

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA KHOA HỌC VÀ KỸ THUẬT THÔNG TIN



BÁO CÁO SEMINAR
MÔN KỸ THUẬT LẬP TRÌNH PYTHON

**Đề tài: Ứng dụng mô hình học sâu DenseNet121 cho nhận dạng
động vật qua hình ảnh**

GVHD: TS. Nguyễn Tất Bảo Thiện

Nhóm sinh viên thực hiện:

- | | |
|-------------------------|----------------|
| 1. Đinh Hoàng Thùy Linh | MSSV: 21522900 |
| 2. Trương Quang Đạt | MSSV: 21522922 |
| 3. Hà Thị Thu Thủy | MSSV: 21522914 |
| 4. Võ Tuấn Nam | MSSV: 21522906 |

NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

This image shows a full page of primary-ruled paper. It features multiple sets of horizontal dashed lines spaced evenly down the page, providing a guide for handwriting practice. The background is white, and there are no margins or other markings present.

....., ngày.....tháng.....năm 20...

Người nhận xét

(Ký tên và ghi rõ họ tên)

BẢNG PHÂN CÔNG, ĐÁNH GIÁ THÀNH VIÊN:

Bảng 1: Phân công, đánh giá thành viên.

Họ và tên	MSSV	Phân công	Đánh giá
Đinh Hoàng Thùy Linh	21522900	<ul style="list-style-type: none"> • Tìm hiểu lý thuyết và kiến trúc của mô hình DenseNet121. • Thực hiện tăng cường dữ liệu đầu vào. • Huấn luyện các model. • Viết mã input đầu ra và tạo những biểu đồ so sánh hiệu suất các model huấn luyện với nhau. 	Hoàn thành
Trương Quang Đạt	21522922	<ul style="list-style-type: none"> • Tìm kiếm bộ dữ liệu các con vật để phục vụ cho việc huấn luyện. • Huấn luyện các model. Fix bug, sửa các lỗi liên quan trong quá trình huấn luyện. 	Hoàn thành
Hà Thị Thu Thủy	21522914	<ul style="list-style-type: none"> • Tìm kiếm bộ dữ liệu các con vật để phục vụ cho việc huấn luyện. • Hỗ trợ huấn luyện các model. 	Hoàn thành
Võ Tuấn Nam	21522906	<ul style="list-style-type: none"> • Thực hiện tổng hợp tài liệu, tổng hợp Word và thực hiện báo cáo. • Hỗ trợ huấn luyện các model. 	Hoàn thành

LỜI MỞ ĐẦU

Trong thập kỷ qua, sự phát triển vượt bậc của trí tuệ nhân tạo (AI) và học sâu (deep learning) đã mang lại nhiều tiến bộ đáng kể trong nhiều lĩnh vực khác nhau, bao gồm nhận dạng hình ảnh, xử lý ngôn ngữ tự nhiên, và phân tích dữ liệu lớn. Đặc biệt, việc ứng dụng học sâu trong nhận dạng hình ảnh đã mở ra nhiều cơ hội mới cho các nhà nghiên cứu và các chuyên gia trong việc tự động hóa các nhiệm vụ phức tạp mà trước đây chỉ có thể thực hiện bởi con người.

Nhận dạng động vật qua hình ảnh là một trong những ứng dụng nổi bật của học sâu, giúp các nhà nghiên cứu sinh học, bảo tồn động vật, và nông dân có thể dễ dàng nhận dạng và theo dõi các loài động vật trong tự nhiên cũng như trong môi trường nuôi nhốt. Việc áp dụng các mô hình học sâu như DenseNet121 không chỉ cải thiện độ chính xác của việc nhận dạng mà còn tiết kiệm thời gian và công sức so với các phương pháp truyền thống.

DenseNet121 (Densely Connected Convolutional Networks) là một kiến trúc mạng nơ-ron tích chập hiện đại được biết đến với khả năng nhận dạng hình ảnh vượt trội nhờ vào cấu trúc mạng nơ-ron dày đặc. Mô hình này cho phép mỗi lớp nhận thông tin từ tất cả các lớp trước đó, giúp cải thiện việc lan truyền thông tin và gradient, đồng thời giảm số lượng tham số cần thiết.

Trong báo cáo này, chúng tôi sẽ trình bày chi tiết về lý thuyết của mô hình DenseNet121, quá trình huấn luyện, tăng cường dữ liệu, kiểm thử, và ứng dụng thực tiễn của mô hình này trong nhận dạng động vật qua hình ảnh. Bên cạnh đó, chúng tôi cũng sẽ so sánh hiệu suất của DenseNet121 với các mô hình học sâu phổ biến khác như VGG16, EfficientNetB7 và MobileNetV3, nhằm đánh giá khả năng ứng dụng của DenseNet121 trong các bối cảnh thực tiễn. Qua đó, chúng tôi hy vọng báo cáo này sẽ mang lại cái nhìn tổng quan về khả năng và tiềm năng của DenseNet121 trong lĩnh vực nhận dạng hình ảnh động vật.

DANH MỤC CÁC BẢNG, HÌNH ẢNH

Danh mục các bảng:

Bảng 1: Phân công, đánh giá thành viên.	3
Bảng 2: Kết luận so sánh hiệu suất các mô hình học sâu trong huấn luyện và kiểm tra.	21
Bảng 3: So sánh các chỉ số hiệu suất của các mô hình học sâu.	22

Danh mục hình ảnh:

Hình 1-1: Kiến trúc mạng học sâu DenseNet121.	9
Hình 1-2: Mô tả Kiến trúc của DenseNet trong dự đoán hình ảnh.	9
Hình 2-1: Biểu đồ phân phối lớp mục tiêu trong tập huấn luyện.	13
Hình 2-2: Minh họa bộ dữ liệu data training.	14
Hình 2-3: tăng cường dữ liệu.	15
Hình 2-4: Quá trình biên dịch và huấn luyện mô hình.	16
Hình 3-1: So sánh kết quả dự đoán và nhãn thực tế.	18
Hình 3-2: Kết quả Test thử trên 1 mẫu thử.	19
Hình 3-3: Biểu đồ so sánh độ chính xác trong huấn luyện và kiểm tra.	20
Hình 3-4: Biểu đồ so sánh độ mất mát trong huấn luyện và kiểm tra.	21
Hình 3-5: Biểu đồ so sánh các chỉ số độ chính xác, độ chính xác dự đoán, độ nhạy và F1-score của các mô hình.	23
Hình 3-6: Ma trận nhầm lẫn của các mô hình học sâu.	25

MỤC LỤC

LỜI MỞ ĐẦU	4
DANH MỤC CÁC BẢNG, HÌNH ẢNH.....	5
Chương 1: GIỚI THIỆU VỀ MÔ HÌNH DENSENET121.....	8
1.1 Tổng quan về DenseNet121	8
1.2 Cấu trúc của DenseNet121	8
1.2.1 DenseNet121 block.....	9
1.2.2 Transition Layer.....	10
1.2.3 Cấu trúc tổng thể của DenseNet121	10
1.3 Ưu và Nhược điểm của DenseNet121	10
1.3.1 Ưu điểm:	10
1.3.2 Nhược điểm:	11
Chương 2: HUẤN LUYỆN, TĂNG CƯỜNG, KIỂM THỬ VÀ ĐỌC GHI PHÂN LOẠI CON VẬT QUA HÌNH ẢNH ỨNG DỤNG MÔ HÌNH DENSENET121	12
2.1 Chuẩn bị dữ liệu.....	12
2.1.1. Bộ dữ liệu	12
2.1.2. Tải lưu trữ và hiển thị thử 5 hình ảnh ngẫu nhiên.	13
2.1.3. Tải và xử lý dữ liệu	15
2.1.4. Tăng cường dữ liệu.....	15
2.2 Xây dựng mô hình DenseNet121.....	15
2.2.1 Sử dụng mô hình DenseNet121	15
2.2.2 Biên dịch và huấn luyện mô hình	16
2.3 Kiểm thử và đánh giá mô hình.....	16
2.3.1 Kiểm thử mô hình	16
2.3.2 Đánh giá các chỉ số	16
2.4 Đọc ghi và phân loại hình ảnh động vật	17
Chương 3: ĐÁNH GIÁ HIỆU SUẤT VÀ SO SÁNH.....	18
3.1 Hiển thị kết quả của DenseNet121 trên tập test.....	18
3.2 Test thử 1 tấm ảnh.....	19
3.3 So sánh với các model khác.....	19
3.3.1 So sánh Accuracy và Loss	20
3.3.2 So sánh Accuracy, Precision, Recall, F1-score	22
3.3.3 So sánh Confusion Matrix	24

**Chương 4: ỨNG DỤNG THỰC TIỄN, HƯỚNG PHÁT TRIỂN VÀ KẾT
LUẬN 27**

4.1Ứng dụng thực tiễn	27
4.2Hướng phát triển	27
4.3Kết luận	28
TÀI LIỆU THAM KHẢO	29

Chương 1: GIỚI THIỆU VỀ MÔ HÌNH DENSENET121

1.1 Tổng quan về DenseNet121

DenseNet121, một phiên bản của kiến trúc DenseNet (Densely Connected Convolutional Networks), được phát triển bởi Gao Huang và các cộng sự, là một trong những tiến bộ quan trọng trong lĩnh vực học sâu cho thị giác máy tính. Được thiết kế để khắc phục các hạn chế của các kiến trúc CNN truyền thống, DenseNet121 tận dụng tối đa các đặc trưng đã học từ tất cả các lớp trước đó để cải thiện khả năng học và hiệu quả tính toán của mô hình.

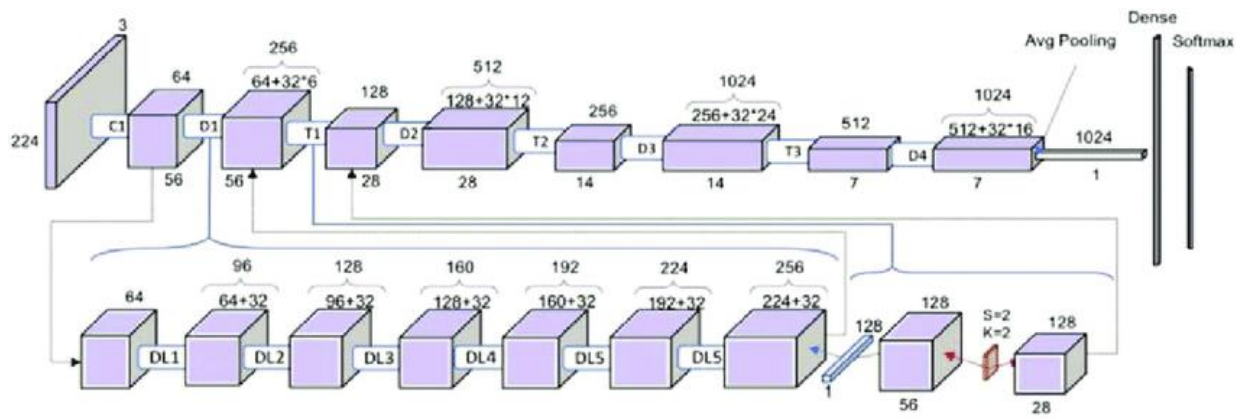
DenseNet121 không chỉ cải thiện sự truyền gradient giữa các lớp mà còn tăng cường việc tái sử dụng các đặc trưng thông qua **kết nối dày đặc** (dense connectivity). Điều này có nghĩa là đầu vào của mỗi lớp bao gồm cả đầu ra của tất cả các lớp trước đó trong cùng một khối (block), giúp tối ưu hóa quá trình học sâu và giảm thiểu số lượng tham số.

Lợi Ích Chính của DenseNet121:

- **Kết Nối Dày Đặc:** Mỗi lớp kết nối trực tiếp với tất cả các lớp trước đó, tạo ra một luồng thông tin rộng khắp mô hình, giúp tránh tình trạng mất thông tin và tăng cường việc học các đặc trưng.
- **Tái Sử Dụng Đặc Trưng:** Các đặc trưng đã học được tái sử dụng trên toàn bộ mô hình, làm giảm số lượng tham số và tránh học lại những đặc trưng đã học từ các lớp trước.
- **Hiệu Quả Tính Toán:** Mô hình yêu cầu ít tham số hơn so với các kiến trúc khác có hiệu suất tương đương như ResNet, làm cho DenseNet121 trở nên hiệu quả hơn về mặt tính toán và bộ nhớ.

1.2 Cấu trúc của DenseNet121

DenseNet121 có một cấu trúc phân cấp bao gồm các **Dense Block** và **Transition Layer**, mỗi thành phần có vai trò cụ thể trong việc học và giảm kích thước của bản đồ đặc trưng.



Hình 1-1: Kiến trúc mạng học sâu DenseNet121.

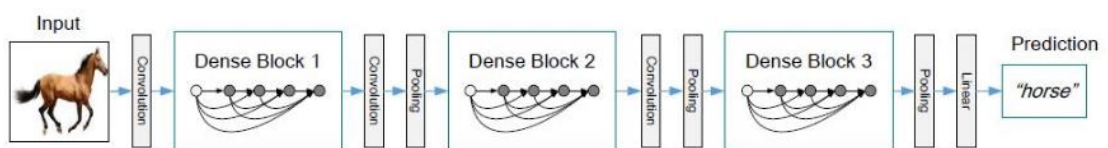
1.2.1 DenseNet121 block

Một **Dense Block** bao gồm nhiều lớp tích chập, trong đó mỗi lớp nhận đầu vào từ tất cả các lớp trước đó. Đầu ra của mỗi lớp trong Dense Block được kết hợp với đầu ra của tất cả các lớp trước đó thông qua phép nối kênh (channel-wise concatenation), và sau đó được đưa vào lớp kế tiếp.

- **Cấu Trúc Dense Block:**

- Mỗi lớp bao gồm Batch Normalization, ReLU, và tích chập 3x3.
- Growth Rate: Mỗi lớp thêm vào một số lượng kênh cố định, gọi là growth rate (thường là 32).

- **Ví dụ:**



Hình 1-2: Mô tả Kiến trúc của DenseNet trong dự đoán hình ảnh.

Trong hình trên, một mô hình DenseNet với ba khối dense được hiển thị. Các lớp giữa hai khối liền kề là các lớp chuyển đổi (transition layers), chúng thực hiện giảm kích thước (downsampling) của các bản đồ đặc trưng thông qua các phép tích chập và gộp (convolution và pooling). Trong khi đó, bên trong mỗi khối dense, kích thước của các bản đồ đặc trưng được giữ nguyên để cho phép nối kết các đặc trưng.

1.2.2 Transition Layer

Transition Layer được sử dụng để điều chỉnh kích thước của bản đồ đặc trưng giữa các Dense Block. Mỗi Transition Layer bao gồm một lớp tích chập 1x1 để giảm số lượng kênh và một lớp gộp trung bình (average pooling) 2x2 để giảm kích thước không gian của bản đồ đặc trưng.

- **Cấu Trúc Transition Layer:**
 - Batch Normalization, ReLU.
 - Tích chập 1x1 để giảm số lượng kênh.
 - Gộp trung bình 2x2 để giảm kích thước không gian.

1.2.3 Cấu trúc tổng thể của DenseNet121

DenseNet121 bắt đầu với một lớp tích chập và gộp để giảm kích thước dữ liệu đầu vào, sau đó là sự xen kẽ giữa các Dense Block và Transition Layer. Cuối cùng, mô hình kết thúc với một lớp Global Average Pooling và một lớp Fully Connected để thực hiện phân loại.

- **Lớp Tích Chập Đầu Tiên:** Tích chập 7x7, stride 2.
- **Max Pooling:** 3x3, stride 2.
- **Dense Block 1:** 6 lớp tích chập.
- **Transition Layer 1:** Tích chập 1x1 và gộp trung bình 2x2.
- **Dense Block 2:** 12 lớp tích chập.
- **Transition Layer 2:** Tương tự như Transition Layer 1.
- **Dense Block 3:** 24 lớp tích chập.
- **Transition Layer 3:** Tương tự như Transition Layer 1.
- **Dense Block 4:** 16 lớp tích chập.
- **Global Average Pooling:** Tính trung bình toàn bộ bản đồ đặc trưng.
- **Fully Connected Layer:** Kết nối đầy đủ cho phân loại cuối cùng.

1.3 Ưu và Nhược điểm của DenseNet121

1.3.1 Ưu điểm:

- **Truyền Thông Tin Hiệu Quả:** Các lớp kết nối dày đặc cho phép truyền thông tin giữa các lớp một cách hiệu quả, giúp tránh tình trạng mất thông tin.
- **Tái Sử Dụng Đặc Trưng:** Đặc trưng đã học được tái sử dụng trên toàn bộ mô hình, làm giảm thiểu số lượng tham số và tối ưu hóa khả năng học.

- **Giảm Số Lượng Tham Số:** Số lượng tham số ít hơn so với các mô hình như VGG16, giúp giảm thiểu yêu cầu tài nguyên tính toán mà vẫn đạt được hiệu suất cao.

1.3.2 Nhược điểm:

- **Yêu Cầu Bộ Nhớ:** Mặc dù số lượng tham số ít hơn, nhưng yêu cầu bộ nhớ cao do phải lưu trữ các bản đồ đặc trưng từ tất cả các lớp trước đó.
- **Phức Tạp Hóa Kết Nối:** Cách kết nối dày đặc có thể làm cho cấu trúc mô hình trở nên phức tạp hơn và yêu cầu tính toán nhiều hơn so với các kiến trúc khác.

Chương 2: HUẤN LUYỆN, TĂNG CƯỜNG, KIỂM THỬ VÀ ĐỌC GHI PHÂN LOẠI CON VẬT QUA HÌNH ẢNH ỨNG DỤNG MÔ HÌNH DENSENET121

2.1 Chuẩn bị dữ liệu

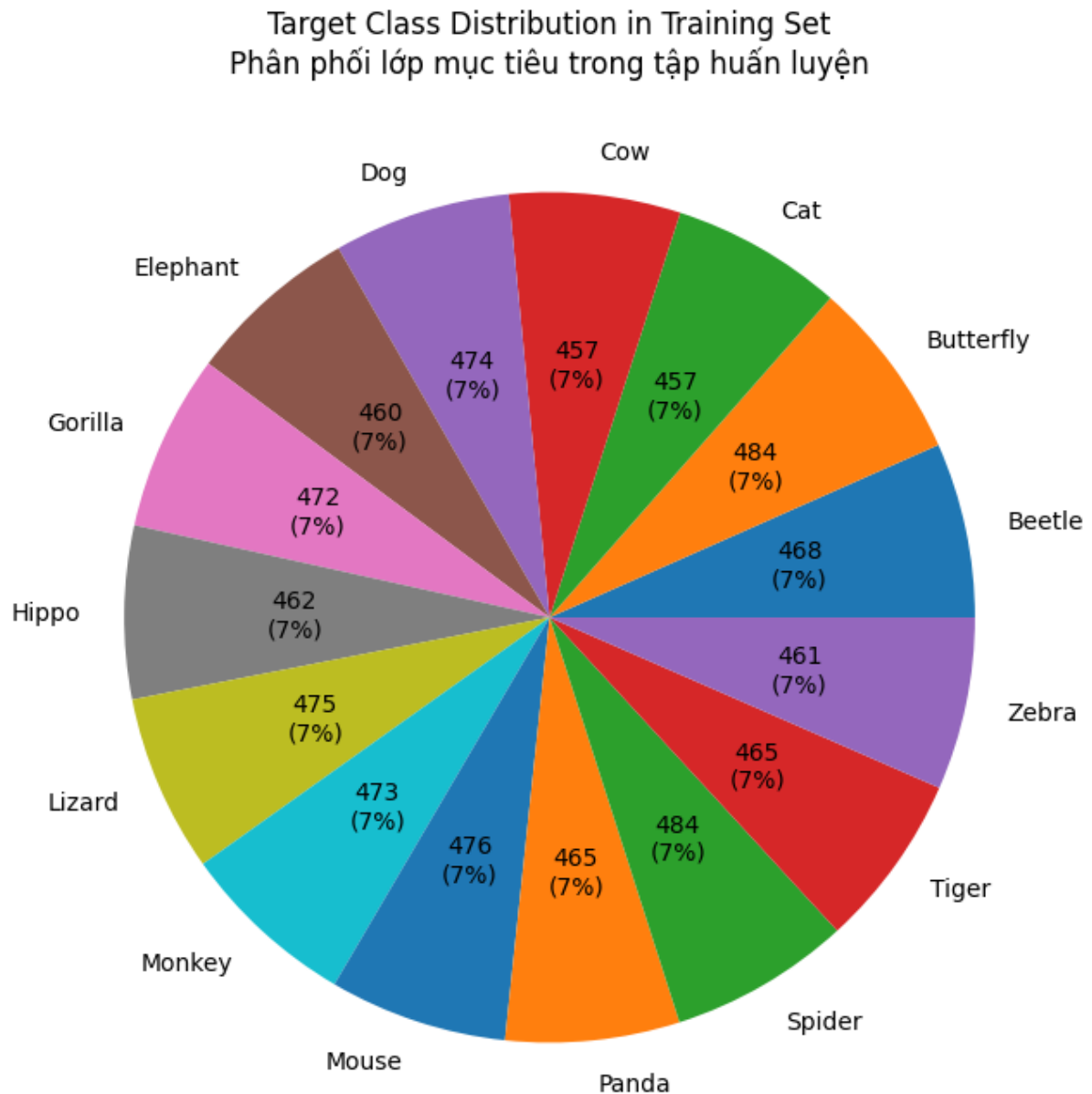
Dữ liệu là yếu tố quan trọng trong việc huấn luyện mô hình học sâu. Đối với bài toán nhận dạng động vật qua hình ảnh, chúng tôi sử dụng một tập dữ liệu gồm các hình ảnh động vật được tổ chức trong các thư mục khác nhau, mỗi thư mục tương ứng với một loại động vật.

2.1.1. Bộ dữ liệu

Bộ dữ liệu mà nhóm đã chọn bao gồm tổng cộng 9750 tệp hình ảnh, được chia thành hai tập chính: tập huấn luyện và tập kiểm thử. Tập huấn luyện chứa 8775 hình ảnh thuộc 15 lớp động vật khác nhau, trong khi tập kiểm thử chứa 975 hình ảnh thuộc cùng 15 lớp này.

Danh sách các lớp động vật trong bộ dữ liệu bao gồm:

1. Beetle (Bọ cánh cứng)
2. Butterfly (Bướm)
3. Cat (Mèo)
4. Cow (Bò)
5. Dog (Chó)
6. Elephant (Voi)
7. Gorilla (Khỉ đột)
8. Hippo (Hà mã)
9. Lizard (Thằn lằn)
10. Monkey (Khỉ)
11. Mouse (Chuột)
12. Panda (Gấu trúc)
13. Spider (Nhện)
14. Tiger (Hổ)
15. Zebra (Ngựa vằn)



Hình 2-1: Biểu đồ phân phối lớp mục tiêu trong tập huấn luyện.

2.1.2. Tải lưu trữ và hiển thị thử 5 hình ảnh ngẫu nhiên.

Mục đích lưu trữ và hiển thị một số hình ảnh mẫu từ mỗi lớp trong tập dữ liệu kiểm tra. Mỗi lớp được đại diện bởi 5 hình ảnh ngẫu nhiên từ tập kiểm thử, giúp trực quan hóa các mẫu hình ảnh trong mỗi lớp và đảm bảo rằng dữ liệu đầu vào đã được xử lý đúng cách.

Quá trình thực hiện

1. Xác định số lượng lớp học:

- Sử dụng thuộc tính `class_names` của tập dữ liệu để xác định số lượng lớp.

2. Khởi tạo danh sách lưu trữ hình ảnh:

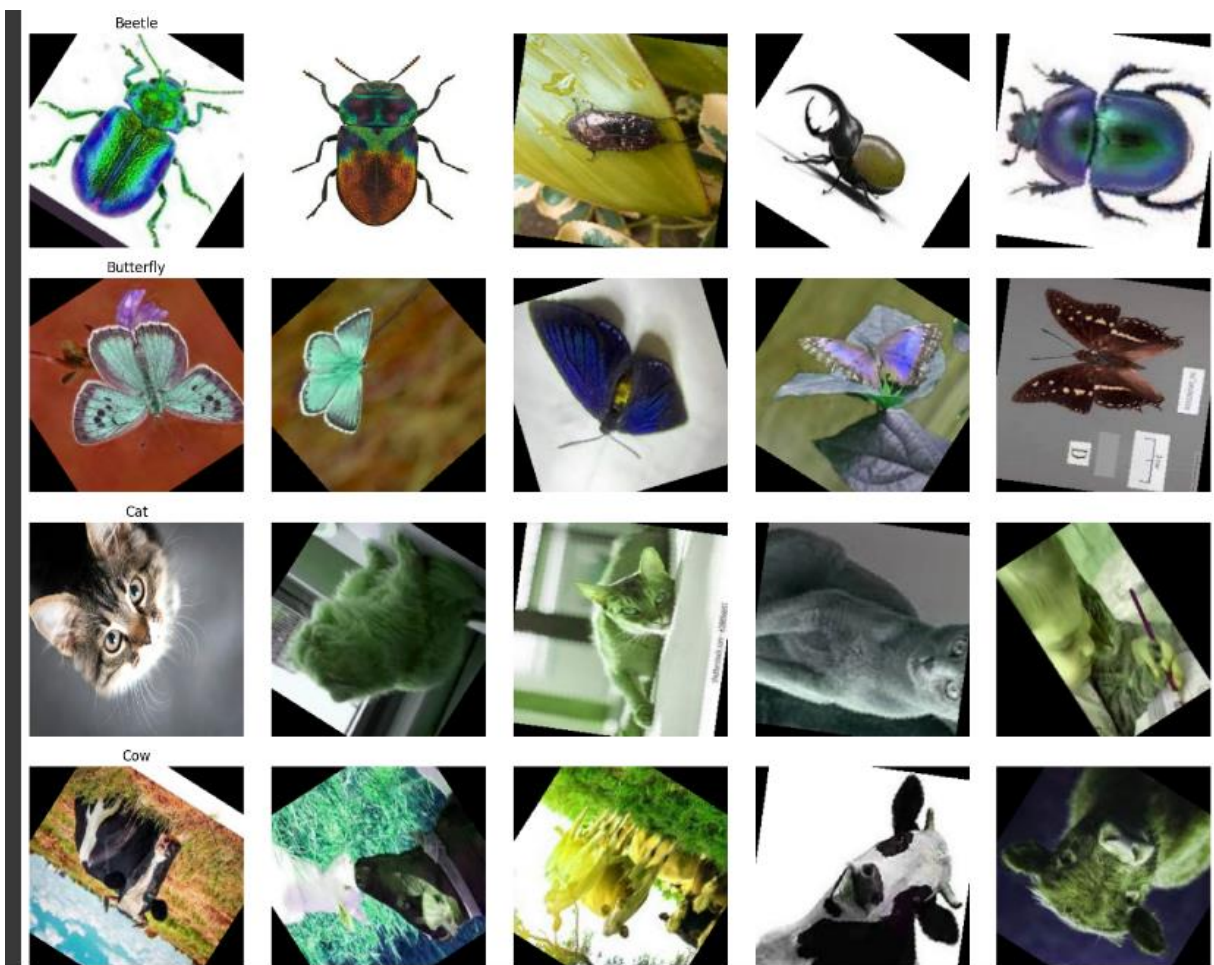
- Tạo một từ điển để lưu trữ các hình ảnh mẫu cho mỗi lớp.

3. Duyệt qua tập dữ liệu kiểm tra:

- Sử dụng vòng lặp để duyệt qua từng batch hình ảnh và nhãn trong tập kiểm tra.
- Chuyển đổi các hình ảnh và nhãn sang dạng numpy để dễ dàng thao tác.
- Xác định lớp của mỗi hình ảnh và lưu trữ hình ảnh vào danh sách tương ứng cho đến khi có đủ 5 hình ảnh cho mỗi lớp.
- Kiểm tra điều kiện để dừng vòng lặp khi đã thu thập đủ hình ảnh cho tất cả các lớp.

4. Hiển thị các hình ảnh:

- Sử dụng thư viện Matplotlib để tạo một lưới biểu đồ với số hàng tương ứng với số lượng lớp và 5 cột.
- Duyệt qua từng lớp và hiển thị 5 hình ảnh tương ứng, đặt tiêu đề cho mỗi hàng là tên lớp.



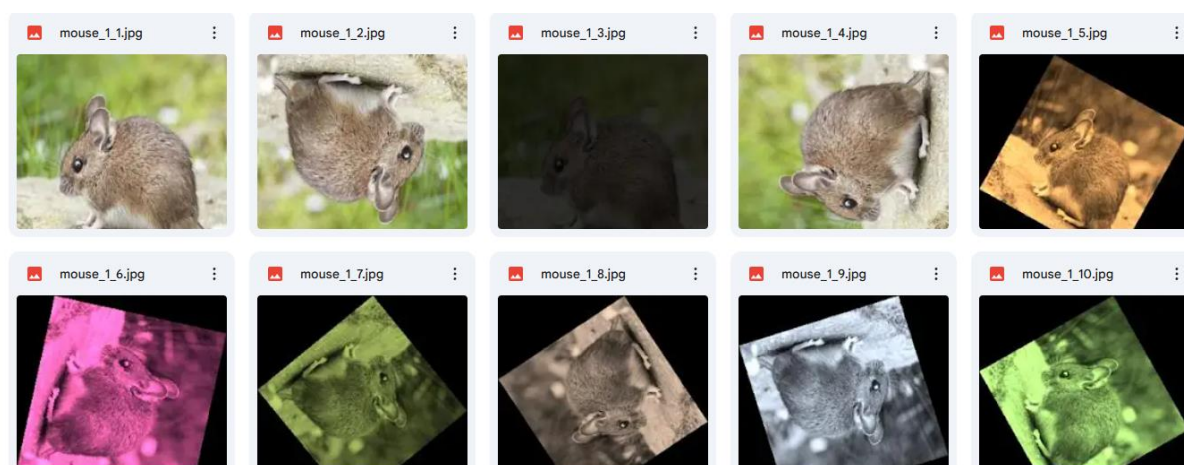
Hình 2-2: Minh họa bộ dữ liệu data training.

2.1.3. Tải và xử lý dữ liệu

Chúng tôi sử dụng thư viện Keras để tải dữ liệu từ Google Drive và xử lý chúng. Dữ liệu được chia thành hai tập: tập huấn luyện và tập kiểm thử. Tập huấn luyện bao gồm các hình ảnh được sử dụng để huấn luyện mô hình, trong khi tập kiểm thử được sử dụng để đánh giá hiệu suất của mô hình sau khi huấn luyện.

2.1.4. Tăng cường dữ liệu

Để mô hình học sâu hoạt động hiệu quả và không bị overfitting, chúng tôi áp dụng các kỹ thuật tăng cường dữ liệu. Các kỹ thuật này bao gồm tô đen, tăng màu, xoay, và thay đổi màu sắc của ảnh để tạo ra các biến thể khác nhau từ ảnh gốc. Những thay đổi này giúp tạo ra nhiều mẫu huấn luyện khác nhau từ một ảnh gốc, làm phong phú thêm tập dữ liệu và cải thiện khả năng tổng quát hóa của mô hình.



Hình 2-3: tăng cường dữ liệu.

2.2 Xây dựng mô hình DenseNet121

2.2.1 Sử dụng mô hình DenseNet121

DenseNet121 là một mô hình học sâu mạnh mẽ với kiến trúc đặc biệt, cho phép mỗi lớp tích hợp nhận đầu vào từ tất cả các lớp trước đó. Điều này giúp tăng cường việc truyền thông tin và gradient trong mạng. Chúng tôi tải mô hình DenseNet121 đã được tiền huấn luyện trên tập dữ liệu ImageNet và tùy chỉnh các lớp đầu ra cho phù hợp với bài toán nhận dạng động vật. Các lớp của DenseNet121 được đặt không trainable để giữ lại các đặc trưng đã học từ ImageNet, và chúng tôi thêm các lớp fully connected mới để phù hợp với số lượng lớp động vật cần nhận dạng.

2.2.2 Biên dịch và huấn luyện mô hình

Chúng tôi sử dụng bộ tối ưu Adam và hàm mất mát categorical crossentropy để biên dịch mô hình. Quá trình huấn luyện được thực hiện trên tập dữ liệu đã được tăng cường. Trong quá trình huấn luyện, mô hình sẽ học cách phân loại các hình ảnh động vật dựa trên các đặc trưng được rút trích từ hình ảnh.

```
# Load mô hình đã được train trước
model_file = '/content/gdrive/MyDrive/Nhan_Dang_Con_Vat/convat_model/model3.h5'
model3 = keras.models.load_model(model_file)

# Tiếp tục huấn luyện mô hình
history3 = model3.fit(train_ds, batch_size=32, epochs=10,
                      validation_data=val_ds)
```

```
Epoch 1/10
209/209 [=====] - 2301s 11s/step - loss: 1.0095 - accuracy: 0.7975 - val_loss: 0.6404 - val_accuracy: 0.8490
Epoch 2/10
209/209 [=====] - 7s 35ms/step - loss: 0.2894 - accuracy: 0.9305 - val_loss: 0.6686 - val_accuracy: 0.8671
Epoch 3/10
209/209 [=====] - 7s 33ms/step - loss: 0.0846 - accuracy: 0.9803 - val_loss: 0.6633 - val_accuracy: 0.8808
Epoch 4/10
209/209 [=====] - 7s 34ms/step - loss: 0.0296 - accuracy: 0.9934 - val_loss: 0.6384 - val_accuracy: 0.8914
Epoch 5/10
209/209 [=====] - 7s 34ms/step - loss: 0.0077 - accuracy: 0.9976 - val_loss: 0.6151 - val_accuracy: 0.9068
Epoch 6/10
209/209 [=====] - 7s 33ms/step - loss: 0.0046 - accuracy: 0.9990 - val_loss: 0.5727 - val_accuracy: 0.9089
Epoch 7/10
209/209 [=====] - 7s 34ms/step - loss: 2.8333e-04 - accuracy: 1.0000 - val_loss: 0.5676 - val_accuracy: 0.9100
Epoch 8/10
209/209 [=====] - 7s 33ms/step - loss: 9.0938e-05 - accuracy: 1.0000 - val_loss: 0.5663 - val_accuracy: 0.9100
Epoch 9/10
209/209 [=====] - 7s 34ms/step - loss: 7.2569e-05 - accuracy: 1.0000 - val_loss: 0.5660 - val_accuracy: 0.9110
Epoch 10/10
209/209 [=====] - 7s 34ms/step - loss: 6.1369e-05 - accuracy: 1.0000 - val_loss: 0.5659 - val_accuracy: 0.9110
```

Hình 2-4: Quá trình biên dịch và huấn luyện mô hình.

2.3 Kiểm thử và đánh giá mô hình

2.3.1 Kiểm thử mô hình

Sau khi hoàn thành quá trình huấn luyện, chúng tôi sử dụng tập kiểm thử để đánh giá độ chính xác của mô hình. Đánh giá mô hình trên tập kiểm thử cho phép chúng tôi kiểm tra khả năng tổng quát hóa của mô hình đối với các dữ liệu mới không có trong tập huấn luyện.

2.3.2 Đánh giá các chỉ số

Chúng tôi sử dụng các chỉ số như độ chính xác, precision, recall và F1-score để đánh giá hiệu suất của mô hình. Độ chính xác đo lường tỷ lệ các dự đoán đúng trên tổng số dự đoán, precision đo lường tỷ lệ các dự đoán đúng trong số các dự đoán là đúng, recall đo lường tỷ lệ các dự đoán đúng trong số các mẫu thực sự đúng, và F1-score là trung bình hài hòa của precision và recall. Bên cạnh đó, ma trận nhầm lẫn (confusion matrix) cũng được sử dụng để đánh giá chi tiết hơn về khả năng phân loại của mô hình. Ma trận nhầm lẫn cho thấy số lượng dự đoán đúng và sai của mô hình cho từng lớp, giúp chúng tôi hiểu rõ hơn về hiệu suất của mô hình đối với từng loại động vật.

2.4 Đọc ghi và phân loại hình ảnh động vật

Chúng tôi xây dựng một hàm để dự đoán và phân loại động vật từ các hình ảnh mới. Hàm này sẽ nhận đầu vào là đường dẫn đến một hình ảnh, thực hiện các bước tiền xử lý để chuẩn bị dữ liệu đầu vào cho mô hình, sau đó sử dụng mô hình đã được huấn luyện để dự đoán loại động vật trong hình ảnh. Kết quả dự đoán sẽ được lưu lại để sử dụng trong các ứng dụng thực tế. Chúng tôi cũng đảm bảo rằng quá trình dự đoán và lưu kết quả được thực hiện một cách hiệu quả và chính xác, để có thể ứng dụng trong các hệ thống nhận dạng động vật thực tế.

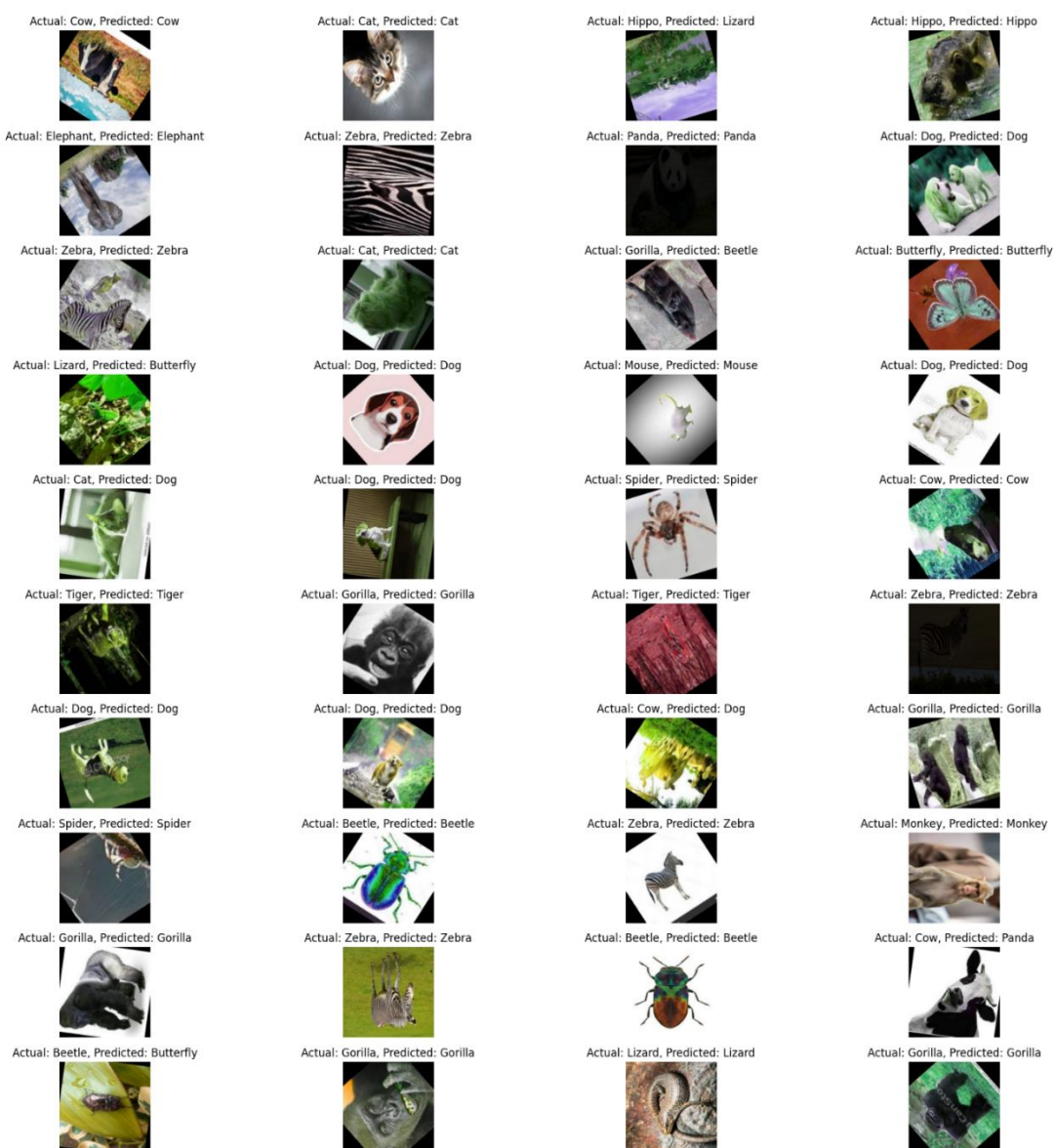
Với các bước trên, chúng tôi đã xây dựng thành công một mô hình DenseNet121 để nhận dạng động vật qua hình ảnh, từ việc chuẩn bị và tăng cường dữ liệu, xây dựng và huấn luyện mô hình, đến kiểm thử và đánh giá hiệu suất. Mô hình này có thể áp dụng thực tiễn để phân loại động vật trong các ứng dụng nghiên cứu sinh học và bảo tồn động vật.

Chương 3: ĐÁNH GIÁ HIỆU SUẤT VÀ SO SÁNH

3.1 Hiện thị kết quả của DenseNet121 trên tập test

Sau khi hoàn thành quá trình huấn luyện mô hình DenseNet121, chúng tôi tiến hành kiểm tra hiệu suất của mô hình trên tập dữ liệu kiểm thử. Mục đích của việc này là để đánh giá khả năng của mô hình trong việc nhận dạng động vật trên các hình ảnh mà mô hình chưa từng thấy trước đó.

Kết quả kiểm tra bao gồm các chỉ số quan trọng như độ chính xác (accuracy), độ mất mát (loss), precision, recall và F1-score. Các chỉ số này giúp chúng tôi hiểu rõ hơn về hiệu suất của mô hình trên tập dữ liệu kiểm thử và khả năng tổng quát hóa của mô hình.



Hình 3-1: So sánh kết quả dự đoán và nhãn thực tế.

3.2 Test thử 1 tấm ảnh

Để đánh giá khả năng của mô hình trong việc dự đoán trên các hình ảnh mới, chúng tôi đã thực hiện việc thử nghiệm dự đoán trên một tấm ảnh cụ thể. Quá trình này bao gồm các bước sau:

1. **Tiền xử lý hình ảnh:** Chuyển đổi hình ảnh về kích thước và định dạng phù hợp với mô hình.
2. **Dự đoán:** Sử dụng mô hình DenseNet121 để dự đoán lớp động vật của hình ảnh.
3. **Hiển thị kết quả:** Hiển thị hình ảnh cùng với lớp động vật được dự đoán để kiểm tra tính chính xác

• **dog.4035.jpg**(image/jpeg) - 27688 bytes, last modified: 9/29/2019 - 100% done
Saving dog.4035.jpg to dog.4035.jpg



```
1/1 [=====] - 0s 36ms/step
1/1 [=====] - 0s 47ms/step
1/1 [=====] - 0s 24ms/step
1/1 [=====] - 0s 30ms/step
VGG16 Predicted class: Dog
EfficientNetB7 Predicted class: Dog
MobileNetV3 Predicted class: Dog
Dense Predicted class: Dog
```

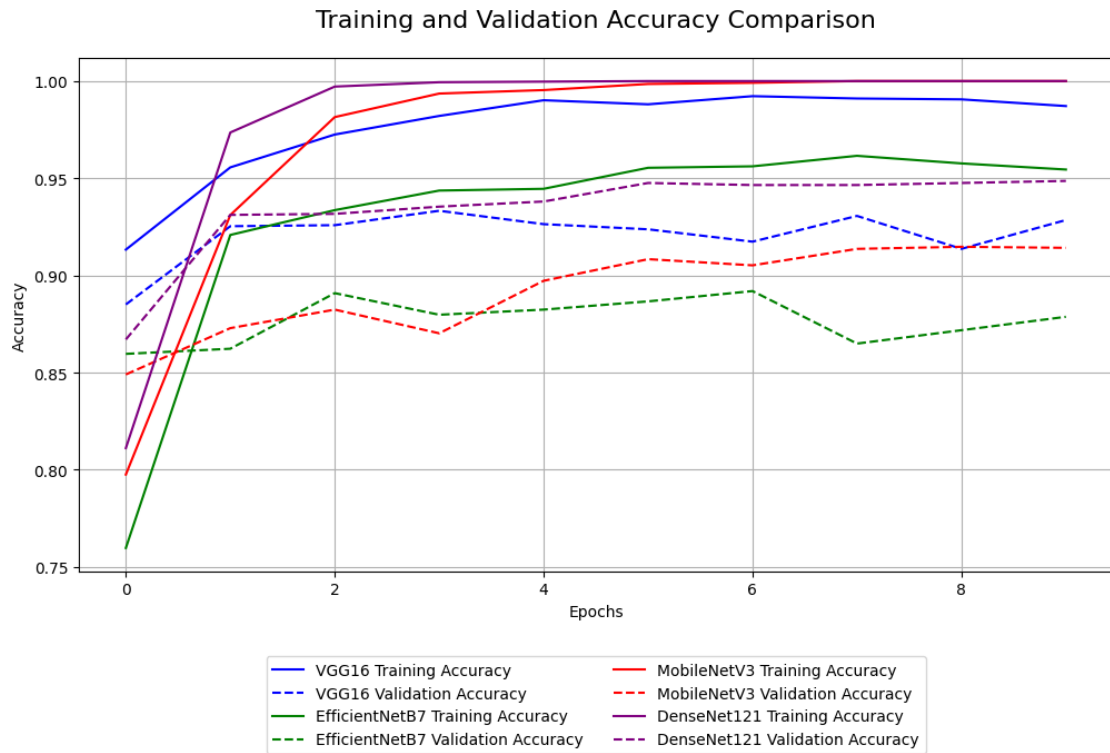
Hình 3-2: Kết quả Test thử trên 1 mẫu thử.

3.3 So sánh với các model khác

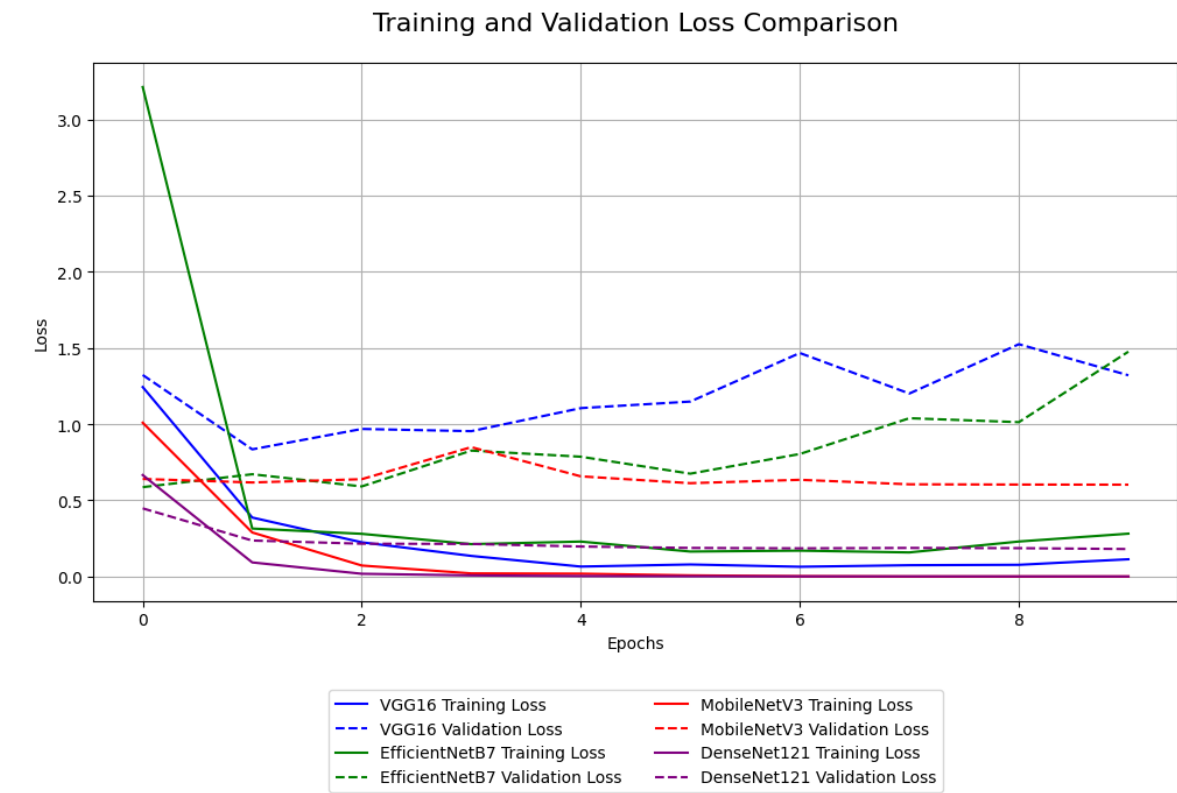
Để đánh giá hiệu suất của DenseNet121 một cách toàn diện, chúng tôi tiến hành so sánh kết quả của mô hình này với các mô hình học sâu khác như VGG16, EfficientNetB7 và MobileNetV3. Các mô hình này đều được huấn luyện trên cùng một tập dữ liệu và được đánh giá trên cùng một tập kiểm thử.

3.3.1 So sánh Accuracy và Loss

Chúng tôi so sánh độ chính xác (accuracy) và độ mất mát (loss) của các mô hình trên tập kiểm thử. Độ chính xác là tỷ lệ phần trăm các dự đoán đúng trên tổng số dự đoán, trong khi độ mất mát đo lường sự khác biệt giữa các giá trị dự đoán của mô hình và các giá trị thực tế.



Hình 3-3: Biểu đồ so sánh độ chính xác trong huấn luyện và kiểm tra.



Hình 3-4: Biểu đồ so sánh độ mất mát trong huấn luyện và kiểm tra.

Bảng 2: Kết luận so sánh hiệu suất các mô hình học sâu trong huấn luyện và kiểm tra.

	VGG16	EfficientNetB7	MobileNetV3	DenseNet121
Độ chính xác huấn luyện	Tăng nhanh chóng và đạt đỉnh gần 1.0 ở epoch 2	Tăng nhanh và ổn định, đạt gần 0.95 ở epoch 1, sau đó tiếp tục tăng nhẹ	Tăng đều nhưng thấp hơn so với VGG16 và EfficientNetB7, đạt khoảng 0.9 ở epoch 2 và duy trì	Tăng rất nhanh, đạt đỉnh gần 1.0 ở epoch 1 và duy trì mức độ cao nhất
Độ chính xác kiểm tra	Tăng dần nhưng dao động quanh mức 0.9 từ epoch 2 đến epoch 9.	Tăng ban đầu nhưng sau đó dao động, không đạt đến mức độ ổn định cao.	Tăng ban đầu sau đó dao động, nhưng có xu hướng giảm nhẹ từ epoch 2 đến 9	Tăng đều và ổn định, dao động nhẹ quanh mức 0.95 từ epoch 2 đến 9

Độ mất mát huấn luyện	Giảm nhanh chóng từ epoch 0 đến epoch 1, sau đó duy trì ổn định.	Giảm nhanh chóng từ epoch 0 đến epoch 1, sau đó duy trì mức thấp.	Giảm nhanh chóng và duy trì ổn định ở mức thấp sau epoch 1	Giảm nhanh chóng và duy trì ổn định ở mức thấp nhất.
Độ mất mát kiểm tra	Giảm ban đầu nhưng sau đó tăng nhẹ và dao động, không ổn định.	Tăng và dao động sau epoch 1, không ổn định	Duy trì ổn định sau khi giảm ban đầu, không dao động nhiều.	Duy trì ở mức thấp và ổn định, ít dao động

Nhận xét tổng quát:

- **DenseNet121** cho thấy kết quả tốt nhất với độ chính xác cao và ổn định cả trong quá trình huấn luyện và kiểm tra, cùng với độ mất mát thấp và ít dao động.
- **VGG16** cũng có hiệu suất khá tốt nhưng có một số dao động trong độ chính xác kiểm tra và độ mất mát kiểm tra.
- **EfficientNetB7** và **MobileNetV3** cho thấy dao động lớn hơn trong độ chính xác kiểm tra và độ mất mát kiểm tra, có thể ít ổn định hơn trong các tập dữ liệu kiểm tra

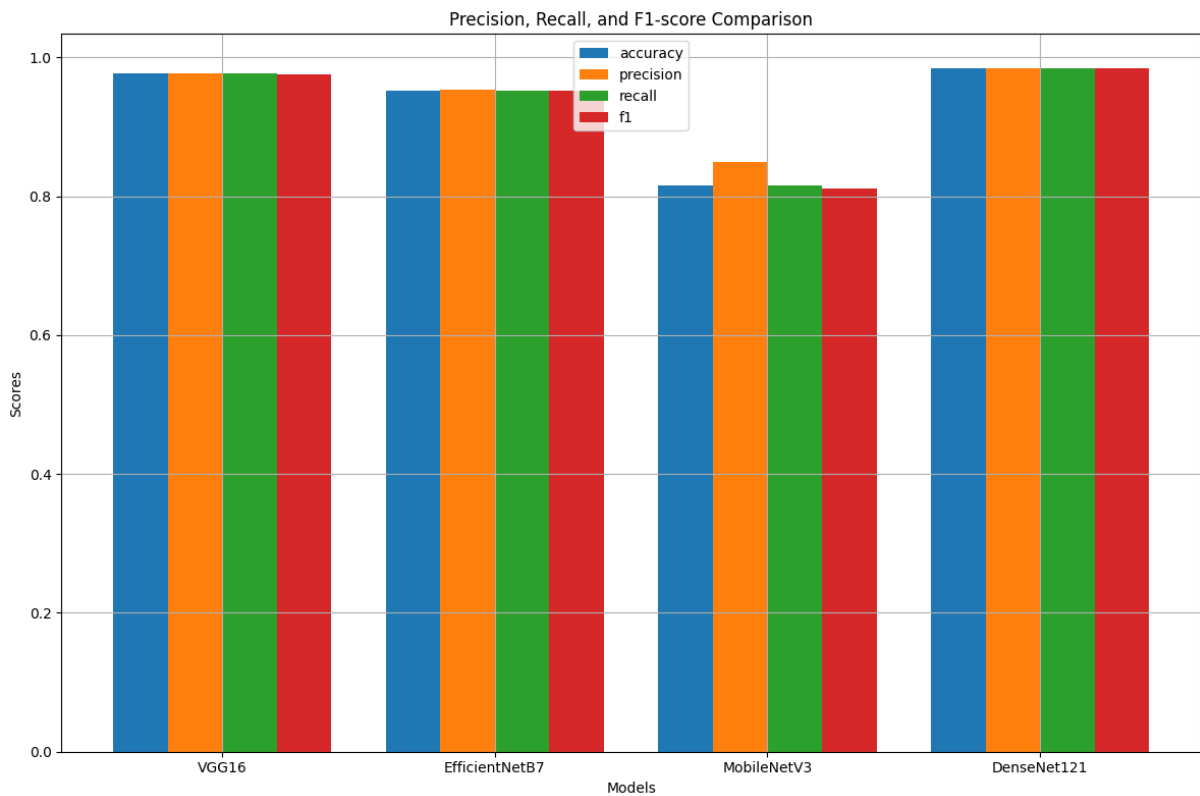
3.3.2 So sánh Accuracy, Precision, Recall, F1-score

Các chỉ số khác như precision, recall và F1-score cũng được sử dụng để đánh giá hiệu suất của mô hình:

- **Precision:** Tỷ lệ các dự đoán đúng trong số các dự đoán là đúng.
- **Recall:** Tỷ lệ các dự đoán đúng trong số các mẫu thực sự đúng.
- **F1-score:** Trung bình hài hòa của precision và recall.

Bảng 3: So sánh các chỉ số hiệu suất của các mô hình học sâu.

Model	Accuracy	Precision	Recall	F1-score
VGG16	0.9764	0.9768	0.9764	0.9763
EfficientNetB7	0.9518	0.9539	0.9518	0.9518
MobileNetV3	0.8154	0.8491	0.8154	0.8117
DenseNet121	0.9846	0.9847	0.9846	0.9846



Hình 3-5: Biểu đồ so sánh các chỉ số độ chính xác, độ chính xác dự đoán, độ nhạy và F1-score của các mô hình.

Dựa vào Bảng trên nhóm có một số nhận xét:

1. DenseNet121:

- Đây là mô hình có hiệu suất tốt nhất trong số các mô hình được so sánh, với độ chính xác (Accuracy), độ chính xác (Precision), độ nhạy (Recall), và điểm F1 (F1-score) đều đạt 0.9846.
- Điều này cho thấy DenseNet121 không chỉ dự đoán chính xác mà còn duy trì cân bằng tốt giữa độ chính xác và độ nhạy.

2. VGG16:

- Mô hình này cũng có hiệu suất rất cao với các chỉ số Accuracy, Precision, Recall và F1-score đều xấp xỉ 0.976.
- So với DenseNet121, VGG16 có hiệu suất thấp hơn một chút, nhưng vẫn rất mạnh mẽ và đáng tin cậy.

3. EfficientNetB7:

- Mô hình này có các chỉ số xấp xỉ 0.952, cho thấy hiệu suất khá cao nhưng vẫn thấp hơn so với VGG16 và DenseNet121.

- EfficientNetB7 vẫn là một lựa chọn tốt, đặc biệt khi cần một mô hình hiệu quả về mặt tài nguyên tính toán.

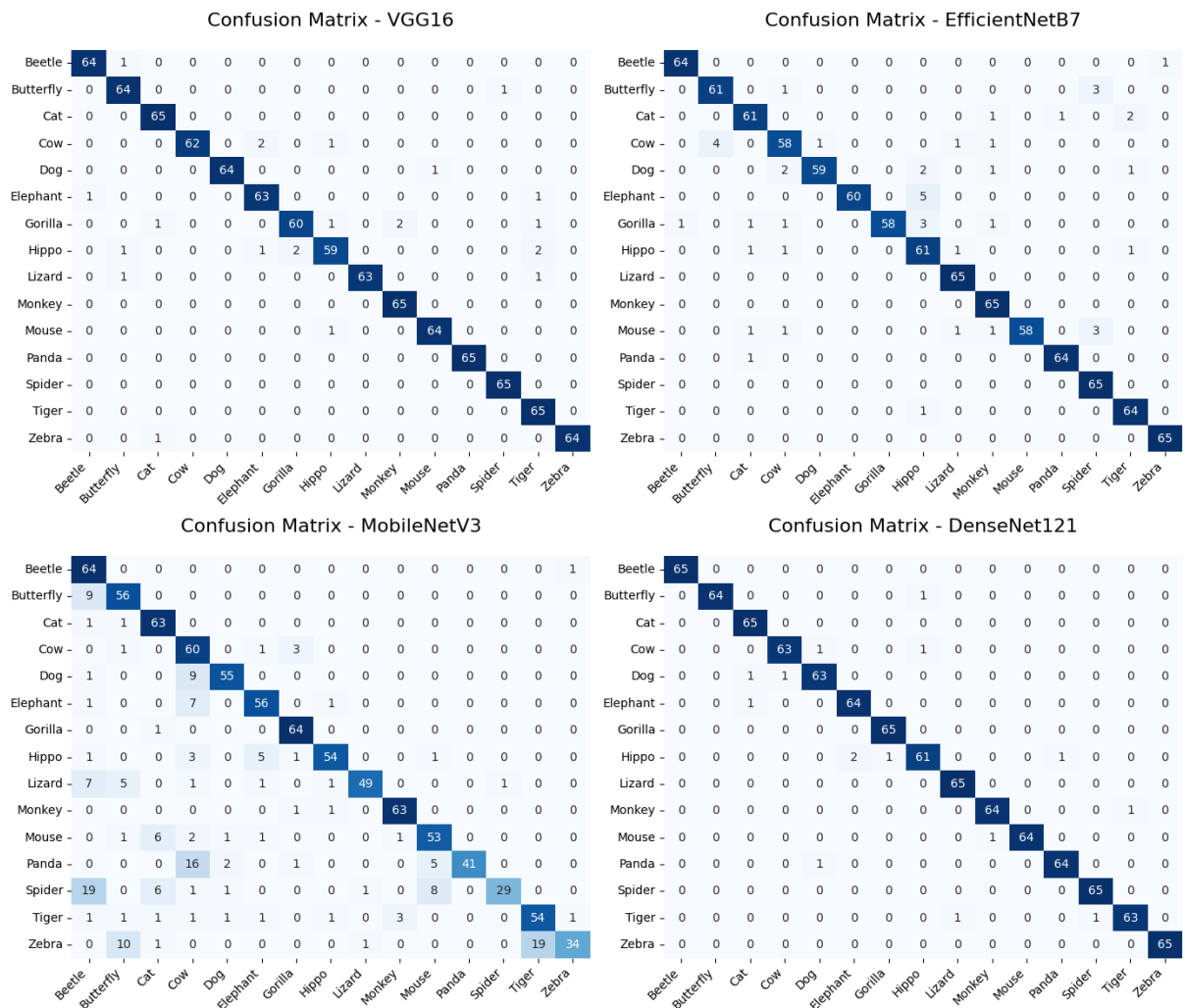
4. MobileNetV3:

- Đây là mô hình có hiệu suất thấp nhất trong số các mô hình được so sánh, với Accuracy và Recall chỉ đạt 0.8154 và F1-score là 0.8117.
- Tuy nhiên, MobileNetV3 có ưu điểm là nhẹ và nhanh, phù hợp cho các ứng dụng yêu cầu tài nguyên tính toán thấp hoặc cần triển khai trên thiết bị di động

Tóm lại thì model DenseNet121 cho hiệu suất cao nhất trong việc training

3.3.3 So sánh Confusion Matrix

Ma trận nhầm lẫn (confusion matrix) cho thấy số lượng dự đoán đúng và sai của mô hình cho từng lớp. Ma trận này giúp chúng tôi hiểu rõ hơn về khả năng phân loại của mô hình đối với từng loại động vật và xác định những lớp nào mô hình còn gặp khó khăn trong việc phân loại chính xác.



Hình 3-6: Ma trận nhầm lẫn của các mô hình học sâu.

1. DenseNet121:

- Là mô hình có hiệu suất tốt nhất với Confusion Matrix thể hiện số lượng dự đoán chính xác rất cao cho tất cả các lớp.
- Gần như không có lỗi hoặc rất ít lỗi ở các lớp, minh chứng cho khả năng phân loại xuất sắc của DenseNet121.
- Các ô chính diagonal đều có giá trị rất cao, cho thấy mô hình này dự đoán chính xác nhất.

2. VGG16:

- Hiệu suất rất tốt, các ô diagonal cũng có giá trị cao, chỉ thua kém DenseNet121 một chút.
- Một vài lỗi nhỏ ở các lớp như "Cow" và "Gorilla" nhưng tổng thể vẫn rất chính xác.

3. EfficientNetB7:

- Hiệu suất tốt nhưng có một số lỗi dự đoán không chính xác hơn so với VGG16 và DenseNet121.
- Một số lớp như "Cow" và "Gorilla" có số lượng lỗi cao hơn.
- Các ô diagonal vẫn chủ yếu có giá trị cao, chứng tỏ mô hình vẫn hoạt động tốt.

4. MobileNetV3:

- Hiệu suất thấp nhất trong số các mô hình được so sánh, thể hiện rõ ràng qua Confusion Matrix với nhiều lỗi dự đoán.
- Các lớp như "Panda", "Tiger", và "Zebra" có số lượng lỗi rất cao, điều này cho thấy mô hình gặp khó khăn trong việc phân loại chính xác các đối tượng này.
- Các ô diagonal có giá trị thấp hơn nhiều so với các mô hình khác, cho thấy mô hình không mạnh bằng.

Tổng kết thì DenseNet121 là model vượt trội nhất, với số lượng lỗi ít và độ chính xác cao nhất.

Chương 4: ỨNG DỤNG THỰC TIỄN, HƯỚNG PHÁT TRIỂN VÀ KẾT LUẬN

4.1 Ứng dụng thực tiễn

Mô hình DenseNet121 có thể được ứng dụng rộng rãi trong nhiều lĩnh vực thực tiễn, bao gồm:

- **Nghiên cứu sinh học và bảo tồn động vật:** Giúp các nhà nghiên cứu phân loại và theo dõi các loài động vật trong tự nhiên, từ đó có những biện pháp bảo tồn và bảo vệ các loài quý hiếm hoặc có nguy cơ tuyệt chủng.
- **Nông nghiệp thông minh:** Sử dụng mô hình để nhận dạng và quản lý các loài động vật gây hại trong nông nghiệp, giúp nâng cao hiệu quả sản xuất và bảo vệ mùa màng.
- **Công nghệ giám sát và an ninh:** Áp dụng trong các hệ thống giám sát tự động để nhận dạng và theo dõi các loài động vật trong các khu vực bảo tồn, công viên quốc gia, và các khu vực đô thị.
- **Ứng dụng giáo dục và giải trí:** Phát triển các ứng dụng giáo dục về động vật học, cung cấp thông tin và kiến thức về các loài động vật cho học sinh, sinh viên và người yêu thích thiên nhiên.

4.2 Hướng phát triển

Để phát huy hết tiềm năng của mô hình DenseNet121 và cải thiện hơn nữa hiệu suất của mô hình, các hướng phát triển trong tương lai bao gồm:

- **Tăng cường dữ liệu:** Tiếp tục thu thập và sử dụng các kỹ thuật tăng cường dữ liệu để tạo ra tập dữ liệu phong phú và đa dạng hơn, giúp mô hình học tốt hơn và cải thiện khả năng tổng quát hóa.
- **Tối ưu hóa mô hình:** Áp dụng các kỹ thuật tối ưu hóa và giảm tham số để cải thiện tốc độ và hiệu quả sử dụng tài nguyên của mô hình, giúp mô hình hoạt động hiệu quả hơn trên các thiết bị có tài nguyên hạn chế.
- **Triển khai và thử nghiệm trên các tập dữ liệu khác:** Thử nghiệm mô hình trên các tập dữ liệu khác nhau để đánh giá khả năng mở rộng và tính linh hoạt của mô hình trong các bối cảnh ứng dụng khác nhau.
- **Phát triển ứng dụng thực tế:** Xây dựng các ứng dụng thực tế dựa trên mô hình DenseNet121 để khai thác tiềm năng của mô hình trong các lĩnh vực nghiên cứu, bảo tồn, nông nghiệp, giám sát, giáo dục và giải trí.

- **Hợp tác nghiên cứu và phát triển:** Tăng cường hợp tác với các tổ chức nghiên cứu, trường đại học, và các công ty công nghệ để tiếp tục cải tiến và phát triển các ứng dụng dựa trên mô hình học sâu.

4.3 Kết luận

Qua quá trình nghiên cứu và triển khai, mô hình học sâu DenseNet121 đã chứng minh được hiệu quả vượt trội trong việc nhận dạng động vật qua hình ảnh. Cụ thể:

- **Độ chính xác cao:** Mô hình DenseNet121 đạt độ chính xác cao hơn so với các mô hình khác như VGG16, EfficientNetB7 và MobileNetV3. Điều này khẳng định khả năng của DenseNet121 trong việc học và phân loại các đặc trưng phức tạp từ hình ảnh động vật.
- **Khả năng tổng quát hóa tốt:** Mô hình cho thấy khả năng tổng quát hóa tốt khi hoạt động ổn định trên tập dữ liệu kiểm thử, bao gồm các hình ảnh mà mô hình chưa từng thấy trong quá trình huấn luyện.
- **Hiệu quả sử dụng tài nguyên:** Với số lượng tham số ít hơn nhưng vẫn duy trì được hiệu suất cao, DenseNet121 là một lựa chọn lý tưởng cho các ứng dụng yêu cầu tài nguyên tính toán hạn chế.

Những kết quả đạt được từ quá trình huấn luyện và kiểm thử mô hình đã khẳng định tiềm năng lớn của DenseNet121 trong lĩnh vực nhận dạng hình ảnh động vật.

TÀI LIỆU THAM KHẢO

1. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2018). Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Retrieved from <https://arxiv.org/abs/1608.06993>
2. Chollet, F. (2017). *Deep Learning with Python*. Manning Publications.
3. Scikit-learn Documentation. *Metrics and scoring: quantifying the quality of predictions*. Retrieved from https://scikit-learn.org/stable/modules/model_evaluation.html
4. Gao Huang et al., “Densely Connected Convolutional Networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
5. Truong-Minh Le, Bao-Thien Nguyen-Tat, Vuong M. Ngo. “Automated evaluation of Tuberculosis using Deep Neural Networks”. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 2022.
6. Tan, Mingxing, and Quoc V. Le. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.” in Proceedings of the International Conference on Machine Learning (ICML), 2019.