P-2 Computer Vision (CV)

- Techniques

  - Convolutional Neural Networks (CNN) → Backbone of CV tasks
  - transfer learning → using pre trained models to save time/resources
  - Data Augmentation → Rotation, Flips, color changes to improve training
  - Attention Mechanisim → letting model focus on important regions
  - Vision transformers (ViTs) → Apply transformer architecture to vision tasks

- Evaluation Metrics

  - Accuracy (classification)
  - precision, Recall, F1 score (detection, segmentation)
  - mAP (mean Avg precision)
  - IoU (intersection over union)
  - PSNR/SSIM (img quality)

- Datasets: ImageNet, MNIST - Libraries: TensorFlow/Keras, Pytorch

- NOTE: things Like CNN fall under deep learning. and things Like GANs/diffusion models are generative models but CLIP is a Multimodal Model.

- NOTE: All of CV is not img gen and obj recognition (ML) alot of it is AI Like img filters, img compression and is more of classic CV which includes things Like edge detection, Feature extraction and is more pre deep learning Era. img gen, Face reg, 89 obj reg came mostly after deep learning