Face-Based Entry / Exit Attendance System Workflow Documentation

System Overview

This system implements a face-only entry/exit attendance pipeline using SCRFD for face detection, DeepSORT for face tracking, and ArcFace for identity recognition. The system does not use any body detector such as YOLO. Instead, it heuristically estimates a full-body region from detected face bounding boxes. Entry and exit events are determined using a line-based finite state machine (FSM) and logged into an SQLite database.

Initialization Phase

- Check GPU availability using ONNX Runtime and log whether CUDA acceleration is enabled.
- Load SCRFD face detection model for face bounding boxes and landmarks.
- Load ArcFace model for generating 512-dimensional face embeddings.
- Load known face embeddings from the known_faces directory (.npy files).
- Initialize the DeepSORT tracker configured for face-only tracking using embeddings.
- Initialize Multi-Embedding Fusion (MEF) buffers for stabilizing recognition.
- Initialize SQLite database and attendance table if not already present.
- Interactively draw a single red line used to separate OUTSIDE and INSIDE regions.

Runtime Frame Processing Workflow

- Read a frame from the video source (camera or video file).
- Optionally resize the frame for performance optimization.
- Skip face detection on some frames based on PROCESS_EVERY_N_FRAMES to improve throughput.
- If detection is enabled, run SCRFD to detect faces and extract landmarks.
- Generate ArcFace embeddings for each detected face.
- Pass face bounding boxes and embeddings to DeepSORT to update or create tracks.
- Maintain per-track metadata including detection count, last seen frame, and bounding boxes.
- Estimate a heuristic full-body bounding box from each face bounding box.
- Accumulate embeddings per track and compute a fused embedding using MEF.
- Perform face recognition by comparing fused embeddings with known face embeddings.
- Assign a permanent stable ID once a face is successfully recognized.

Entry / Exit Decision Logic (FSM)

- Only recognized tracks with sufficient detection confidence are eligible for entry/exit evaluation.
- Compute a motion reference point using the bottom-center of the estimated body bounding box.
- Determine which side of the red line the reference point lies on using a cross-product test.
- Initialize a per-person FSM state (INSIDE or OUTSIDE) based on the first observed position.
- Detect a crossing when the sign of the side-of-line value changes.
- Trigger ENTRY when the state transitions from OUTSIDE to INSIDE.
- Trigger EXIT when the state transitions from INSIDE to OUTSIDE.
- Apply cooldown frames to prevent duplicate or oscillating events.
- Log valid entry or exit events into the SQLite database.

Visualization and Cleanup

- Draw face bounding boxes or estimated body bounding boxes on the output frame.
- Overlay person name, stable ID, and recognition confidence.
- Draw the entry/exit red line and system statistics (entries, exits, active faces).
- Remove stale tracks based on time since last detection.
- Preserve stable IDs permanently to allow re-identification when a person reappears.

Key Characteristics and Limitations

This system prioritizes identity persistence and recognition stability using face embeddings. However, because it relies on face-derived motion rather than full-body detection, line-based entry/exit detection can be sensitive to camera angle, line orientation, and head movement. For higher reliability, zone-based logic or body detectors are recommended.