Multi-Class Urban Issue Identification System

Team Name: MlArtist

1. Introduction

This project presents a robust, multi-class image classification system designed to identify and categorize various urban issues from user-submitted images. The primary goal is to create a reliable tool that can assist municipal authorities by automating the initial analysis of citizen reports, covering problems such as road damage, vandalism, illegal parking, and public cleanliness. The mandatory anonymization pipeline, a crucial part of this system, obscures faces and license plates before any classification is done, safeguarding both property and individual privacy.

2. Data Sourcing and Analysis

Data Sources

A wide range of data was compiled from nine different sources, mostly from the Roboflow Universe platform, in order to create a thorough model. To improve data diversity, a deliberate choice was made to locate several sources for the minority classes (Parking, Potholes, Trash). The last sources were:

- **Graffiti:** Graffiti-1 (hruts-workspace)
- **Damaged Signs:** Damaged-Signs-Hind-2 (road-inspection)
- **Damaged Construction:** Concrete-3 (road-ywxxe)
- **Illegal Parking:** Illegal-Parking-5 (parking-amu50) & illegal-parking-1 (hao-61fh1)
- **Potholes:** Road-holes-1 (road-holes) & pothole-ek-1 (evansworkspace)
- **Trash:** Domestic-trash-2 (trash-drone) & Domestic-Trash--1 (datacluster-labs-agryi)

Initial Data Analysis & Insights

The initial aggregation of these 8,946 images revealed a significant class imbalance. The "Damaged Sign" and "Parking" classes contained over 2,200 images each, while the "Trash" and "Pothole" classes had fewer than 710. This insight was critical, as training a model on this imbalanced data would lead to a heavily biased system that performs poorly on minority classes.

The initial distribution was as follows:

Label	Image Count
Road_Issues_Damaged_Sign	2409
Parking_Issues_Illegal_Parking	2217
Vandalism_Graffiti	2128
Infrastructure_Damage_Concrete	990
Road_Issues_Pothole	706
Domestic_trash	496

3. Data Preprocessing and Enrichment

Preprocessing Methods

To guarantee uniformity and model compatibility, a standardized preprocessing pipeline was used on every image.

- **Anonymization:** Using OpenCV, a required anonymization phase was put in place to safeguard privacy in accordance with the challenge requirements. Each image's faces and

Russian license plates were identified using two pre-trained Haar Cascade models. These sensitive areas were then obscured by adding a strong Gaussian blur.

- **Image Resizing & Normalization:** All images were resized to a uniform dimension of 128x128 pixels. After that, pixel values were normalized from the [0, 255] range to the [0, 1] range by dividing by 255.
- **Data Augmentation:** To avoid overfitting and enhance model generalization, a series of random augmentations were applied to the training data via Keras' ImageDataGenerator. Rotations, zooms, shifts, and horizontal flips were among them.

Data Enrichment (Balancing)

To address the class unbalanced, we enriched the dataset using an **oversampling** technique. Making use of Scikit-learn's resample tool, Until each category had the same quantity of images as the largest class (2,409 number of images), samples from the minority classes were duplicated at random. This produced a 14,454 number of image dataset that was perfectly balanced, guaranteeing that the model gives each category the same amount of weight during training.

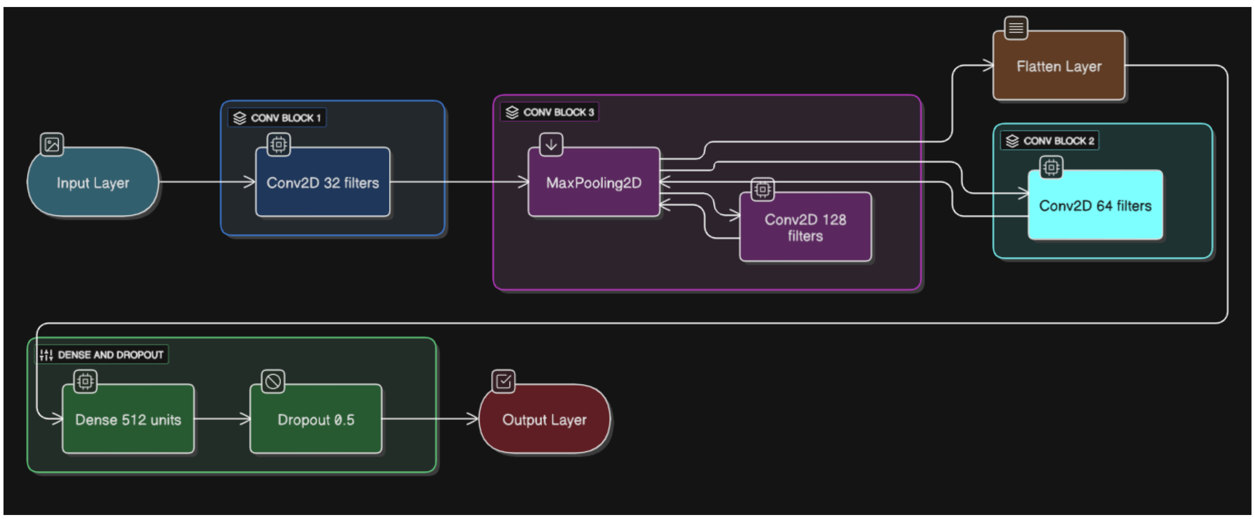
The successful result of this balancing process is shown below:

Label	Image Count
Vandalism_Graffiti	2409
Road_Issues_Damaged_Sign	2409
Infrastructure_Damage_Concrete	2409
Parking_Issues_Illegal_Parking	2409
Road_Issues_Pothole	2409

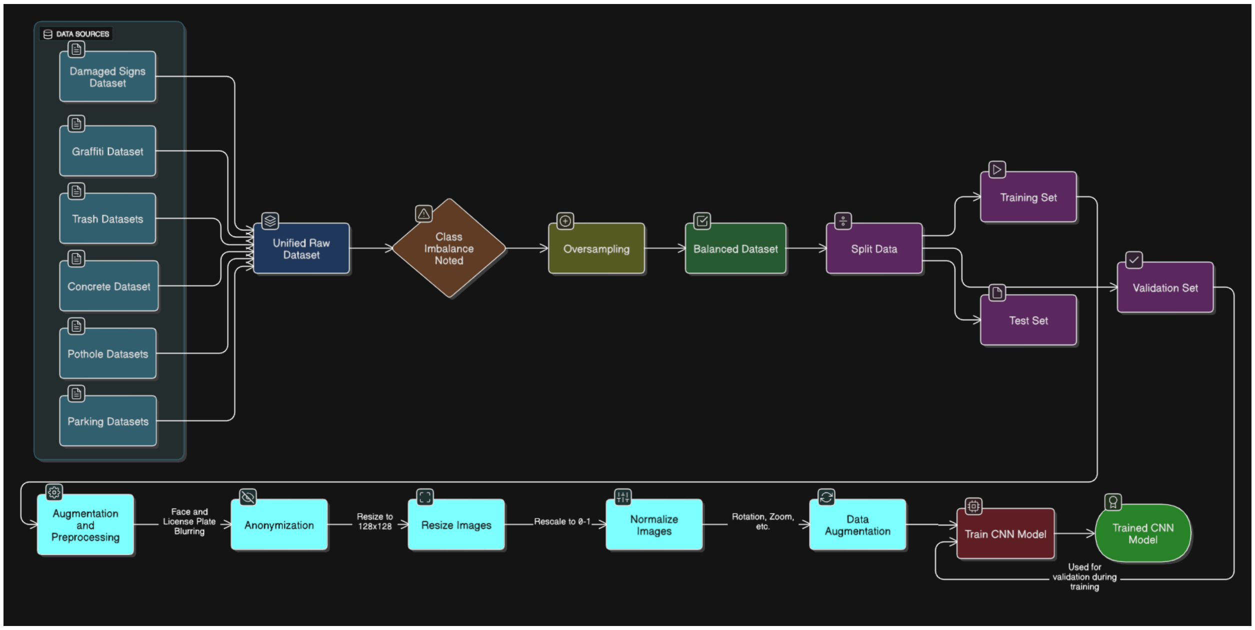
Domestic_trash	2409

4. Model Architecture

The model is a Sequential Convolutional Neural Network (CNN) built with TensorFlow/Keras. The architecture includes of three convolutional blocks, each one contain a Conv2D layer with ReLU activation and a MaxPooling2D layer for down-sampling. To prevent overfitting, a Dropout layer (with a rate of 0.5) was included after the flattening stage. The final output layer uses a softmax activation function to classify an image into one of the 6 distinct categories.



This model is trained on a Unified Raw Dataset compiled from six different sources, including Damaged Signs, Graffiti, and Pothole datasets. Before training, the images undergo a rigorous preprocessing pipeline. This involves anonymization (blurring faces and license plates), resizing all images to a uniform 128x128 pixels, and normalizing pixel values to a 0-1 range. To enhance the model's ability to generalize, data augmentation techniques like rotation and zooming are applied to the training set. The final, preprocessed images are then used to train the CNN, with a separate validation set used to monitor performance and prevent overfitting during the training process.



5. Rationale for Model Evaluation Metric/s

While overall accuracy is a useful metric, it can be misleading for multi-class problems. A model could achieve high accuracy by simply performing well on the majority classes while failing completely on the minority classes.

For this reason, we chose a more thorough set of metrics to evaluate the model fairly across all categories:

- **Validation Loss:** Used as the primary monitor for our EarlyStopping and ReduceLROnPlateau callbacks, as it provides a smooth measure of how well the model is generalizing to unseen data.
- **Validation Accuracy:** Used as a general benchmark and for saving the best version of the model during training via ModelCheckpoint.
- **Classification Report (Precision, Recall, and F1-Score):** This was our primary tool for detailed performance analysis on the final, unseen test set. The **F1-Score** is particularly important, as it provides a harmonic mean of Precision and Recall for each class. By

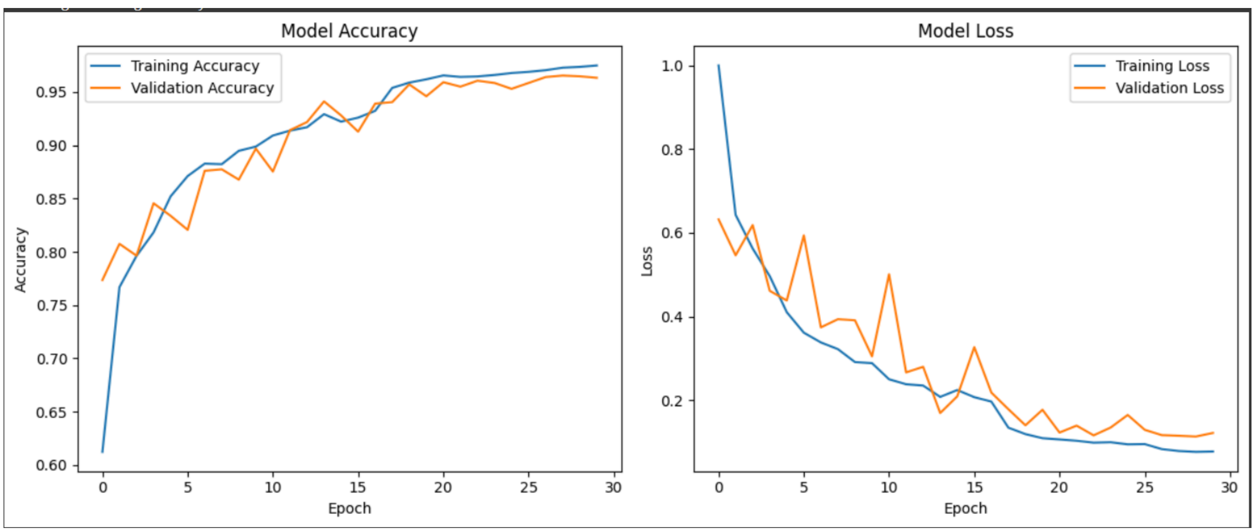
achieving high F1-scores across all categories, we can confirm that the model is both precise (not making false accusations) and has high recall (not missing actual issues), leading to a reliable and trustworthy system.

6. Results and Conclusion

Results

The final model was trained for a total of **21 epochs** before the EarlyStopping callback decided to take over due to lack of validation loss improvement. The model achieved a final **Test Accuracy of 97.72%**, with a final **Test Loss of 0.0730**.

The training history below illustrates the learning curves, showing a healthy convergence where validation accuracy closely tracks training accuracy, indicating a well-generalized model.



The classification report provides more detailed information on the test data and verifies that the model does indeed score 0.95 or above on each individual category. This results concludes that the data balancing strategy used to balance the class distribution was effective and accurate for high precision on both majority and minority classes.

	Precision	Recall	F1-Score

Domestic_trash	0.99	1.00	1.00
Infrastructure_Damage_Concrete	0.97	1.00	0.99
Parking_Issues_Illegal_Parking	0.99	0.98	0.99
Road_Issues_Damaged_Sign	1.00	0.97	0.98
Road_Issues_Pothole	0.96	0.98	0.97
Vandalism_Graffiti	0.95	0.94	0.95
Overall	**0.98**	**0.98**	**0.98**

Conclusion

A test accuracy of 97.72% was obtained on the final model, indicating that the CNN can classify a variety of urban problems accurately and reliably. Underlying the successes was multi-modal data fusion at scale and, critically, oversampling in an enrichment step. The new system is a powerful, unbiased proof of concept that meets all the essential requirements from the datathon challenge with strong results and positions being ready-to-implement for practical use.

Team members-

Nithila Nimraka Thawalampitiya

Imasha Nilupul Jayarathne

Thamindu Kulasekara