Computer Vision Internship – Take-Home Assignment

Task: Fine-tune Gemma for Littering Behavior Detection (Object Detection)

Overview

As part of our internship selection process, this take-home assignment is designed to evaluate your ability to work on applied computer vision tasks involving large language models (LLMs). Your goal is to fine-tune **Google's Gemma** model to identify **people littering** in images by drawing **bounding boxes** around them.

Objective

To leverage the power of open-source LLMs and vision models in a unified pipeline to solve a **visual object detection task** focused on identifying **people** in the act of **littering**.

Dataset Details

- **Source**: Roboflow Universe
- **Dataset**: *Littering*
- **Link**: https://universe.roboflow.com/thesis-kztn8/littering-whlsk/browse

You may augment or filter the dataset to emphasize human figures engaged in the act of littering.

Assignment Tasks

1. **Problem Definition:** Formulate and analyze the problem at hand.
2. **Data Preparation:**
   - Inspect the dataset and verify the presence of person + littering context.
   - If needed, use pseudo-labeling, manual filtering, or rule-based heuristics to create bounding boxes around individuals actively littering.
   - Convert the dataset into a suitable format (YOLO, COCO, etc.) for training.
3. **Modeling:** Use Gemma for finetuning on colab notebook, Your model must output bounding boxes around people who are littering.
4. **Training and Fine-tuning:** Clearly state the approach used for training/fine-tuning Gemma.
5. **Evaluation**
   - Use mAP (mean Average Precision), precision/recall, or IoU metrics to evaluate detection performance.
   - Include confusion matrices or error analysis on "littering vs not littering".

6. **Optional Bonus**

- ○ Insights into the model's performance, its strengths, weaknesses and steps to improve.
- ○ Steps taken to improve the performance and its impact on the model.

Deliverables

- **Technical Report** (PDF or Markdown):
  - ○ Problem understanding and approach
  - ○ Training methodology
  - ○ Results and analysis

- **Code (Colab notebook)**

Submission Guidelines

- **Deadline**: **May 29, 2025 : 11:59 AM**
- Submit via shared Google Drive link with access to Colab Notebook.
- Ensure all code, instructions, and assets are self-contained and reproducible.

If you have questions or need clarifications, feel free to reach out to us.