Mechanistic Interpretability for Vision Models Optimization

23th July 2025

Computer Vision's course project

Authors:

Camilla Giuliani [1883207] & **Pietro D'Annibale** [1917211]

The project goal

Starting from a baseline ViT observe the model behavior and try to find a good compromise while trying to get a better inference time at the cost of a worse model accuracy.

Overview

- ▲ Goald and ideas
- ▲ Dataset
- ▲ Model Architecture?
- ▲ Training
- ▲ ACDC
- ▲ Final results and evaluation

Dataset

Tiny ImageNet
- 200 classes
- 64x64 images

Augmentations applied:
- Random horizontal flip
- Random resized crop
- Random rotation
- Gaussian noise
- Random erasing
- Normalization with ImageNet mean & std