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

SUBMITTED BY,

Avanthikalakshmi Vinod

TABLE OF CONTENT

OBJECTIVE

The objective of Milestone 3 is to evaluate the trained crowd counting model on a real-world crowd dataset containing images and videos, and to analyze its performance in estimating crowd size using density maps.

This milestone focuses on model inference, result visualization, and count estimation, rather than training.

DATASET USED

For testing and evaluation, a crowd dataset containing dense crowd scenes was used.

- Dataset type: Crowd images and videos
- Source: Publicly available crowd datasets (e.g., Mall dataset / Kaggle crowd datasets)
- Nature of scenes:
- Dense crowd regions
- Varying illumination
- Partial occlusions
- Different crowd densities

The dataset helps in validating the model's ability to generalize to unseen real-world scenarios.

MODEL USED

Model architecture: CSRNet (Convolutional Neural Network for Crowd Counting)

- Trained weights: model50.pth
- Framework: PyTorch

CSRNet is a density-map-based crowd counting model, where the output is a density map instead of explicit person detection.

IMPLEMENTATION DETAILS

•Programming Language: Python

- Libraries used:
- PyTorch
- OpenCV
- NumPy
- PIL
- Gradio (for interface and visualization)

Both image-based and video-based inference were implemented.

METHODOLOGY

Step 1: Input Preprocessing

- Input frames/images are resized and converted to RGB format.
- Images are normalized and converted into tensors before being passed to the model.

Step 2: Density Map Generation

- The trained CSRNet model processes the input image.
- The model outputs a density map, where higher intensity regions indicate higher crowd density.

Step 3: Crowd Count Estimation

- The total crowd count is calculated by integrating (summing) the density map values.
- Since density values are not direct human counts, a scale factor is applied to calibrate the output.

$$\text{Final Count} = \sum (\text{Density Map}) \times \text{Scale Factor}$$

This calibration ensures more realistic crowd estimates.

Step 4: Visualization

- Density maps are converted into heatmaps using color mapping.
- Heatmaps are overlaid on the original image for visual interpretation.
- The final estimated crowd count is displayed on the output image.

RESULTS AND OBSERVATIONS

- The model successfully generated meaningful density maps for dense crowd scenes.
    - Crowd counts increased proportionally with crowd density.
    - Heatmap visualization clearly highlighted high-density regions.
    - Minor deviations in exact count were observed due to:
    - Occlusions
    - Perspective variations
    - Extremely dense regions

However, overall performance was consistent and reliable for crowd estimation.

CONCLUSION

Milestone 3 successfully demonstrated the practical application of the trained crowd counting model on real-world datasets.

The model was able to generate density maps, estimate crowd size, and visualize results effectively for both images and videos.

This milestone validates the robustness and applicability of the system for real-time crowd analysis and monitoring.