CMP722 –Assignment

Title:

Self-Supervised Learning or Fine-Tuning for Visual Understanding

Objective:

Train and evaluate a computer vision model using fine-tuning and self-supervised learning on a public dataset. The goal is to explore how representation learning or transfer learning can improve model performance compared to training from scratch.

Instructions:

1. Choose a public dataset such as CIFAR-10, STL-10, ImageNet subset, COCO, or a domain-specific dataset (e.g., medical, satellite, or industrial vision).
2. Select a model architecture (e.g., ResNet, ViT, CLIP, or SimCLR).
3. Implement both of the followings:
    - Self-supervised pretraining (e.g., contrastive, masked image modeling).
    - Fine-tuning a pretrained model on the chosen dataset.
4. Compare results with a baseline model trained from scratch.
5. Evaluate using metrics like accuracy, F1-score, or mAP.
6. Submit a report (max 5 pages) including:
    - Introduction and motivation
    - Dataset description
    - Model and training details
    - Experimental results and analysis
    - Conclusion and future work

Deliverables:

- Source code (GitHub link or notebook)
- Final report (PDF)

Deadline:

15-12-2025