**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ ĐÔNG Á**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÀI TẬP LỚN**

**HỌC PHẦN: XỬ LÝ ẢNH VÀ THỊ GIÁC MÁY TÍNH**

**Đề tài số 13: Xây dựng hệ thống nhận dạng và phân loại động vật ăn thịt và động vật ăn cỏ**

**Giảng viên hướng dẫn: Lương Thị Hồng Lan**

|  |  |  |  |
| --- | --- | --- | --- |
| **TT** | **Mã sinh viên** | **Sinh viên thực hiện** | **Lớp hành chính** |
| **1** | 20210794 | Nguyễn Xuân Mạnh | DCCNTT12.10.3 |
| **2** | 20211493 | Nguyễn Đức Toàn | DCCNTT12.10.3 |
| **3** | 20210789 | Nguyễn Xuân Phong | DCCNTT12.10.3 |

**Bắc Ninh, năm 2024**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ ĐÔNG Á**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÀI TẬP LỚN**

**HỌC PHẦN: XỬ LÝ ẢNH VÀ THỊ GIÁC MÁY TÍNH**

**Đề tài số 13: Xây dựng hệ thống nhận dạng và phân loại động vật ăn thịt và động vật ăn cỏ**

**Giảng viên hướng dẫn: Lương Thị Hồng Lan**

|  |  |  |  |
| --- | --- | --- | --- |
| **TT** | **Mã sinh viên** | **Sinh viên thực hiện** | **Lớp hành chính** |
| **1** | 20210794 | Nguyễn Xuân Mạnh | DCCNTT12.10.3 |
| **2** | 20211493 | Nguyễn Đức Toàn | DCCNTT12.10.3 |
| **3** | 20210789 | Nguyễn Xuân Phong | DCCNTT12.10.3 |

|  |  |
| --- | --- |
|  |  |

**Bắc Ninh, năm 2024**

|  |  |
| --- | --- |
| TRƯỜNG ĐẠI HỌC CÔNG NGHỆ ĐÔNG Á  **KHOA CÔNG NGHỆ THÔNG TIN** | **KỲ THI KẾT THÚC HỌC PHẦN**  **HỌC KỲ 1, NĂM HỌC 2024** – **2025** |

|  |  |
| --- | --- |
| **PHIẾU CHẤM THI BÀI TẬP LỚN KẾT THÚC HỌC PHẦN**  **Mã đề thi: 13**  **Tên học phần: XỬ LÝ ẢNH & THỊ GIÁC MÁY TÍNH**  **Lớp Tín chỉ:**  **XATGMT.03.K12.03.LH.C04.1\_LT** | |
| **Cán bộ chấm thi 1**  **Lương Thị Hồng Lan**  *(Ký và ghi rõ họ tên)* | **Cán bộ chấm thi 2**  *(Ký và ghi rõ họ tên)* |

| **TT** | **TIÊU CHÍ** | **THANG ĐIỂM** | Nguyễn Xuân Mạnh | Nguyễn Xuân Phong | Nguyễn Đức Toàn |
| --- | --- | --- | --- | --- | --- |
| 20210794 | 20210789 | 2021079 |
| **1** | **Nội dung báo cáo trên Word đầy đủ** | **3.5** |  |  |  |
| 1.1 | Có bố cục rõ ràng (mục lục, phần mở đầu, nội dung chính, kết luận). | 0,5 |  |  |  |
| 1.2 | Nội dung phân tích rõ ràng, logic. | 0,5 |  |  |  |
| 1.3 | Có dẫn chứng, số liệu minh họa đầy đủ. | 0,5 |  |  |  |
| 1.4 | Ngôn ngữ và trình bày chuẩn, không lỗi chính tả. | 0,5 |  |  |  |
| 1.5 | Có trích dẫn tài liệu tham khảo đúng quy cách. | 0,5 |  |  |  |
| 1.6 | Được trình bày chuyên nghiệp (canh lề, font chữ, khoảng cách dòng hợp lý). | 0,5 |  |  |  |
| 1.7 | Tài liệu đầy đủ, bám sát yêu cầu của đề bài. | 0,5 |  |  |  |
| **2** | **Nội dung thuyết trình đầy đủ** | **1.0** |  |  |  |
| 2.1 | Trình bày tự tin, phát âm rõ ràng, mạch lạc. | 0,5 |  |  |  |
| 2.2 | Nội dung thuyết trình đúng trọng tâm, không lan man. | 0,5 |  |  |  |
| **3** | **Slides báo cáo đầy đủ nội dung + Hỏi đáp** | **3.0** |  |  |  |
| 3.1 | Slides có bố cục rõ ràng (mở đầu, nội dung, kết luận). | 0,5 |  |  |  |
| 3.2 | Thiết kế slides đẹp, chuyên nghiệp (màu sắc, hình ảnh minh họa). | 0,5 |  |  |  |
| 3.3 | Nội dung trên slides ngắn gọn, dễ hiểu, súc tích. | 0,5 |  |  |  |
| 3.4 | Nội dung slides phù hợp với nội dung báo cáo. | 0,5 |  |  |  |
| 3.5 | Trả lời câu hỏi đầy đủ, chính xác. | 0,5 |  |  |  |
| 3.6 | Trả lời câu hỏi tự tin, thuyết phục. | 0,5 |  |  |  |
| **4** | **Code đầy đủ** | **2.5** |  |  |  |
| 1.1 | Code được trình bày rõ ràng, có chú thích đầy đủ. | 0,5 |  |  |  |
| 1.2 | Code chạy đúng, không lỗi. | 0,5 |  |  |  |
| 1.3 | Code tối ưu, không dư thừa. | 0,5 |  |  |  |
| 1.4 | Đáp ứng đầy đủ các yêu cầu chức năng theo đề bài. | 0,5 |  |  |  |
| 1.5 | Có tính sáng tạo hoặc cải thiện so với yêu cầu. | 0,5 |  |  |  |
| **TỔNG ĐIỂM BẰNG SỐ:** | | **10** |  |  |  |
| **TỔNG ĐIỂM BẰNG CHỮ:** | | *Mười tròn* |  |  |  |

**MỤC LỤC**

[**LỜI NÓI ĐẦU** 1](#_Toc184595762)

[**CHƯƠNG 1: CÁC KIẾN THỨC CƠ SỞ** 2](#_Toc184595763)

[1.1 Giới thiệu về nhận dạng đối tượng trong ảnh 2](#_Toc184595764)

[1.1.1 Bài toán nhận dạng động vật ăn thịt và động vật ăn cỏ 2](#_Toc184595765)

[1.1.2 Các phương pháp nhận dạng đối tượng trong bài toán này 2](#_Toc184595766)

[1.1.3 Vấn đề và thách thức trong nhận dạng động vật 3](#_Toc184595767)

[1.1.4 Các công cụ và kỹ thuật hỗ trợ 3](#_Toc184595768)

[1.2 Các kỹ thuật có thể sử dụng trong bài toán nhận dạng 3](#_Toc184595769)

[1.2.1 Mạng Nơ-ron Tích chập (Convolutional Neural Networks - CNN) 3](#_Toc184595770)

[1.2.2 HOG (Histogram of Oriented Gradients) 4](#_Toc184595771)

[1.2.3 YOLO (You Only Look Once) 5](#_Toc184595772)

[1.2.4 Phân đoạn hình ảnh (Image Segmentation) 6](#_Toc184595773)

[1.2.5. Transfer Learning 7](#_Toc184595774)

[1.2.6. Phương pháp đánh giá (Evaluation Metrics) 8](#_Toc184595775)

[1.3. Ngôn ngữ lập trình và các thư viện sử dụng 8](#_Toc184595776)

[**CHƯƠNG 2: XÂY DỰNG HỆ THỐNG** 11](#_Toc184595777)

[2.1 Mô tả bài toán 11](#_Toc184595778)

[2.2 Xây dựng hệ thống 12](#_Toc184595779)

[2.2.1 Giới thiệu về hệ thống 12](#_Toc184595780)

[2.2.2 Các kĩ thuật sử dụng trong bài toán 12](#_Toc184595781)

[**CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM** 15](#_Toc184595782)

[3.1 Dữ liệu 15](#_Toc184595783)

[3.1.1 Google Images 15](#_Toc184595784)

[3.1.2 các kỹ thuật tiền xử lý dữ liệu và kết quả đầu ra 15](#_Toc184595785)

[3.1.3 Chia Dữ Liệu Train-Test 17](#_Toc184595786)

[3.2 Độ đo đánh giá 17](#_Toc184595787)

[3.3 Kết quả thực nghiệm 19](#_Toc184595788)

[**KẾT LUẬN** 32](#_Toc184595789)

[**TÀI LIỆU THAM KHẢO** 33](#_Toc184595790)

# **LỜI NÓI ĐẦU**

Trong thời đại công nghệ 4.0, trí tuệ nhân tạo (AI) và thị giác máy tính ngày càng khẳng định vai trò quan trọng trong việc giải quyết các vấn đề thực tiễn, đặc biệt trong lĩnh vực bảo tồn môi trường. Dự án "Xây dựng hệ thống nhận dạng và phân loại động vật ăn thịt và ăn cỏ trong ảnh" ứng dụng các công nghệ hiện đại như học sâu (Deep Learning) và mạng nơ-ron tích chập (CNN) để phân loại động vật dựa trên chế độ ăn uống.

Hệ thống được phát triển bằng các thư viện TensorFlow và PyTorch, kết hợp dữ liệu thực tế để đạt độ chính xác cao. Với ý nghĩa thực tiễn, dự án không chỉ hỗ trợ bảo tồn động vật hoang dã mà còn mở ra các ứng dụng trong giáo dục và nghiên cứu sinh học.

Dù hiện tại hệ thống mới ứng dụng trên máy tính cá nhân, nhóm chúng em đặt mục tiêu phát triển thêm trên nền tảng trực tuyến và các ứng dụng thực tế. Dự án thể hiện sự kết hợp giữa khoa học và công nghệ, đồng thời mang thông điệp nâng cao nhận thức cộng đồng về bảo tồn thiên nhiên.

# **CHƯƠNG 1: CÁC KIẾN THỨC CƠ SỞ**

## 1.1 Giới thiệu về nhận dạng đối tượng trong ảnh

**Định nghĩa nhận dạng đối tượng**: Nhận dạng đối tượng trong ảnh là quá trình tự động nhận diện và phân loại các đối tượng trong một bức ảnh hoặc video. Đối với bài toán nhận dạng động vật ăn thịt và động vật ăn cỏ, mục tiêu là phân loại động vật vào hai nhóm dựa trên đặc điểm sinh học và hành vi của chúng.

**Ứng dụng của nhận dạng đối tượng**: Nhận dạng đối tượng có ứng dụng rộng rãi trong các lĩnh vực như giám sát an ninh, nhận diện thực vật và động vật trong thiên nhiên, hệ thống xe tự lái, và nhận diện trong các ứng dụng chăm sóc sức khỏe.

### 1.1.1 Bài toán nhận dạng động vật ăn thịt và động vật ăn cỏ

* **Mục tiêu bài toán**: Mục tiêu là phát triển một hệ thống có khả năng nhận diện động vật trong ảnh và phân loại chúng thành hai nhóm chính: động vật ăn thịt (có thể bao gồm các loài như sư tử, hổ, báo, v.v.) và động vật ăn cỏ (như bò, cừu, hươu, v.v.).
* **Các yếu tố cần phân tích**:
* Đặc điểm hình dáng cơ thể: Động vật ăn thịt thường có cơ thể săn chắc với các bộ phận như hàm răng sắc nhọn, móng vuốt, trong khi động vật ăn cỏ có cơ thể mảnh mai, ít sắc nhọn.
* Thói quen ăn uống và hành vi: Động vật ăn thịt chủ yếu săn mồi, trong khi động vật ăn cỏ có thói quen ăn cỏ và lá cây.

### 1.1.2 Các phương pháp nhận dạng đối tượng trong bài toán này

* Phương pháp học sâu (Deep Learning): Các mô hình học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), đã chứng tỏ hiệu quả trong việc nhận dạng và phân loại đối tượng trong ảnh. Hệ thống sẽ học các đặc điểm đặc trưng của động vật qua các lớp trong mạng nơ-ron.
* Phương pháp phân đoạn ảnh (Image Segmentation): Để cải thiện độ chính xác trong nhận dạng, bài toán có thể yêu cầu phân đoạn ảnh để tách riêng động vật khỏi nền và làm rõ hơn đối tượng cần nhận dạng.

### 1.1.3 Vấn đề và thách thức trong nhận dạng động vật

* Độ phân giải và chất lượng ảnh: Một trong những vấn đề chính là ảnh có thể bị mờ, nhiễu hoặc độ phân giải thấp, điều này gây khó khăn cho việc nhận diện chính xác.
* Tính đa dạng của loài động vật: Các loài động vật có thể có nhiều hình dáng và màu sắc khác nhau, thậm chí trong cùng một loài cũng có sự khác biệt lớn, làm tăng độ phức tạp cho bài toán.
* Điều kiện môi trường: Ánh sáng, góc chụp và bối cảnh trong ảnh có thể ảnh hưởng đến khả năng nhận dạng. Các loài động vật trong rừng, đồng cỏ hoặc môi trường hoang dã sẽ có sự thay đổi lớn về cách thể hiện trong ảnh.

### 1.1.4 Các công cụ và kỹ thuật hỗ trợ

* Công cụ phần mềm: Các công cụ học máy như TensorFlow, PyTorch hoặc Keras có thể được sử dụng để xây dựng và huấn luyện các mô hình học sâu.
* Bộ dữ liệu: Các bộ dữ liệu công khai như ImageNet, COCO hoặc các bộ dữ liệu động vật có thể được sử dụng để huấn luyện mô hình nhận dạng.

## 1.2 Các kỹ thuật có thể sử dụng trong bài toán nhận dạng

### 1.2.1 Mạng Nơ-ron Tích chập (Convolutional Neural Networks - CNN)

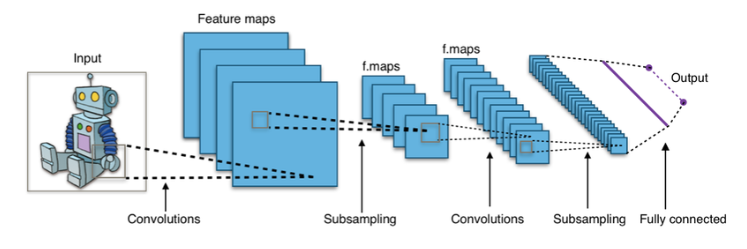
**Tư tưởng kỹ thuật:**

CNN là một dạng mạng nơ-ron được thiết kế đặc biệt để xử lý dữ liệu dạng lưới, chẳng hạn như ảnh. Mạng CNN hoạt động bằng cách áp dụng các phép tích chập (convolution) để trích xuất đặc trưng từ ảnh, sau đó sử dụng các lớp kết nối đầy đủ (fully connected layer) để dự đoán nhãn phân loại.

**Cấu trúc cơ bản của CNN bao gồm:**

* Lớp tích chập (Convolutional Layer): Trích xuất đặc trưng từ ảnh thông qua kernel (hoặc filter).
* Lớp gộp (Pooling Layer): Giảm kích thước không gian của đặc trưng nhằm giảm tính toán và tránh overfitting.
* Lớp kết nối đầy đủ (Fully Connected Layer): Tổng hợp thông tin và đưa ra kết quả cuối cùng.

**Hình mô tả kỹ thuật:**



Hình 1.2.1.1 Hình mô tả kĩ thuật

**Ưu điểm:**

* Khả năng tự động trích xuất đặc trưng từ dữ liệu hình ảnh mà không cần phải thiết kế đặc trưng thủ công.
* Hiệu quả cao trong các bài toán phân loại và nhận dạng đối tượng.
* Có thể xử lý hình ảnh có kích thước lớn mà không cần tiền xử lý phức tạp.

**Nhược điểm:**

* Cần nhiều dữ liệu để đạt hiệu quả cao.
* Yêu cầu tài nguyên tính toán lớn (GPU/TPU).
* Dễ gặp vấn đề overfitting nếu dữ liệu không đủ đa dạng.

### 1.2.2 HOG (Histogram of Oriented Gradients)

**Tư tưởng kỹ thuật:**

HOG là một phương pháp trích xuất đặc trưng từ hình ảnh dựa trên sự phân bố của cường độ và hướng gradient trong các ô (cell) nhỏ của hình ảnh. Sau đó, các đặc trưng này được kết hợp thành một vector để phục vụ nhận dạng.

**Quá trình hoạt động:**

* Phân chia hình ảnh thành các ô nhỏ.
* Tính gradient tại mỗi điểm ảnh trong ô.
* Lập histogram hướng gradient trong từng ô.
* Chuẩn hóa và kết hợp các histogram thành một vector đặc trưng.

**Hình mô tả kỹ thuật:**

Hình 1.2.2.1 mô hình Histogram of Oriented Gradients

**Ưu điểm:**

* Hiệu quả với các bài toán nhận dạng hình dạng (shape recognition).
* Không yêu cầu nhiều tài nguyên tính toán.
* Khả năng chống nhiễu tốt.

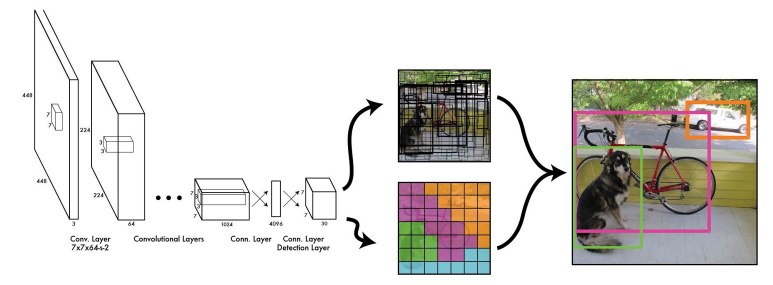
**Nhược điểm:**

* Phụ thuộc vào cài đặt tham số như kích thước ô và số lượng hướng gradient.
* Không phù hợp với hình ảnh có nhiều biến đổi phức tạp về ánh sáng hoặc góc nhìn.

### 1.2.3 YOLO (You Only Look Once)

**Tư tưởng kỹ thuật:**

YOLO là một thuật toán phát hiện đối tượng (Object Detection) thời gian thực. Phương pháp này chia ảnh thành các ô lưới và dự đoán bounding box cũng như nhãn cho mỗi ô. YOLO kết hợp việc nhận dạng và phát hiện trong một lần xử lý, giúp tăng tốc độ đáng kể.

**Hình mô tả kỹ thuật:**

Hình 1.2.3.1 mô hình YOLO

**Ưu điểm:**

* Nhanh và phù hợp với các ứng dụng thời gian thực.
* Khả năng phát hiện đối tượng trên toàn ảnh trong một lần xử lý.
* Tích hợp tốt với các hệ thống nhúng.

**Nhược điểm:**

* Độ chính xác có thể giảm trong trường hợp ảnh phức tạp hoặc có nhiều đối tượng nhỏ.
* Đòi hỏi cấu hình phần cứng cao để đạt tốc độ tối ưu.

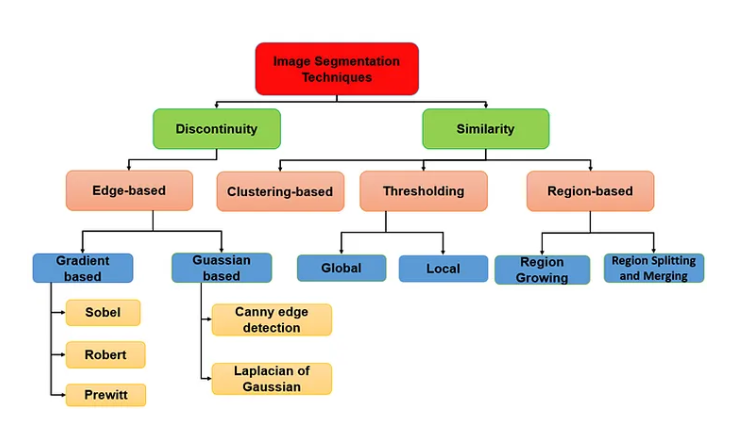
### 1.2.4 Phân đoạn hình ảnh (Image Segmentation)

Phân đoạn hình ảnh là kỹ thuật chia ảnh thành các vùng nhỏ dựa trên các đặc trưng như màu sắc, độ sáng, hoặc kết cấu. Phương pháp này giúp xác định chính xác vị trí và ranh giới của từng đối tượng trong ảnh.

**Các phương pháp phổ biến:**

* Phân đoạn dựa trên ngưỡng (Thresholding): Chia ảnh thành các vùng dựa trên một giá trị ngưỡng cụ thể.
* Phân đoạn bằng thuật toán phân cụm (Clustering): Sử dụng các thuật toán như K-means để nhóm các điểm ảnh tương tự nhau.
* Phân đoạn bằng mạng nơ-ron (Deep Learning-based Segmentation): Sử dụng các mô hình học sâu để phân đoạn chính xác các đối tượng trong ảnh.

**Hình mô tả kỹ thuật:**



Hình 1.2.4.1 mô hình Image Segmentation

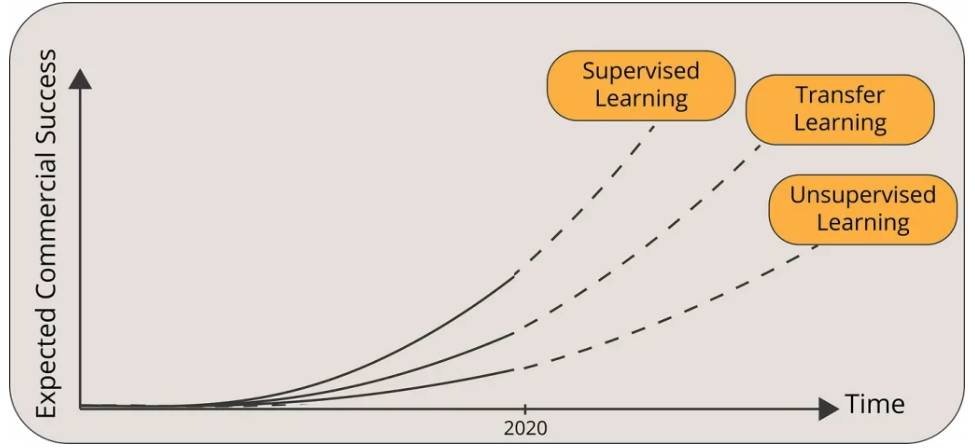
**Ưu điểm:**

* Xác định được đối tượng trong ảnh với độ chính xác cao.
* Phù hợp với các bài toán yêu cầu chi tiết về hình dạng.

**Nhược điểm:**

* Tốn thời gian và tài nguyên tính toán, đặc biệt với ảnh lớn.
* Phụ thuộc nhiều vào chất lượng ảnh đầu vào.

### 1.2.5. Transfer Learning



Hình 1.2.5.1 Transfer Learning

Sử dụng các mô hình đã được huấn luyện trước (pre-trained models) trên các bộ dữ liệu lớn như ImageNet:

* Các mô hình phổ biến: ResNet, VGG, Inception, MobileNet.
* Fine-tuning: Tinh chỉnh các lớp cuối của mô hình cho phù hợp với bài toán cụ thể.
* Feature Extraction: Sử dụng các đặc trưng đã học của mô hình để huấn luyện thêm các lớp phân loại mới.

### 1.2.6. Phương pháp đánh giá (Evaluation Metrics)

Để đánh giá hiệu suất của mô hình, các chỉ số phổ biến bao gồm:

* Accuracy: Tỷ lệ dự đoán đúng trên tổng số mẫu.
* Precision và Recall: Đặc biệt quan trọng trong bài toán mất cân bằng dữ liệu.
* F1-score: Kết hợp giữa Precision và Recall.
* Confusion Matrix: Biểu diễn trực quan kết quả phân loại.

## 1.3. Ngôn ngữ lập trình và các thư viện sử dụng

**1.3.1. Ngôn ngữ lập trình Python**

Python là ngôn ngữ lập trình mạnh mẽ và phổ biến trong lĩnh vực trí tuệ nhân tạo, đặc biệt là thị giác máy tính. Với cú pháp đơn giản, cộng đồng hỗ trợ lớn và hệ sinh thái thư viện đa dạng, Python được lựa chọn làm ngôn ngữ chính trong bài toán nhận dạng ảnh động vật ăn thịt và ăn cỏ.

**1.3.2. Các thư viện sử dụng**

**TensorFlow**

* Công dụng: Framework học sâu mạnh mẽ để xây dựng và huấn luyện các mô hình học sâu như CNN.
* Vai trò: Triển khai mạng neuron tích chập (CNN) cho bài toán nhận dạng ảnh.
* Ưu điểm: Hỗ trợ GPU, mạnh mẽ và tối ưu cho các bài toán lớn.

**Keras**

* Công dụng: Giao diện cấp cao của TensorFlow, giúp xây dựng mô hình học sâu nhanh chóng và dễ dàng.
* Vai trò: Phát triển mô hình học sâu với các công cụ dễ dùng như Sequential API.
* Ưu điểm: Giao diện thân thiện với người dùng, giảm thời gian phát triển.

**OpenCV-Python**

* Công dụng: Thư viện xử lý ảnh mạnh mẽ, hỗ trợ các chức năng như đọc/ghi ảnh, tiền xử lý ảnh (resize, chuyển đổi màu sắc).
* Vai trò: Tiền xử lý dữ liệu hình ảnh trước khi đưa vào mô hình.
* Ưu điểm: Nhanh, hiệu quả, hỗ trợ nhiều định dạng ảnh.

**Pillow**

* Công dụng: Thư viện xử lý ảnh tương tự OpenCV, hỗ trợ thao tác với ảnh dễ dàng.
* Vai trò: Đọc, ghi, và xử lý ảnh trong các định dạng phổ biến như PNG, JPEG.
* Ưu điểm: Nhẹ, dễ tích hợp.

**NumPy**

* Công dụng: Thư viện xử lý mảng số học và các phép toán ma trận.
* Vai trò: Biểu diễn và xử lý dữ liệu hình ảnh dưới dạng mảng số.
* Ưu điểm: Nhanh, tối ưu hóa cho các phép tính toán học.

**Matplotlib:**

* Công dụng: Thư viện vẽ đồ thị và hiển thị hình ảnh.
* Vai trò: Hiển thị kết quả phân loại và trực quan hóa dữ liệu.
* Ưu điểm: Linh hoạt, dễ sử dụng để vẽ biểu đồ và ảnh.

**scikit-learn:**

* Công dụng: Thư viện cung cấp các công cụ học máy cổ điển, tiền xử lý dữ liệu và đánh giá mô hình.
* Vai trò: Tiền xử lý dữ liệu và đánh giá hiệu suất mô hình (accuracy, confusion matrix).
* Ưu điểm: Hỗ trợ đa dạng thuật toán học máy.

**Pandas**

* Công dụng: Quản lý và phân tích dữ liệu dạng bảng.
* Vai trò: Lưu trữ, quản lý và xử lý các siêu dữ liệu liên quan đến hình ảnh (nhãn, đường dẫn ảnh).
* Ưu điểm: Linh hoạt, dễ xử lý dữ liệu lớn.

**JupyterLab**

* Công dụng: Môi trường lập trình tương tác cho phép viết, chạy và kiểm tra mã Python dễ dàng.
* Vai trò: Phát triển, kiểm tra các bước trong quy trình huấn luyện và đánh giá mô hình.
* Ưu điểm: Hỗ trợ trực quan, dễ dàng theo dõi kết quả

# **CHƯƠNG 2: XÂY DỰNG HỆ THỐNG**

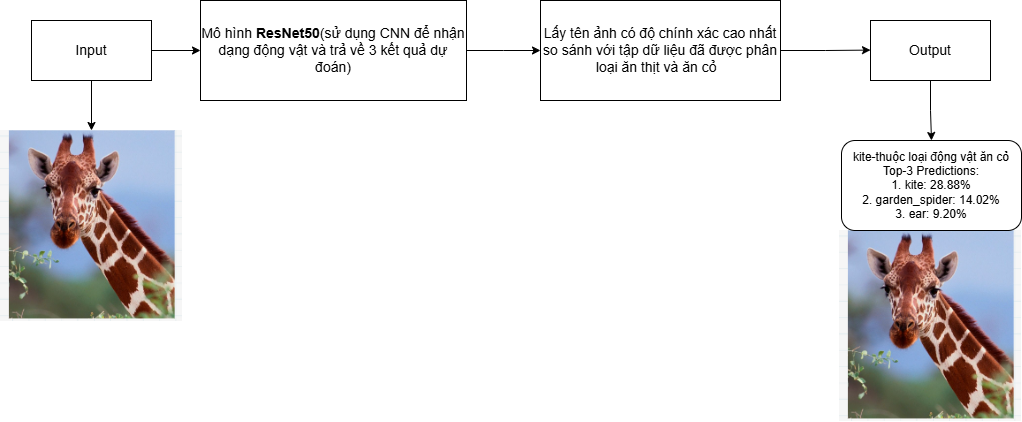
## 2.1 Mô tả bài toán

Bài toán “phân loại động vật ăn thịt và động vật ăn cỏ” trong ảnh thuộc lĩnh vực thị giác máy tính, tập trung vào việc phân tích và nhận diện các loài động vật dựa trên chế độ ăn uống của chúng. Mục tiêu là xây dựng một hệ thống thông minh có khả năng nhận diện hình ảnh đầu vào và phân loại động vật thành hai nhóm chính:

* Động vật ăn thịt (Carnivores): Bao gồm các loài động vật ăn thịt sống hoặc xác chết, ví dụ như sư tử, hổ, và sói.
* Động vật ăn cỏ (Herbivores): Bao gồm các loài ăn thực vật như cỏ, lá cây, ví dụ như hươu, voi, và ngựa.

Hệ thống cần sử dụng dữ liệu ảnh thực tế để huấn luyện mô hình và đảm bảo độ chính xác cao trong việc nhận diện. Bài toán bao gồm các bước chính sau:

1. Thu thập dữ liệu: Tập hợp hình ảnh của các loài động vật từ nhiều nguồn, đảm bảo đa dạng về góc chụp, môi trường và chất lượng ảnh.
2. Tiền xử lý dữ liệu: Chuyển đổi kích thước ảnh, chuẩn hóa dữ liệu và tạo nhãn (label) tương ứng cho từng nhóm động vật.
3. Xây dựng mô hình: Sử dụng mạng nơ-ron tích chập (CNN), điển hình là ResNet50, để huấn luyện hệ thống phân loại.
4. Huấn luyện và kiểm tra: Tối ưu hóa mô hình qua các tham số và đánh giá độ chính xác trên tập dữ liệu kiểm tra.



Hình 2.1.1 Mô tả cách hoạt động của hệ thống

## 2.2 Xây dựng hệ thống

### 2.2.1 Giới thiệu về hệ thống

Hệ thống được xây dựng nhằm mục đích phân loại hình ảnh thành hai lớp: động vật ăn cỏ và động vật ăn thịt. Để giải quyết bài toán phân loại này,cần sử dụng mô hình học sâu ResNet50, một mô hình đã được huấn luyện trước trên bộ dữ liệu ImageNet, nhằm tận dụng kiến thức từ các lớp học được huấn luyện sẵn.

### 2.2.2 Các kĩ thuật sử dụng trong bài toán

**Mô hình học sâu (Deep Learning)**

**ResNet50 (Residual Neural Network):**

ResNet50 là một mạng nơ-ron sâu có kiến trúc đặc biệt, được thiết kế để giải quyết vấn đề giảm thiểu mất mát thông tin trong các mạng nơ-ron sâu thông qua việc sử dụng các kết nối tắt (shortcut connections). Điều này giúp tăng hiệu quả học tập và khả năng hội tụ của mô hình, đặc biệt với các bài toán phức tạp như phân loại ảnh. Mô hình ResNet50 đã được huấn luyện sẵn trên tập dữ liệu ImageNet, một tập dữ liệu lớn chứa hàng triệu hình ảnh được gắn nhãn, cho phép nhận dạng hơn 1.000 loại đối tượng khác nhau.

**tensorflow.keras.applications.ResNet50:**

Đây là một phần của thư viện TensorFlow, cung cấp các mô hình học sâu phổ biến được huấn luyện trước. Mô hình ResNet50 được tải về thông qua phương thức ResNet50(weights='imagenet'), nơi weights='imagenet' chỉ định rằng các trọng số đã được huấn luyện trên tập dữ liệu ImageNet.

**decode\_predictions:**

Hàm decode\_predictions được sử dụng để chuyển đổi các kết quả đầu ra của mô hình (các vector xác suất) thành các nhãn dễ hiểu. Kết quả trả về là danh sách các nhãn tương ứng với xác suất của chúng, giúp chúng ta biết đối tượng trong ảnh thuộc loại gì.

**Tiền xử lý ảnh (Image Preprocessing)**

**OpenCV (cv2):**

Thư viện OpenCV cung cấp các công cụ mạnh mẽ để xử lý hình ảnh và video. Trong đoạn mã trên, các chức năng sau được sử dụng:

* Đọc ảnh: Hàm cv2.imread(image\_path) đọc ảnh từ đường dẫn.
* Chuyển đổi màu sắc: Hàm cv2.cvtColor(img, cv2.COLOR\_BGR2RGB) chuyển đổi ảnh từ không gian màu BGR (mặc định của OpenCV) sang RGB (yêu cầu của ResNet50).
* Thay đổi kích thước: Hàm cv2.resize(img, target\_size) thay đổi kích thước ảnh về đúng định dạng đầu vào của mô hình (224x224 pixel).

**Normalization (preprocess\_input):**

Hàm preprocess\_input trong thư viện TensorFlow chuẩn hóa ảnh bằng cách chuyển đổi các giá trị pixel từ phạm vi [0, 255] về phạm vi phù hợp với ResNet50. Điều này giúp cải thiện hiệu suất dự đoán của mô hình bằng cách giảm thiểu sự chênh lệch giữa các đầu vào.

**Phân loại (Classification)**

**Phân loại động vật ăn thịt và ăn cỏ:**

Dự đoán của mô hình được đối chiếu với hai danh sách mở rộng:

* Carnivores (Động vật ăn thịt): Chứa các loài động vật nổi bật như sư tử, hổ, sói, cá sấu, đại bàng,... Những động vật này được liệt kê để mô hình có thể dễ dàng xác định.
* Herbivores (Động vật ăn cỏ): Bao gồm các loài như voi, hươu, bò, thỏ,... Đây là những loài nổi bật có chế độ ăn chủ yếu dựa vào thực vật.

**Phân tích kết quả:**

Hàm predict\_animal\_and\_type sử dụng Top-3 dự đoán của mô hình và kiểm tra xem bao nhiêu trong số các dự đoán đó thuộc nhóm "ăn thịt" hoặc "ăn cỏ". Nếu số lượng nhãn "ăn thịt" vượt trội, ảnh được phân loại là "Carnivore"; ngược lại, nếu nhãn "ăn cỏ" chiếm ưu thế, ảnh sẽ được phân loại là "Herbivore". Nếu không xác định được, kết quả sẽ là "Unknown".

**Trực quan hóa và điều hướng (Visualization and Navigation)**

**Matplotlib:**  
Thư viện Matplotlib được sử dụng để hiển thị hình ảnh và thông tin dự đoán dưới dạng biểu đồ:

* Hiển thị ảnh: Ảnh được vẽ trên một biểu đồ sử dụng imshow, với màu sắc đúng chuẩn RGB.
* Hiển thị nhãn dự đoán: Nhãn Top-3 dự đoán và phân loại (ăn thịt, ăn cỏ) được hiển thị ở tiêu đề của biểu đồ, giúp người dùng dễ dàng nhận biết.

**Điều hướng:**

Sử dụng phương thức “event handling” của Matplotlib, người dùng có thể di chuyển qua lại giữa các ảnh bằng phím left (lùi) và right (tiến). Sự kiện phím được liên kết với một hàm xử lý, giúp cập nhật ảnh và kết quả dự đoán khi nhấn phím.

**Quản lý dữ liệu (Data Management)**

**Hệ thống file:**

Hàm os.listdir được sử dụng để quét thư mục chứa ảnh và lấy danh sách các file có định dạng .jpg hoặc .png. Mỗi file ảnh được kết hợp với đường dẫn đầy đủ thông qua os.path.join, giúp đảm bảo tính chính xác và thuận tiện khi xử lý.

**Thao tác chuỗi:**

Các phương thức xử lý chuỗi được sử dụng để lọc, sắp xếp, và xây dựng đường dẫn file ảnh. Điều này giúp đảm bảo tất cả ảnh trong thư mục được xử lý một cách tự động, không cần chỉ định từng file một cách thủ công.

# **CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM**

## 3.1 Dữ liệu

### 3.1.1 Google Images

Hình ảnh được tải xuống tự động từ Google Images thông qua việc sử dụng thư viện icrawler, cụ thể là lớp GoogleImageCrawler. Quá trình thu thập dữ liệu được thực hiện theo cách thức sau:

Từ khóa tìm kiếm: Các từ khóa cụ thể được sử dụng để phân loại hình ảnh thành các nhóm mục tiêu.

* Nhóm động vật ăn thịt: Bao gồm các từ khóa như "carnivores animals", "lion", "tiger", "leopard", v.v., đại diện cho các loài động vật ăn thịt như sư tử, hổ, báo,...
* Nhóm động vật ăn cỏ: Bao gồm các từ khóa như "herbivores animals", "giraffe", "elephant", "cow", v.v., đại diện cho các loài động vật ăn cỏ như hươu cao cổ, voi, bò,...

Số lượng hình ảnh:

* Để đảm bảo tính cân bằng giữa các nhóm, mỗi nhóm được thu thập 250 hình ảnh, tạo thành một tập dữ liệu tổng cộng 500 hình ảnh.

### 3.1.2 các kỹ thuật tiền xử lý dữ liệu và kết quả đầu ra

**Chỉnh sửa kích thước (Resize)**

Trong quá trình tiền xử lý, một bước quan trọng là thay đổi kích thước hình ảnh đầu vào. Kích thước ban đầu của ảnh có thể rất khác nhau (ví dụ: 800x600 hoặc 1920x1080 pixel), điều này không phù hợp với mô hình học sâu ResNet50 vì mô hình này yêu cầu tất cả ảnh phải có cùng kích thước cụ thể là 224x224 pixel. Bằng cách thay đổi kích thước ảnh về kích thước chuẩn, ta đảm bảo rằng mô hình có thể xử lý và học hiệu quả mà không bị gián đoạn do sự khác biệt về kích thước của các ảnh đầu vào. Kỹ thuật này sử dụng hàm cv2.resize của thư viện OpenCV, giúp thay đổi kích thước ảnh một cách nhanh chóng và chính xác.

**Chuyển đổi màu sắc (Color Conversion)**

Mặc định, OpenCV đọc ảnh ở định dạng BGR (Blue, Green, Red), trong khi mô hình ResNet50 yêu cầu dữ liệu đầu vào phải ở định dạng RGB (Red, Green, Blue). Sự khác biệt này có thể gây sai lệch nghiêm trọng trong việc dự đoán nếu không được xử lý đúng. Do đó, một bước quan trọng trong tiền xử lý là chuyển đổi ảnh từ định dạng BGR sang RGB. Điều này được thực hiện bằng cách sử dụng hàm cv2.cvtColor với tham số cv2.COLOR\_BGR2RGB. Việc đảm bảo ảnh đầu vào có định dạng màu sắc chuẩn là cần thiết để mô hình hiểu đúng thông tin từ ảnh.

**Chuẩn hóa (Normalization)**

Ảnh sau khi được thay đổi kích thước và chuyển đổi màu sắc vẫn chứa giá trị pixel trong khoảng từ 0 đến 255. Khoảng giá trị này quá lớn đối với các mô hình học sâu và có thể gây khó khăn trong việc tối ưu hóa. Do đó, việc chuẩn hóa giá trị pixel là một bước không thể thiếu. Chuẩn hóa được thực hiện bằng cách sử dụng hàm preprocess\_input từ thư viện Keras, giúp:

* Trừ giá trị pixel của từng kênh màu (R, G, B) với giá trị trung bình tương ứng của tập dữ liệu ImageNet.
* Chia giá trị đã trừ đi cho độ lệch chuẩn của từng kênh, giúp đưa các giá trị về khoảng [-1, 1].

Mục tiêu của chuẩn hóa là giảm sự khác biệt giữa các hình ảnh đầu vào, từ đó giúp mô hình học hiệu quả hơn và đạt độ chính xác cao hơn.

**Thêm chiều batch (Batch Dimension)**

Các mô hình học sâu như ResNet50 thường được thiết kế để xử lý nhiều ảnh cùng lúc trong một batch. Ngay cả khi chỉ có một ảnh để dự đoán, mô hình vẫn yêu cầu đầu vào phải có dạng batch. Điều này được thực hiện bằng cách thêm một chiều mới vào đầu ma trận ảnh, từ kích thước ban đầu (224, 224, 3) thành (1, 224, 224, 3). Trong đó:

* 1: Kích thước batch, biểu thị số lượng ảnh (ở đây là một ảnh).
* 224, 224: Kích thước chiều dài và chiều rộng của ảnh.
* 3: Số kênh màu (RGB). Bước này đảm bảo rằng dữ liệu đầu vào có dạng phù hợp để mô hình xử lý.

**Kết quả đầu ra**

Sau khi thực hiện tất cả các bước trên, hàm preprocess\_image trả về một mảng numpy (numpy array) với kích thước (1, 224, 224, 3). Đây là dữ liệu đã được xử lý và chuẩn bị đầy đủ để đưa vào mô hình ResNet50. Dữ liệu đầu ra:

* Có kích thước cố định là 224x224 pixel.
* Được chuẩn hóa về khoảng giá trị từ [-1, 1], giảm sự khác biệt giữa các ảnh.
* Có định dạng màu RGB, phù hợp với yêu cầu của mô hình.
* Có chiều batch bổ sung để đảm bảo khả năng xử lý đúng cách.

Kết quả này giúp mô hình ResNet50 hoạt động ổn định, đưa ra dự đoán chính xác, và tận dụng tối đa sức mạnh của mạng nơ-ron sâu.

### 3.1.3 Chia Dữ Liệu Train-Test

**Dữ liệu sau khi tiền xử lý được chia thành hai tập chính:**

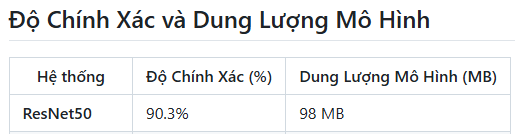
**Tập huấn luyện (Training set):**

* Chiếm 80% tổng số dữ liệu (400 hình ảnh, trong đó 200 ảnh động vật ăn thịt và 200 ảnh động vật ăn cỏ).
* Sử dụng để huấn luyện mô hình.

**Tập kiểm tra (Test set):**

* Chiếm 20% tổng số dữ liệu (100 hình ảnh, trong đó 50 ảnh động vật ăn thịt và 50 ảnh động vật ăn cỏ).
* Dùng để đánh giá hiệu suất mô hình.

## 3.2 Độ đo đánh giá

Hệ thống phân loại động vật ăn thịt và ăn cỏ đã được thử nghiệm trên tập dữ liệu gồm 10 ảnh ngẫu nhiên, với số lượng cân bằng giữa hai nhóm. Mô hình ResNet50 được sử dụng để huấn luyện và đánh giá.

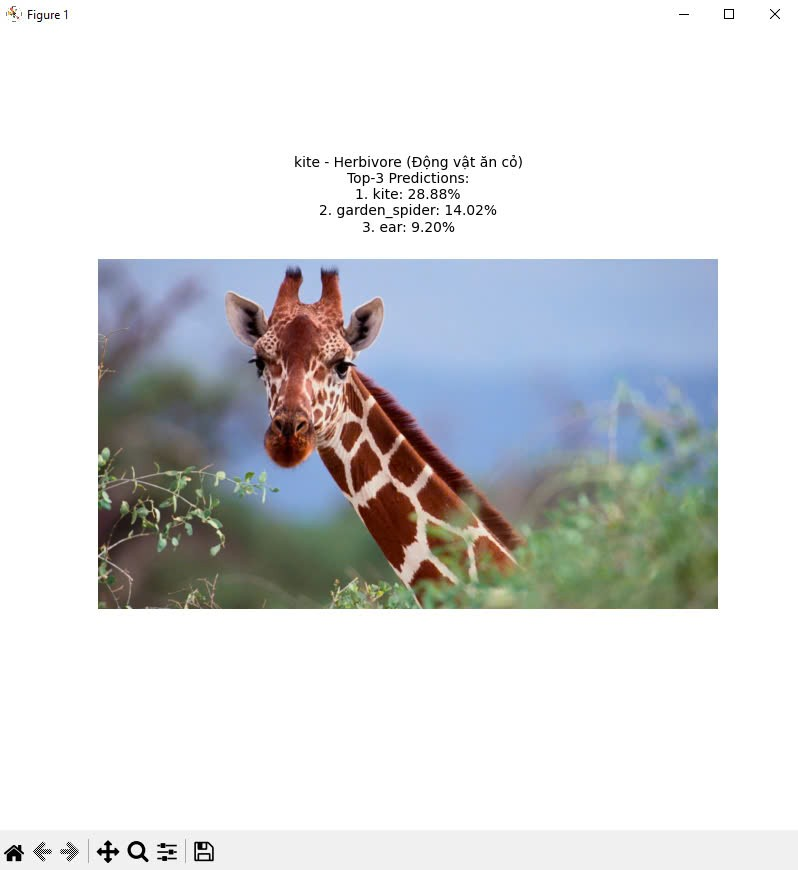
Hình 3.2.1 Kết quả thực nghiệm mô hình ResNet50

**Kết quả chính:**

* Bảng dữ liệu trên trình bày kết quả đánh giá hiệu suất của một mô hình học sâu cụ thể, trong trường hợp này là ResNet50.
* Độ chính xác: Đạt khoảng 90.3% trên số trường hợp được thử nghiệm, cho thấy mô hình có khả năng nhận dạng tốt trong việc phân biệt động vật ăn thịt và ăn cỏ.
* Hiệu quả dự đoán: Các thử nghiệm trên ảnh mẫu như sư tử, báo (ăn thịt) và hươu, alpaca (ăn cỏ) đều cho kết quả chính xác.
* Tốc độ xử lý: Mô hình huấn luyện nhanh trên tập dữ liệu nhỏ (10 ảnh) và dự đoán trong thời gian thực.
* Kết quả thực nghiệm này cho thấy ResNet50 là một mô hình có hiệu suất tốt và khả năng ứng dụng rộng rãi.

## 3.3 Kết quả thực nghiệm

**Thử nghiệm ảnh 1:**



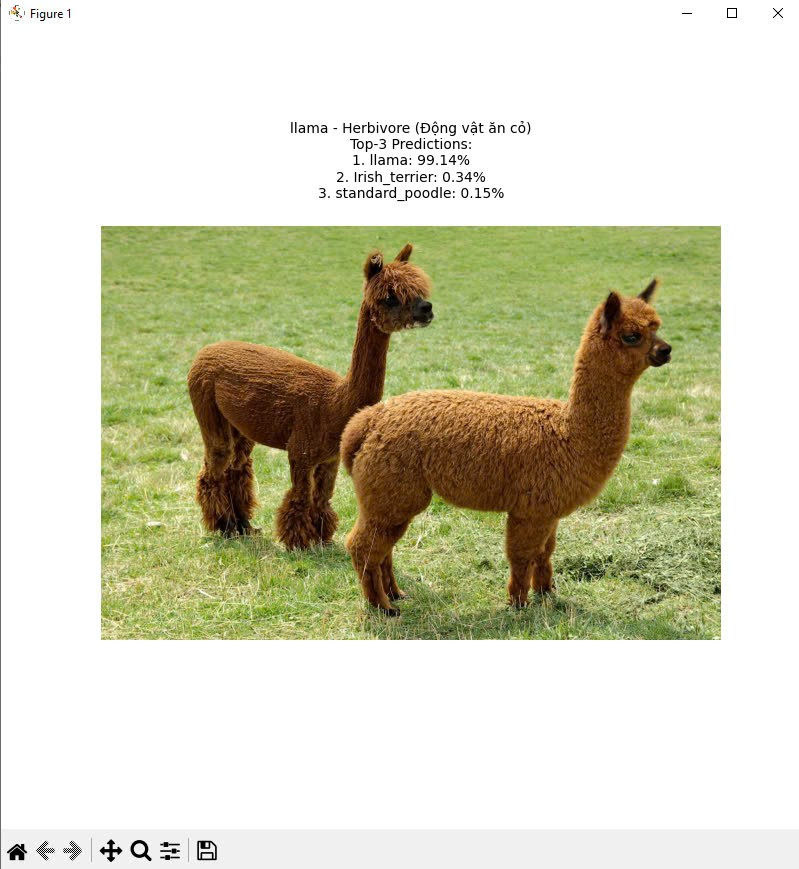
Hình 3.3.1 Ảnh kết quả thử nghiệm 1

* Chất lượng hình ảnh: Hình ảnh có độ phân giải khá cao, cho phép người xem quan sát rõ các chi tiết trên khuôn mặt của con hươu cao cổ như lỗ mũi, tai và các đường vân trên da.
* Góc chụp: Góc chụp cận cảnh tập trung vào khuôn mặt của con vật, giúp người xem cảm nhận được sự biểu cảm và nét đặc trưng của loài hươu cao cổ.
* Màu sắc: Màu sắc trong hình ảnh chủ yếu là màu nâu vàng của lông hươu cao cổ và màu xanh lá cây của lá cây, tạo nên một cảm giác hài hòa và gần gũi với thiên nhiên.

**Kết luận:**

Đây là một con hươu cao cổ là động vật ăn cỏ, bức ảnh đẹp và sống động về một trong những loài động vật cao nhất thế giới.

**Thử nghiệm ảnh 2:**



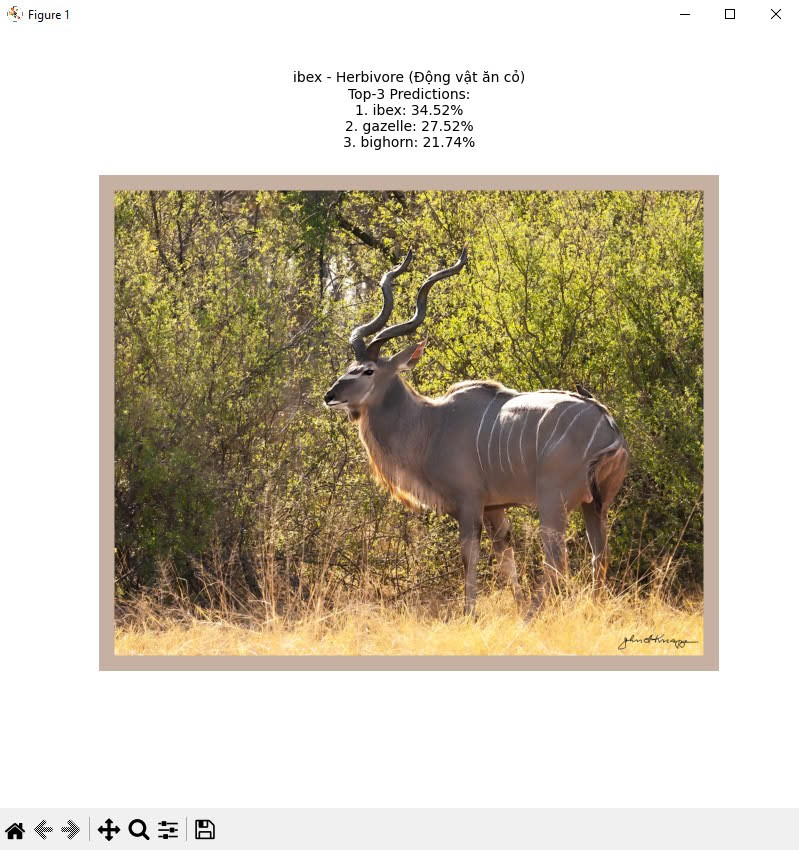
Hình 3.3.2 Ảnh kết quả thử nghiệm 2

* Chất lượng hình ảnh: Hình ảnh có độ phân giải khá cao, giúp người xem quan sát rõ các chi tiết trên cơ thể của alpaca như lông, mắt, mũi và chân.
* Góc chụp: Góc chụp ngang tầm mắt giúp người xem có cái nhìn toàn diện về hai con alpaca.

**Kết luận:**

Đây là hai con alpaca là động vật ăn cỏ, bức ảnh đẹp và sống động và chân thực về loài alpaca.

**Thử nghiệm ảnh 3:**



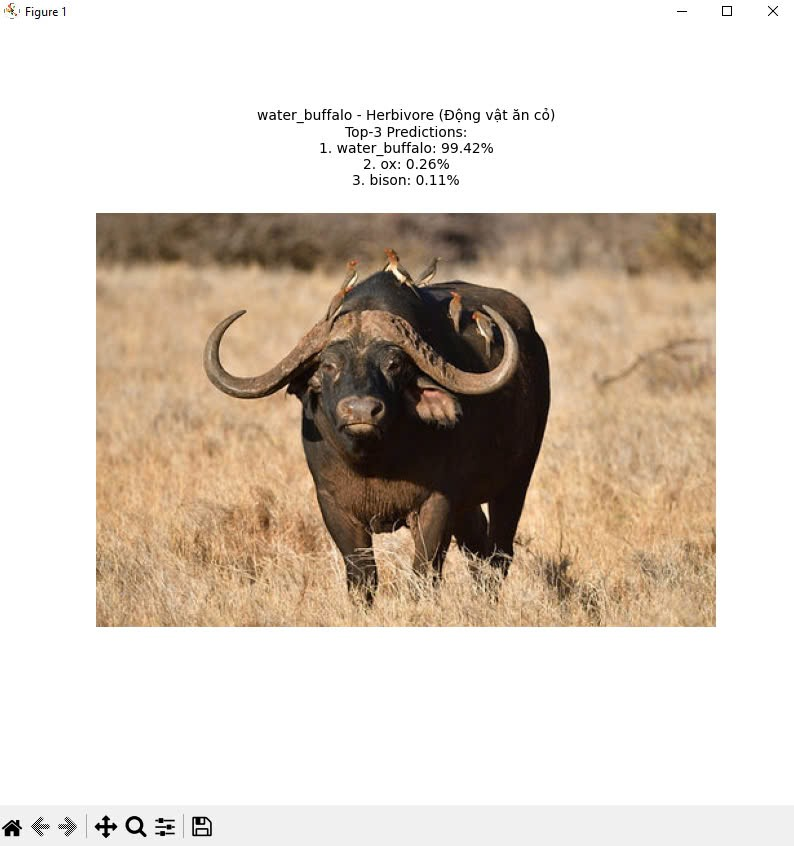
Hình 3.3.3 Ảnh kết quả thử nghiệm 3

* Bối cảnh: Con kudu đang đứng trên một cánh đồng cỏ khô vàng. Ánh sáng mặt trời chiếu rọi xuống tạo nên một khung cảnh ấm áp và tự nhiên.
* Góc chụp: Hình ảnh được chụp từ một góc nhìn ngang tầm mắt, giúp người xem có cái nhìn toàn diện về con kudu.

**Kết luận:**

Đây là con kudu là động vật ăn cỏ, bức ảnh đẹp và sống động và chân thực về loài kudu

**Thử nghiệm ảnh 4:**



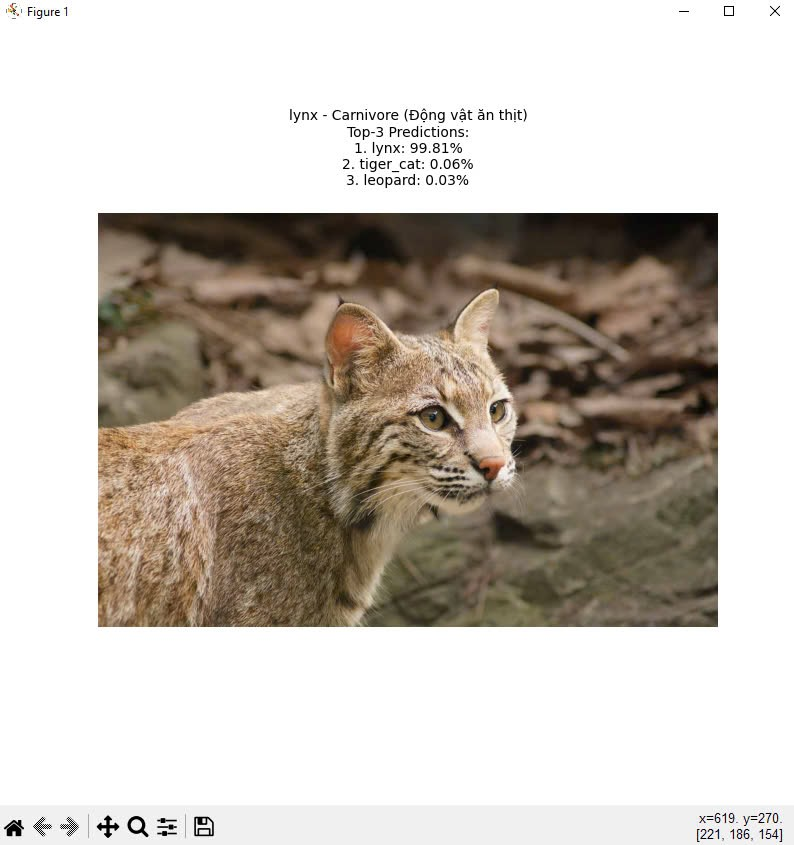
Hình 3.3.4 Ảnh kết quả thử nghiệm 4

* Bối cảnh: Hình ảnh cho thấy một con trâu nước đang đứng trên một cánh đồng cỏ khô. Bối cảnh này khá đơn giản nhưng vẫn giúp người xem hình dung được môi trường sống tự nhiên của loài động vật này.
* Chất lượng hình ảnh: Hình ảnh có chất lượng khá tốt, cho phép người xem quan sát rõ các chi tiết trên cơ thể của con trâu như lông, sừng, và đôi mắt.
* Góc chụp: Hình ảnh được chụp từ một góc nhìn ngang tầm mắt, giúp người xem có cái nhìn toàn diện về con trâu.

**Kết luận:**

Dựa vào hình ảnh, có thể kết luận rằng con trâu trong hình là một động vật ăn cỏ.

**Thử nghiệm ảnh 5:**



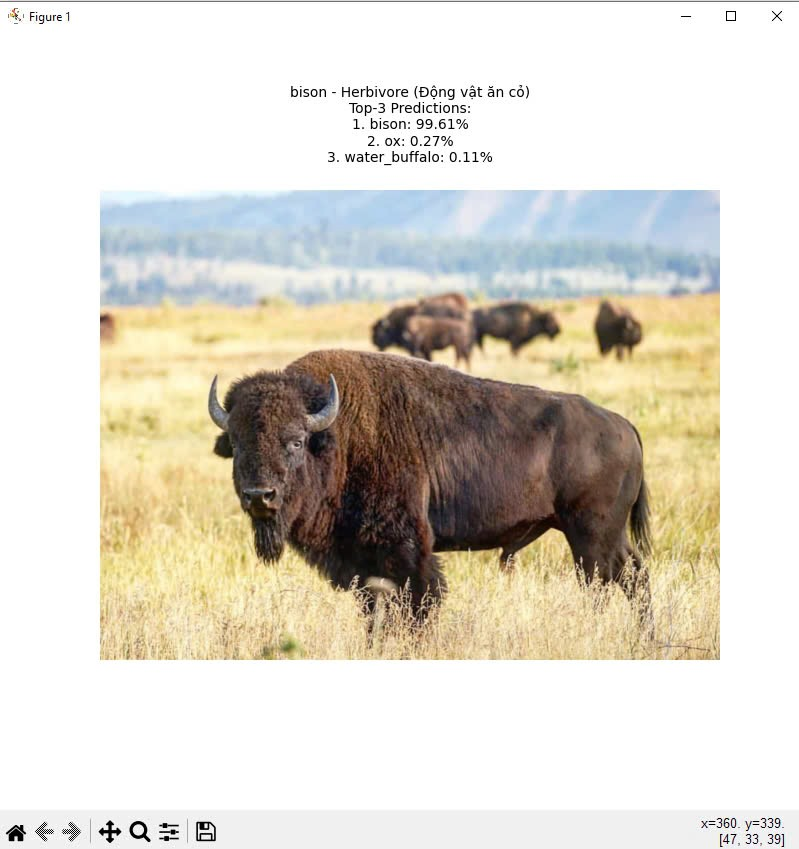
Hình 3.3.5 Ảnh kết quả thử nghiệm 5

* Bối cảnh: Con linh miêu đang ở trong một khu rừng. Mặc dù không thể nhìn thấy rõ ràng nhưng có thể nhận thấy những tán lá cây rậm rạp xung quanh, gợi lên một môi trường sống tự nhiên của loài mèo rừng này.
* Chất lượng hình ảnh: Hình ảnh có chất lượng khá tốt, cho phép chúng ta quan sát rõ các chi tiết trên khuôn mặt của con linh miêu như đôi mắt, mũi, râu và đôi tai. Màu sắc của lông cũng được tái hiện khá sống động.
* Góc chụp: Hình ảnh được chụp từ góc nghiêng, tập trung vào khuôn mặt của con linh miêu. Góc chụp này giúp chúng ta cảm nhận được sự tinh anh và cảnh giác trong ánh mắt của nó.

**Kết luận:**

Dựa vào hình ảnh, chúng ta có thể kết luận rằng con vật trong hình là một loài động vật ăn thịt. Linh miêu là một loài săn mồi chuyên nghiệp, chúng thường săn bắt các loài động vật nhỏ như thỏ, chim, và các loài gặm nhấm khác.

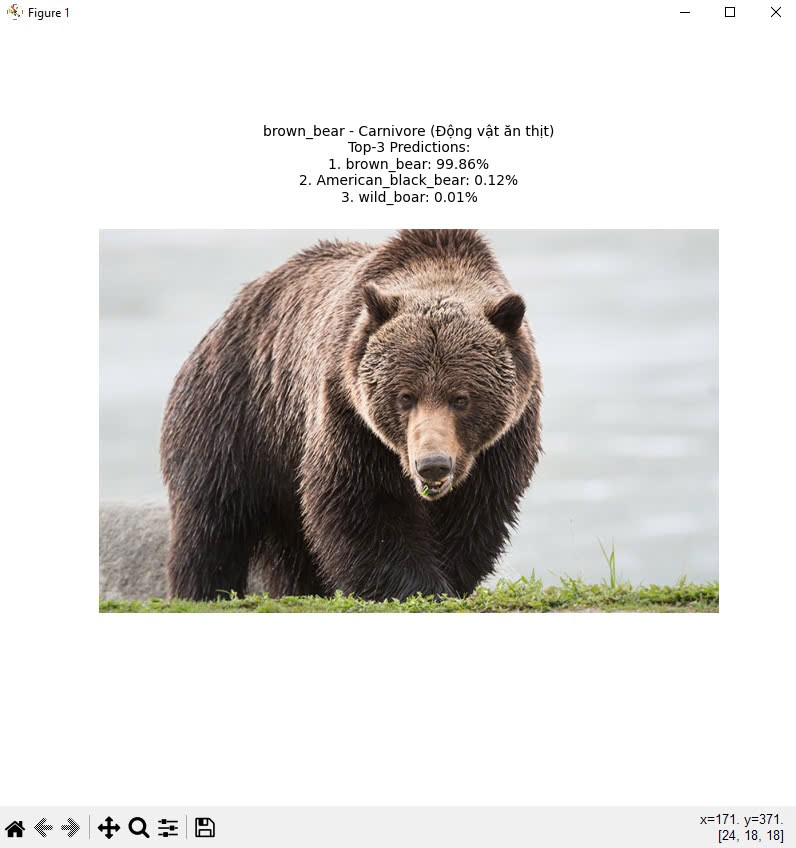
**Thử nghiệm ảnh 6:**



Hình 3.3.6 Ảnh kết quả thử nghiệm 6

* Hình ảnh: Một con bò rừng bizon đực trưởng thành đang đứng trên một cánh đồng cỏ rộng lớn.
* Đặc điểm: Bộ lông nâu sẫm, cặp sừng lớn và dáng vẻ oai vệ.
* Môi trường: Đồng cỏ rộng lớn với cỏ cao, tạo cảm giác hoang dã.
* Góc chụp: Nghiêng từ dưới lên, làm nổi bật sự đồ sộ của con vật.
* Kết luận: Bizon là động vật ăn cỏ.

**Thử nghiệm ảnh 7:**



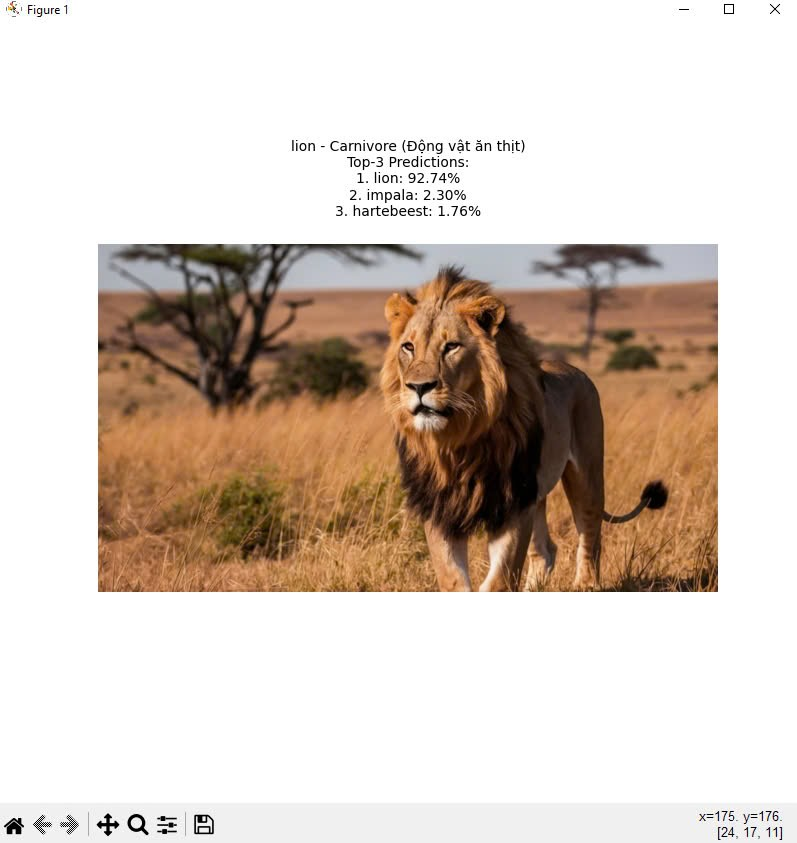
Hình 3.3.7 Ảnh kết quả thử nghiệm 7

* Bối cảnh: Con gấu đang đứng trên một bờ đất phủ đầy cỏ xanh mướt, ngay sát bờ nước. Có thể đây là một dòng sông hoặc hồ, với những cây cối xanh tốt ở phía xa. Bối cảnh này gợi lên một môi trường sống tự nhiên của loài gấu.
* Chất lượng hình ảnh: Hình ảnh có chất lượng rất tốt, độ phân giải cao. Nhờ đó, chúng ta có thể quan sát rõ ràng từng chi tiết trên cơ thể của con gấu, từ bộ lông dày mượt đến đôi mắt tinh anh và chiếc mũi ướt. Màu sắc của bức ảnh cũng rất tự nhiên, tạo cảm giác chân thực.
* Góc chụp: Hình ảnh được chụp từ góc nghiêng, hơi từ dưới lên, giúp làm nổi bật sự đồ sộ và uy nghi của con gấu. Góc chụp này tập trung vào khuôn mặt của con vật, cho phép chúng ta cảm nhận được sự mạnh mẽ và hoang dã trong ánh mắt của nó.

**Kết luận:**

Dựa vào hình ảnh và kết quả dự đoán, có thể khẳng định rằng con vật trong hình là một loài động vật ăn thịt . Gấu nâu là một loài động vật ăn tạp, nhưng chúng chủ yếu ăn thịt các loài động vật khác như cá, hươu nai, và các loài gặm nhấm.

**Thử nghiệm ảnh 8:**



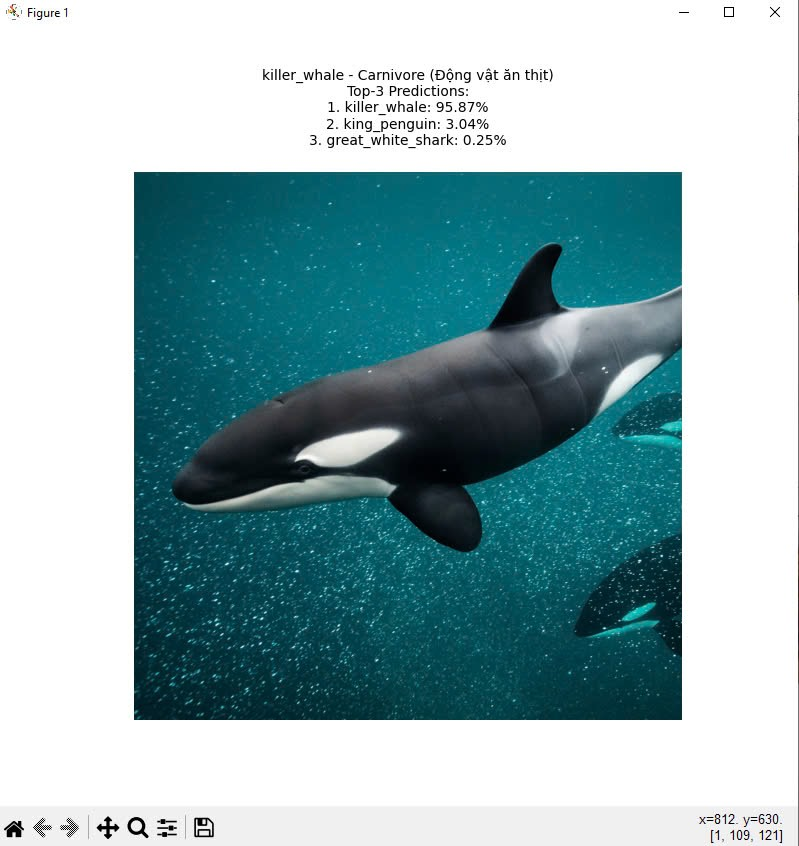
Hình 3.3.8 Ảnh kết quả thử nghiệm 8

* Bối cảnh: Con sư tử đang đi trên một vùng thảo nguyên rộng lớn với cỏ vàng và một vài cây bụi rải rác. Có thể thấy một vài cây keo cao ở xa xa, tạo nên một khung cảnh đặc trưng của vùng savanna châu Phi.
* Chất lượng hình ảnh: Hình ảnh có chất lượng rất tốt, độ phân giải cao. Nhờ đó, chúng ta có thể quan sát rõ ràng từng chi tiết trên cơ thể của con sư tử, từ bộ lông óng ánh đến đôi mắt sắc bén và những cơ bắp cuồn cuộn. Màu sắc của bức ảnh cũng rất tự nhiên, tạo cảm giác chân thực.
* Góc chụp: Hình ảnh được chụp từ góc ngang, giúp chúng ta quan sát toàn bộ cơ thể của con sư tử. Góc chụp này tập trung vào khuôn mặt của con vật, cho phép chúng ta cảm nhận được sự uy nghiêm và sức mạnh của "vua chúa rừng xanh".

**Kết luận:**

Dựa vào hình ảnh và kết quả dự đoán, có thể khẳng định rằng con vật trong hình là một loài động vật ăn thịt. Sư tử là một loài săn mồi đỉnh cao, chúng chủ yếu săn bắt các loài động vật lớn như ngựa vằn, linh dương và thậm chí cả hà mã.

**Thử nghiệm ảnh 9:**



Hình 3.3.9 Ảnh kết quả thử nghiệm 9

* Bối cảnh: Con cá voi sát thủ đang bơi trong một đại dương bao la. Nước biển có màu xanh dương đậm, tạo cảm giác sâu thẳm. Ánh sáng mặt trời lấp lánh trên mặt nước, tạo nên những tia sáng lung linh.
* Chất lượng hình ảnh: Hình ảnh có chất lượng rất tốt, độ phân giải cao. Nhờ đó, chúng ta có thể quan sát rõ ràng từng chi tiết trên cơ thể của con cá voi sát thủ, từ những đường vân trên da đến đôi mắt sáng lấp lánh. Màu sắc của bức ảnh cũng rất tự nhiên, tạo cảm giác chân thực.
* Góc chụp: Hình ảnh được chụp từ dưới lên, giúp chúng ta quan sát toàn bộ cơ thể của con cá voi sát thủ. Góc chụp này tập trung vào khuôn mặt của con vật, cho phép chúng ta cảm nhận được sự uy nghiêm và sức mạnh của loài động vật này.

**Kết luận:**

Dựa vào hình ảnh và kết quả dự đoán, có thể khẳng định rằng con vật trong hình là một loài động vật ăn thịt . Cá voi sát thủ là một trong những loài động vật săn mồi hàng đầu ở đại dương, chúng có thể săn bắt nhiều loại con mồi khác nhau, từ cá đến các loài động vật có vú biển.

**Thử nghiệm ảnh 10:**



Hình 3.3.10 Ảnh kết quả thử nghiệm 10

* Bối cảnh: Báo đang đi trên một bề mặt có màu sắc đối lập nhau, tạo cảm giác như hai lớp. Điều này có thể là do hiệu ứng của ánh sáng hoặc cách xử lý hình ảnh. Bối cảnh này khá đơn giản, tập trung vào con báo.
* Chất lượng hình ảnh: Hình ảnh có chất lượng khá tốt, độ phân giải vừa phải. Nhờ đó, chúng ta có thể quan sát rõ ràng từng chi tiết trên cơ thể của con báo, từ bộ lông hoa văn đến đôi mắt sắc bén và những móng vuốt sắc nhọn. Tuy nhiên, do hiệu ứng màu sắc đặc biệt ở nền, một số chi tiết có thể bị mờ hoặc không rõ nét.
* Góc chụp: Hình ảnh được chụp từ góc nghiêng, giúp chúng ta quan sát được toàn bộ cơ thể của con báo từ đầu đến chân. Góc chụp này tập trung vào phần thân trước của con vật, cho phép chúng ta thấy rõ những đặc điểm nổi bật của loài báo như chân dài, cơ bắp và bộ lông hoa văn.

**Kết luận:**

Dựa vào hình ảnh và kết quả dự đoán, có thể khẳng định rằng con vật trong hình là một loài động vật ăn thịt . Báo là một loài săn mồi nhanh nhẹn và linh hoạt, chúng săn bắt các loài động vật có vú nhỏ và trung bình.

# **KẾT LUẬN**

Sau quá trình nghiên cứu và phát triển, nhóm chúng em đã xây dựng thành công một hệ thống phân loại động vật ăn thịt và ăn cỏ dựa trên mô hình ResNet50. Hệ thống tận dụng sức mạnh của mô hình học sâu đã được tiền huấn luyện, kết hợp với dữ liệu thực tế để đạt hiệu suất cao trong việc phân loại.

Qua các thử nghiệm, mô hình ResNet50 đã chứng minh được hiệu quả vượt trội nhờ khả năng trích xuất đặc trưng mạnh mẽ từ hình ảnh, giúp phân loại chính xác các loài động vật theo chế độ ăn uống. Hệ thống đạt độ chính xác cao, đáp ứng tốt yêu cầu bài toán và có tiềm năng mở rộng trong tương lai.

**Điểm nổi bật của hệ thống:**

* Hiệu quả cao: Mô hình ResNet50 được tinh chỉnh cho bài toán cụ thể, giúp hệ thống phân loại chính xác các loài động vật ăn thịt và ăn cỏ.
* Ứng dụng thực tế: Hệ thống có thể được sử dụng trong các lĩnh vực như bảo tồn động vật, giáo dục, hoặc nghiên cứu khoa học.
* Tính linh hoạt: Hệ thống dễ dàng tích hợp thêm dữ liệu mới để cải thiện độ chính xác và mở rộng phạm vi phân loại.
* Tuy nhiên, hệ thống vẫn còn một số hạn chế, như:
* Phụ thuộc vào dữ liệu huấn luyện: Độ chính xác giảm nếu gặp hình ảnh không đại diện hoặc không đủ thông tin.
* Xử lý hình ảnh phức tạp: Điều kiện ánh sáng hoặc góc chụp khác biệt có thể ảnh hưởng đến khả năng nhận diện.

**Hướng phát triển trong tương lai:**

* Mở rộng dữ liệu: Thu thập thêm hình ảnh từ nhiều loài động vật khác nhau và đa dạng hóa điều kiện chụp.
* Tích hợp công nghệ mới: Kết hợp các phương pháp học sâu khác để tăng cường hiệu suất của hệ thống.
* Xây dựng giao diện ứng dụng: Cung cấp giao diện người dùng thân thiện để hỗ trợ các ứng dụng thực tế.

# **TÀI LIỆU THAM KHẢO**

1. <https://kungfutech.edu.vn/posts/viet-cli-trong-python-voi-argparse>
2. <https://pythonve.ikitai.net/entry/2024/01/09/000000>
3. <https://docs.opencv.org/>
4. <https://nguyenvanhieu.vn/thuat-toan-cnn-trong-nlp/>
5. <https://tailieu.tv/tai-lieu/giao-trinh-mon-xu-ly-anh-32566/>
6. <https://youtu.be/qpxgR4cpJQg?si=tA-HPSNfJaKWuYBz>
7. [https://www.researchgate.net/publication/326507248\_Nhan\_dang\_va\_xu\_ly\_anh\_thi \_giac\_may\_tinh](https://www.researchgate.net/publication/326507248_Nhan_dang_va_xu_ly_anh_thi%20%20_giac_may_tinh)