Natural Images Project

Prepared by:

DIYA' ASHRAF JAMAL AL. QAISSE

Table of Contents

Project Documentation: Natural Images Classification

1. Project Overview

This project focuses on developing a high-performance multiclass image classifier using the Natural Images dataset. The primary objective is to leverage transfer learning techniques to categorize images into eight distinct classes accurately. The project encompasses the entire machine learning workflow, from data acquisition and rigorous preprocessing to model training with state-of-the-art architecture and final evaluation.

2. Dataset Description

The Natural Images dataset serves as the foundation for this study. It is a curated collection designed to benchmark the robustness of deep neural networks against various image degradations.

2.1 Dataset Composition

The dataset contains 6,899 images across 8 classes, aggregated from diverse and reputable sources in the computer vision community.

Class	Image Count	Primary Source
Airplane	727	PASCAL VOC
Car	968	Stanford Cars Dataset
Cat	885	Kaggle Dogs vs. Cats
Dog	702	Kaggle Dogs vs. Cats
Flower	843	ImageNet
Fruit	1,000	Fruits-360 Dataset
Motorbike	788	PASCAL VOC

Class	Image Count	Primary Source
Person	986	PubFig83-LFW Dataset

2.2 Technical Specifications

The dataset occupies approximately 179.7 MB and is distributed under the CC BY-NC-SA 4.0 license. Each class is stored in its respective subdirectory, facilitating easy programmatic access.

---

3. Data Preprocessing

To ensure optimal model performance and stability, a standardized preprocessing pipeline was implemented using the Python Imaging Library (PIL).

3.1 Preprocessing Objectives

The pipeline transforms raw, heterogeneous images into a uniform format suitable for deep learning models. The key configurations include:

- Target Dimensions: 256 x 256 pixels.
- Color Space: RGB (3 channels).
- Padding Strategy: Centered with black padding (RGB: 0, 0, 0).

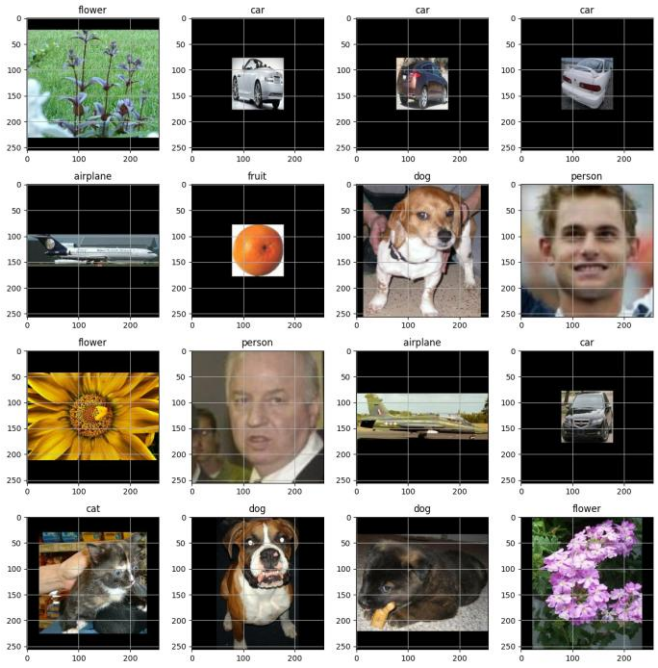
3.2 Transformation Pipeline

The transformation process follows a strict sequence to maintain image integrity:

1. Orientation Correction: EXIF metadata is used to transpose images, ensuring they are correctly oriented.
2. Channel Standardization: Alpha channels are blended onto black backgrounds, and grayscale images are converted to RGB.
3. Aspect Ratio Preservation: Images are scaled using Lanczos resampling so the largest dimension fits the target size without geometric distortion.
4. Canvas Centering: The scaled image is centered on a blank 256x256 canvas, ensuring the model learns object features rather than padding artifacts.

3.3 Preprocessed Image Samples

Below are examples of images after undergoing the resizing, aspect ratio preservation, and canvas centering steps:



---

4. Model Architecture and Training

The project utilizes Transfer Learning with the MobileNetV2 architecture, which provides a balance between computational efficiency and high classification accuracy.

4.1 Architecture Details

The model is built by stacking a custom classifier head onto a pre-trained base:

- Base Model: MobileNetV2 pre-trained on ImageNet, with weights frozen to retain low-level feature extraction capabilities.
- Classifier Head: Consists of a Global Average Pooling layer, a 50% Dropout layer for regularization, a 128-unit Dense layer with ReLU activation, and an 8-unit Softmax output layer.

4.2 Model Architecture Diagram

Layer (type)	Output Shape	Param #
rescale (Rescaling)	(None, 256, 256, 3)	0
data_augmentation (Sequential)	(None, 256, 256, 3)	0
mobilenetv2_1.00_224 (Functional)	(None, 8, 8, 1280)	2,257,984
global_average_pooling2d_4 (GlobalAveragePooling2D)	(None, 1280)	0
dropout_5 (Dropout)	(None, 1280)	0
dense_12 (Dense)	(None, 128)	163,968
dense_13 (Dense)	(None, 8)	1,032

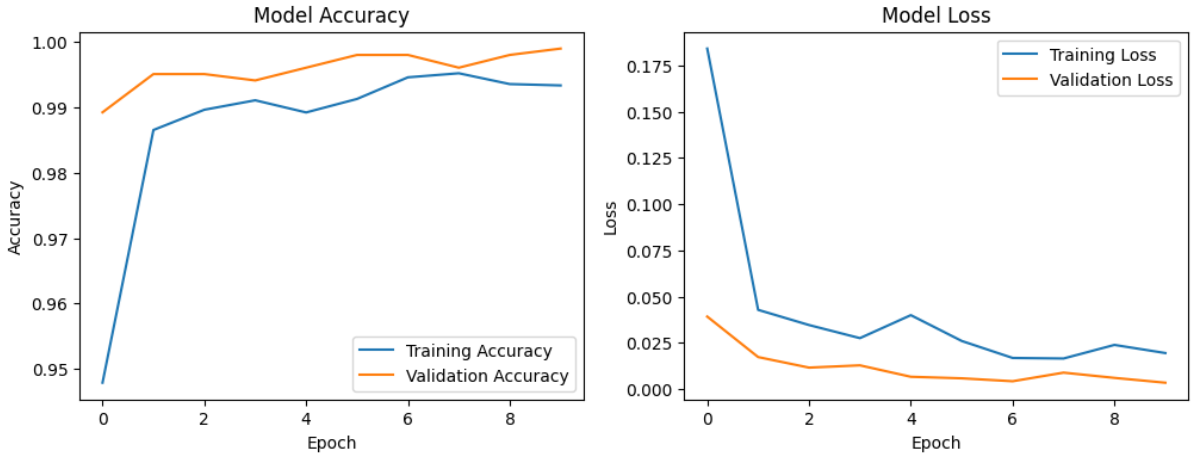
Total params: 2,422,984
Trainable params: 165,000
Non-trainable params: 2,257,984

4.3 Training Strategy

The training process incorporates on-the-fly data augmentation and optimized data pipelines:

- Data Augmentation: Random horizontal flips, rotations (6%), and zooms (6%) are applied to the training set to prevent overfitting.
- Optimization: The model uses the Adam optimizer (Learning Rate: 0.001) and Sparse Categorical Crossentropy loss.
- Data Splitting: The dataset is partitioned into Training (70%), Validation (15%), and Testing (15%) sets.

4.4 Training Results Plots



---

5. Evaluation and Results

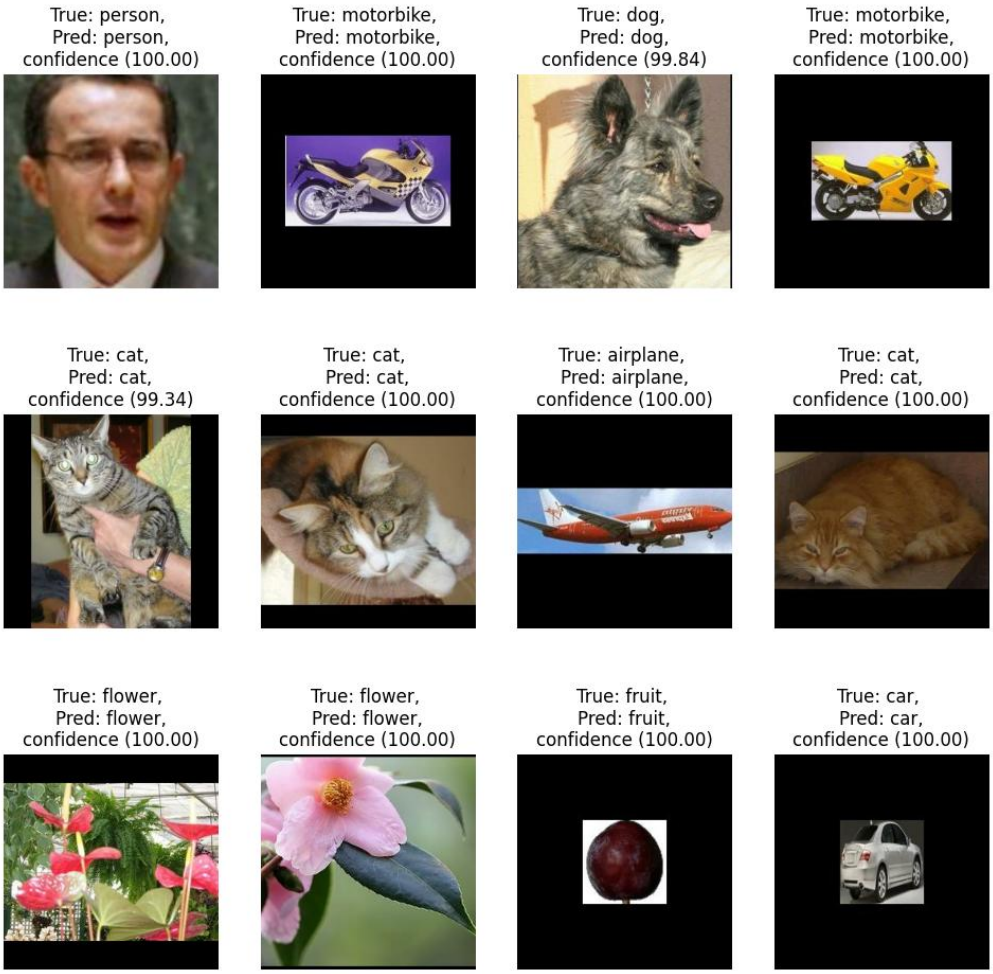
The final model performance was evaluated on the unseen test partition to ensure an unbiased measure of its generalization capability.

5.1 Performance Metrics

The evaluation yielded exceptional results, demonstrating the effectiveness of the transfer learning approach and the quality of the preprocessing pipeline.

Metric	Value
Test Accuracy	99.71%
Test Loss	0.0101

5.2 Testing Visualizations



5.3 Conclusion

The model achieved near-perfect classification on the Natural Images dataset. The combination of MobileNetV2's robust feature extraction and the meticulous aspect-ratio-preserving preprocessing allowed the network to distinguish between the eight classes with high confidence. This documentation confirms the project's success in meeting its technical and academic objectives.