

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

BÁO CÁO GIỮA KÌ

MÔN HỌC: NHẬP MÔN THỊ GIÁC MÁY TÍNH

ĐỀ TÀI: PHÂN LOẠI MẶT ĐỘNG VẬT (ANIMAL FACE)

Giảng viên hướng dẫn: Ths.Lê Minh Hưng

Sinh viên thực hiện: Đỗ Trọng Khánh – 19521676
Võ Phạm Duy Đức – 19521383
Trịnh Công Danh – 19521326

Lớp: CS231.M13.KHCL

I. MÔ TẢ BỘ DỮ LIỆU

- Bộ dữ liệu của của nhóm gồm 6 lớp: dog, cat, face, wolf, lion, tiger.
- Tổng số hình nhóm thu thập được là: 2763 hình → Trung bình mỗi lớp khoảng 460 hình

Dog face: 461 hình



Wolf face: 436 hình

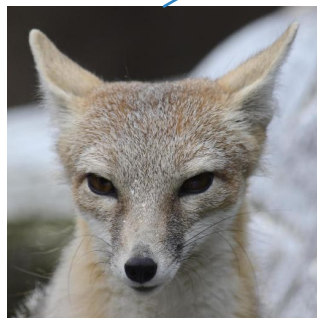


Lion face: 470 image

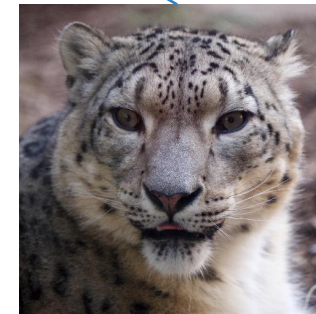


ANIMAL
FACE

Cat face: 460 hình



Fox face: 477 hình



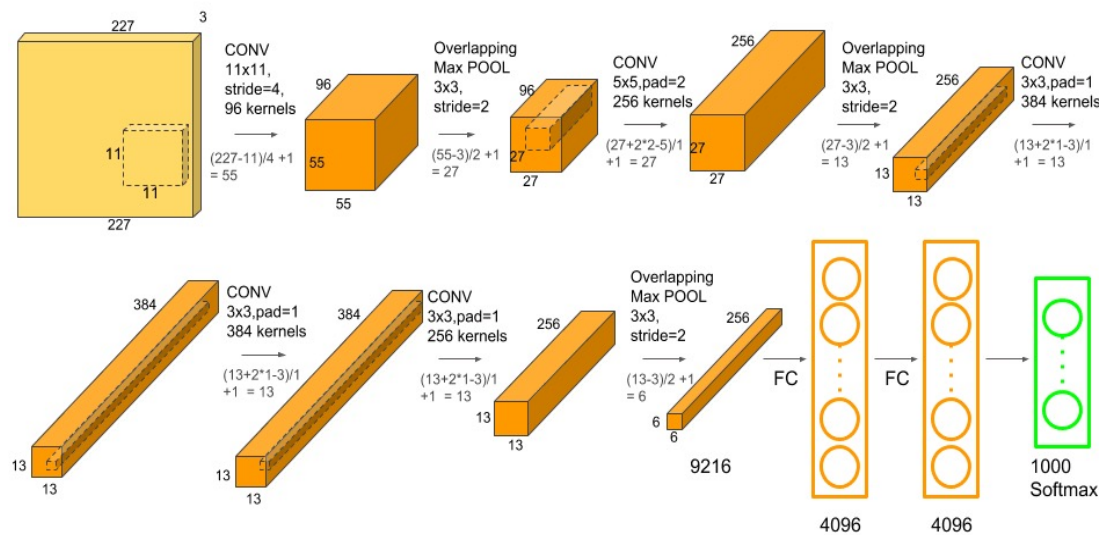
Tiger face: 459 hình

II. TRÍCH XUẤT ĐẶC TRƯNG VÀ HUẤN LUYỆN MODEL

1. MÔ HÌNH MẠNG CNN-AlexNet

a. Giới thiệu về mô hình và Ý tưởng

- Mạng AlexNet bao gồm 8 lớp (tính luôn lớp input là 9)
- Sử dụng kernel có size $11 \times 11 \times 3$
- Nhược điểm của mô hình mạng AlexNet:
 - + Dùng kernel size lớn (11×11) dẫn đến số lượng trọng số nhiều và vùng quét thông tin rộng → Không lấy được nhiều thông tin chi tiết.
 - + Số lượng trọng số ở lớp Fully-connected Layer quá lớn trên tổng số lượng trọng số của mô hình mạng.

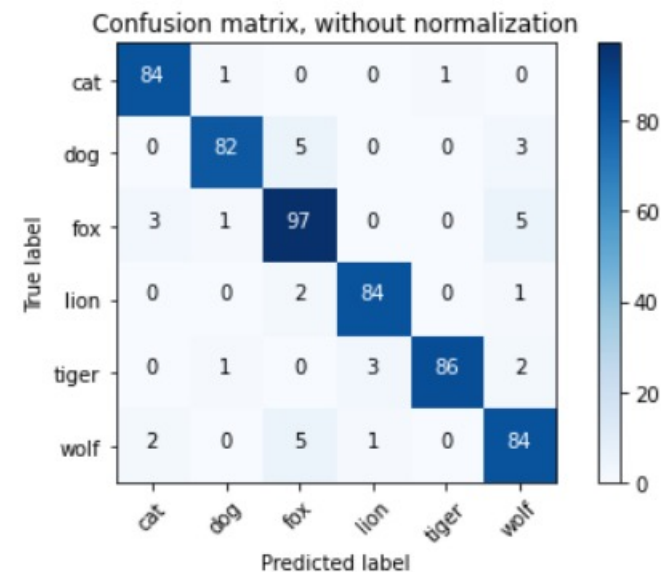


- Với model Alexnet, nhóm sẽ bỏ lớp Softmax và lấy đặc trưng của lớp Fully connected cuối cùng.
→ Lấy vector đặc trưng đó đưa vào mô hình để huấn luyện

b. Huấn luyện và Đánh giá mô hình

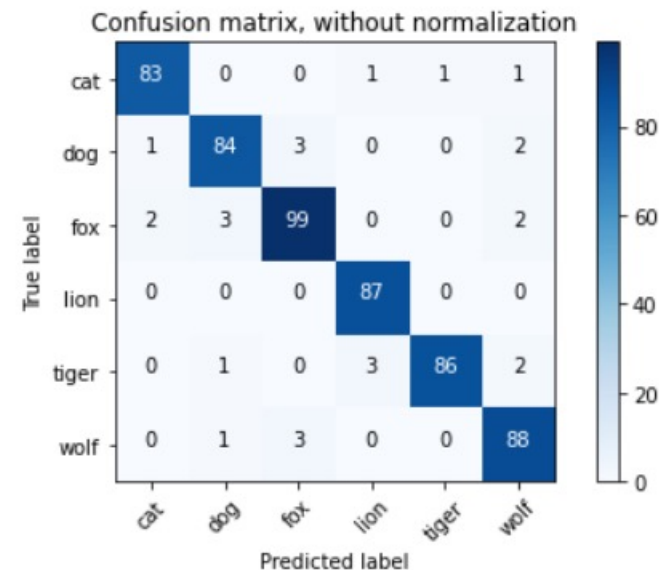
Model Support Vector Machine (SVM)

	precision	recall	f1-score	support
cat	0.94	0.98	0.96	86
dog	0.96	0.91	0.94	90
fox	0.89	0.92	0.90	106
lion	0.95	0.97	0.96	87
tiger	0.99	0.93	0.96	92
wolf	0.88	0.91	0.90	92
accuracy			0.93	553
macro avg	0.94	0.94	0.94	553
weighted avg	0.94	0.93	0.94	553



Model Logistic Regression

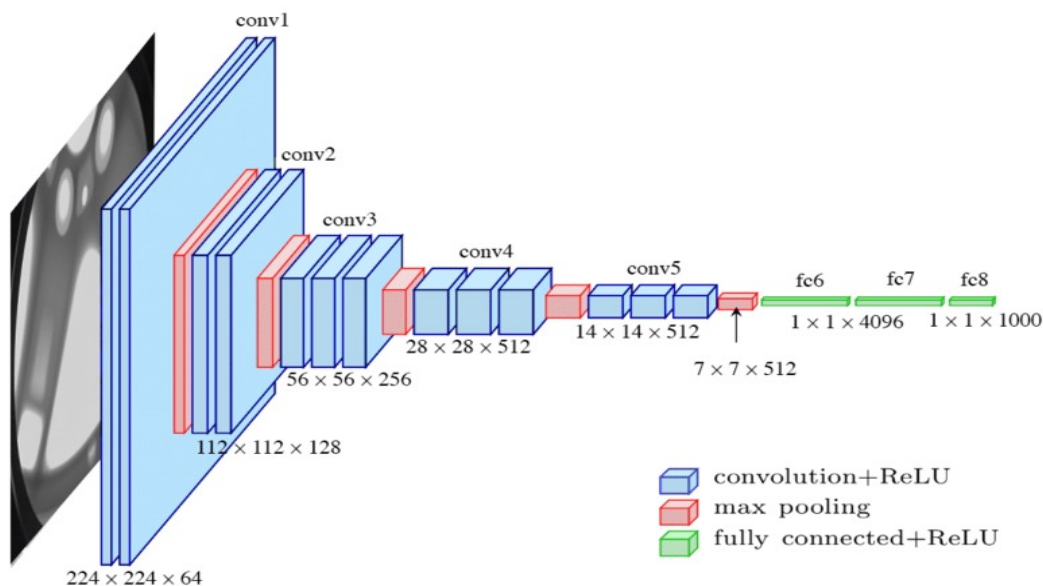
	precision	recall	f1-score	support
cat	0.97	0.97	0.97	86
dog	0.94	0.93	0.94	90
fox	0.94	0.93	0.94	106
lion	0.96	1.00	0.98	87
tiger	0.99	0.93	0.96	92
wolf	0.93	0.96	0.94	92
accuracy			0.95	553
macro avg	0.95	0.95	0.95	553
weighted avg	0.95	0.95	0.95	553



2. MÔ HÌNH MẠNG CNN-VGG16

a. Giới thiệu về mô hình và Ý tưởng

- **VGG16** sử dụng 3 lớp **conv layer 3x3**
- Nhờ việc sử dụng 3 lớp **conv layer 3x3** làm tăng tính **effective receptive field** hơn so với lớp **conv layer 11x11** của mô hình **AlexNet**. Ngoài ra nó sẽ giúp tăng tính **non-linearities**

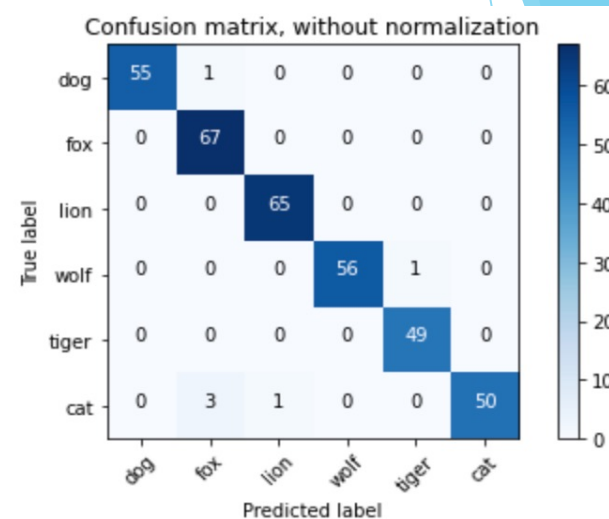


- Trong **VGG16** ở lớp **FC7 features** thì có tính tổng quát hoá cho nhiều đặc vụ khác.
- Do đó nhóm sẽ sử dụng VGG16 để trích xuất đặc trưng ở lớp **FC7 features**.
- Sau đó sẽ Flatten các đặc trưng thành vector.
- Cuối cùng là đem những vector đó vào mô hình để huấn luyện.

b. Huấn luyện và Đánh giá mô hình

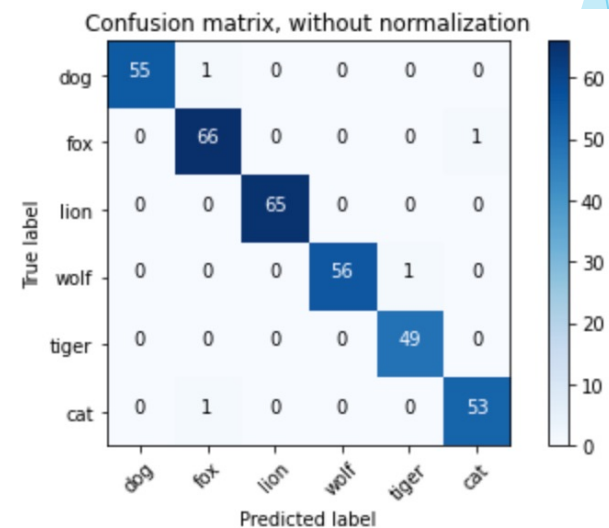
Model Support Vector Machine (SVM)

	precision	recall	f1-score	support
dog	1.00	0.98	0.99	56
fox	0.94	1.00	0.97	67
lion	0.98	1.00	0.99	65
wolf	1.00	0.98	0.99	57
tiger	0.98	1.00	0.99	49
cat	1.00	0.93	0.96	54
accuracy			0.98	348
macro avg	0.98	0.98	0.98	348
weighted avg	0.98	0.98	0.98	348



Model Logistic Regression

	precision	recall	f1-score	support
dog	1.00	0.98	0.99	56
fox	0.97	0.99	0.98	67
lion	1.00	1.00	1.00	65
wolf	1.00	0.98	0.99	57
tiger	0.98	1.00	0.99	49
cat	0.98	0.98	0.98	54
accuracy			0.99	348
macro avg	0.99	0.99	0.99	348
weighted avg	0.99	0.99	0.99	348



3. TRÍCH XUẤT ĐẶC TRƯNG SIFT TRONG OPEN CV

a. Giới thiệu tổng quan về phương pháp trích xuất đặc trưng SIFT

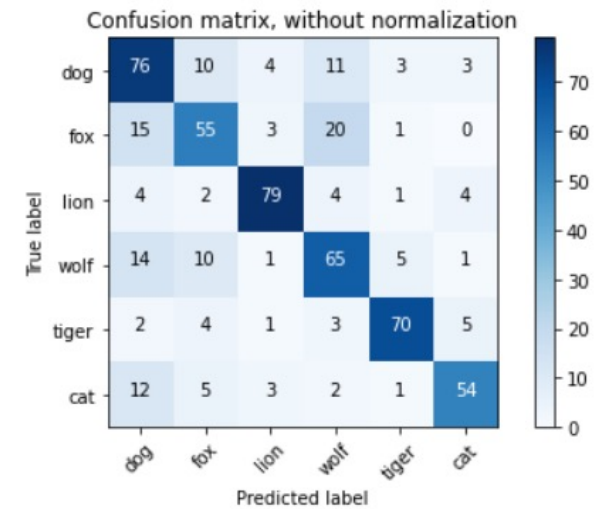
- SIFT là một thuật toán có hiệu quả khá cao vì dựa theo các cục bộ bất biến trong ảnh. Đặc trưng được trích chọn trong SIFT là các điểm đặc biệt keypoints
- Phương pháp trích chọn điểm đặc trưng cục bộ bất biến SIFT gồm các bước:
 - + Dò tìm các điểm cực trị (Scale-space Extrema Detection)
 - + Lọc và trích xuất các điểm đặc biệt (Keypoint Localization)
 - + Xác định hướng cho các điểm nổi bật
 - + Mô tả các điểm nổi bật



b. Huấn luyện và Đánh giá mô hình

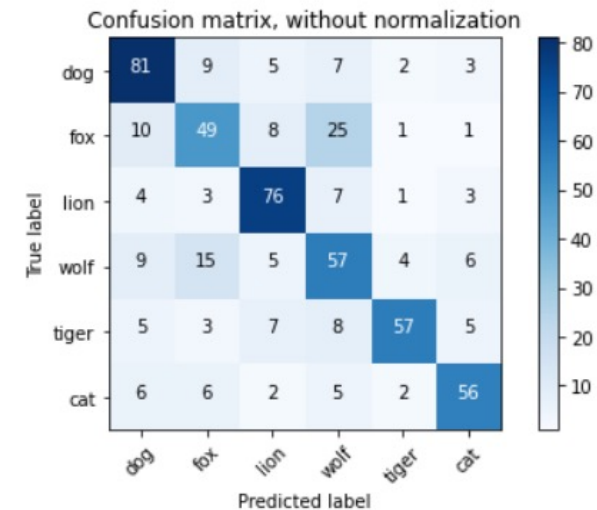
Model Support Vector Machine (SVM)

	precision	recall	f1-score	support
dog	0.62	0.71	0.66	107
fox	0.64	0.59	0.61	94
lion	0.87	0.84	0.85	94
wolf	0.62	0.68	0.65	96
tiger	0.86	0.82	0.84	85
cat	0.81	0.70	0.75	77
accuracy			0.72	553
macro avg	0.74	0.72	0.73	553
weighted avg	0.73	0.72	0.72	553



Model Logistic Regression

	precision	recall	f1-score	support
dog	0.70	0.76	0.73	107
fox	0.58	0.52	0.55	94
lion	0.74	0.81	0.77	94
wolf	0.52	0.59	0.56	96
tiger	0.85	0.67	0.75	85
cat	0.76	0.73	0.74	77
accuracy			0.68	553
macro avg	0.69	0.68	0.68	553
weighted avg	0.69	0.68	0.68	553



III. NHẬN XÉT

Model Support Vector Machine (SVM)

	Dog	Cat	Fox	Lion	Tiger	Wolf
Alexnet	94%	96%	90%	96%	96%	90%
VGG16	99%	96%	97%	99%	99%	99%
SIFT	66%	75%	61%	85%	84%	65%

Model Logistic Regression

	Dog	Cat	Fox	Lion	Tiger	Wolf
Alexnet	94%	97%	94%	98%	96%	94%
VGG16	99%	98%	98%	100%	99%	99%
SIFT	73%	74%	55%	77%	75%	56%

- Model VGG16 cho kết quả tốt nhất ở cả hai mô hình SVM và Logistic Regression
- Trích xuất đặc trưng SIFT cho kết quả thấp nhất ở cả hai mô hình SVM và Logistic Regression
 - + Lớp Fox bị dự đoán qua lớp Wolf khá nhiều

THANKS
FOR
WATCHING!