Tab 1

Đại Học Quốc Gia Thành Phố Hồ Chí Minh

Trường Đại học Công Nghệ Thông Tin



Môn học: Nhập Môn Thị Giác Máy Tính

Mã môn: CS231.P21

Học kỳ II (2024-2025)

Chủ Đề: Phân Loại Hình Ảnh Xe Bốn Bánh

Giảng Viên: TS.Mai Tiến Dũng

Nhóm sinh viên thực hiện

| Họ và Tên | MSSV |
| --- | --- |
| Đặng Chí Nguyên | 22520963 |
| Đường Thị Mộng Thúy | 22521454 |
| Mã Kim Phát | 22521071 |

| MSSV | Họ và Tên | Nhiệm vụ | Mức độ hoàn thành |
| --- | --- | --- | --- |
| 22520963 | Đặng Chí Nguyên | -Quay và gán nhãn dữ liệu  -Thực hiện phương pháp Dominant Color, LBP, PCA | 100% |
| 22521454 | Đường Thị Mộng Thúy | -Quay và gán nhãn dữ liệu  -Thực nghiệm các phương pháp chiết xuất lên mô hình học máy | 100% |
| 22521071 | Mã Kim Phát | -Quay và gán nhãn dữ liệu  -Thực hiện phương pháp HOG, chiết xuất đặc trưng dựa trên mô hình VGG16, ResNet50 | 100% |

Mục Lục

[**I. Giới thiệu và lý do chọn bài toán 2**](#_4ryhotxxz3ly)

[**II. Phát biểu bài toán 2**](#_78odt0x30q2x)

[**III. Phương pháp thực hiện 3**](#_8xhhj1su6otr)

[A. Phương pháp chiết xuất đặc trưng truyền thống 3](#_7bfn6lj0hyvh)

[1. Chiết xuất đặc trưng màu chủ đạo (Dominant Color) 3](#_m38wc36w2ea7)

[a) Khái niệm 3](#_e9sj15zu0fs)

[b) Các bước thực hiện 3](#_xls5j6u88gqs)

[c) Thực nghiệm 3](#_wlwgghrtw3yr)

[2. Chiết xuất đặc trưng HOG 4](#_xnqa0w2luwu9)

[a) Khái niệm 4](#_kl28rpqzvhrh)

[b) Các bước hiện hiện 4](#_rf2xhextt3by)

[c) Thực nghiệm 4](#_itiq0033cayw)

[3. Chiết xuất đặc trưng LBP 5](#_e48qcl6ec9k)

[a) Khái niệm 5](#_fjirgljokyl2)

[b) Các bước thực hiện 5](#_hp45jnzbk547)

[c) Thực nghiệm 5](#_on7bmpep1mtp)

[B. Phương pháp chiết xuất đặc trung học sâu 6](#_iyxlfl57ng33)

[1. Chiết xuất đặc trưng bằng mô hình VGG16 6](#_w1gseytozit5)

[a) Khái niệm 6](#_kzny9i9ziyga)

[b) Các bước thực hiện 6](#_qp1xttx1h7yd)

[2. Chiết xuất đặc trưng bằng mô hình ResNet50 7](#_yzc6cyhfzxo3)

[a) Khái niệm 7](#_5eyhpzf72pcq)

[b) Các bước thực hiện 7](#_qlpox0peeefc)

[C. Principal component analysis (PCA) 7](#_ect5vlkyc399)

[a) Khái niệm 7](#_tx55zo9aumbh)

[b) Các bước thực hiện 8](#_za5qaevzv2yk)

[**IV. Kết quả thực nghiệm 8**](#_7i2dygp44dka)

[A. Bộ dữ liệu 8](#_4jx8mkl579z9)

[1. Tổng quan 8](#_u2h1g7s9v2dz)

[2. Phương pháp thu thập và các bước tiền xử lý 9](#_svpllmgj27vw)

[a. Thu thập dữ liệu thô 9](#_11kjpdau66ho)

[b. Tiền xử lý 9](#_jf41nnidctml)

[B. Các độ đo nhóm lựa chọn 9](#_6t5dc18t8qme)

[C. So sánh giữa các đặc trưng 10](#_7sanvtz9al25)

[D. Thực Nghiệm 11](#_ifj1hz7cfc3)

[**V. Đánh giá 13**](#_jnztzzf1dnzn)

[**Tài liệu tham khảo 13**](#_8b7fbfiw0wu0)

# Giới thiệu và lý do chọn bài toán

Trong thời đại công nghệ phát triển, trí tuệ nhân tạo (AI) và thị giác máy tính (Computer Vision) ngày càng được ứng dụng rộng rãi, đặc biệt trong lĩnh vực giao thông thông minh. Một trong những ứng dụng thiết thực là nhận diện và phân loại phương tiện giao thông từ hình ảnh nhằm hỗ trợ quản lý đô thị, điều phối giao thông và nâng cao hiệu quả giám sát.

Tại Việt Nam, giao thông phức tạp với nhiều loại phương tiện 4 bánh như ô tô con, xe tải, xe buýt, xe bán tải,... khiến việc phân loại trở nên cần thiết trong các hệ thống như: giám sát giao thông, bãi đỗ xe thông minh, và phân tích hành vi giao thông.

Vì vậy, nhóm chọn đề tài này với mục tiêu xây dựng hệ thống phân loại phương tiện từ hình ảnh một cách chính xác và phù hợp với thực tế tại Việt Nam.

# Phát biểu bài toán

-Trong bối cảnh phát triển mạnh mẽ của trí tuệ nhân tạo và thị giác máy tính, việc tự động nhận diện và phân loại các phương tiện giao thông từ hình ảnh ngày càng trở nên quan trọng. Một trong những ứng dụng thực tiễn là việc phân loại các loại xe 4 bánh (chẳng hạn như ô tô con, xe tải, xe buýt, xe bán tải...) từ hình ảnh thu được từ camera giám sát giao thông, bãi đỗ xe thông minh hoặc hệ thống quản lý giao thông.

-Mục tiêu của bài toán là xây dựng một hệ thống có thể:

* Tiền xử lý và trích xuất đặc trưng phù hợp từ hình ảnh đầu vào.
* Huấn luyện một mô hình phân loại với độ chính xác cao, có khả năng nhận diện chính xác các loại xe dù cho tính chất giao thông hỗn loạn ở Việt Nam

- Bài toán nhận đầu vào và đầu ra lần lượt là:

**Input**:

1.Tập dữ liệu gồm có N phần tử là ảnh số, mỗi ảnh chỉ có 1 phương tiện 4 bánh duy nhất có góc chụp từ trên cao trong điều kiện ánh sáng đầy đủ (ban ngày) và đã được gán nhãn:

Tập các nhãn của là: {0, 1, 2, 3, 4}

Trong đó:

* ﻿0: Container
* ﻿1: Truck-half\_truck
* ﻿2: Bus
* ﻿3: SUV-Crossover-MPV
* ﻿4: Sedan

﻿﻿Trong mỗi phần tử của :

* là ảnh số 𝓘 hoặc là một vector đặc trưng có kích thước d chiều: 𝓘 =
* ﻿﻿ là giá trị nhãn (label) được gán cho

2. Ảnh 𝓘

**Output:**

• : giá trị nhãn dự đoán của ảnh

# Phương pháp thực hiện

## **Phương pháp chiết xuất đặc trưng truyền thống**

### **Chiết xuất đặc trưng màu chủ đạo (Dominant Color)**

#### **Khái niệm**

-Chiết xuất đặc trưng màu chủ đạo là một phương pháp chiết xuất đặc trưng màu. Đây là quá trình rút trích ra một hoặc nhiều màu đại diện cho hình ảnh. Những màu này phản ánh đặc trưng thị giác tổng thể của ảnh.

-Phương pháp chiết xuất đặc trưng màu chủ đạo nổi bật nhờ khả năng giảm chiều dữ liệu trong khi vẫn giữ lại bản chất thị giác của ảnh một cách đơn giản và hiệu quả. Nhờ vậy, phương pháp chiết xuất màu chủ đạo được áp dụng rộng rãi trong các bài toàn phân loại cảnh vật, gợi ý thời trang,...

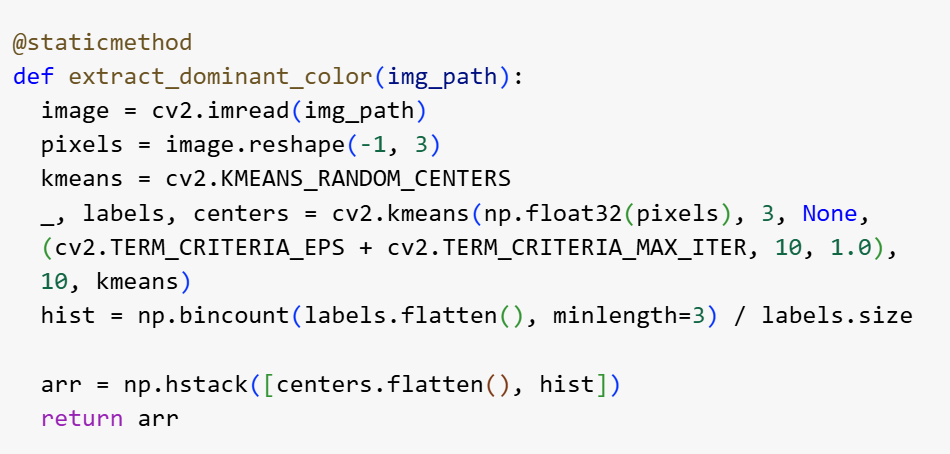
#### **Các bước thực hiện**

-Các bước thực hiện chiết xuất màu chủ đạo:

* Làm phẳng ảnh thành 1 mảng 2 chiều gồm các điểm ảnh, mỗi điểm ảnh là 1 vector đặc trưng màu.
* Sử dụng thuật toán k-means để nhóm các điểm ảnh thành các cụm màu. Số lượng cụm được xác định trước biểu thị cho số lượng màu chủ đạo muốn chiết xuất.
* Sau khi phân cụm, các điểm ảnh sẽ được gán nhãn là cụm màu mà nó thuộc về. Từ đó, ta có thể tính được tỉ lệ phần trăm của từng màu chủ đạo trong toàn bộ ảnh.

#### **Thực nghiệm**

-Chúng tôi đã thực nghiệm ứng dụng của phương pháp chiết xuất đặc trưng với thuật toán k-means phân cụm 3 màu chủ đạo.



### **Chiết xuất đặc trưng HOG**

#### **Khái niệm**

**-HOG (Histogram of Oriented Gradients)** là một kỹ thuật chiết xuất đặc trưng trong thị giác máy tính, được sử dụng phổ biến để mô tả hình dạng và cấu trúc của đối tượng trong hình ảnh. Phương pháp này dựa trên ý tưởng rằng hình dạng và đường viền của một vật thể có thể được đặc trưng tốt bằng cách phân tích hướng của gradient cục bộ tại các vùng nhỏ trong ảnh.

-Phương pháp chiết xuất đặc trưng HOG nổi bật bởi khả năng **giữ lại thông tin cấu trúc quan trọng của ảnh** trong khi vẫn có tính bất biến tương đối đối với thay đổi ánh sáng và hình ảnh nhiễu. Nhờ đó, HOG được áp dụng rộng rãi trong các bài toán như phát hiện người, nhận dạng phương tiện giao thông, và phân loại hình ảnh.

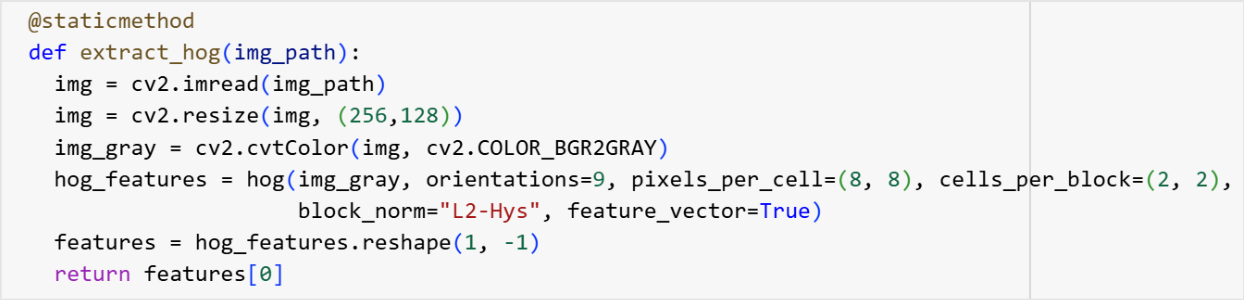
#### **Các bước hiện hiện**

Để thực hiện phương pháp chiết xuất đặc trưng HOG, ta cần thực hiện lần lượt các bước sau đây:

* Chia ảnh thành các ô nhỏ (cells) có kích thước cố định
* Tính toán gradient tại mỗi điểm ảnh để xác định độ lớn và hướng của sự thay đổi cường độ
* Lập biểu đồ tần suất (histogram) các hướng của gradient trong mỗi ô, phản ánh sự phân bố của các cạnh (edges) trong khu vực đó.
* Chuẩn hóa các biểu đồ theo khối (blocks) để giảm sự ảnh hưởng do thay đổi ánh sáng hoặc độ tương phản
* Kết hợp các histogram từ toàn bộ ảnh để tạo thành vector đặc trưng đầu ra đại diện cho hình dạng tổng thể

#### **Thực nghiệm**

Sau đây là các tham số mà nhóm sử dụng khi thực hiện phương pháp chiết xuất đặc trưng HOG:



### **Chiết xuất đặc trưng LBP**

#### **Khái niệm**

-Chiết xuất đặc trưng Local Binary Pattern là phương pháp mô tả kết cấu trong ảnh xám. Phương pháp này dựa trên việc so sánh cường độ của các điểm ảnh lân cận với điểm ảnh trung tâm để tạo ra một chuỗi nhị phân đại diện cho kết cấu tại từng vị trí.

-Phương pháp LBP nổi bật nhờ khả năng tính toán nhanh, gọn nhẹ. Áp dụng tốt trong các bài toán phân loại ảnh vì nó giữ lại được đặc trưng kết cấu của ảnh.

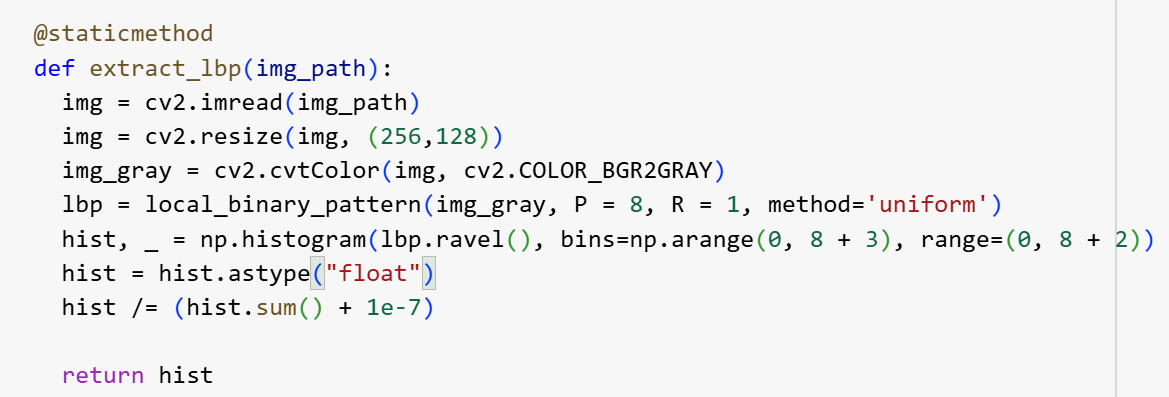
#### **Các bước thực hiện**

-Các bước thực hiện chiết xuất đặc trưng LBP:

* Ảnh được chuyển về màu xám nếu là ảnh màu.
* Với mỗi điểm ảnh, thực hiện so sánh giá trị cường độ của các điểm ảnh lân cận với pixel xung quanh. Nếu giá trị của điểm ảnh xung quanh lớn hơn hoặc bằng điểm ảnh trung tâm, gán là 1, ngược lại gán là 0.
* Sau khi thực hiện so sánh trên, thu được chuỗi N số nhị phân. Sau đó chuyển chuỗi nhị phân này thành số thập phân thu được giá trị LBP tại pixel trung tâm.
* Toàn bộ các pixel của ảnh được duyệt qua để thu được ảnh đặc trưng LBP.
* LBP có thể được biểu diễn dưới dạng tần suất xuất hiện của các mã LBP dùng làm vector đặc trưng cho ảnh.

#### **Thực nghiệm**

-Trong phần chiết xuất đặc trưng LBP, chúng tôi đã sử dụng số pixel xung quanh là 8 với khoảng cách từ pixel trung tâm đến các pixel xung quanh là 1.



## **Phương pháp chiết xuất đặc trung học sâu**

### **Chiết xuất đặc trưng bằng mô hình VGG16**

#### **Khái niệm**

**-VGG16** là một mô hình mạng nơ-ron tích chập sâu (Convolutional Neural Network – CNN) nổi tiếng trong lĩnh vực thị giác máy tính, được phát triển bởi nhóm nghiên cứu Visual Geometry Group (VGG) thuộc Đại học Oxford và công bố trong cuộc thi **ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014**.

-Mô hình VGG16 gồm 16 lớp có trọng số, bao gồm:

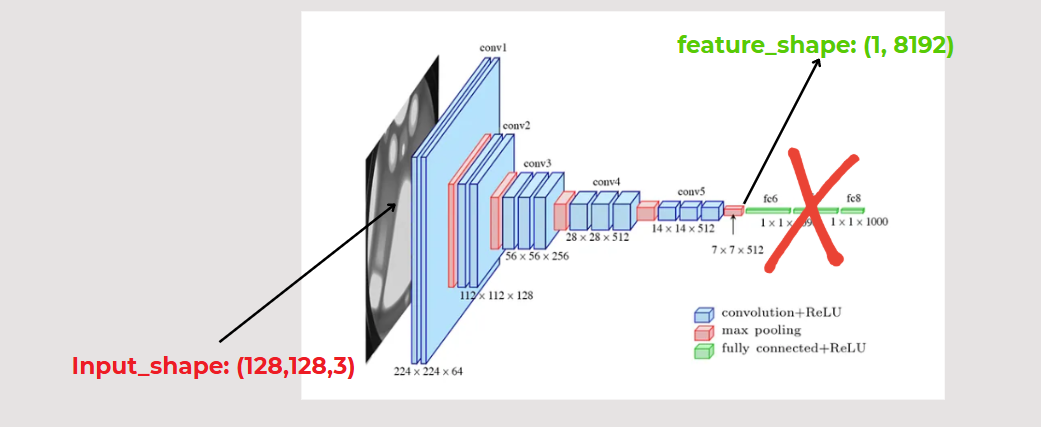
* 13 lớp Convolutional
* 3 lớp Fully connected

-**VGG16** được huấn luyện trên **tập dữ liệu ImageNet** với hơn 1 triệu ảnh thuộc 1000 lớp đối tượng khác nhau, và đã đạt hiệu suất rất cao trong nhiều tác vụ phân loại hình ảnh. Cùng với tính đơn giản trong kiến trúc, phù hợp cho các thí nghiệm hạn chế về mặt tài nguyên mà vẫn cho ra kết quả chính xác hơn các mô hình học máy truyền thống trong tác vụ phân loại ảnh.

#### **Các bước thực hiện**

-Để sử dụng mô hình **VGG16** để chiết xuất đặc trưng, chúng tôi chỉ sử dụng các **13 lớp Convolutional** để lấy các đặc trưng của ảnh mà không sử dụng **các lớp Fully Connected**. Sau đó chúng tôi sẽ sử dụng **vector đặc trưng** này đi huấn luyện trong mô hình học máy truyền thống.

-Dưới đây là hình ảnh trực quan và cách mà chúng tôi ứng dụng mô hình **VGG16** vào chiết xuất đặc trưng:



### **Chiết xuất đặc trưng bằng mô hình ResNet50**

#### **Khái niệm**

**-ResNet-50** là một mô hình mạng nơ-ron tích chập sâu (CNN) thuộc họ **Residual Network (ResNet)**, được giới thiệu bởi nhóm nghiên cứu của Microsoft trong bài báo "Deep Residual Learning for Image Recognition" tại hội nghị CVPR năm 2015. Mô hình này đặc biệt nổi bật khi giành chiến thắng tại cuộc thi **ImageNet 2015** với độ chính xác vượt trội

-ResNet-50 bao gồm **50 lớp có trọng số**, trong đó sử dụng cấu trúc đặc biệt gọi là **residual block**. Khác với các mạng sâu truyền thống, ResNet đưa vào các kết nối tắt (skip connections) cho phép thông tin **nhảy qua** một hoặc nhiều lớp. Nhờ đó, mô hình có thể học hiệu quả ngay cả khi mạng rất sâu, tránh được vấn đề **gradient biến mất** (vanishing gradient) – một trở ngại lớn khi huấn luyện các mạng sâu.

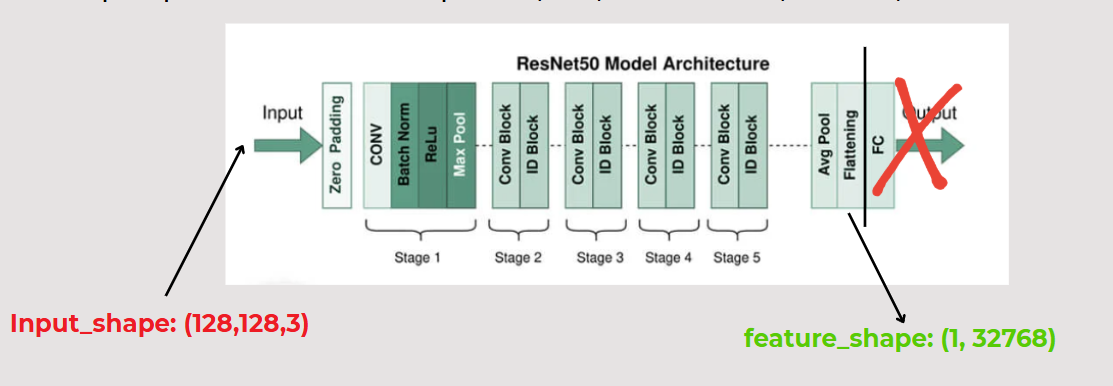
-Mô hình ResNet-50 có 50 lớp có trọng số, bao gồm:

* 49 lớp Convolutional
* 1 lớp fully connected

#### **Các bước thực hiện**

-Tương tự như mô hình VGG16, chúng tôi chỉ sử dụng các lớp Convolutional chỉ để chiết xuất đặc trưng mà không dùng đến lớp fully connected. Sau khi có được kết quả của vector đặc trưng, chúng tôi dùng vector này để huấn luyện bằng các mô hình học máy truyền thống.

-Dưới đây là hình ảnh trực quan và cách mà chúng tôi ứng dụng mô hình ResNet-50 vào chiết xuất đặc trưng:



## **Principal component analysis (PCA)**

#### **Khái niệm**

-Principal Component Analysis (PCA) là một phương pháp giảm chiều dữ liệu tuyến tính. PCA giúp chuyển đổi tập dữ liệu ban đầu sang một không gian mới sao cho các trục tọa độ chứa nhiều thông tin nhất có thể.

-PCA nổi bật trong khả năng giảm chiều dữ liệu nhằm tăng tốc độ xử lý nhưng vẫn giữ được hầu hết thông tin quan trọng của tập dữ liệu ban đầu.

#### **Các bước thực hiện**

-Các bước thực hiện PCA nhằm giảm chiều dữ liệu như sau:

* Dữ liệu đầu vào ban đầu sẽ được chuẩn hóa để có trung bình bằng 0 và độ lệch chuẩn bằng 1(chuẩn hóa z-score).
* Sau đó sẽ tính toán ma trận hiệp phương sai từ dữ liệu chuẩn hóa nhằm thể hiện mối quan hệ tuyến tính giữa các chiều dữ liệu.
* Từ ma trận hiệp phương sai, bắt đầu tính toán giá trị riêng và vector riêng nhằm xác định mức độ phân tán của dữ liệu.
* Chọn ra một số vector riêng ứng với các giá trị riêng lớn nhất. Số lượng thành phần được chọn dựa vào mục tiêu giữ lại bao nhiêu % phương sai.
* Chiếu dữ liệu ban đầu lên không gian mới được tạo bởi các vector riêng đã được chọn để thu được dữ liệu với số chiều thấp hơn.

# Kết quả thực nghiệm

Sau khi áp dụng các phương pháp chiết xuất đặc trưng truyền thống (như HOG, LBP, v.v.) và các đặc trưng học sâu (Deep Features từ các mô hình như VGG16, ResNet), nhóm đã tiến hành huấn luyện và đánh giá hiệu suất của hai mô hình phân loại: SVM (Support Vector Machine) và KNN (K-Nearest Neighbors). Kết quả tổng hợp được thể hiện qua các tiêu chí đánh giá phổ biến như độ chính xác (accuracy) và F1-score.

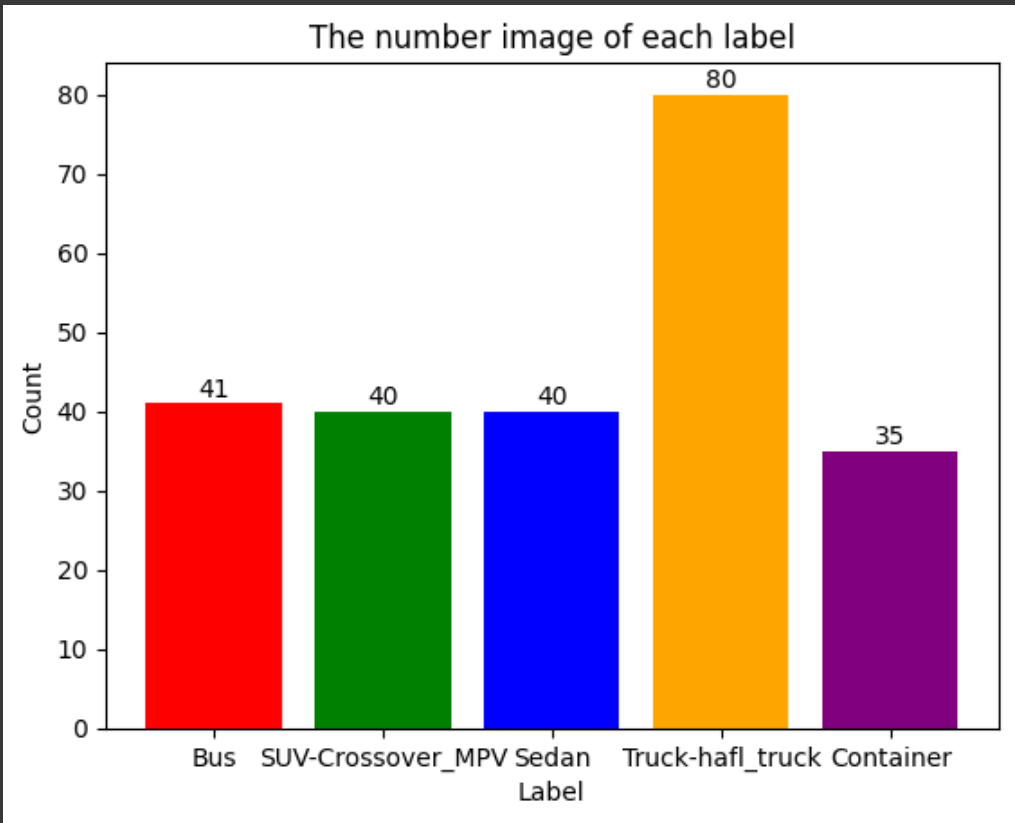
## **Bộ dữ liệu**

### **Tổng quan**

Bộ dữ liệu được sử dụng bao gồm 236 ảnh số. Mỗi ảnh thuộc về 1 trong 5 lớp được quy định bao gồm:

* **SUV-Crossover-MPV**: Gồm các xe gầm cao hoặc Minivan. Công suất chở người tối đa lên đến 16 chỗ.
* **Truck-Half-Truck**: Gồm các xe tải hoặc bán tải.
* **Bus**: Gồm các xe chuyên dùng để chở số lượng lớn hành khách, công suất chở người lớn hơn SUV-Crossover-MPV
* **Container**: xe có kéo rơ moóc hoặc sơ mi rơ moóc phía sau dùng để chuyên chở hàng hóa.
* **Sedan**: Loại xe gầm thấp chuyên dùng để chở người. Công suất chuyên chở tối đa lên đến 5 người.

Phân bố của các nhãn trong bộ dữ liệu:



### **Phương pháp thu thập và các bước tiền xử lý**

#### **Thu thập dữ liệu thô**

Dữ liệu thô ban đầu là các đoạn video được quay từ điện thoại thông minh theo góc từ trên cao. Các đoạn video được thu thập tại nhiều tuyến đường khác nhau trên địa bàn TPHCM.

#### **Tiền xử lý**

Các đoạn video sau đó được tiền xử lý bằng YOLO để nhận diện và cắt ảnh thành các ảnh chứa đối tượng là phương tiện 4 bánh. Các ảnh này sau đó được phân loại thủ công vào 1 trong 5 lớp như đã đề cập ở trên. Sau đó, tùy vào phương pháp trích xuất đặc trưng khác nhau mà điều chỉnh lại kích cỡ của ảnh sao cho phù hợp.

## **Các độ đo nhóm lựa chọn**

* Trong bài toán này, do không đặt nặng yêu cầu đặc biệt như việc ưu tiên phát hiện đúng các trường hợp bệnh (đòi hỏi độ đo recall cao), nhóm quyết định lựa chọn hai độ đo mang tính tổng quát và phổ biến là Accuracy và F1-score. Hai độ đo này giúp đánh giá toàn diện hiệu quả của mô hình, vừa phản ánh được độ chính xác tổng thể (Accuracy), vừa cân bằng giữa Precision và Recall thông qua F1-score.
* **Độ chính xác (Accuracy)** được tính theo công thức:

Accuracy = =

* **F1-score** được tính theo công thức:

F1-score = 2 x

## **So sánh giữa các đặc trưng**

* Dưới đây là kết quả nhóm chúng tôi thực nghiệm các phương pháp chiết xuất đặc trưng lên hai mô hình KNN và SVM với lần lượt các tham số được khảo sát trên phương pháp Grid Search như sau:

**KNN:**

* "n\_neighbors" : [i + 1 for i in range(1,25) if i % 2 == 0],
* 'weights' : ['uniform', 'distance'],
* 'algorithm' : ['auto', 'ball\_tree', 'kd\_tree', 'brute']

**SVM:**

* 'C': [0.001, 0.01, 0.1, 1, 10, 100, 1000],
* 'kernel': ['linear', 'poly', 'rbf', 'sigmoid' ],
* 'gamma': ['scale', 'auto']
* Sau đó phương pháp cho ra config tốt nhất như sau:
* KNN: algorithm='auto', n\_neighbors=3, weights='distance'
* SVM: (C=10, gamma='scale', kernel='linear'

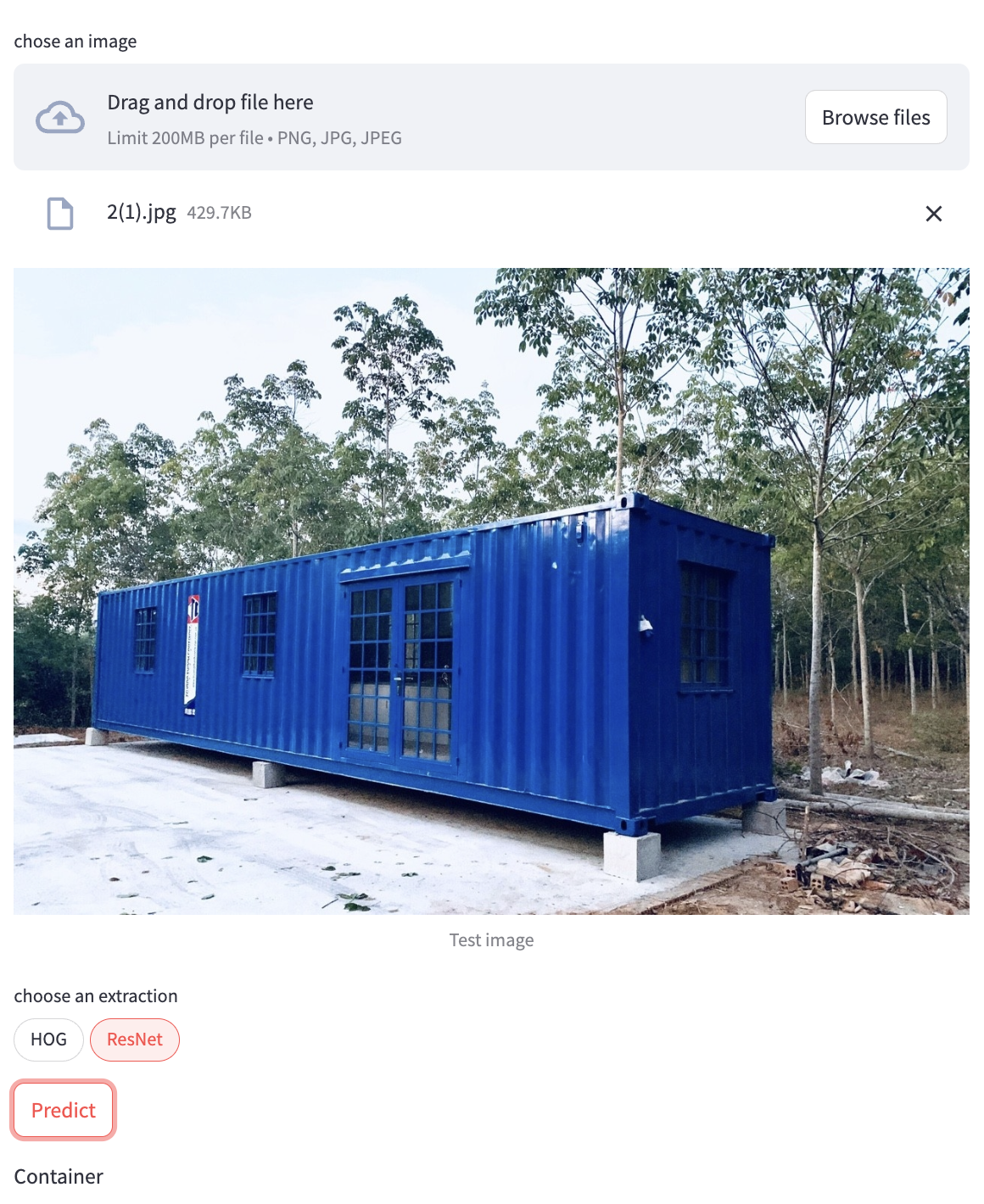
|  | KNN | | SVM | |
| --- | --- | --- | --- | --- |
| **Đặc trưng truyền thống** | Accuracy | F1\_score | Accuracy | F1\_score |
| dominant\_color | 52% | 51% | 48% | 45% |
| HOG | **67%** | **64%** | 81% | 80% |
| LBP | 44% | 41% | 29% | 14% |
| HOG+LBP | 67% | 64% | 81% | 80% |
| HOG+dominant\_color | 52% | 51% | 73% | 71% |
| HOG (PCA) | 65% | 62% | **85%** | **85%** |
| HOG (PCA) + LBP | 62% | 60% | 77% | 76% |
| **Đặc trưng Học Sâu** |  |  |  |  |
| ResNet50 | 71% | 69% | **90%** | **89%** |
| VGG16 | 69% | 63% | 85% | 85% |

## **Thực Nghiệm**

* Theo yêu cầu bài toán ( sử dụng 1 hình ảnh xe 4 bánh để nhận diện kết quả dự đoán ra là tên của loại xe đó )



* Theo yêu cầu của thầy ( sử dụng 1 hình ảnh nhà container để áp dụng cho bài toán dự đoán )



**Nhận xét :** Mô hình đã phân loại hình ảnh là "Container" thay vì một loại xe 4 bánh. Điều này cho thấy kết quả không đúng với mục tiêu bài toán, dù có thể đúng về mặt đối tượng trong ảnh.

**Đánh giá :**

* Ảnh đầu vào không phải là xe 4 bánh mà là một nhà container tĩnh.
* Cả xe và container đều có nhiều cạnh, góc và hình dạng hình hộp tương tự. HOG sẽ trích xuất những đặc trưng này.
* ResNet có thể học các đặc trưng chung của "vật thể hình hộp" nếu dữ liệu huấn luyện không đa dạng.
* Không có cơ chế nào để kiểm tra xem ảnh đầu vào có phải là xe hay không trước khi đưa vào mô hình chính.

# Đánh giá

**Kết quả đạt được (bên cạnh những hạn chế):** Mặc dù còn sai sót trong phân loại trường hợp cụ thể, nhóm đã đạt được:

* Xây dựng thành công hệ thống nhận diện xe 4 bánh, thể hiện khả năng ứng dụng thị giác máy tính vào thực tiễn.
* Phát triển ứng dụng (app) dễ sử dụng, nâng cao trải nghiệm tương tác với mô hình.
* Ứng dụng mô hình SVM với ưu điểm vượt trội: Khả năng xử lý hiệu quả dữ liệu chiều cao, tổng quát hóa tốt và ít bị overfitting, đóng góp vào tính ổn định của hệ thống.

**Hướng cải thiện :** Để nâng cao độ chính xác và tin cậy, hệ thống cần tập trung khắc phục các hạn chế. Cụ thể, cần cải thiện chất lượng và đa dạng dữ liệu huấn luyện (bổ sung mẫu âm tính, tăng cường dữ liệu) và tích hợp bước tiền xử lý/lọc ảnh đầu vào thông minh. Sự kết hợp giữa mô hình mạnh mẽ như SVM và quy trình dữ liệu chặt chẽ sẽ là chìa khóa để đạt hiệu suất tối ưu và ứng dụng thành công trong thực tiễn.

# Tài liệu tham khảo

<https://machinelearningcoban.com/2017/06/15/pca/#3-principal-component-analysis> [15.5.2025]

<https://machinelearningcoban.com/2017/06/21/pca2/> [15.5.2025]

https://arxiv.org/html/2407.18665v1 [17.5.2025]

https://www.geeksforgeeks.org/vgg-16-cnn-model/ [14.5.2025]

https://www.geeksforgeeks.org/vgg-16-cnn-model/[14.5.2025]