SmolVLM for Dense Video Captioning

Week 3: Internship Update

- Added PySceneDetect to SmolVLM 2B (video_processing.py)
    - Differen thresholds and min_scene_len
    - HistogramDetect
    - Should add tracking of the number of frames
    - Tested on my video + some videos from VideoFeedback

- Fine-tuned SmolVLM 2B
  - The notebook is on 500M
  - Optional LoRA & QLoRA
    - went over LoRa & QLoRA theory and papers
  - Fine-tuned on VideoFeedback
  - Used GoogleColab
  - Tried on HuggingFaceFV/finevideo: too big dataset and streaming was not supported
  - VideoFeedback is good for captioning only!