1. Introduction

This report outlines the development of a citizen-driven platform designed to streamline the reporting of common urban issues such as potholes, littering, and vandalism. The solution leverages a robust computer vision pipeline to automatically identify, classify, and protect the privacy of images submitted by the public. The system's core capabilities include identifying multiple urban issues within a single image, classifying each issue into a predefined category, and ensuring the anonymity of individuals and properties.

2. Methodology and Approach

Our approach to building this system was structured into three key phases: data preparation, model training, and a two-stage inference pipeline for anonymization and classification.

2.1. Data Collection and Annotation

To train a model capable of accurately identifying diverse urban issues, a comprehensive dataset was essential. The data was sourced from publicly available datasets to cover a range of issues. The following categories were defined for the model:

- broken_road_signs
- damaged_buildings
- littering
- parking_issues
- potholes
- vandalism

Using the open-source annotation tool **LabelImg**, each image in the dataset was meticulously annotated. Bounding boxes were drawn around every instance of a defined issue, and a corresponding class label was assigned. This process generated the necessary .txt files in the YOLO format, which served as the ground truth for our object detection model.

2.2. Image Anonymization

To address the privacy requirements of the challenge, we implemented a dedicated anonymization step before any classification was performed. The methodology for this was inspired by the approach outlined in the public GitHub repository, https://github.com/rakiiibul/LF_Anonymization.git. This involved using a pre-trained object detection model to specifically identify personally identifiable information (PII) such as faces and license plates within a submitted image. Once these objects were detected, a blurring filter was applied to the bounding box regions, effectively obscuring them while preserving the rest of the image content. This ensures that sensitive data is protected before the image is passed to the main classification model.

2.3. Model Training and Architecture

For the classification and multi-issue identification task, we chose **YOLOv8**, a state-of-the-art object detection model known for its speed and accuracy. The model was trained on our custom, annotated dataset. The training process leveraged a powerful GPU environment to accelerate convergence, allowing the model to learn the patterns and features associated with each urban issue category.

The final model is capable of performing multi-label classification and object detection in a single pass. For a given image, it can detect all instances of the urban issues we defined, draw bounding boxes around them, and assign the correct class label to each.

3. Results and Model Evaluation

The model was evaluated using standard object detection metrics. The following table summarizes the key performance indicators from our training results.

Metric	Value	Interpretation
metrics/mAP50(B)	0.45217	The model has a mean Average Precision of 45.2% at an Intersection over Union (IoU) threshold of 0.5. This indicates a moderate ability to correctly identify and localize the urban issues.
metrics/mAP50-95(B)	0.17981	The mAP averaged across various IoU thresholds from 0.5 to 0.95 is 18.0%. The significant gap between mAP50 and mAP50-95 suggests the model is good at roughly locating objects but struggles with precise bounding box localization.
metrics/precision(B)	0.59385	Of all the bounding boxes predicted by the model, 59.4% were correct.
metrics/recall(B)	0.39870	The model successfully identified 39.9% of all the actual urban issues present in the dataset.

Class-Specific Performance (maps):

The model's performance varied across different issue categories:

- **Vandalism:** mAP of 0.3209
- **Potholes:** mAP of 0.19276
- **Broken Road Signs:** mAP of 0.14493
- **Damaged Buildings:** mAP of 0.14246
- **Littering:** mAP of 0.14498
- **Parking Issues:** mAP of 0.13283

The model demonstrated the best performance in detecting **vandalism**, likely due to distinctive visual features. Performance for other categories was lower, indicating areas for improvement.

Inference Speed:

The model's inference speed was measured as follows:

- **Preprocess:** 0.39ms
- **Inference:** 4.75ms
- **Postprocess:** 4.98ms

The inference time of approximately 4.75ms (on a CPU in this run) is very fast, making the model suitable for real-time applications.

4. Conclusion and Future Work

The developed system successfully demonstrates a working end-to-end pipeline for urban issue reporting. It effectively performs image anonymization and classifies multiple issues using a trained YOLOv8 model. The results indicate that while the model is proficient at general object detection (mAP50), there is a need to improve its ability to precisely localize issues (mAP50-95).

To enhance the system's performance, the following steps are recommended for future work:

- **Dataset Expansion:** Collect a larger and more diverse dataset for the lower-performing categories like 'parking_issues' and 'damaged_buildings'.
- **Annotation Refinement:** Re-evaluate and refine the bounding box annotations to ensure greater precision, which would directly improve the mAP50-95 score.
- **Hyperparameter Tuning:** Experiment with different YOLOv8 training hyperparameters (e.g., learning rate, batch size) to optimize performance for the specific dataset.
- **Model Architecture:** Consider exploring other versions of YOLO (e.g., YOLOv9, YOLO-NAS) or other model architectures that may offer better localization accuracy.