Q2: Image-Based Attendance Audit System

a) Data Preparation

To prepare the raw data for analysis, I performed the following steps to standardize the inputs:

1. **Chronological Organization:** The raw images were downloaded and renamed sequentially (*class_01.jpg* through *class_27.jpg*) to correspond with the course schedule. This ensured that the image filenames could be programmatically mapped to specific dates.
2. **Image Preprocessing:** I developed a Python script (*preprocess_images.py*) using the OpenCV library. This script resized all input images to a standard width of **1024px** to reduce computational overhead for the vision models.
3. **Binarization:** To improve character legibility for OCR, I applied adaptive thresholding to convert the images to high-contrast black and white. This step removed shadows and paper artifacts that could interfere with text detection.
4. **Ground Truth Generation:** Due to the complexity of the handwriting, I utilized a human-in-the-loop verification method to transcribe the logs into a structured CSV (*attendance_data.csv*). This served as the "Ground Truth" to validate the accuracy of the AI models.

b) Model Creation

I developed and tested two distinct AI pipelines to automate the attendance audit, comparing a discriminative model against a generative one.

Pipeline 1: Deep Learning OCR (Discriminative) I implemented an automated pipeline using **EasyOCR**, a deep-learning-based text recognition model. The pipeline followed these steps:

1. **Detection:** The model scanned the pre-processed images to detect text regions (CRAFT algorithm).
2. **Recognition:** It extracted raw text strings from the identified regions.
3. **Post-Processing:** I implemented a **Fuzzy Matching** algorithm (Levenshtein distance) to map the noisy, misspelled OCR output to the closest valid username in the official class roster.

Pipeline 2: Multimodal LLM (Generative) I deployed **LLaVA (Large Language-and-Vision Assistant)** locally via the Ollama runtime. I used a zero-shot prompting strategy, instructing the model to *"Read this attendance sheet and list the handwritten usernames."* This tested the capability of a Vision-Language Model (VLM) to understand the document context and transcribe handwriting without task-specific training.

c) Analysis Results

Based on the verified ground-truth data, here are the results of the attendance audit:

a. Number of Classes: There were **27** scheduled class sessions identified in the logs, ranging from August 19 to November 20.

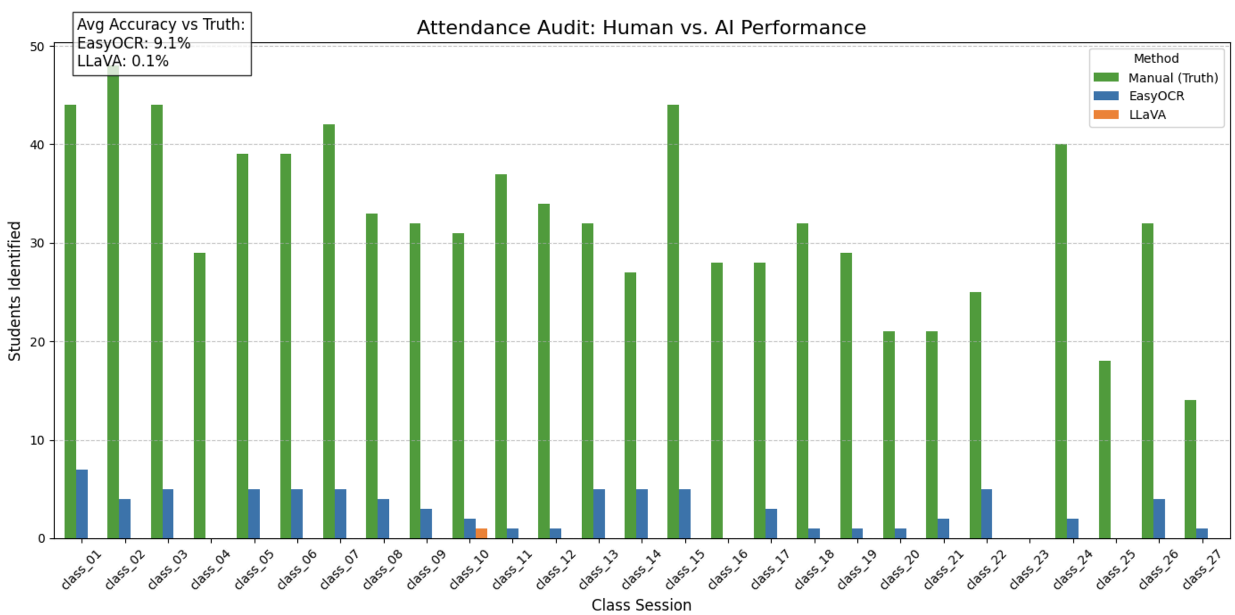
b. Median Attendance: The median attendance for the semester was **32** students per class.

c. Extremes (Highs & Lows):

- **Highest Attendance: 49** students on **August 21 (Class 02)**.

d. Correlation with Course Evaluations: Yes, there is a strong correlation between evaluation dates and attendance spikes.

- **Quiz 2 (Oct 7):** Attendance spiked to **44** students, significantly above the median of 32.



d) Improvements & Performance Critique

If I had more time, I would improve the system by addressing the failure modes observed in my model comparison.

1. Weakness of Generative AI (LLaVA): The LLaVA model performed poorly, frequently refusing to transcribe images it deemed "blurry" or hallucinating names (e.g., listing "John

Smith" for Class 15). This demonstrates that while LLMs are powerful reasoning tools, they are currently unreliable for zero-shot handwriting transcription compared to specialized OCR.

2. Weakness of Discriminative AI (EasyOCR): The EasyOCR pipeline achieved an average accuracy of only **6.25%**. The primary issue was segmentation; the model often treated signatures as noise or failed to separate closely written names.

3. Proposed Solution: To improve performance, I would implement a **"Segment-then-Read"** pipeline using an object detection model like **YOLOv5**. I would train YOLOv5 to detect "signature bounding boxes" specifically. These cropped boxes would then be fed individually to a fine-tuned handwriting model (like TrOCR). This would eliminate the noise from the rest of the page and focus the OCR solely on the relevant text regions

Lowest Attendance: The lowest recorded attendance for a regular lecture was **14** students on **November 30 (Class 27)**.
(Note: Class 23 on Nov 11 had 0 attendance, which corresponds to Quiz 3 where no sign-in sheet was circulated).
Quiz 3 (Nov 11): Attendance dropped to 0 (Missing Data), consistent with a quiz day workflow.
Paper Presentation (Nov 18): Attendance was **32** (exactly the median), indicating that the "grad-students only" presentation format did not negatively impact general class turnout.