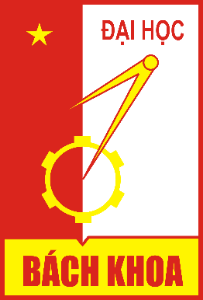
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

VIỆN ĐIỆN TỬ VIỄN THÔNG

----- □  □ -----



**BÀI TẬP LỚN: XỬ LÝ ẢNH SỐ**

***Đề tài:***

**Xây dựng chương trình tự động**

**phân loại ong mang phấn hoa**

Giảng viên hướng dẫn: TS. Lê Thị Lan

Nhóm sinh viên thực hiện: Nhóm 7

|  |  |
| --- | --- |
| Họ và tên | MSSV |
| Ngô Văn Minh | 20193012 |
| Phạm Đức Mạnh | 20193002 |
| Lê Hoàng Tùng | 20193179 |
| Viên Đình Anh | 20192700 |

Hà Nội, 7 năm 2023

MỤC LỤC

[CHƯƠNG 1. GIỚI THIỆU CHUNG 5](#_Toc140964109)

[1.1. Giới thiệu về đề tài 5](#_Toc140964110)

[1.2. Mục tiêu của đề tài 5](#_Toc140964111)

[1.3. Phương pháp đề xuất 6](#_Toc140964112)

[1.4. Phân chia nhiệm vụ trong nhóm 6](#_Toc140964113)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 7](#_Toc140964114)

[2.1. Tìm hiểu về bài toán phân loại ảnh 7](#_Toc140964115)

[2.2. Đặc trưng của ong mang phấn và ong không mang phấn 8](#_Toc140964116)

[a) Đặc trưng màu sắc 8](#_Toc140964117)

[b) Đặc trưng hình dáng, vị trí túi phấn 8](#_Toc140964118)

[2.3. Đề xuất lựa chọn đặc trưng 9](#_Toc140964119)

[a) Đối với đặc trưng màu sắc: 9](#_Toc140964120)

[b) Đối với đặc trưng hình dạng và vị trí túi phấn: 9](#_Toc140964121)

[2.4. Xây dựng mô hình 9](#_Toc140964122)

[a) Tiền xử lý hình ảnh 9](#_Toc140964123)

[b) Trích xuất đặc trưng 10](#_Toc140964124)

[c) Phân loại bằng SVM 13](#_Toc140964125)

[CHƯƠNG 3. THỰC HIỆN VÀ KẾT QUẢ 16](#_Toc140964126)

[3.1. Tiền xử lý hình ảnh 16](#_Toc140964127)

[a) Giảm kích thước ảnh 16](#_Toc140964128)

[b) Áp dụng bộ lọc Gaussian 16](#_Toc140964129)

[c) Chọn vùng ảnh cần thiết 17](#_Toc140964130)

[3.2. Trích xuất đặc trưng 18](#_Toc140964131)

[a) Trích xuất đặc trưng HOG 18](#_Toc140964132)

[b) Trích xuất đặc trưng color histogram (local) 19](#_Toc140964133)

[3.3. Áp dụng bộ phân loại SVM, KNN, NB 21](#_Toc140964134)

[3.4. Kết quả và đánh giá 23](#_Toc140964135)

[a) Kết quả về độ chính xác 23](#_Toc140964136)

[b) Thử nghiệm mô hình 23](#_Toc140964137)

[CHƯƠNG 4. KẾT LUẬN 24](#_Toc140964138)

[4.1. Kết luận 24](#_Toc140964139)

[4.2. Hướng phát triển sau này 24](#_Toc140964140)

DANH MỤC HÌNH VẼ

[Hình 2.1 : Các kênh màu RGB 10](#_Toc140930532)

[Hình 2.2 : Các kênh màu HSV 11](#_Toc140930533)

[Hình 2.3 : Công thức tính tương tự về màu sắc giữa hai ảnh 12](#_Toc140930534)

[Hình 2.4 : Hai ảnh khác nhau nhưng có cùng lược đồ màu 12](#_Toc140930535)

[Hình 2.5 : Phân bố và hình dạng vài biến thể HOG 13](#_Toc140930536)

[Hình 2.6 : Ảnh minh họa hàm phân loại SVM 14](#_Toc140930537)

[Hình 3.1: Hàm giảm kích thước ảnh 16](#_Toc140930538)

[Hình 3.2 : Hàm bộ lọc gaussian 17](#_Toc140930539)

[Hình 3.3: Ảnh sau khi qua bộ lọc gaussian 17](#_Toc140930540)

[Hình 3.4 : Hàm trích xuất đặc trưng HOG 18](#_Toc140930541)

[Hình 3.5 : Hàm trích xuất đặc trưng color histogram 19](#_Toc140930542)

[Hình 3.6 : Hàm tìm kích thước vùng tối ưu 19](#_Toc140930543)

[Hình 3.7 : Bộ phân loại SVM, KNN, NB 21](#_Toc140930544)

# GIỚI THIỆU CHUNG

## Giới thiệu về đề tài

Trên toàn cầu, sự suy thoái của quần thể ong mang mật đã gây ra một loạt vấn đề đáng lo ngại, đặc biệt là trong lĩnh vực nông nghiệp và môi trường. Với vai trò quan trọng trong quá trình thụ phấn của cây trồng và duy trì hệ sinh thái tự nhiên, ong mang mật đóng góp không nhỏ vào sự đa dạng sinh học và sự phát triển bền vững. Để giám sát và nghiên cứu hiệu quả các quần thể ong mang mật, phân loại chính xác ong mang mật từ hình ảnh đã trở thành một phương pháp quan trọng.

Hy vọng rằng nghiên cứu này sẽ đóng góp vào việc bảo vệ và quản lý hiệu quả các quần thể ong mang mật. Kết quả của chúng tôi có thể hỗ trợ các nhà nghiên cứu, chuyên gia và những người quan tâm trong việc theo dõi và giám sát quần thể ong mang mật, từ đó đưa ra các biện pháp bảo vệ và duy trì sự đa dạng sinh học của chúng. Hi vọng rằng công trình này sẽ góp phần vào nghiên cứu và ứng dụng công nghệ hình ảnh trong việc bảo vệ và bảo tồn ong mang mật, góp phần quan trọng vào sự phát triển bền vững của hệ sinh thái tự nhiên và nông nghiệp.

## Mục tiêu của đề tài

* Tìm hiểu bài toán phân loại ảnh : nắm vững các phương pháp và thuật toán học máy để hiểu, phân tích và phân loại ảnh dựa trên đặc trưng và nội dung của chúng. Mục tiêu là xây dựng mô hình chính xác và hiệu quả để tự động phân loại ảnh trong các ứng dụng thực tế.
* Tìm hiểu đặc trưng của ong mang phấn và ong không mang phấn : gồm các đặc trưng màu sắc, kết cấu, hình dáng, độ sâu, …
* Đề xuất lựa chọn đặc trưng ảnh sử dụng : Sau khi tìm hiểu đặc trưng của ong mang phấn và ong không mang phấn, bọn em sẽ tìm hiểu ưu nhược điểm của từng đặc trưng, đánh giá các đặc trưng và lựa chọn đặc trưng phù hợp với bài toán.
* Xây dựng mô hình để phân loại ong mang mật: Sau khi lựa chọn đặc trưng ảnh sử dụng, bọn em sẽ xây dựng mô hình để sử dụng đặc trưng đó trong bài toán. Mô hình này sẽ được huấn luyện trên tập dữ liệu đã chuẩn bị để học cách phân loại đúng các hình ảnh ong mang mật.
* Đánh giá hệ thống: Chúng em sẽ đánh giá hiệu suất của hệ thống phân loại bằng cách sử dụng các phương pháp đánh giá chính xác như ma trận nhầm lẫn, độ chính xác, độ phủ và độ đo F1. Qua đó, chúng em sẽ đánh giá khả năng phân loại chính xác của hệ thống và đo lường độ hiệu quả của mô hình đã chọn trong việc phân loại ong mang mật.
* Triển khai hệ thống phân loại: Sau khi xây dựng và đánh giá hệ thống, chúng em sẽ triển khai nó để có thể sử dụng trong các ứng dụng thực tế. Hệ thống phân loại sẽ được thiết kế để nhận đầu vào là một hình ảnh ong mang mật và trả về kết quả phân loại chính xác về loại ong mang mật đó.

## Phương pháp đề xuất

Sử dụng các thuật toán trích xuất đặc trưng cơ bản của xử lý ảnh

* Trích xuật đặc trưng : HOG
* Kĩ thuật phân loại : SVM

Trong đề tài này, chúng tôi tập trung vào việc phát triển một hệ thống phân loại ong mang mật sử dụng hai thuật toán chính: HOG (Histogram of Oriented Gradients) và SVM (Support Vector Machine). HOG là một phương pháp phân tích hình ảnh mạnh mẽ và phổ biến trong việc trích xuất đặc trưng từ hình ảnh. Phương pháp này dựa trên việc tính toán và biểu diễn độ lớn và hướng của các gradient trong ảnh. Bằng cách tạo ra một biểu đồ đặc trưng dựa trên hướng và phân bố của các gradient, HOG giúp chúng ta nắm bắt được các đặc điểm quan trọng của ong mang mật trong hình ảnh. SVM là một thuật toán học máy phân loại được sử dụng rộng rãi. Nó dựa trên việc xác định các đường phân chia tối ưu giữa các lớp dữ liệu khác nhau. SVM có khả năng tìm ra một siêu phẳng tốt nhất để phân chia dữ liệu thành các lớp, từ đó giúp chúng ta phân loại các ảnh ong mang mật một cách chính xác và đáng tin cậy. Bằng cách kết hợp HOG và SVM, chúng tôi nhằm xây dựng một hệ thống phân loại ong mang mật chính xác và hiệu quả. Quá trình phát triển của chúng tôi bao gồm các bước từ việc thu thập và tiền xử lý dữ liệu hình ảnh ong mang mật cho đến việc trích xuất đặc trưng HOG và xây dựng mô hình SVM để phân loại. Chúng tôi sử dụng một tập dữ liệu đa dạng và lớn để huấn luyện và đánh giá hiệu suất của hệ thống.

## Phân chia nhiệm vụ trong nhóm

|  |  |
| --- | --- |
| Ngô Văn Minh *(Leader)* | - Điều phối hoạt động, công việc các thành viên  - Tìm hiểu các kĩ thuật mà bài toán cần sử dụng |
| Phạm Đức Mạnh | - Triển khai code thuật toán HOG, SVM  - Slide báo cáo |
| Lê Hoàng Tùng, Viên Đình Anh | - Tìm hiểu lý thuyết thuật toán color histogram, HOG và SVM  - Code phần color histogram  - Báo cáo |

# CƠ SỞ LÝ THUYẾT

## Tìm hiểu về bài toán phân loại ảnh

Bài toán phân loại ảnh là một bài toán trong lĩnh vực trí tuệ nhân tạo, có mục tiêu phân loại các hình ảnh vào các nhóm tương ứng dựa trên nội dung, đặc trưng hoặc chủ đề của chúng. Đây là một vấn đề quan trọng trong xử lý ảnh và có ứng dụng rộng rãi trong nhiều lĩnh vực như công nghệ thông tin, y tế, thị giác máy tính và ngành quảng cáo.

Bài toán phân loại ảnh đặt ra một số thách thức cần được vượt qua. Dưới đây là một số thách thức chính và các phương pháp giải quyết thường được sử dụng:

* Đa dạng và phức tạp của dữ liệu: Ảnh có thể có độ phân giải, tỷ lệ khung cảnh, màu sắc và chi tiết khác nhau. Các đặc trưng có thể thay đổi tùy thuộc vào ngữ cảnh và vấn đề cụ thể.
* Dữ liệu huấn luyện hạn chế: Để xây dựng một mô hình phân loại ảnh chính xác, cần có một tập dữ liệu huấn luyện đủ lớn và đa dạng. Tuy nhiên, việc thu thập và gán nhãn cho dữ liệu có thể tốn kém và tốn thời gian. Các phương pháp giải quyết bao gồm sử dụng các tập dữ liệu công khai có sẵn và kỹ thuật mở rộng dữ liệu bằng cách áp dụng các phép biến đổi như xoay, thu phóng và lật ảnh để tăng cường độ đa dạng.
* Overfitting: Khi mô hình quá tinh chỉnh cho dữ liệu huấn luyện cụ thể, nó có thể không tổng quát hoá tốt với dữ liệu mới. Để giảm thiểu overfitting, các phương pháp như dropout, regularization và cross-validation được áp dụng để đảm bảo rằng mô hình không chỉ học thuộc lòng dữ liệu huấn luyện mà còn hiểu được các đặc trưng chung.
* Số lượng lớn các lớp phân loại: Trong bài toán phân loại ảnh, có thể có hàng trăm hoặc hàng nghìn lớp phân loại khác nhau. Để đối phó với số lượng lớp lớn, một phương pháp phổ biến là sử dụng mô hình học sâu có kiến trúc nhiều tầng (deep architecture) như mạng nơ-ron tích chập đa tầng (multi-layer CNN) và mạng nơ-ron chồng chéo (stacked neural networks) để xây dựng một mô hình phân loại mạnh mẽ.

Quy trình các bước thông thường để giải quyết bài toán phân loại ảnh :

1. Thu thập dữ liệu: Thu thập một tập dữ liệu đủ lớn và đa dạng các ảnh đại diện cho các lớp phân loại khác nhau. Dữ liệu này có thể được thu thập từ các nguồn công khai hoặc tự tạo.
2. Tiền xử lý dữ liệu: Tiền xử lý dữ liệu bao gồm các bước như điều chỉnh kích thước ảnh, chuẩn hóa giá trị pixel, cắt tỉa hoặc đối xứng ảnh để chuẩn bị dữ liệu cho quá trình huấn luyện.
3. Chia tập dữ liệu: Chia tập dữ liệu thành ba phần: tập huấn luyện, tập xác thực và tập kiểm tra. Tập huấn luyện được sử dụng để huấn luyện mô hình, tập xác thực được sử dụng để điều chỉnh các siêu tham số và tập kiểm tra để đánh giá hiệu suất cuối cùng của mô hình.
4. Trích chọn đặc trưng: Sử dụng các phương pháp trích chọn đặc trưng để tạo ra các biểu diễn số học hoặc đặc trưng từ dữ liệu ảnh ban đầu. Các phương pháp này có thể bao gồm phân tích thành phần chính (PCA), trích xuất đặc trưng bằng mạng nơ-ron sâu hoặc sử dụng các kỹ thuật học sâu như transfer learning.
5. Chọn bộ phân loại: Chọn bộ phân loại phù hợp để phân loại ảnh dựa trên các đặc trưng đã được trích chọn. Các bộ phân loại phổ biến có thể là Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest hoặc mạng nơ-ron.
6. Huấn luyện mô hình: Huấn luyện mô hình bằng cách sử dụng tập dữ liệu huấn luyện và áp dụng bộ phân loại đã chọn. Quá trình này bao gồm việc tối ưu hóa các siêu tham số của bộ phân loại để tạo ra mô hình tốt nhất.
7. Đánh giá và điều chỉnh: Đánh giá hiệu suất của mô hình sử dụng tập xác thực và điều chỉnh các siêu tham số của bộ phân loại để cải thiện hiệu suất.
8. Kiểm tra và đánh giá cuối cùng: Sử dụng tập kiểm tra để đánh giá hiệu suất cuối cùng của mô hình trên dữ liệu mới. Đánh giá bằng các độ đo như độ chính xác (accuracy), độ phủ (recall), độ chính xác dương dương (precision), và F1-score.
9. Tinh chỉnh và triển khai: Nếu kết quả không đạt yêu cầu, có thể tiến hành tinh chỉnh mô hình bằng cách thay đổi kiến trúc, các tham số đầu vào hoặc thêm dữ liệu. Sau khi đạt được mô hình tốt, nó có thể được triển khai và sử dụng để phân loại ảnh trong ứng dụng thực tế.

Các ứng dụng của bài toán phân loại ảnh là rất đa dạng, bao gồm nhận dạng khuôn mặt, phát hiện đối tượng, phân loại sản phẩm, phân loại chủ đề, gợi ý hình ảnh và nhiều hơn nữa. Việc phân loại ảnh tự động giúp tăng tốc độ xử lý dữ liệu, tối ưu hóa quy trình làm việc và cung cấp thông tin quan trọng từ ảnh một cách nhanh chóng và chính xác.

## Đặc trưng của ong mang phấn và ong không mang phấn

### Đặc trưng màu sắc

Ong mang phấn có thêm túi phấn màu sáng, nên tổng quan của hình ảnh ong mang phấn sẽ sáng hơn so với ong không mang phấn.

### Đặc trưng hình dáng, vị trí túi phấn

Vị trí của túi phấn thường được nằm ở bụng, nên ong mang phấn sẽ có hình dạng toàn bộ cơ thể khác với ong không mang phấn.

## Đề xuất lựa chọn đặc trưng

### Đối với đặc trưng màu sắc:

Có thể sử dụng phân đoạn màu sắc để phân tách vùng màu của ong mang phấn và ong không mang phấn.

Áp dụng các phương pháp tính toán màu sắc trung bình hoặc histogram màu sắc trong hệ màu RGB hoặc HSV để so sánh màu trung bình của hai loại ong.

### Đối với đặc trưng hình dạng và vị trí túi phấn:

Sử dụng các kỹ thuật phát hiện đối tượng và phân đoạn hình dạng để tìm và xác định vị trí của túi phấn.

Phân tích tỷ lệ và hình dạng của vùng bụng để phân biệt hai loại ong. Do đó có thể sử dụng phương pháp như HOG (Histogram of Oriented Gradients) để trích xuất đặc trưng hình dạng và vị trí túi phấn.

## Xây dựng mô hình

### Tiền xử lý hình ảnh

**Nhiễu Gauss**: Nhiễu này có được do bản chất rời rạc của bức xạ (hệ thống ghi ảnh bằng cách đếm các photon (lượng tử ánh sáng). Mỗi pixel trong ảnh nhiễu là tổng giá trị pixel đúng và pixel ngẫu nhiên .

**Bộ lọc gaussian :** Xét trên miền tần số thì bộ lọc Gaussian sẽ giảm thành phần tần số cao trong hình ảnh hay nói cách khách bộ lọc Gaussian là bộ lọc thông cao ( chỉ giữ lại thành phần tần số thấp).Trong toán học, việc ứng dụng Gaussian Blur cho một hình cũng chính là tính tích chập (Convolution) hình đó với hàm Gaussian.Một cách trực quan, đây được xem như là phương pháp làm mờ mịn cũng giống như hiệu ứng hình ảnh được đặt dưới một lớp màn trong suốt bị mờ. Nó không giống với trường hợp hình ảnh bị mờ do hậu quả của ống kính bị mất tiêu điểm (out of focus) hay do bóng của đối tượng dưới ánh sáng thường. Nó có thể giúp làm giảm nhiễu (Noise) và mức độ chi tiết (không mong muốn) của hình ảnh.

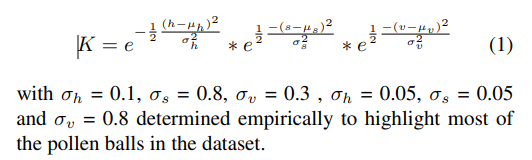
Dưới đây là phương trình hàm Gaussian trong không gian 1 chiều và 2 chiều :

A math equations on a white background

Description automatically generated

Hình 2.1 : Phương trình hàm gaussian trong không gian 1 chiều và 2 chiều

**Áp dụng bộ lọc gaussian vào bài toán :** qua xem xét một số bài báo tương tự về tiền xử lý hình ảnh trong phân loại ong mang phấn, nhóm chúng em nhận thấy áp dụng bộ lọc gaussian với các tham số này đem lại kết quả tốt nhất trong khử nhiễu hình ảnh :



(trích từ bài báo Recognition of Pollen-bearing Bees from Video using Convolutional Neural Network của Ivan F. Rodriguez , Remi M ´ egret , Edgar Acuna˜ , Jose L. Agosto-Rivera . Tugrul Giray.)

### Trích xuất đặc trưng

**Trích xuất đặc trưng màu sắc :**

***Hệ màu RGB :***

RGB là không gian màu rất phổ biến được dùng trong đồ hoạ máy tính và nhiều kĩ thuật số khác. Ý tưởng chính của không gian màu này là sự kết hợp của ba màu sắc cơ bản: màu đỏ (R, Red), xanh lục (G, Green), xanh lơ (B, Blue) để mô tả tất cả các màu sắc khác.

Nếu như một ảnh số được mã hoá bằng 24bit nghĩa là 8bit cho kênh R, 8bit cho kênh G, 8bit cho kênh B thì mỗi kênh này màu sẽ nhận giá trị từ 0-255. Với mỗi giá trị khác nhau của các kênh màu kết hợp với nhau ta sẽ nhận được một màu khác nhau, như vậy ta sẽ có tổng cộng 255255255 = 1.66 triệu màu sắc.

Ví dụ màu đen là sự kết hợp của các kênh màu (R, G, B) với các giá trị tương ứng là (0, 0, 0), màu vàng (255, 255, 0) … Nếu ta dùng 16bit để mã hoá một kênh màu thì dải màu sẽ rất rộng: 3\*2^16…. Một con số rất lớn.

A color wheel with different colors

Description automatically generated

Hình 2.2 : Các kênh màu RGB

***Hệ màu HSV***

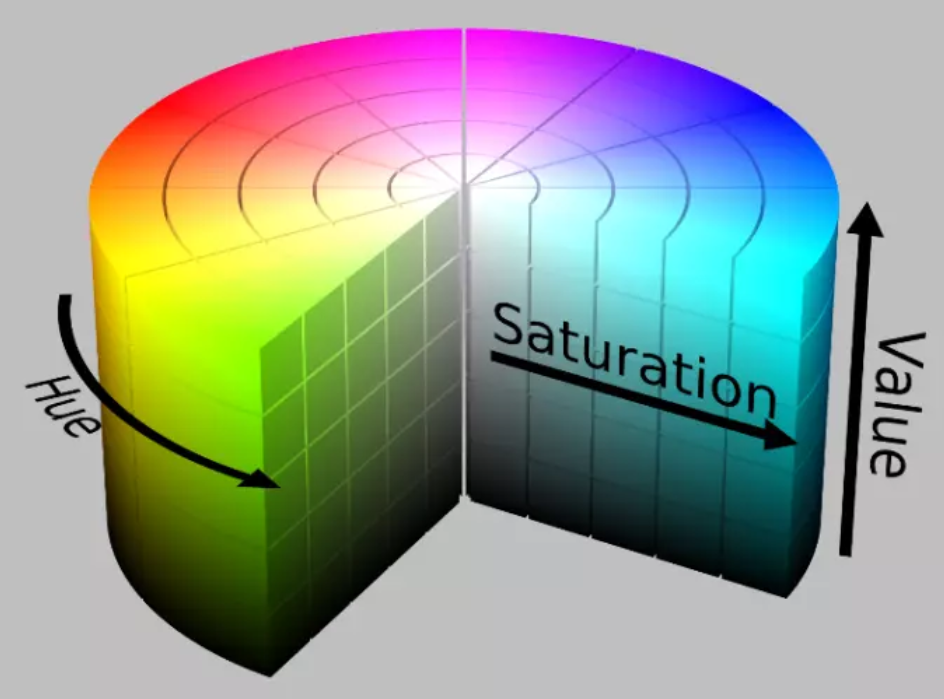
Hệ màu HSV cũng gần tương tự với HSL là không gian màu được dùng nhiều trong lĩnh vực chỉnh sửa ảnh, phân tích ảnh và là một phần của lĩnh vực thị giác máy tính. Hệ không gian này dựa vào ba thông số để mô tả màu sắc:

H = Hue: Màu sắc.

S = Saturation: Độ đậm đặc, sự bão hoà.

V = Value: Giá trị cường độ sáng.

Không gian màu này thường được biểu diễn dưới dạng hình trụ hoặc nón. Theo đó, đi theo hình vòng tròn từ 0 - 360 độ là trường biểu diễn màu sắc (Hue). Trường này bắt đầu từ màu đỏ đầu tiên (red primary) tới màu xanh lục đầu tiên (gree primary) nằm trong khoảng 0 - 120 độ, từ 120 - 240 độ là màu xanh lục tới màu xanh lơ (green primary – blue primary), từ 240 - 360 độ là từ màu đen tới lại màu đỏ.



Hình 2.3 : Các kênh màu HSV

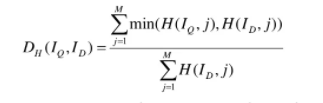
#### Hình 2.1 : Không gian màu HSV trong trích xuất đặc trưng màu sắc

Số bin sử dụng: 64 bins, 128 bins, 256 bins.

Sử dụng 2 dạng histogram:

* Global: Tính toán Histogram trên toàn ảnh
* Local : Tính toán Histogram trên từng vùng

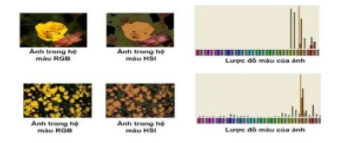
Công thức tính tương tự về màu sắc giữa lược đồ màu của ảnh truy vấn và lược đồ màu của ánh trong CSDL ảnh được định nghĩa:



Hình 2.4 : Công thức tính tương tự về màu sắc giữa hai ảnh

Công thức trên cho ta thấy, tính tương tự về màu sắc được tính bằng phần giao của 2 lược đồ màu ảnh truy vấn và ảnh trong cơ sở dữ liệu ảnh. Kết quả sẽ là một lược đồ màu thể hiện dọ giống nhau giữa 2 ảnh trên.

Tuy nhiên vì lược đồ màu chỉ thể hiện tính phân bố màu toàn cục của ảnh mà không xét đến tính phân bố cục bộ của điểm ảnh nên có thể có 2 ảnh trong rất khác nhau nhưng lại có cùng lược đồ màu:



Hình 2.5 : Hai ảnh khác nhau nhưng có cùng lược đồ màu

Hình 2.3: Phân bố cục bộ của Pixel ản

Để khắc phục được tình trạng này, chúng ta dùng phân hoạch lưới ô vuông trên ảnh. Lược đồ màu của ảnh là không duy nhất.

Trích xuất đặc trưng màu sắc vừa có ưu điểm và nhược điểm:

Ưu điểm:

* Phản ánh phân bố màu sắc trong ảnh.
* Bất biến với phép quay ảnh (không làm méo).
* Bất biến với phép dịch ảnh.

Nhược điểm:

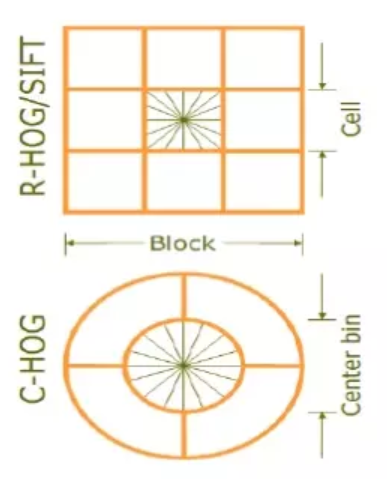
* Không phản ánh tính không gian.
* Nhạy với phép thay đổi ánh sáng.

**Trích xuất đặc trưng HOG (**Histogram of gradient):

Histogram of gradient (HOG) là đặc trưng được dùng nhiều trong lĩnh vực phát hiện đối tượng, được đề xuất bởi Bill Triggs và Navel Dalal vào năm 2005 tại viện nghiên cứu INRIA.

Ý tưởng chính của phương pháp là hình dạng và trạng thái xuất hiện của vật có thể được đặc trưng bằng sự phân bố về cường độ và hướng của cạnh. Đặc trưng này được phát triển dựa trên SIFT, đặc trưng HOG được tính trên cả một vùng. Do sự biến thiên màu sắc trong các vùng là khác nhau, kết quả là các vùng sẽ cho ta vector đặc trưng của nó. Vì vậy, để có được đặc trưng của toàn bộ cửa sổ (window) ta phải kết hợp nhiều vùng liên tiếp lại với nhau.

Đặc trưng HOG có một vài biến thể thường gặp như: R-HOG, R2-HOG, C-HOG, …Các đặc trưng này khác nhau ở các phân bố và hình dạng của các ô như trong hình:

****

Hình 2.6 : Phân bố và hình dạng vài biến thể HOG

### Phân loại bằng SVM

SVM (Support Vector Machine) là một thuật toán máy học sử dụng trong bài toán phân loại và hồi quy. Nó dựa trên một khái niệm cơ bản, đó là tìm một siêu phẳng phân chia tốt nhất giữa các lớp dữ liệu khác nhau trong không gian đặc trưng.

Siêu phẳng là một siêu mặt phẳng (hyperplane) có số chiều bằng số chiều của không gian đặc trưng. Trong bài toán phân loại, SVM tìm cách xác định siêu phẳng sao cho khoảng cách từ siêu phẳng đến các điểm dữ liệu gần nhất (các vectơ hỗ trợ) là lớn nhất có thể. Các vectơ hỗ trợ là các mẫu dữ liệu quan trọng nhất để xác định siêu phẳng phân chia.

A graph of a line and a hyperplane

Description automatically generated

Hình 2.7 : Ảnh minh họa hàm phân loại SVM

Mục tiêu chính của SVM là tối đa hóa margin (khoảng cách) giữa siêu phẳng và các điểm dữ liệu gần nhất. Margin lớn tương đương với việc mô hình phân loại có khả năng tốt hơn đối với các điểm dữ liệu mới.

Trong trường hợp không thể tìm được một siêu phẳng hoàn hảo phân chia hai lớp dữ liệu tuyến tính, SVM sử dụng khái niệm kernel để biến đổi không gian đặc trưng ban đầu sang một không gian khác có thể phân chia tốt hơn. Kernel giúp SVM xử lý các bài toán phân loại phi tuyến.

Đầu vào của SVM có thể là các đặc trưng của ong, chẳng hạn như hình dạng, kích thước, màu sắc hoặc các đặc trưng được trích xuất từ ảnh của ong. Với mỗi mẫu dữ liệu (ong), các đặc trưng này sẽ được biểu diễn dưới dạng một vector đặc trưng.

SVM sẽ học từ các mẫu dữ liệu đã được gán nhãn với thông tin về ong mang phấn và ong không mang phấn. Thuật toán SVM sẽ cố gắng tìm ra một siêu phẳng phân chia tốt nhất giữa hai lớp dữ liệu này. Siêu phẳng này sẽ cố gắng tối đa hóa margin, tức là khoảng cách từ siêu phẳng đến các vectơ hỗ trợ gần nhất (support vectors) là lớn nhất có thể.

Các vectơ hỗ trợ là các mẫu dữ liệu quan trọng nhất trong việc xác định siêu phẳng phân chia. SVM sẽ xác định các vectơ hỗ trợ và sử dụng chúng để xác định siêu phẳng tốt nhất phân chia giữa hai lớp.

Khi mô hình SVM đã được huấn luyện, ta có thể sử dụng nó để phân loại các mẫu dữ liệu mới. SVM sẽ dự đoán xem mỗi mẫu dữ liệu mới là ong mang phấn hay không dựa trên các đặc trưng của nó và siêu phẳng phân chia đã được xác định.

Với SVM, ta có thể tinh chỉnh các tham số và kernel để cải thiện hiệu suất phân loại và đạt được độ chính xác tốt nhất cho bài toán phân loại ong mang phấn.

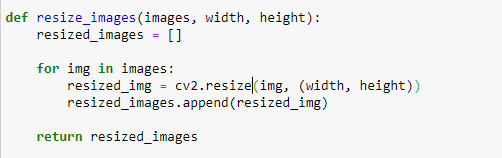
# THỰC HIỆN VÀ KẾT QUẢ

## Tiền xử lý hình ảnh

### Giảm kích thước ảnh

Ở đây bọn em đưa toàn bộ ảnh về đồng kích thước150\*90 sẽ đảm bảo rằng tất cả các hình ảnh trong tập dữ liệu đều có kích thước như nhau trước khi đưa vào SVM để phân loại áp dụng có hiệu quả.

Hàm resize\_images :



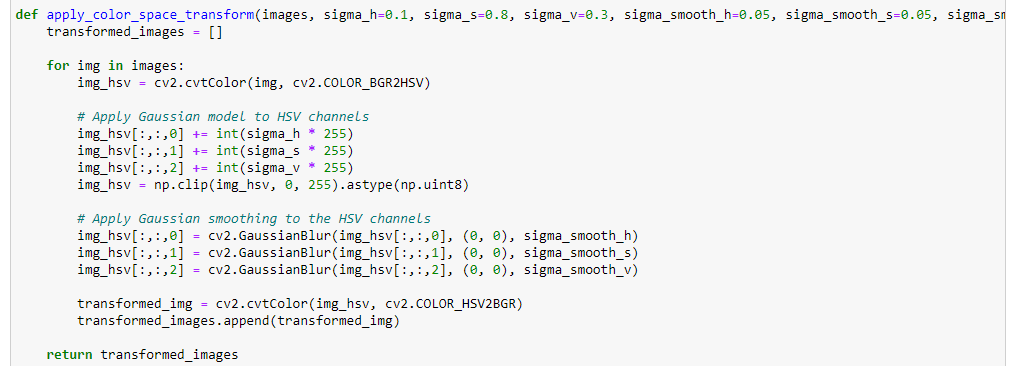
Hình 3.1: Hàm giảm kích thước ảnh

### Áp dụng bộ lọc Gaussian

Áp dụng bộ lọc Gaussian để giảm nhiễu trước khi sử dụng SVM trong bài toán phân loại ảnh có thể cải thiện hiệu suất của mô hình SVM. Bộ lọc Gaussian là một trong những bộ lọc thông dụng trong xử lý ảnh để làm mờ ảnh và giảm nhiễu bằng cách áp dụng một ma trận lõm lên ảnh.

Bước tiền xử lý này có thể giúp loại bỏ những chi tiết không cần thiết hoặc nhiễu trong ảnh, giúp mô hình SVM tập trung vào những đặc trưng quan trọng hơn và tăng độ chính xác của phân loại.

Dưới đây là hàm bộ lọc gausian :

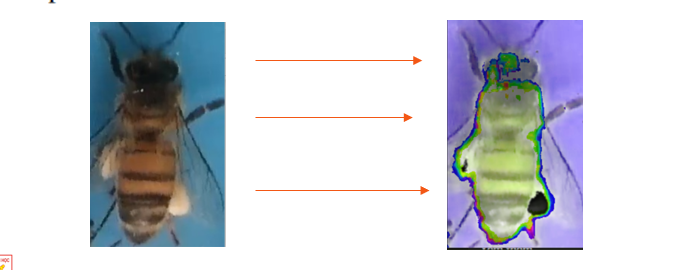


Hình 3.2 : Hàm bộ lọc gaussian

Các tham số dùng trong hàm lọc gaussian đã được chọn ở mục 2.4.a)

Kết quả sau khi qua bộ lọc Gaussian sẽ là ảnh mới có mức độ làm mờ cao hơn so với ảnh gốc, và những chi tiết nhỏ như nhiễu, điểm chấm, hay các đặc trưng không cần thiết khác sẽ bị giảm đi. Điều này giúp loại bỏ nhiễu trong ảnh và làm cho các đặc trưng chính của ảnh trở nên rõ ràng hơn, dễ dàng hơn trong việc xử lý hoặc phân loại ảnh.

Dưới đây là ảnh sau khi qua bộ lọc gaussian :

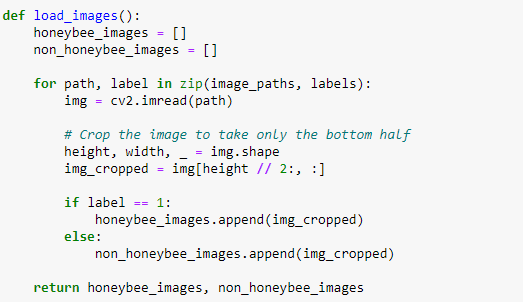


Hình 3.3: Ảnh sau khi qua bộ lọc gaussian

### Chọn vùng ảnh cần thiết

Vì mục tiêu của bài toán là phân loại ong mang phấn và ong không mang phấn, bọn em nhận thấy rằng ảnh ong mang phấn thì vùng chứa phấn nằm ở một nửa dưới của ảnh. Cùng với gợi ý từ cô giáo, bọn em đã tiến hành chọn một nửa vùng ảnh phía dưới để trích xuất đặc trưng và so sánh kết quả phân loại với trích xuất đặc trưng từ toàn bộ ảnh.

Hàm tải ảnh từ thư mục và chọn vùng ảnh cần thiết (cắt ½ ảnh) :



Hình 3.4 : Chọn vùng ảnh cần thiết

## Trích xuất đặc trưng

### Trích xuất đặc trưng HOG

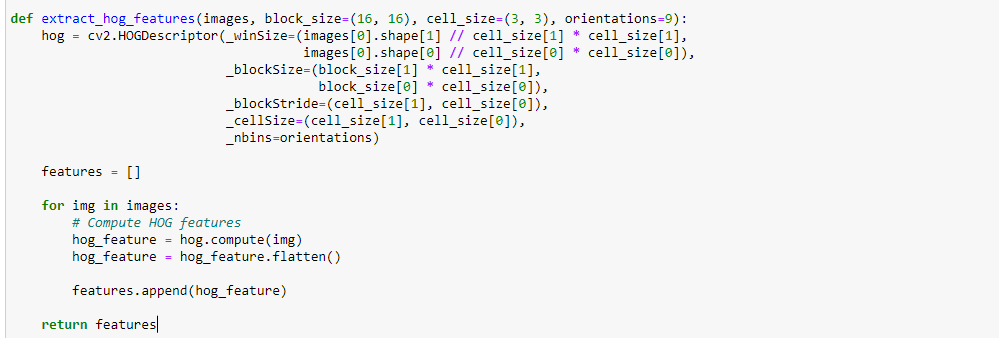
Sau khi tiền xử lý hình ảnh, bọn em áp dụng hàm trích xuất đặc trưng HOG để tạo vecto đặc trưng HOG với các tham số :

**block\_size=(16, 16)**: Kích thước của khối (block) trong ảnh. Trong phương pháp HOG, ảnh được chia thành các khối (block), và các histogram gradient sẽ được tính toán cho mỗi khối. Trong trường hợp này, kích thước của khối được đặt là 16x16 pixel. Kích thước này phụ thuộc vào đặc trưng của ảnh và nên được lựa chọn sao cho đủ lớn để bao phủ thông tin cần thiết của ảnh.

**cell\_size=(3, 3)**: Kích thước của cell trong ảnh. Cell là đơn vị cơ bản để tính toán hướng gradient và tính toán histogram. Trong trường hợp này, kích thước của cell được đặt là 3x3 pixel. Kích thước cell cũng là một tham số quan trọng và nên được chọn cẩn thận. Kích thước nhỏ hơn có thể giúp phát hiện các chi tiết nhỏ hơn, trong khi kích thước lớn hơn có thể giúp tăng tính tổng quát của đặc trưng.

**orientations=9**: Số lượng bin trong histogram, tức là số lượng khoảng hướng gradient sẽ được chia để tính toán. Trong trường hợp này, số lượng bin được đặt là 9. Số lượng bin càng lớn, đặc trưng HOG càng chi tiết và phức tạp hơn, nhưng cũng dễ dẫn đến hiện tượng overfitting.

Hàm trích xuất đặc trưng HOG :



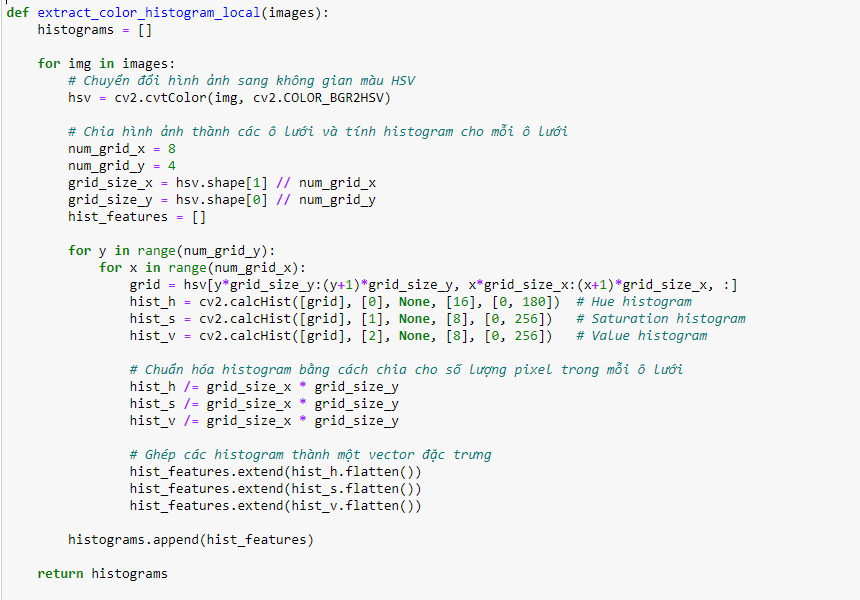
Hình 3.5 : Hàm trích xuất đặc trưng HOG

### Trích xuất đặc trưng color histogram (local)

Ở đây, bọn em chia ảnh thành các vùng con (patches) có kích thước 15x7 pixel. Đối với mỗi vùng con, tính toán histogram màu của nó. Histogram màu là một biểu đồ thống kê số lượng pixel có giá trị màu trong các phạm vi màu cụ thể. Để trích xuất histogram màu, ta có thể chia không gian màu của ảnh thành các bins, sau đó đếm số lượng pixel rơi vào mỗi bin. Việc này sẽ tạo ra một vector đặc trưng biểu thị sự phân bố màu sắc trong vùng con.

Cuối cùng, ta sẽ nhận được một vector đặc trưng cho mỗi vùng con 15x7 pixel trong ảnh ban đầu.

Hàm trích xuất đặc trưng color histogram :



Hình 3.6 : Hàm trích xuất đặc trưng color histogram

Để tìm được kích thước vùng là 15\*7 tối ưu kết quả phân loại đúng. Bọn em đã thử nghiệm kích thước vùng từ 0->20, và nhận thấy với kích thước vùng là 15\*7 cho kết quả phân loại đúng cao nhất.

Hàm tìm kích thước vùng để tối ưu kết quả phân loại :



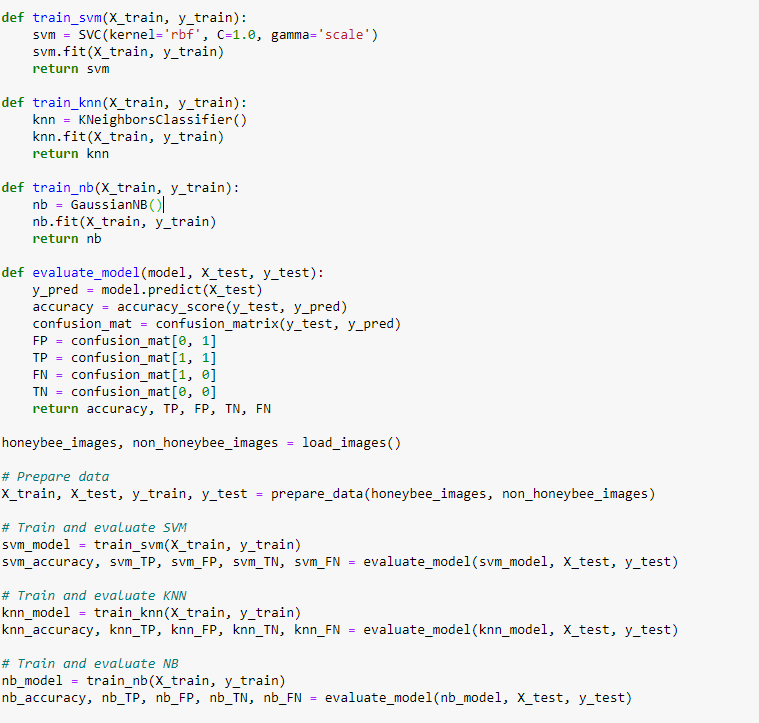
Hình 3.7 : Hàm tìm kích thước vùng tối ưu

## Áp dụng bộ phân loại SVM, KNN, NB

Sau khi tiền xử lý hình ảnh, trích xuất đặc trưng từ hình ảnh, bọn em đã áp dụng 3 bộ phân loại khác nhau để so sánh kết quả.

Với SVM, bọn em đã chọn các tham số bộ phân loại SVM với kernel = ‘rbf’, C = 1.0, gamma = ‘scale ‘. SVM sẽ sử dụng hàm kernel RBF để tính độ tương đồng giữa các điểm dữ liệu trong không gian đặc trưng. Điều này giúp SVM tạo ra các đường ranh giới phân chia phi tuyến giữa các lớp dữ liệu. Với c = 1.0, nó xác định mức độ mà SVM cho phép các điểm dữ liệu bị phân loại sai trong quá trình tối ưu hóa. Khi sử dụng gamma="scale", tham số gamma sẽ tự động thay đổi tùy theo số chiều đặc trưng của dữ liệu. Điều này có ý nghĩa là khi bọn em thay đổi số chiều đặc trưng, tham số gamma cũng sẽ tự động thay đổi tương ứng, giúp điều chỉnh mức độ lan tỏa của hàm kernel RBF phù hợp với dữ liệu mới.

Với KNN, bọn em áp dụng KNN với k = 1, tức huật toán chỉ sử dụng duy nhất một điểm láng giềng gần nhất để quyết định lớp của điểm dữ liệu mới. Khi có một điểm dữ liệu mới cần được phân loại, KNN sẽ tìm điểm dữ liệu gần nhất từ tập dữ liệu huấn luyện (dựa trên khoảng cách Euclidean hoặc các phương pháp khoảng cách khác) và sau đó gán cho điểm mới lớp của điểm láng giềng đó.



Hình 3.8 : Bộ phân loại SVM, KNN, NB

## Kết quả và đánh giá

### Kết quả về độ chính xác

|  |  |  |  |
| --- | --- | --- | --- |
| Đặc trưng | SVM | KNN | NB |
| Gaussian  Color histogram  HOG | 88.1%  83.2 % | 83.2%  79.7% | 80%  72.7% |
| Without gaussian Color histogram  HOG | 93%  82.5% | 80.4%  77.6% | 77.6%  76.9% |
| Gausian  Cắt ½ ảnh  Color histogram  HOG | 89.5%  88.1% | 82.5%  74.8% | 78.3%  78.3% |
| Without gaussian  Cắt ½ ảnh  Color histogram  HOG | 88.8%  81.8% | 79%  65% | 67.8%  76.9% |

Sau khi chạy mô hình, chúng em đã thống kê lại bảng kết quả đánh giá, dựa vào bảng trên, chúng em có một số nhận xét về mô hình như sau :

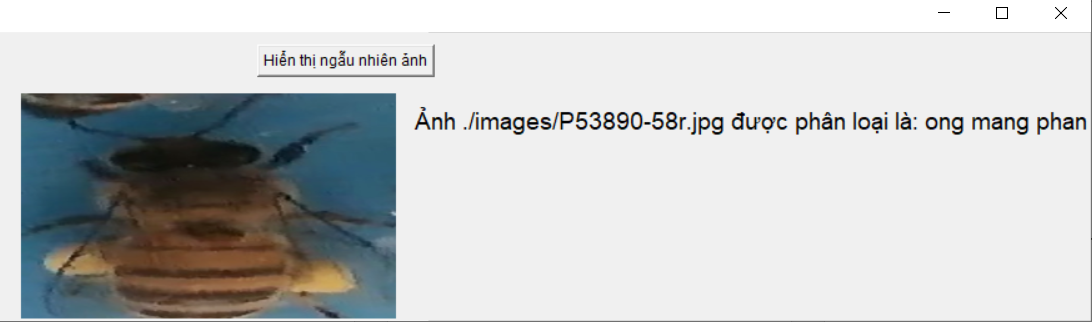
* Bộ phân loại SVM tốt hơn so với KNN và NB trong các trường hợp.
* Áp dụng thêm bộ lọc Gaussian vào tiền xử lý ảnh thì cải thiện độ chính xác thêm 2-5% so với không áp dụng bộ lọc.
* Đặc trưng color histogram cho độ chính xác tốt hơn so với đặc trưng HOG trong đa số trường hợp.
* Áp dụng cắt ½ ảnh khi áp dụng gaussian thì độ chính xác của thuật toán

SVM tăng lên 2-3% so với không áp dụng cắt ảnh.

* Những ảnh cho kết quả sai thì rơi nhiều vào false positive (ảnh ong mang phấn hoa mà dự đoán sai). Theo bọn em đánh giá thì nguyên nhân bởi vì đặc trưng ong mang phấn hoa thì những ảnh ong mang phấn hoa này có đặc trưng màu tối hơn nên dễ gây nhầm lẫn với các ảnh khác.

### Thử nghiệm mô hình

Ở đây, bọn em có thử nghiệm mô hình bằng giao diện thân thiện, chọn hình ảnh và hiển thị ra kết quả phân loại ong mang phấn.



# KẾT LUẬN

## Kết luận

Nội dung của bài tập lớn đã trình bày được một số các khái niệm về tìm kiếm ảnh người dựa vào truy vấn, kĩ thuật trích chọn đặc trưng HOG và color histogram, tuy nhiên kết quả vẫn còn một vài vấn đề hạn chế, như là tốc độ tra cứu chưa được nhanh, hay kết quả tra cứu chưa chính xác 100%.

Mặc dù đã hoàn thành mục tiêu bài tập lớn nhưng do điều kiện về thời gian có hạn mà lĩnh vực cần tìm hiểu cũng tương đối rộng nên những gì tìm hiểu được trong bài tập sẽ khó tránh khỏi những thiếu xót. Chương trình thử nghiệm cũng chưa thực sự hoàn thiện nhưng đó cũng là một kết quả khả quan. Trong thời gian tới nếu có điều kiện chúng em sẽ xây dựng lại chương trình tìm kiếm ảnh người một cách hoàn chỉnh hơn, tối ưu thuật toán nhăm tăng tốc độ tra cứu và độ chính xác của kết quả hiện thị.

## Hướng phát triển sau này

Mặc dù áp dụng các bộ phân loại với các đặc trưng HOG và color histogram cho kết quả độ chính xác khá cao, nhưng vẫn có thể cải tiến và phát triển thêm để đáp ứng tốt các nhu cầu ngày càng tang :

* Áp dụng các đặc trưng mới có thể phù hợp hơn
* Kết hợp các loại đặc trưng hỗn hợp với nhau
* Áp dụng cho phân loại ong mang phấn trong video
* Áp dụng cho nhiều loại ong khác nhau