Chapter 3

Methodology

This chapter outlines the systematic approach taken to address the problem of classifying and assessing road defects using artificial intelligence. The methodology is divided into key phases: data collection, data preprocessing, algorithm selection, system implementation, and testing. Each phase is essential to accurately classify and assess road defects.

3.1 System Architecture

3.1.1 Overview of the System

The system receives video or images from the drone, preprocess them to enhance the quality, and uses YOLOv5-based model to detect and classify road defects. Severity assessment is in accordance to the official road maintenance manual and guidelines of the Department of Public Works and Highways (DPWH), and the integration of the multi-view depth estimation using epipolar spatio-temporal networks. The results are visualized in a user-friendly interface for assessment.

3.1.2 System Flow Diagram

The system's operational flow is outlined as follows:

1. Road Segment Image Capture
2. Preprocessing (image enhancement and alignment)
3. Defect Detection (YOLOv5 model)
4. Defect Depth Estimation (ESTDepth model)
5. Severity Assessment (following DPWH criteria)
6. Visualization of Results (interface for assessment)

3.2 Data Collection and Preprocessing

3.2.1 Data Sources

The custom YOLOv5 model was trained using a public dataset that was acquired from Roboflow. The data for testing is collected using the DJI Mavic 3 Multispectral drone, and essential hyperparameters, such as batch size, epoch count, and input dimensions, were tuned for optimal model performance.

3.3 Algorithm Selection

3.3.1 YOLOv5 Architecture

YOLOv5 was selected for its balance of real-time processing capability and accuracy, essential for detecting road defects in dynamic environments.

3.3.2 Severity Assessment

The Multi-view Depth Estimation using Epipolar Spatio-Temporal Networks was selected due to the high cost and limited accessibility of LiDAR technology. By applying epipolar geometry and temporal consistency across sequential frames, this approach provides an accurate depth estimation from a standard video footage (Long et al., 2021).

3.4 System Implementation

3.4.1 Development

The system is developed using Python.

3.4.2 Process Flow

3.4.3 Code Implementation

3.5 Testing and Validation

3.5.1 Test and Validation Plan

The system is tested using data gathered from ground truthing which involves manual inspection and measuring of road defects to verify the type, shape, and dimensions of the defect. These manual observations serve as a baseline reference to measure the system's accuracy in detecting, classifying, and severity assessment of road defects.

3.5.2 Challenges and Limitations

One major limitation is the availability of local labeled datasets, which affects the model's training, as most datasets available are those captured from foreign countries only.

3.6 Tools and Technologies

3.6.1 Programming Languages

The system was developed in Python

3.6.2 Libraries and Frameworks

Key libraries and frameworks include YOLOv5 for object detection, OpenCV for image processing, and PyTorch for model training and evaluation.

3.6.3 Software

Python, OpenCV, PyTorch, YOLOv5, CUDA, and Visual Studio Code IDE.

3.6.4 Hardware

The data was captured from a DJI Mavic 3 Multispectral Drone and the model was trained on a desktop computer with the following specifications:

- Memory: 16GB
- Processor: AMD® Ryzen 7 5700G @ 3.8Ghz
- Graphics: NVIDIA® GeForce GTX 1660 SUPER
- OS: Windows 11 Pro
- Disk: 1TB NVMe SSD