This document includes information about course project for the undergraduate course Deep Learning, from School of Computer Science and Engineering, Sun Yat-sen University, 2025.

First Project: Fine-grained classification

The challenge in fine-grained classification is that images from different classes often look similar and images from a single class could look different. Try your best to achieve the performance at least with accuracy 90% on the following dataset. You may learn more recent advances in fine-grained classification from https://paperswithcode.com/task/fine-grained-image-classification .

CUB200 Bird Dataset: http://www.vision.caltech.edu/datasets/cub_200_2011/

Project requirements:

In the first course project, you need to develop a deep learning model to solve one visual classification task. In particular, you are required to finish the following sub-tasks.

- (10%) Train a ResNet50 model to overfit the training dataset, and show the model performance curve with respect to training epochs on the training dataset.
- (50%) Explore effects of 5 or more training tips and tricks on model performance (see first 4 class slides), e.g., data augmentations, batch normalization, ensemble model, fine-tuning pre-trained model, etc. You need to compare the model performance when using each tip/trick with that when not using the tip/trick. Model performance is measured mainly by classification accuracy. Confusion matrix may also be used for some experiments.
- (10%) Interpretability of the model: to visualize and analyze correct and incorrect predictions
- (10%) Use auxiliary loss: besides cross-entropy to train the classifier, explore whether the supervised contrastive loss as an auxiliary loss term can help improve the classifier performance.
- (20%) Report writing: organize and write clearly, including all report sections (Title, Abstract, Introduction, Method, Experiments, Conclusion, Reference list).

Beside the must-do items listed above, you can get bonus by trying to finishing one or more items below (Note: all items are optional and only for those students who would like to challenge themselves):

- Compare the performance of ResNet50 with other model backbone(s)
- Use generative models (e.g., GAN, diffusion model) to generate synthetic images and explore whether the synthetic images can help improve classifier performance.
- Vision-language model (VLM) to help improve performance

Important date: **Week 11 (2025.05.11)**, submit project report together with source code.