Transfer Learning for Image Classification using Pretrained Model

1 Group members

- **Ville Sebastian Olsson:** My goal is to become skilled at transfer learning and fine-tuning techniques in PyTorch, and explore and implement semi-supervised learning and model adaptation.
- **Silpa Soni Nallacheruvu:** Understand Vision Transformers, fine-tuning, and model compression techniques. Apply semi-supervised learning methods. Enhance skills in designing and evaluating deep learning experiments.
- **Joakim Axnér:** Gain hands-on experience in transfer learning and fine-tuning techniques using PyTorch. Learn to implement and evaluate different fine-tuning strategies, including layer-wise adaptation and learning rate scheduling. Hopefully do some semi-supervised learning.
- **Raahitya Botta:** Understand best practices for transfer learning and fine tuning, Learn to adapt the pretrained models using PyTorch, and gain more experience in the regularization techniques and hyperparameter tuning.

2 Project Type and Desired Grade

We will complete Default Project 1 + extension. We are aiming for an A.

3 Problem description

The problem consists of two stages: first, a binary classification task (cat vs. dog), and second, a more challenging multi-class classification problem involving 37 different cat and dog breeds. We will begin by fine-tuning the final layer of a pre-trained ResNet model and then progressively investigate more sophisticated fine-tuning strategies.

To deepen our understanding and improve performance, we will explore extensions that address practical challenges in transfer learning — such as tuning efficiency, label scarcity, and domain shift. These include semi-supervised learning techniques to leverage unlabeled data and parameter-efficient fine-tuning methods such as LoRA. Our goal is not just to achieve strong predictive performance, but to analyze how different fine-tuning strategies trade off between accuracy, computational cost, and generalization.

Our exploration of model compression is inspired by the paper by Song Han et al. (2016), which outlines effective methods for reducing model size and computational cost. We aim to implement and evaluate these methods on fine-tuned Vision Transformer models.

4 Data

Oxford-IIIT Pet dataset. It will randomly be partitioned into training/validation/test sets of appropriate size.

5 Deep Learning Packages

`torch`, `torchvision` (for ResNet models). We will implement the training and evaluation code ourselves.

6 Experiments

The binary classifier and multi-class classifier we build in the first two tasks will serve as baselines, with all model layers untouched except the final one. The initial experiments will be fine-tuning the last l layers simultaneously (Strategy 1), gradual unfreezing (Strategy 2) and the impact of data augmentation.

7 Milestones

E grade: Basic binary and multi-class classification pipeline with $\geq 99\%$ and $\geq 90\%$ accuracy, respectively. The task about imbalanced classes is also addressed. **D–C range:** In addition, some of the $E \to B/A$ avenues are implemented. The findings are presented, but the results are not satisfactory. **B–A range:** 3-5 of the $E \to B/A$ avenues are implemented, with some/many satisfactory results.