CITS4402 Computer Vision

Project: Histograms of Oriented Gradients for Human Detection

Due Date: Friday, 23rd May 2025, 11:59 PM (NO EXTENSIONS)

Grouping and Submission

- Form groups of up to **three** members.
- Each group must submit **one** project report via LMS by the deadline.
- All group members will receive the **same** mark, assuming equal contribution.

Project Overview

Human detection is a fundamental task in computer vision with applications in surveillance, autonomous driving, and pedestrian tracking. In this project, you will implement a **human detection system** using the **Histograms of Oriented Gradients (HOG) feature descriptor** combined with a **Support Vector Machine (SVM) classifier**. The method was introduced in the **CVPR 2005 paper: "Histograms of Oriented Gradients for Human Detection"** by Dalal and Triggs [1].

The project is divided into **four phases**, guiding you through data collection, feature extraction, model training, evaluation, and an ablation study to explore the effects of different HOG parameters.

Phase One: Dataset Building

The first phase involves acquiring and preparing a dataset suitable for training and testing the HOG-based human detector. A collection of public datasets can be found at:

GitHub - ViswanathaReddyGajjala/Datasets (Note: Some datasets may no longer be available.)

Dataset Requirements:

- **Training Set:** At least **500** images consisting of both **human (positive) and non-human (negative) samples**.
- **Testing Set:** Exactly **100 positive** (human) and **100 negative** (non-human) images.

Phase Two: Feature Extraction and SVM Training

In this phase, extract **HOG features** and train an **SVM classifier** to distinguish between human and non-human images. You can follow [1] for HOG feature extraction properties (as listed below). After conducting the ablation studies, you might find that changing some properties results in higher performance. If you find that modifying certain properties improves performance, update the settings in this section accordingly. The final reported properties in Phase Two should reflect those that yield the best performance, as they will be evaluated on a hidden dataset.

HOG Feature Extraction Properties from [1]:

- **Preprocessing:** No gamma correction, use raw **grayscale** images.
- **Gradient Computation:** Use the **[-1, 0, 1] filter** (without smoothing).
- **Orientation Binning:** Use **9 bins**, covering angles **0°–180°**.

- **Cell Size: 8×8** pixels.
- **Block Size: 16×16** pixels (comprising **four** 8×8 cells).
- **Block Normalization: L2-Hys normalization** (Lowe-style clipped L2 norm).
- **Block Stride: 8-pixel spacing** (4-fold coverage of each cell).
- **Detection Window Size: 64×128 pixels**.
- **Classifier: Linear SVM** for binary classification.

Implementation Steps:

1. **Extract HOG Features** from each image in the training dataset.
2. **Train a Linear SVM** using extracted HOG features with positive (human) and negative (non-human) labels.
3. **Evaluate Model Performance** on the testing set using appropriate metrics.
    - You may use **accuracy, precision, and recall**, or adopt the **miss rate vs. false positive per window (FPPW)** metric used in [1].

Phase Three: Ablation Study

An ablation study helps analyze how changes in HOG properties affect detection performance. Each group must conduct at least **one ablation study per group member** (e.g., a group of three must conduct three studies). The focus of this project is on feature engineering rather than tuning the classifier. Hence, the ablation studies should be conducted on HOG feature extraction rather than the SVM classifier.

Possible Ablation Studies:

- Changing the **number of orientation bins** (e.g., 6, 9, 12).
- Altering the **block size** (e.g., 16×16 vs. 32×32).
- Using **different gradient filters** (e.g., Sobel vs. [-1, 0, 1]).
- Modifying **normalization techniques** (e.g., L1 vs. L2-Hys).

Implementation Steps:

1. Modify **one** HOG property at a time.
2. Retrain the SVM classifier.
3. Compare performance with the original setup.
4. Visualize the results (as shown in **Figure 4 of [1]**).

💡 *Tip:* Clearly present ablation study findings in **graphs and tables** for better interpretation.

Phase Four: Reporting & GUI Implementation (10%)

Report Guidelines:

Prepare a structured report covering **phases 1–3**, with a maximum of **500 words per phase**. Each section should provide enough detail to allow for **reproducibility**.

Report Sections:

1. **Dataset Collection:** Explain dataset selection, preprocessing, and storage format.
2. **Feature Extraction & Model Training:** Describe HOG settings, SVM training, and evaluation metrics.

3. **Ablation Study:** Present experimental results and their impact on detection performance.

Graphical User Interface (GUI) Implementation:

Groups must develop a **GUI** to visualize the classifier's predictions.

GUI Requirements:

- Load a directory containing **10 positive and 10 negative** images.
- Display each image **one by one**.
- Show the **prediction (human or non-human)** below each image.
- Save the predictions in a file named **predictions.xlsx**, placed in the **primary directory** of your submission. The spreadsheet should look like this:

	A	B
1	filename	prediction
2	001.png	1
3	002.png	0
4	003.png	1
5	...	...
6	100.png	1
7		

Specifically, column A should take the filenames, while column B should save the binary predictions.

Mark Distribution

- Qualitative assessment of reproducibility, rigor, and thoroughness of the project (80%)
  Marks are assigned based on the quality of implementation and the clarity of the report.
  - Phase One: 24%
  - Phase Two: 32%
  - Phase Three: 24%
- Quantitative assessment of the accuracy of the trained classifier (20%)
  The trained classifier will be evaluated on a 'hidden' dataset. Marks given will be proportional to the accuracy achieved on this 'hidden' dataset.

Final Notes

- Ensure your code is **well-documented** with clear comments.
- Use **meaningful variable names** and follow **best coding practices**.
- At the very beginning of your report, note down the full name and student number of each group member.
- All files should be zipped into a .zip file for submission. After unzipping, the primary directory should include:
  - **Report.pdf** – Your project report.
  - A script file named **GUI** (with the appropriate file extension). This file should contain the code for launching the GUI.
  - A folder '**Testing Images**'. This folder should contain the 10 positive and 10 negative images.
  - A README.txt file (optional).

    - Any additional supporting files (e.g., code) should be saved in another folder named '**Others**'.
- Any questions should be addressed during weekly lab sessions.

Reference

[1] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," CVPR, 2005.