng nguồn từ các tỉnh thành khác nhau như Hà Nội, Hải Phòng, Quảng Bình nhằm tăng tính khách quan cho vấn đề phân loại và tính toán độ tương tự hình ảnh. Tổng số 5386 ảnh với 7 nhãn được phân bố như sau:

Bảng 3-1: Thống kê về tập dữ liệu thực nghiệm.

Nhãn	Số lượng
Da cam MY	1080
Đầu chuột màu xám	721
Hồng tròn trái tim	689
Mào gà màu vàng	1129
Hồng tròn chữ M	427
Hồng tròn ô van	560
Hồng tròn	780



Hình 3-9: Một số hình ảnh trong tập dữ liệu.

Các ảnh được phân bố với tỉ lệ 80% để huấn luyện, 20% để kiểm chứng mô hình. Các thực nghiệm được thực hiện trên 2 hệ thống, dành cho 2 loại tác vụ khác nhau.

1. Môi trường thực hiện quá huấn luyện mô hình sử dụng Google Colab:
 - CPU: 1x Single core hyper threaded Xeon Processor 2.3Ghz.
 - GPU: 1x Tesla K80, 12GB GDDR5 VRAM.
 - RAM: 13GB.
 - Disk: 30GB.
2. Môi trường thực hiện quá trình nhận dạng và tìm kiếm:
 - CPU: Intel Core i5-4200H (2 cores, 4 threads) 2.8Ghz.
 - GPU: Nvidia GTX 1080, 8GB GDDR5 VRAM.
 - RAM: 16GB DDR4.
 - Ổ cứng: SSD 128GB.

3.10.2 Kết quả và đánh giá.

Các thực nghiệm được thực hiện sử dụng thư viện Keras. Với mạng VGG16 được Keras cung cấp, độ chính xác được tính toán và so sánh trong 100 vòng (epochs) và mô hình có độ chính xác tốt nhất được lưu lại để làm kết quả thực nghiệm. Đối với hoạt động tính toán véc tơ đại diện cho các ảnh nhằm thực tìm kiếm thông tin mẫu vật ảnh tương đồng, tất cả các ảnh trong tập dữ liệu được đưa qua mô hình mạng đã huấn luyện. Các véc tơ đại diện thu thập được tại lớp kết nối đầy đủ cuối cùng sẽ được lưu lại sử dụng trong quá trình tìm kiếm sau này. Các tham số được thử nghiệm để tối ưu trong việc huấn luyện bao gồm kích thước bộ lọc, số vòng huấn luyện, tốc độ học, thuật toán tối ưu hóa. Các tham số trên được thực hiện tối ưu qua các thực nghiệm và bảng 3-2 cho thấy kết hợp tốt nhất của các tham số.

Bảng 3-2: Giá trị tham số tối ưu

Tham số	Giá trị tối ưu
Kích thước bộ lọc	3x3
Số vòng huấn luyện	100
Tốc độ học	0,001
Thuật toán tối ưu	Adam

Với môi trường thực nghiệm như trình bày ở trên, kết quả phân loại mẫu vật đạt được độ chính xác tổng thể là 84.57%, trong đó loại ma túy da cam MY có độ chính xác tốt nhất (93.57%) và loại ma túy hồng tròn có độ chính xác thấp nhất (66.83%).

Bảng 3-3 cho thấy kết quả phân loại chi tiết cho từng loại ma túy:

Bảng 3-3: Kết quả phân loại

Nhãn	Độ chính xác
Da cam MY	93.57%
Đầu chuột màu xám	81.82%
Hồng tròn trái tim	80.2%
Mào gà màu vàng	90.1%
Hồng tròn chữ M	83.2%
Hồng tròn ô van	81.9%
Hồng tròn	66.83%

Trong luận văn này, em thực hiện đánh giá tìm kiếm mẫu vật theo phương pháp lấy ngẫu nhiên 100 hình ảnh mẫu vật ma túy và dùng làm ảnh đầu vào cho quá trình tìm kiếm. Thu thập 10 kết quả đầu tiên của mỗi ảnh đầu vào, tiến hành đánh giá chủ quan theo thang điểm 0-100 về độ tương tự của nó với ảnh đầu vào. Các kết quả từ 80-100 được xem là tương đồng và nhỏ hơn 80 được xem là không tương đồng. Độ chính xác được tính là tổng số ảnh được đánh giá tương đồng trên tổng số ảnh thu được

$$\text{Độ chính xác} = \frac{\{\text{Số ảnh tương đồng}\}}{\{\text{Tổng số ảnh thu được}\}}$$

Hình 3-11 cho thấy ví dụ về kết quả tìm kiếm với hình ảnh mào gà màu vàng. Với kết quả trong hình 3-11 có thể đánh giá chủ quan là có 8 ảnh tương đồng và 2 ảnh không tương đồng trong 10 ảnh đầu tiên của kết quả tìm kiếm. Kết quả cuối cùng cho thấy độ chính xác tìm kiếm theo đánh giá chủ quan trên toàn bộ 100 mẫu là 75%.

The screenshot shows a web application interface for image search and comparison. It includes a search bar, a file upload section, a comparison tool, and a results table.

Search and Upload Section:

- Chọn ảnh nhận diện: Chon tệp Mau5_7.png
- Tỉ lệ so khớp: 80
- Loại: Viên nén ma túy
- Buttons: Tìm kiếm, Thêm thông tin

Comparison Section:

- Chọn ảnh so sánh: Viên nén ma túy
- Ảnh so sánh: (Image placeholder)
- Buttons: Tìm kiếm

Results Section:

Tổng số bản ghi: 89

#	Ảnh	Thông tin	Đường dẫn	Loại	GD/SS	Thông tin người nhập	Thao tác
1		mao_ga_mau_vang	2_20231019082139978129.png	Viên nén ma túy	Ảnh so sánh	test	Chỉnh sửa Xóa
2		dau_chuot_mau_trang	3_20231019082145031696.png	Viên nén ma túy	Ảnh so sánh	3	Chỉnh sửa Xóa

Hình 3-10: Giao diện tìm kiếm và trả kết quả.

Các kết quả nhận được cho thấy sự khả quan của phương pháp phân loại và tìm kiếm hình ảnh mẫu vật được áp dụng trên tập dữ liệu tương ứng. Mặc dù có sự chênh lệch giữa kết quả phân loại của các loại mẫu vật ma túy khác nhau, nhưng kết quả trung bình cũng như kết quả riêng của hầu hết các loại ma túy là có nhiều tiềm năng trong khoảng chấp nhận được của các ứng dụng phân loại ảnh. Kết quả tìm kiếm cũng có độ chính xác khá tốt, dù vậy mới trải qua đánh giá mang tính chủ quan. Về thời gian chạy, việc huấn luyện mô hình phân loại mất khoảng 60 phút, trong khi thời gian để đưa toàn bộ hơn 5386 ảnh qua mô hình để thu thập véc tơ đại diện là 70 phút. Thời gian để thực hiện tìm kiếm từ khi cung cấp ảnh đầu vào đến khi trả về kết quả là 10 giây sau khi đã có mô hình và có các véc tơ đại diện của các ảnh trong tập ảnh để so sánh.

3.11 Kết luận chương 3

Trong chương này, luận văn đã trình bày chi tiết quá trình phân tích, thiết kế và cài đặt hệ thống nhận dạng và tìm kiếm các mẫu vật. Hệ thống được xây dựng

trên nền tảng phát triển Keras cho chức năng nhận dạng và sử dụng phương pháp so sánh độ tương đồng đặc trưng mẫu vật cho chức năng tìm kiếm. Kết quả thực nghiệm với bộ dữ liệu kiểm thử và trong ứng dụng thực tế đều đạt hiệu quả tốt, chủ yếu nhờ vào việc giới hạn số lượng mẫu vật cần nhận dạng còn 7 loại, một số lượng rất ít so với tổng số mẫu vật mà tủy đã được phát hiện. Hệ thống nhận dạng mẫu vật vẫn cần nhiều cải tiến, đặc biệt trong việc mở rộng phạm vi mẫu vật cũng như nâng cao kích thước và chất lượng của cơ sở dữ liệu hình ảnh huấn luyện. Để cải thiện độ chính xác của mô hình nhận dạng trong tương lai, tôi đề xuất thử nghiệm và đánh giá các mô hình mạng học sâu đã được huấn luyện trước, đặc biệt là những mạng đã đạt được kết quả xuất sắc trong các cuộc thi nhận dạng hình ảnh quy mô lớn do ImageNet tổ chức hàng năm như VGG 19, GoogleNet và Microsoft ResNet (2015).

KẾT LUẬN

I. TÓM TẮT VỀ CÁC KẾT QUẢ ĐẠT ĐƯỢC.

Sau thời gian nghiên cứu và triển khai, tập trung vào việc nhận dạng và tìm kiếm mẫu vật, luận văn đã đạt được các kết quả chủ yếu như sau:

- Đánh giá được các ưu điểm và nhược điểm của các phương pháp học sâu thông qua khảo sát.
- Đã nghiên cứu phát triển một hệ thống nhận dạng và tìm kiếm mẫu vật nói chung, cũng như hệ thống nhận dạng và tìm kiếm thông tin về mẫu vật ma túy cụ thể.
- Hiểu được các phương pháp cơ bản trong quy trình tiền xử lý, trích xuất đặc trưng, và phân loại để giải quyết vấn đề nhận dạng mẫu vật.

Các kết quả trên không chỉ góp phần xây dựng cơ sở lý thuyết mà còn mang lại những ứng dụng tiềm năng trong thực tiễn, đặc biệt trong việc cải thiện hiệu quả quản lý kho tang vật.

II. NHỮNG VẤN ĐỀ CHƯA ĐƯỢC GIẢI QUYẾT BỞI LUẬN VĂN

Do thời gian nghiên cứu luận văn hạn chế nên nhiều vấn đề vẫn còn tồn tại và chưa thể được giải quyết:

- Chương trình mới dừng lại ở mức thử nghiệm, chưa thật sự là một chương trình hoàn chỉnh, đầy đủ các tính năng.
- Việc nhận dạng và tìm kiếm mẫu vật chưa tập trung được nhiều vào các loại mẫu vật khác nhau và hậu xử lý.
- Việc thử nghiệm thuật toán nhận dạng mới chỉ tiến hành trên mẫu vật nhận dạng cơ bản, có hình dạng đơn giản.

III. HƯỚNG PHÁT TRIỂN

Trong tương lai, các hướng nghiên cứu và phát triển sẽ tập trung vào:

1. **Hoàn thiện hệ thống:** Tiếp tục phát triển hệ thống thành một ứng dụng hoàn chỉnh với đầy đủ các tính năng cần thiết, đáp ứng được các yêu cầu thực tiễn trong quản lý và tìm kiếm mẫu vật.
2. **Mở rộng phạm vi ứng dụng:** Nâng cao khả năng nhận dạng và tìm kiếm với nhiều loại mẫu vật khác nhau, đặc biệt là các mẫu vật có độ phức tạp cao hoặc sự biến đổi lớn về đặc trưng hình ảnh.

3. **Cải tiến thuật toán học sâu:** Áp dụng các thuật toán học sâu tiên tiến hơn, tối ưu hóa hiệu suất nhận dạng và cải thiện khả năng phân loại trong môi trường dữ liệu lớn.
4. **Ứng dụng thực tế:** Tiến tới triển khai thử nghiệm thực tế tại các cơ quan quản lý tang vật, từ đó thu thập thêm dữ liệu thực nghiệm để cải thiện hiệu quả hoạt động của hệ thống.

MỤC LỤC THAM KHẢO

- [1] N. T. Đỗ and B. V. Phạm, Giáo trình xử lý ảnh, Thái Nguyên, 2007.
- [2] Y. P. Y. L. a. S. Y. Hanjiang Lai, "Simultaneous Feature Learning and Hash Coding with Deep," 2015.
- [3] N. D. a. B. Triggs, "Histograms of Oriented Gradients for Human Detection," p. 886–893, 2005.
- [4] F. B. Y. L. a. J. Y-Lan Boureau, "Learning Mid-Level Features for Recognition, In," Proc. CVPR (2010).
- [5] V. S. U. S. a. S. B. Gal Chechik, "Large Scale Online Learning of Image Similarity Through," *Journal of Machine Learning Research* 11, p. 1109–1135, 2010.
- [6] Andrej Karpathy. , CS231n Convolutional Neural Networks for Visual Recognition - Transfer Learning, <https://cs231n.github.io/transfer-learning/>.
- [7] F. L., S. S., L. R. and W. S. , Classification of kiwifruit grades based, Sensors (Switzerland), 2016.
- [8] Y.-Y. Ou, A.-C. Tsai, X.-P. Zhou and J.-F. Wang, Automatic drug pills detection based on enhanced feature pyramid network and convolution neural networks, 2020.
- [9] GilPress., Visually Linking AI, Machine Learning, Deep Learning, Big Data, <https://whatsthebigdata.com/2016/10/17/visually-linking-ai-machine-learning-deep-learning-big-data-and-data-science/>, 2016.
- [10] T. T. Nguyễn , Deep Learning cơ bản, 2019.
- [11] D. V. and V. F., A guide to convolution arithmetic for deep, 2016.
- [12] S. Hijazi, K. Rishi and C. Rowen, Using Convolutional Neural Networks, IP Group, Cadence.
- [13] A. Karpathy., CS231n Convolutional Neural Networks for Visual Recognition . <http://cs231n.github.io/classification/>.
- [14] Olshausen, Bruno A., and David J. Field. 1996. "Emergence of simple-cell receptive field properties by learning a sparse code for natural images." *Nature* 381.6583:607-609.
<http://www.nature.com/nature/journal/v381/n6583/abs/381607a0.html>

- [15] Hinton, G. E. and Salakhutdinov, R. R. 2006. "Reducing the dimensionality of data with neural networks." *Science* vol. 313 no. 5786 pp. 504-507.
<http://www.sciencemag.org/content/suppl/2006/08/04/313.5786.504.DC1>.
- [16] Hinton, Geoffrey E. 2009. "Deep belief networks." *Scholarpedia*, 4(5):5947.
http://www.scholarpedia.org/article/Deep_belief_networks
- [17] Hinton, G. E. and Salakhutdinov, R. R. 2006. "Reducing the dimensionality of data with neural networks." *Science* vol. 313 no. 5786 pp. 504-507.
<http://www.sciencemag.org/content/suppl/2006/08/04/313.5786.504.DC1>.
- [18] Hinton, Geoffrey E. 2009. "Deep belief networks." *Scholarpedia*, 4(5):5947.
http://www.scholarpedia.org/article/Deep_belief_networks