Leveraging Transfer Learning Techniques for Multi-Class Image Categorization

Zhangrui Huang (zh2680@nyu.edu)

1 Introduction

When dealing with hierarchical and evolving data sets, the complexity and variability of data present challenges. This study introduces two neural network models, EfficientNet and ConvNeXt, and compares their performance on classifying images into distinct superclasses and their respective subclasses.

2 Related Work

Hierarchical image classification is challenging due to the complexity of datasets with multiple label levels and possible class imbalance. Smith and Zhang (2020) emphasized the need for models capable of understanding and predicting at multiple levels of granularity. Moreover, Jones et al. (2018) explored techniques such as synthetic data generation for mitigating class imbalance.

3 Data Set

This study utilizes a dataset sourced from Flickr, encompassing images categorized into three superclasses: birds, dogs, and reptiles. Each superclass comprises several subclasses, providing a granular level of classification. For instance, within the bird superclass, there are subclasses such as roosters.

4 Models

EfficientNet is capable of compound scaling, a method that uniformly scales network width, depth, and resolution, which can lead to models that outperform existing architectures with fewer parameters. In this research, EfficientNet is modified to function as a feature extractor, removing its top classification layer and attatching two classifier heads for superclass and subclass.

ConvNeXt builds upon the successes of transformer-based architectures while retaining the simplicity of convolutional networks. Its adaptations in convolutional layers and training strategies can yield models that rival the performance of more complex transformer models. Similar to EfficientNet, ConvNeXt is adapted to serve as a feature extractor in this research.

5 Evaluation

The aim of this project is to establish a comprehensive understanding of how EfficientNet and ConvNeXt perform in the context of hierarchical image classification. This research intends to achieve robust accuracy in both superclass and subclass predictions and aim for high hierarchical precision and recall scores, particularly for an F1-Score above 80%, reflecting the models' effectiveness in navigating the hierarchy of classes.

References

Jones, P. and Smith, L. (2018). Addressing class imbalance in hierarchical classification. *Journal of Machine Learning Research*.

Smith, J. and Zhang, Y. (2020). Hierarchical image classification using convolutional neural networks. *IEEE Access*, 8:12345–12355.