Image recognition Using convolutional neural network

Date 18th, May 2024

Abstract
This project aims to develop an image recognition system using convolutional neural networks (CNNs). The model is trained on the Oxford flower dataset to classify images into 17 or 102 categories. Key steps include data preprocessing, model architecture design, training, and evaluation. The model of the 17 flowers dataset achieved an accuracy of about 54%, while the model of the 102 flowers dataset records less than 1% accuracy on the test set. Additionally, when analyzing specific features, the color model improved accuracy by about 30% on the 17 flowers dataset, while the shape model showed an improvement of less than 1%. On the 102 flowers dataset, both models performed poorly, achieving less than 1% accuracy. This highlights the challenge of accuracy with large datasets and suggests a more complex model for future improvements including datasets, data processes, and deeper architectures.

1. Introduction

1.1 Background
Image recognition is a major task in machine learning since it can be applied in numerous ways. The advancements in deep learning, particularly convolutional neural networks (CNNs), have significantly improved the accuracy of image recognition systems.

1.2 Problem Statement
This project aims to develop a CNN-based model to classify images of the flower dataset from the University of Oxford. The goal is to understand the capabilities and limitations of my model given current resources and techniques.

1.3 Objectives
- To design and train a CNN model for image classification.
- To design possible alternative models
- To evaluate the model's performance using appropriate metrics.

1.4 Scope
The project focuses on image classification and does not focus on the collected data itself.

2. Methodology

2.1 Data Collection
Borrowed from Visual Geometry Group, a dataset containing 102 categories and each from 40 and 258 images was used. The dataset with 17 types of flowers with 80 images for each from the group is also used.

2.2 Data Preprocessing
The images were normalized to have pixel values between 0 and 1. Since the number of images for each group is different, the weight of the category was set. For the different testing models, color and shape information is extracted.

2.3 Model Selection
A CNN architecture is provided in the model summary file. It includes the number of layers with passed values.

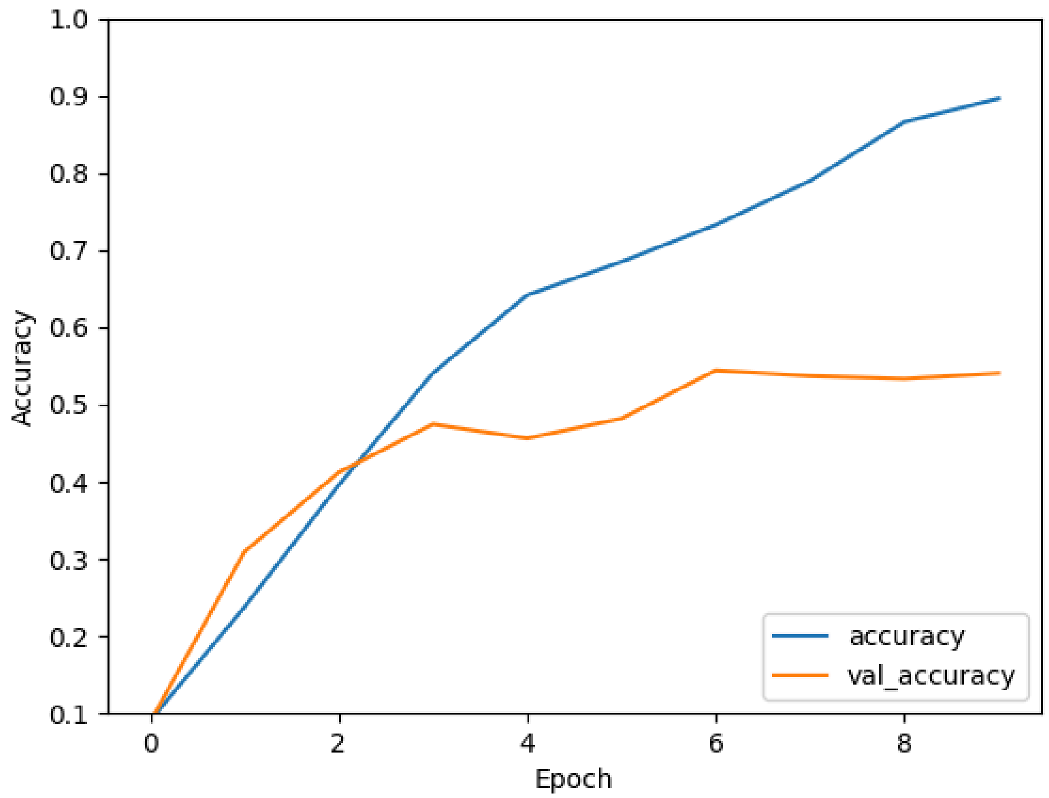
2.4 Training Process
The model was trained using the Adam optimizer with a default learning rate and the training process involved 5 or 10 epochs.

2.5 Evaluation Metrics
Accuracy at each step and final accuracy were used to evaluate the model's performance.

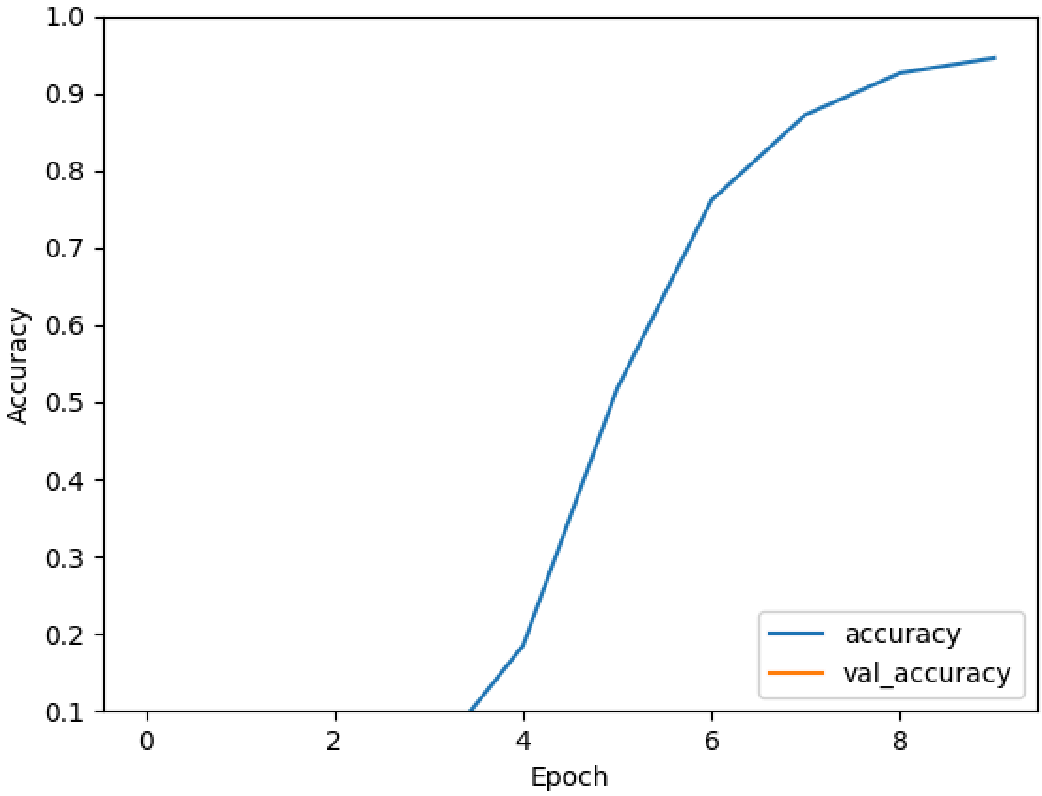
3. Results

3.1 Performance Metrics
The best CNN model for the 17 flowers dataset achieved an accuracy of about 54% on the test set. The color model for the data was 34%, while the shape model was less than 1%. The ensemble model was also about the same as the CNN model. However, on the 102 flowers dataset, no model made it to 1% accuracy.

3.2 Visualizations
The following graphs show the training and validation accuracy and loss over epochs.
The graph below is the CNN model on 17 flowers



The next graph is the CNN model on 102 flowers



3.3 Comparative Analysis

Compared to the 102 flowers model, which achieved less than 1% accuracy, the CNN model for 17 flowers shows that complexity is essential for a large dataset.

4. Discussion

4.1 Interpretation

The results indicate that while the CNN model can learn from the 17 flowers dataset, achieving moderate accuracy, it fails to generalize to the more complex 102 flowers dataset. This suggests that the current model architecture and training strategy are insufficient for handling high variability and complexity. Similarly, when the dataset is small and has obvious color differences, the color model works. However, in complex dataset, the accuracy of color model goes down. The generalized shape model does not do any good job in any case.

4.2 Challenges and Limitations

Key challenges include the increased complexity and diversity of the 102 flowers dataset, which likely requires more sophisticated models and longer training times. The limitations are computational ability and availability of dataset.

4.3 Future Work

Future improvements could involve experimenting with deeper machine learning models and collecting more variety of data. Additionally, leveraging transfer learning from pre-trained models on larger datasets might improve performance.

5. Conclusion

This work suggests the need for more advanced techniques and models when dealing with highly complex and diverse datasets.