Project Title:

Deep Vision Crowd Monitor: AI for Density Estimation and Overcrowding Detection

Sreeja Banoth

1. Introduction

Milestone 3 emphasizes deploying the trained CSRNet model within a live crowd analysis framework.
Instead of working only with static images, this stage enables continuous analysis of video streams obtained from sources such as webcams or recorded footage.
The system processes each video frame, performs inference using the CSRNet model, and visualizes real-time crowd estimates through an interactive interface built using Gradio.

2. Objective

The objective of this milestone is to build and validate a complete real-time crowd monitoring pipeline powered by the trained CSRNet network.
The system simulates real-world surveillance scenarios where continuous video feeds are analyzed to estimate crowd density, detect congestion, and support safety-oriented decision-making.
This setup demonstrates how AI-based crowd counting can be applied in practical monitoring environments.

3. Real-Time Crowd Monitoring Pipeline

3.1 Video Frame Acquisition

A live video stream is obtained using either a laptop webcam or a pre-recorded video file.
OpenCV is utilized to continuously capture frames using cv2.VideoCapture, enabling seamless real-time processing similar to CCTV camera feeds.

3.2 Frame Preprocessing

Each captured frame undergoes preprocessing steps consistent with those applied during model training to ensure reliable predictions:

- Optional resizing to improve inference speed
- Conversion from BGR to RGB color space
- Normalization using ImageNet mean and standard deviation
- Conversion into a tensor format compatible with PyTorch
- Addition of a batch dimension before inference

Maintaining preprocessing consistency between training and deployment is critical for model accuracy.

3.3 CSRNet Inference

The preprocessed frame is passed through the trained CSRNet model.
The model produces a density map representing the spatial distribution of people within the frame.
The estimated crowd count is obtained by summing all values in the predicted density map.

3.4 Visualization Using Gradio

Gradio is employed to create a user-friendly web-based interface that:

- Displays the live video feed
- Performs real-time inference on each frame
- Shows the estimated number of people
- Optionally presents a visual representation of the density map

This interface allows easy interaction and monitoring without complex setup.

3.5 Density Map Overlay (Optional)

To enhance interpretability, the predicted density map can be resized and normalized to align with the original video frame.
By blending the density map with the frame, regions with higher crowd concentration become visually prominent.
This visualization helps analyze crowd distribution patterns and identify congested zones.

3.6 Automated Crowd Alerts

The system supports alert mechanisms based on predefined crowd thresholds.
When the estimated count exceeds the specified limit, the system can:

- Display warning messages on the interface
- Visually highlight the video frame
- Record alert events for further analysis

Such alerts reflect how real-world surveillance systems assist authorities in managing overcrowding.

4. System Execution Flow

1. Launch the Gradio application.

2. Start capturing video from the webcam or file source.
3. For each incoming frame:
    - Apply preprocessing
    - Generate a density map using CSRNet
    - Compute the total crowd count
    - Display the processed frame and predictions
    - Activate alerts if thresholds are crossed
4. Continue real-time monitoring until the application is stopped.

5. Practical Use Cases

- Monitoring crowded public spaces such as railway stations and shopping centers
- Managing queues and waiting areas
- Supporting emergency evacuation planning
- Real-time analytics for smart city surveillance
- Intelligent video-based security systems

6. Conclusion

Milestone 3 highlights the transition of CSRNet from offline experimentation to real-time deployment.
By integrating live video processing, visualization tools, and alert mechanisms, this milestone demonstrates the feasibility of AI-driven crowd monitoring systems in real-world scenarios.
It plays a crucial role in understanding how deep learning models can be effectively applied for intelligent surveillance and public safety applications.