Technical Report: Who's that Pokemon Image Classification

Andrew Sokol

December 4, 2025

Abstract

This paper presents a deep learning approach for classifying Pokemon species using a large scale image dataset containing 1000 classes. I used transfer learning with MobileNetV2 to evaluate the effectiveness of the pretrained convolutional architecture on a multiclass image task. The model is trained ona dataset of Pokemon images sourced from Kaggle, with preprocessing steps including resizing, normalization, batching and data augmentation. My MobileNetV2-based classifier achieves a test accuracy of approximately 85%, demonstrating decent performance given the complexity and visual similarity between many classes.

1 Introduction

Image classification is a core task in computer vision with applications in automated identification systems, content tagging, mobile applications, and augmented reality. Pokemon images present a unique classification challenge due to the large number of visually similar classes, variations in pose and style, and inconsistent dataset quality.

The goal of this project is to build a machine learning model capable of correctly identifying a pokemon species from an input image. To address this problem, we use transfer learning with MobileNetv2, which allows efficient feature extraction while reducing training time and overfitting risk. This report describes my dataset, preprocessing steps, model architecture, training procedure, and results.

2 Methodology

2.1 Data Preprocessing

This dataset, sourced from Kaggle, contains 1000 Pokemon classes, each with color images. Shiny pokemon were also included. All images were stored in a single folder and automatically split into training and validation sets using TensorFlow's image_dataset_from_directory.

Each image underwent the following preprocessing steps:

- Resized to 160x160 pixels
- Normalized pixel values to the [0,1] range
- One-hot encoded labels for multiclass classification
- Bastched into groups of 64 for efficient training.
- Prefetched for GPU pipeline optimization

To improve generalization, I applied data augmentation including random flips, rotations, and zoom transformations.

2.2 Model Architecture

I used MobileNetV2 as the base convolutional neural network due to its:

- Pertaining on ImageNet
- Lightweight, fast design ideal for large class problems
- Strong performance on limitable or variable quality datasets

The full architecture included:

- MobileNetV2 backbone (trainable)
- Global Average Pooling layer
- Dense layer with 128 ReLU units
- Dropout layer (0.3)
- Softmax output layer with 1000 neurons

2.3 Training Setup

Training was performed using:

- Optimizer: Adam
- Loss function: Categorical Crossentropy
- Batch size: 64
- Epochs: 50
- Validation split: 30%

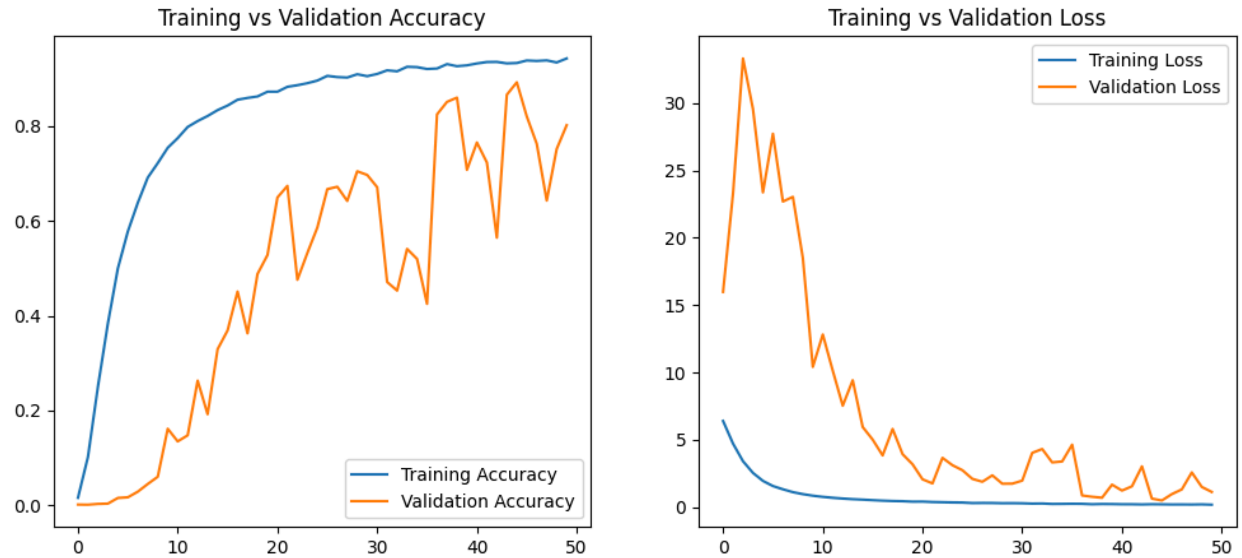
A separate test set was evaluated after training to assess generalization. Predictions were further analyzed using accuracy, F1 scores, and confusion patterns.

3 Experiment results

My MobileNetV2 classifier achieved the following performance:

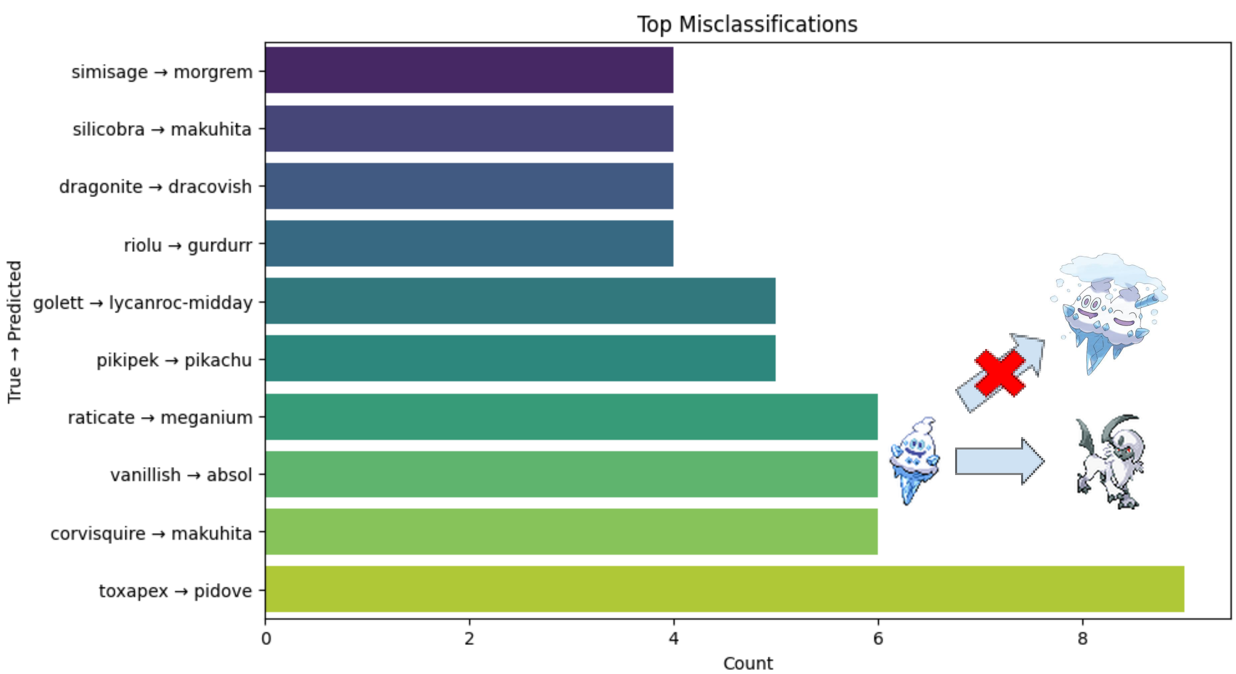
Metric	Value (%)
Test Accuracy	85.11
F1 Macro	82.67
F1 Weighted	84.07
F1 Micro	85.11

3.1 Learning Curves



Training curves show high accuracy and low loss, validation behavior has frequent spikes, indicating overfitting.

3.2 Misclassification Analysis



The most common errors in the model are honestly puzzling. Pokemon who were consistently misclassified bear little to no resemblance to the predicted pokemon. For example, in the third most common misclassification Vanillish is consistently misclassified as Absol despite having an evolution (Vanilluxe) that looks extremely similar.

4 Discussion

The results demonstrate that transfer learning with MobileNetV2 is effective for large-scale Pokemon classification. The model handled 1000 classes with substantial visual overlap, inconsistent dataset quality, and an unbalanced dataset.

Strengths of the approach include:

- Strong performance despite dataset imbalance
- Good generalization due to augmentation and pretrained features
- Efficient training compared to training a CNN from scratch

Limitation include:

- Some classes had very few images, reducing accuracy for these rare pokemon
- Unexplained repeated misclassifications
- Fully unfreezing MobuileNetV2 risked overfitting

Future work would include larger input sizes (MobileNetV2 is optimized for 224x224), and class rebalancing.

5 Conclusion

This project demonstrates that MobileNetV2 based transfer learning can achieve strong performance on a large class Pokemon image classification task. Achieving test accuracy around 95% shows the effectiveness of deep convolutional features even with class imbalance and visually similar categories. Future improvements may involve more data augmentation, larger input sizes, and class rebalancing.

References

[1] *1000 Pokémon Dataset.* 2021. Kaggle. https://www.kaggle.com/datasets/noodulz/pokemon-dataset-1000 Keras+1 Kaggle+1

[2] *MobileNet, MobileNetV2 and MobileNetV3.* 2025. Keras Documentation. https://keras.io/api/applications/mobilenet/