



BIRDS

CLASSIFICATION

Deep learning - CS431.N11

Thành viên



Hồ Mỹ Hạnh

19521470



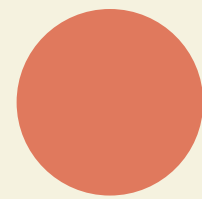
**Đinh Hoàng
Linh Đan**

19521309



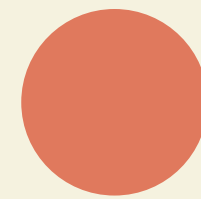
**Trần Nguyễn
Quỳnh Anh**

19521217



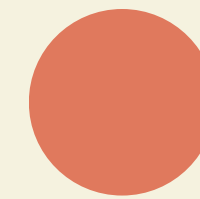
I

**Phát biểu
bài toán**



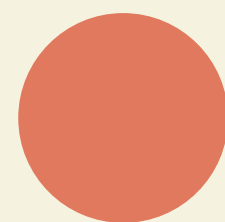
II

Bộ dữ liệu



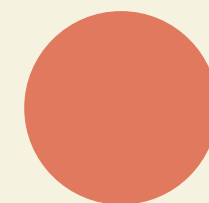
III

Mô hình



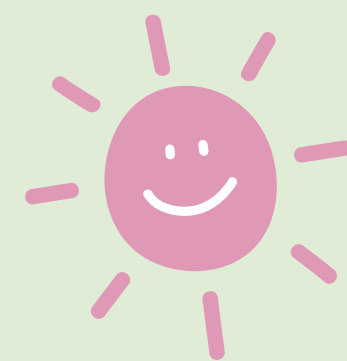
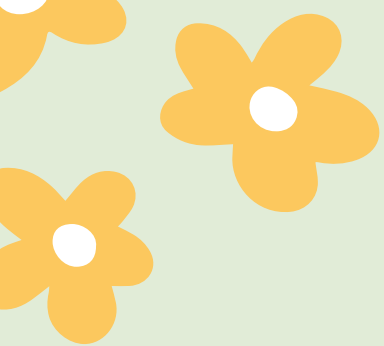
IV

**Độ đo
đánh giá**

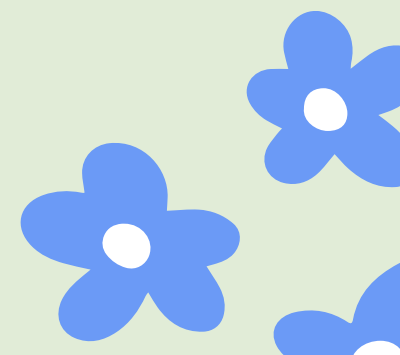


V

**Kết quả
thử
nghiệm**



PHÁT BIỂU BÀI TOÁN

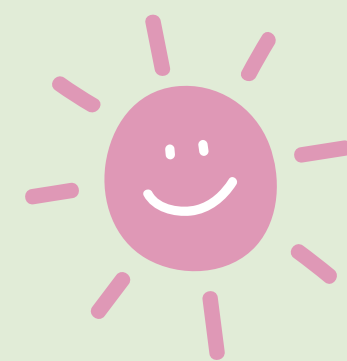
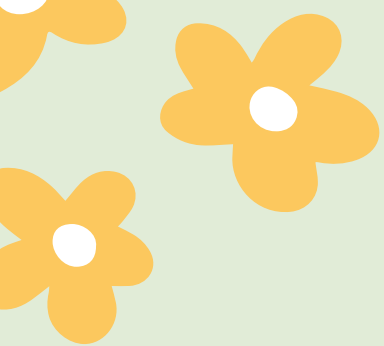


Phát biểu bài toán

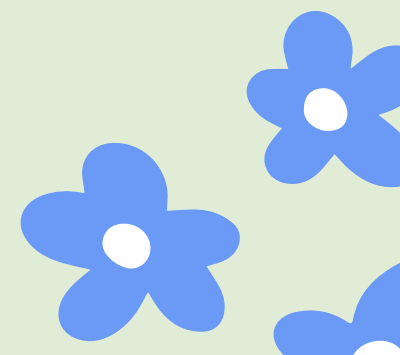
- Bài toán phân loại loài chim giúp phân biệt các loài chim, phục vụ cho quá trình bảo tồn các loài chim, giúp ích cho việc dạy học về những loài chim.
- **Input:** ảnh chứa 1 loài chim
- **Output:** tên của loài chim

Predict label: CAPE GLOSSY STARLING
Original label: CAPE GLOSSY STARLING





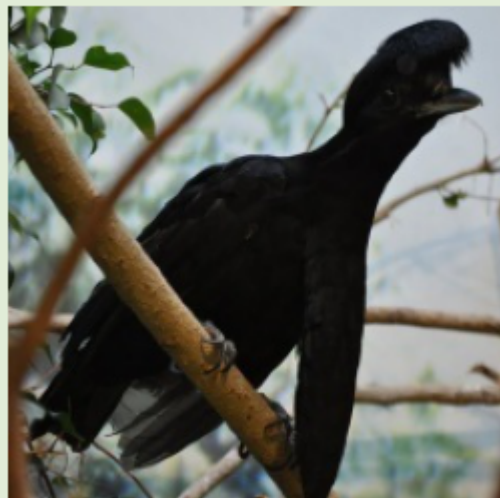
BỘ DỮ LIỆU



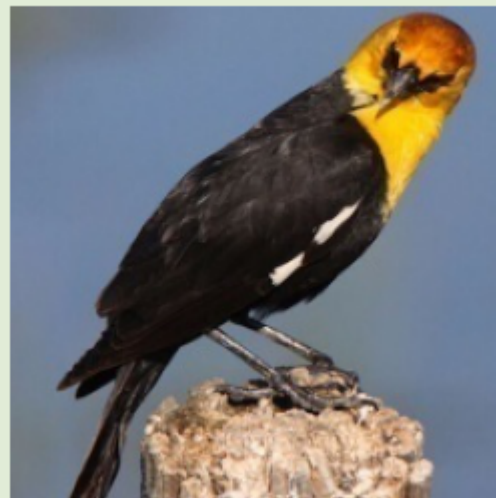
BIRDS 450 SPECIES - IMAGE CLASSIFICATION

Bộ dữ liệu thuộc Kaggle về 450 loài chim, bao gồm 70,626 training images, 2250 test images (5 ảnh mỗi loài) và 2250 validation images (5 ảnh mỗi loài)

UMBRELLA BIRD



YELLOW HEADED BLACKBIRD



INDIAN PITTA



UMBRELLA BIRD

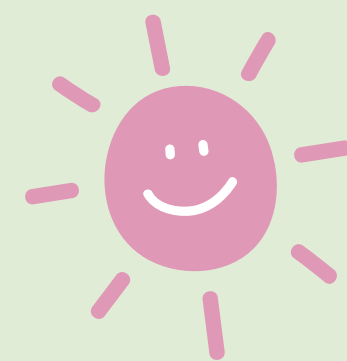
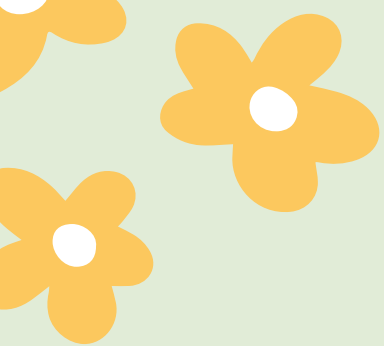


YELLOW HEADED BLACKBIRD

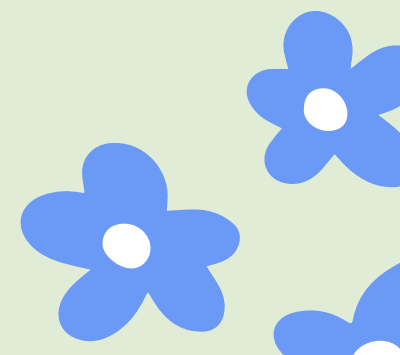


INDIAN PITTA





MÔ HÌNH

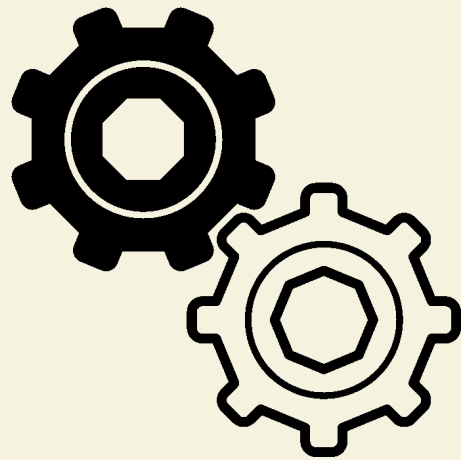


Protocol

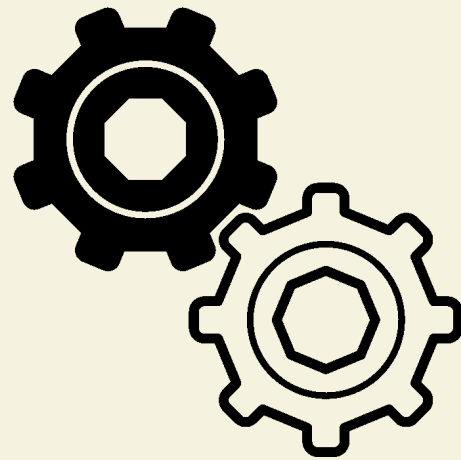
- Learning rate: 0.1, 0.01, 0.001
- Batch size: 32, 64
- Epoch: 5
- Optimizer: Adam, SGD, RMSprop
- Với SGD có sử dụng momentum (0.9, 0.95) và không sử dụng momentum



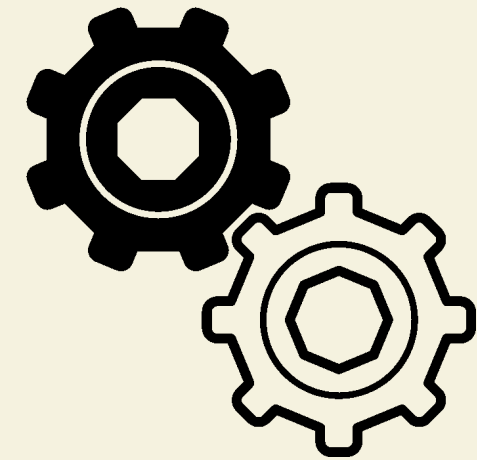
Inception V3



EfficientNet B0



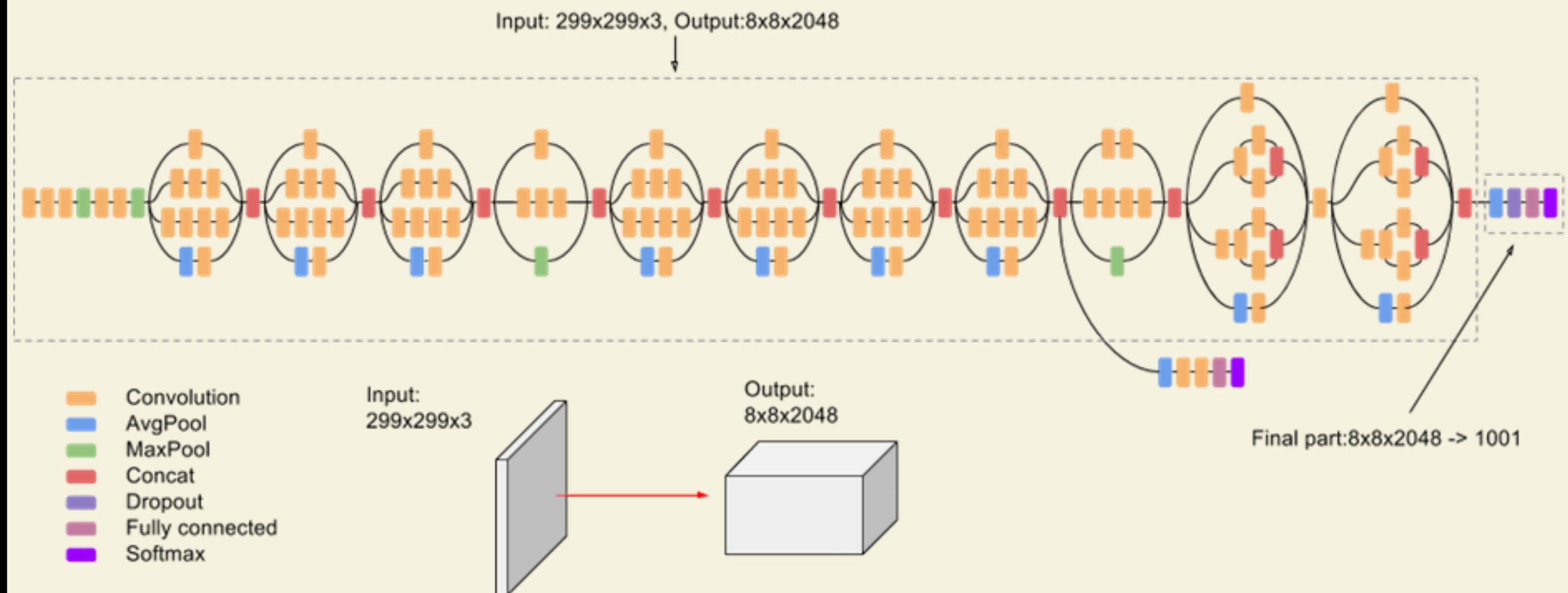
MobileNet V3



01

InceptionV3

- Mô hình Inception v3 dựa trên bài báo gốc: "Rethinking the Inception Architecture for Computer Vision" của Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna được đề xuất 12/2015.
- Trong đồ án, sử dụng mô hình InceptionV3 pretrained của Keras.



| | Adam | | SGD | | | RMSprop | |
|------------|---------------|----------|---------------|----------|----------|---------------|----------|
| batch_size | learning_rate | accuracy | learning_rate | momentum | accuracy | learning_rate | accuracy |
| 32 | 0.01 | 0.8082 | 0.01 | 0.9 | 0.8581 | 0.01 | 0.7886 |
| | | | | 0.95 | 0.8821 | | |
| | | | | x | 0.7353 | | |
| | 0.001 | 0.8848 | 0.001 | 0.9 | 0.7211 | 0.001 | 0.8704 |
| | | | | 0.95 | 0.8323 | | |
| | | | | x | 0.3924 | | |
| | 0.0001 | 0.8363 | 0.0001 | 0.9 | 0.4236 | 0.0001 | 0.7643 |
| | | | | 0.95 | 0.5987 | | |
| | | | | x | 0.1632 | | |
| 64 | 0.01 | 0.8141 | 0.01 | 0.9 | 0.8483 | 0.01 | 0.7834 |
| | | | | 0.95 | 0.871 | | |
| | | | | x | 0.6831 | | |
| | 0.001 | 0.9066 | 0.001 | 0.9 | 0.6822 | 0.001 | 0.8611 |
| | | | | 0.95 | 0.7426 | | |
| | | | | x | 0.2215 | | |
| | 0.0001 | 0.8069 | 0.0001 | 0.9 | 0.2281 | 0.0001 | 0.7779 |
| | | | | 0.95 | 0.4039 | | |
| | | | | x | 0.1474 | | |



02

EfficientNet

B0

- Năm 2019, nhóm nghiên cứu Google gồm Mingxing Tan và Quoc V. Le đã phát hành bài báo liên quan tới họ CNN mới: "*EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*"
- Kiến trúc mạng EfficientNet là một mạng *convolutional neural network*:
 - Tăng accuracy
 - Giảm số parameters và FLOPS.



02

EfficientNet B0

Table 1. EfficientNet-B0 baseline network – Each row describes a stage i with \hat{L}_i layers, with input resolution $\langle \hat{H}_i, \hat{W}_i \rangle$ and output channels \hat{C}_i . Notations are adopted from equation 2.

| Stage i | Operator $\hat{\mathcal{F}}_i$ | Resolution $\hat{H}_i \times \hat{W}_i$ | #Channels \hat{C}_i | #Layers \hat{L}_i |
|--------------|-----------------------------------|--|--------------------------|------------------------|
| 1 | Conv3x3 | 224×224 | 32 | 1 |
| 2 | MBConv1, k3x3 | 112×112 | 16 | 1 |
| 3 | MBConv6, k3x3 | 112×112 | 24 | 2 |
| 4 | MBConv6, k5x5 | 56×56 | 40 | 2 |
| 5 | MBConv6, k3x3 | 28×28 | 80 | 3 |
| 6 | MBConv6, k5x5 | 14×14 | 112 | 3 |
| 7 | MBConv6, k5x5 | 14×14 | 192 | 4 |
| 8 | MBConv6, k3x3 | 7×7 | 320 | 1 |
| 9 | Conv1x1 & Pooling & FC | 7×7 | 1280 | 1 |

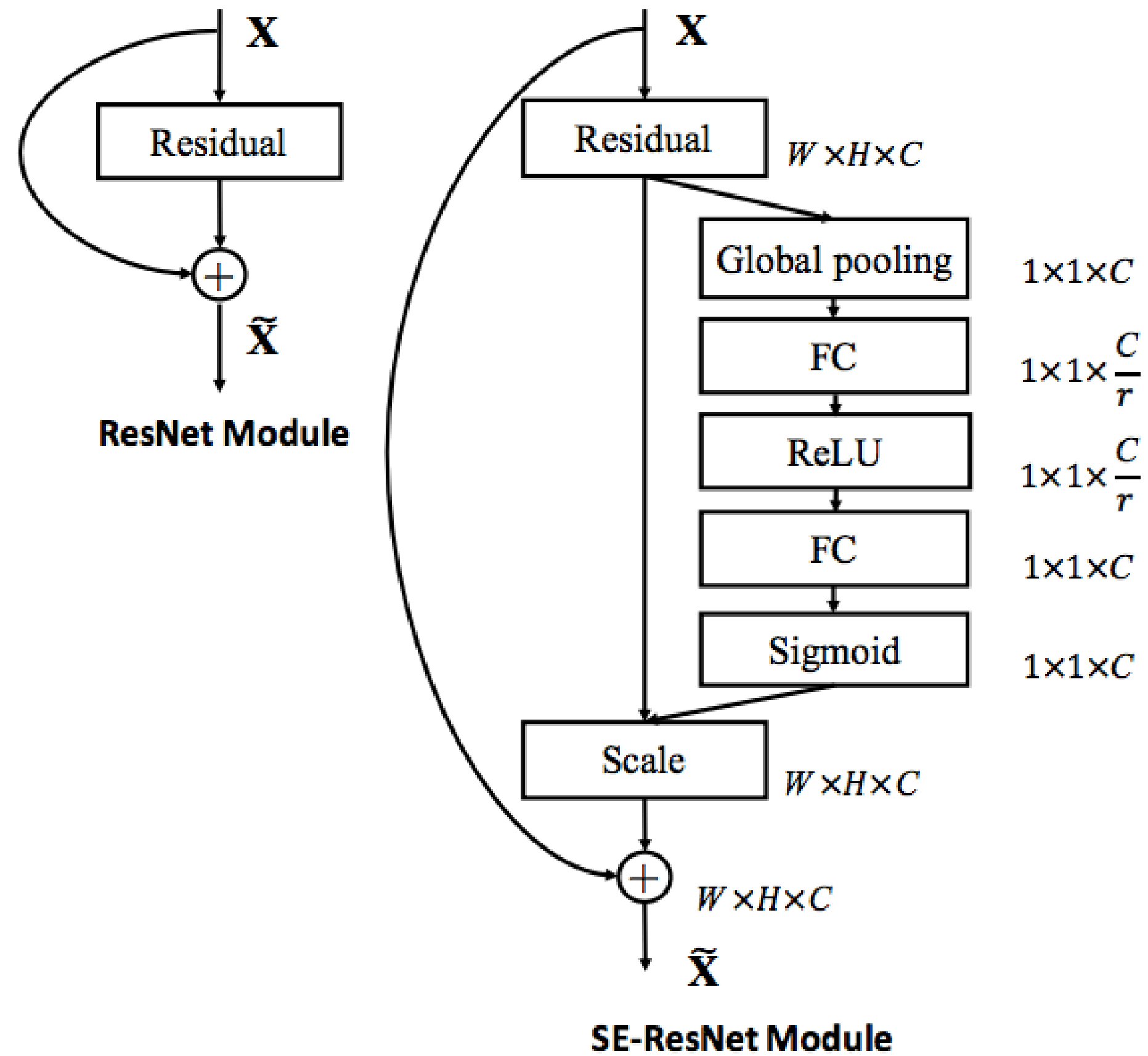
EfficientNet-B0 baseline network

| | Adam | | SGD | | | RMSprop | |
|------------|---------------|----------|---------------|----------|----------|---------------|----------|
| batch_size | learning_rate | accuracy | learning_rate | momentum | accuracy | learning_rate | accuracy |
| 32 | 0.01 | 0.7657 | 0.01 | 0.9 | 0.8013 | 0.01 | 0.7305 |
| | | | | 0.95 | 0.8231 | | |
| | | | | x | 0.612 | | |
| | 0.001 | 0.8446 | 0.001 | 0.9 | 0.6143 | 0.001 | 0.7704 |
| | | | | 0.95 | 0.7054 | | |
| | | | | x | 0.063 | | |
| | 0.0001 | 0.7639 | 0.0001 | 0.9 | 0.065 | 0.0001 | 0.7342 |
| | | | | 0.95 | 0.1985 | | |
| | | | | x | 0.003 | | |
| 64 | 0.01 | 0.7747 | 0.01 | 0.9 | 0.7711 | 0.01 | 0.7642 |
| | | | | 0.95 | 0.8018 | | |
| | | | | x | 0.4472 | | |
| | 0.001 | 0.8448 | 0.001 | 0.9 | 0.447 | 0.001 | 0.8123 |
| | | | | 0.95 | 0.6086 | | |
| | | | | x | 0.0186 | | |
| | 0.0001 | 0.731 | 0.0001 | 0.9 | 0.0148 | 0.0001 | 0.7258 |
| | | | | 0.95 | 0.0604 | | |
| | | | | x | 0.0025 | | |

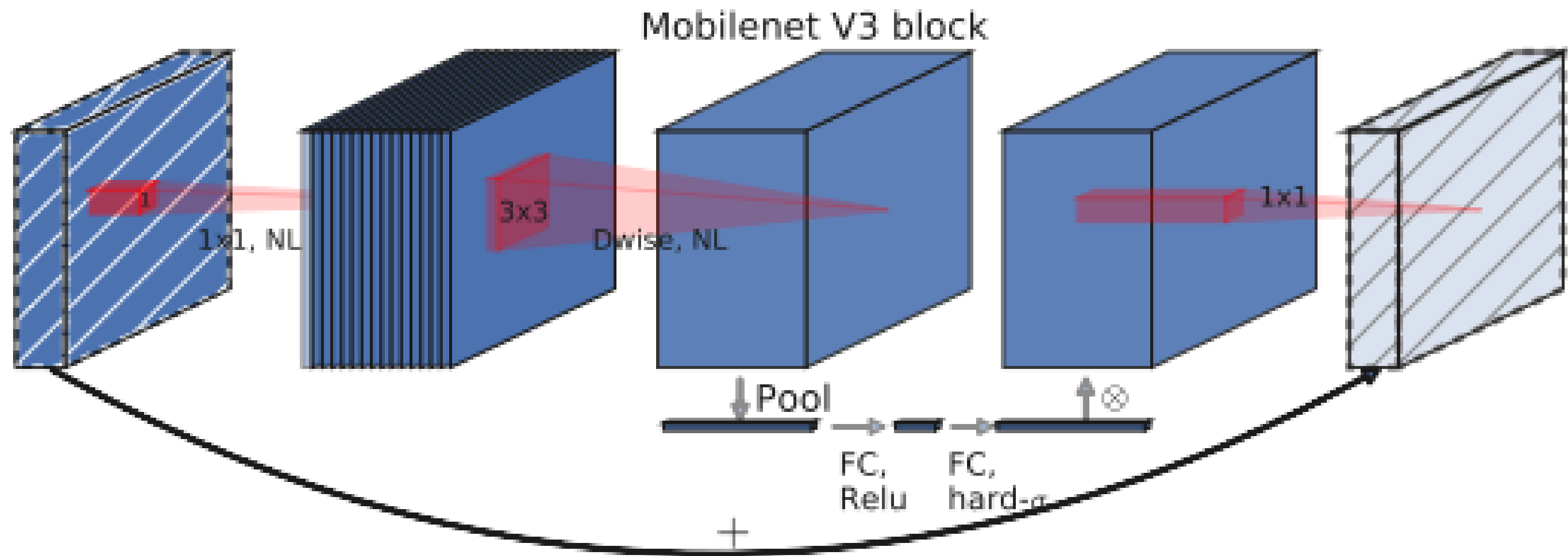
03

MobileNet V3

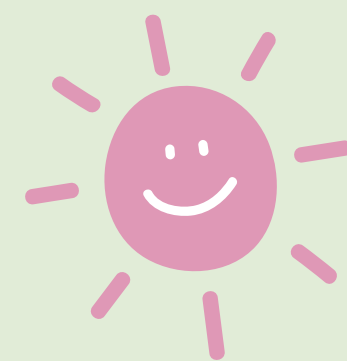
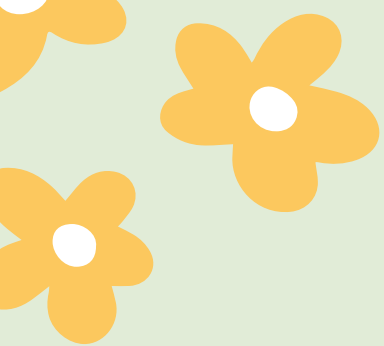
- MobileNet V3 được phát triển bởi đội ngũ Google, được giới thiệu lần đầu năm 2017.
- Sử dụng các tính chập tích tên DSC (Depthwise Separable Convolution) nhằm giảm kích thước mô hình và giảm độ phức tạp tính toán.
- Ý tưởng của Depthwise Separable Convolution là chia phép convolution làm 2 phần: Depthwise convolution & Pointwise convolution



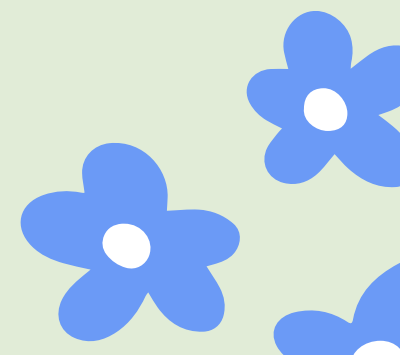
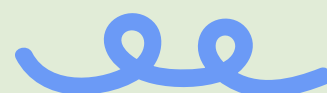
Kiến trúc mạng của MobileNet V3



| | Adam | | SGD | | | RMSprop | |
|------------|---------------|----------|---------------|----------|----------|---------------|----------|
| batch_size | learning_rate | accuracy | learning_rate | momentum | accuracy | learning_rate | accuracy |
| 32 | 0.01 | 0.2289 | 0.01 | 0.9 | 0.8723 | 0.01 | 0.1018 |
| | | | | 0.95 | 0.8801 | | |
| | | | | x | 0.8267 | | |
| | 0.001 | 0.8084 | 0.001 | 0.9 | 0.7876 | 0.001 | 0.7116 |
| | | | | 0.95 | 0.8901 | | |
| | | | | x | 0.2301 | | |
| | 0.0001 | 0.8191 | 0.0001 | 0.9 | 0.2141 | 0.0001 | 0.8784 |
| | | | | 0.95 | 0.423 | | |
| | | | | x | 0.063 | | |
| 64 | 0.01 | 0.1471 | 0.01 | 0.9 | 0.8933 | 0.01 | 0.1547 |
| | | | | 0.95 | 0.8707 | | |
| | | | | x | 0.8329 | | |
| | 0.001 | 0.7449 | 0.001 | 0.9 | 0.8524 | 0.001 | 0.6027 |
| | | | | 0.95 | 0.8502 | | |
| | | | | x | 0.224 | | |
| | 0.0001 | 0.9058 | 0.0001 | 0.9 | 0.2093 | 0.0001 | 0.8142 |
| | | | | 0.95 | 0.416 | | |
| | | | | x | 0.053 | | |



ĐỘ ĐO ĐÁNH GIÁ



ACCURACY

F1-SCORE
Macro

PRECISION
Macro

RECALL
Macro

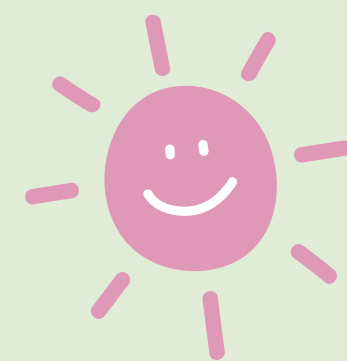
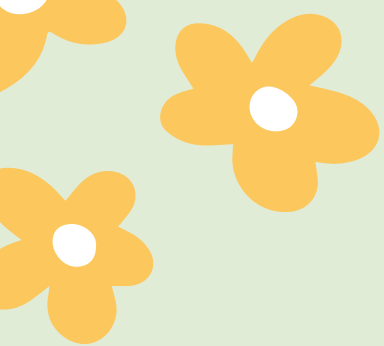
- Accuracy: là tỷ lệ phần trăm dự đoán đúng cho dữ liệu thử nghiệm.

$$accuracy = \frac{\text{correct predictions}}{\text{all predictions}}$$

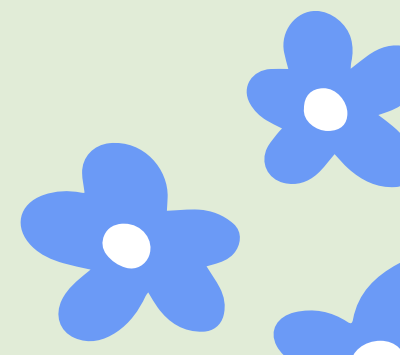
- Macro average metric:

$$. \text{ Macroaveraging Precision } Prc^{macro}(D) = \frac{\sum_{c_i \in \mathcal{C}} Prc(D, c_i)}{|\mathcal{C}|}$$

$$. \text{ Macroaveraging Recall } Rec^{macro}(D) = \frac{\sum_{c_i \in \mathcal{C}} Rcl(D, c_i)}{|\mathcal{C}|}$$



KẾT QUẢ THỬ NGHIỆM



| | Inception V3 | EfficientNet B0 | MobileNet V3 |
|------------|--------------|-----------------|--------------|
| Accuracy | 0.9458 | 0.9689 | 0.9258 |
| F1 - score | 0.9506 | 0.9716 | 0.9335 |
| Precision | 0.9556 | 0.9743 | 0.9414 |
| Recall | 0.9458 | 0.9689 | 0.9258 |

| | Inception V3 | EfficientNet B0 | MobileNet V3 |
|------------------|--------------|-----------------|--------------|
| Số lượng tham số | 22 724 834 | 4 049 571 | 7 486 554 |

Predict label: BREWERS BLACKBIRD
Original label: BREWERS BLACKBIRD



Predict label: BREWERS BLACKBIRD
Original label: BREWERS BLACKBIRD



Predict label: COMMON GRACKLE
Original label: BREWERS BLACKBIRD



Predict label: COMMON GRACKLE
Original label: COMMON GRACKLE



THANK
YOU