GLOBAL ACADEMY OF TECHNOLOGY, BENGALURU

ROAD CONDITION MONITORING FOR SMART CITIES:

POTHOLE DETECTION USING CNN

PRESENTED BY:

CHARAN RK
(1GA12AD014)

ABSTRACT

BACKGROUND:

- India: 3,597 annual deaths due to poor road conditions; potholes contribute to 9% of road accidents.
- Smart Cities initiative aims to improve urban living through technology, including road safety.
- Project introduces CNN-based system for automated pothole detection.

OBJECTIVES

PRIMARY OBJECTIVES:

- Develop a reliable pothole detection system using CNNs.
- Compare the performance of different CNN models (VGG16, VGG19, ResNet50, AlexNet).
- Employ Explainable AI methods to interpret and validate the model outputs.

SECONDARY OBJECTIVES:

- Contribute to enhancing road safety and reducing accidents in Smart Cities.
- Provide a scalable solution that can be integrated into existing road monitoring systems.

METHODOLOGY

DATA COLLECTION:

- This dataset is a mix of images of roads from Google and already present pothole dataset.
- Data includes images of both good and poor road conditions to ensure balanced training.

CNN MODELS:

1. **VGG16 Architecture**:

- 16 Layers (13 convolutional layers and 3 fully connected layers).
- Uses small 3x3 filters throughout the convolutional layers.
- The depth of the network allows for more complex feature extraction.
- Ends with three Fully Connected (FC) layers before the final softmax layer for classification.

METHODOLOGY (CONTINUED)

2 . VGG19 Architecture:

- 19 Layers (16 convolutional layers and 3 fully connected layers).
- Similar to VGG16 but with three additional convolutional layers.
- Offers more in-depth feature extraction but at the cost of increased computational complexity.

3 . ResNet50 Architecture:

- 50 Layers (includes convolutional layers, identity blocks, and fully connected layers)
- Allows for much deeper networks without performance degradation.
- Consists of 48 convolutional layers along with 1 MaxPool and 1 AveragePool layer.

4. AlexNet Architecture:

- 8 Layers (5 convolutional layers and 3 fully connected layers).
- Uses larger filters in the initial layers followed by smaller ones in deeper layers.
- Incorporates dropout layers to reduce overfitting

METHODOLOGY (CONTINUED)

EXPLAINABLE AI (XAI) METHODS:

LIME (Local Interpretable Model-agnostic Explanations):

- Generates local explanations by perturbing input data and observing the change in predictions.
- Useful for understanding individual predictions in the context of CNNs.

SHAP (SHapley Additive exPlanations):

- Provides a unified measure of feature importance based on game theory.
- Explains the contribution of each feature to the final prediction.

Saliency Maps:

- Visualizes the most influential pixels in an image that affect the model's output.
- Highlights areas that the model considers important for decision making.

Grad-CAM (Gradient-weighted Class Activation Mapping):

- Produces coarse localization maps by calculating gradients of the target concept.
- Highlights the regions of the input image that are most important for the model's prediction.

DEPLOYMENT AND INTERPRETATION

1
2
3
4
5
6
7
8
9

DATA PREPROCESSING

- Data collection, cleaning, and preprocessing
- Ensuring data quality and reliability
- Data labeling and annotation

BUILDING THE MODEL

- Various CNN Model architecture.
- Hyperparameter tuning
- Training process and validation

MODEL EVALUATION AND INTERPRETATION

- Pass test data to the best_model.
- Assessing model's performance, including accuracy, precision and recall

RESULTS

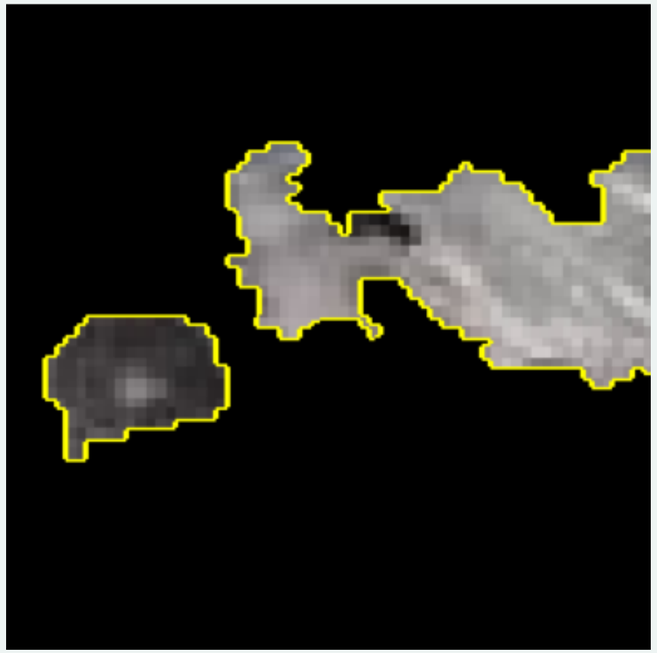
- Upon Hyper-Parameter Tuning on our models, we found AlexNet as the best model with test accuracy around 80% for 200 epoch.

XAI METHODS

LIME (Local Interpretable Model-agnostic Explanations):

Original Image

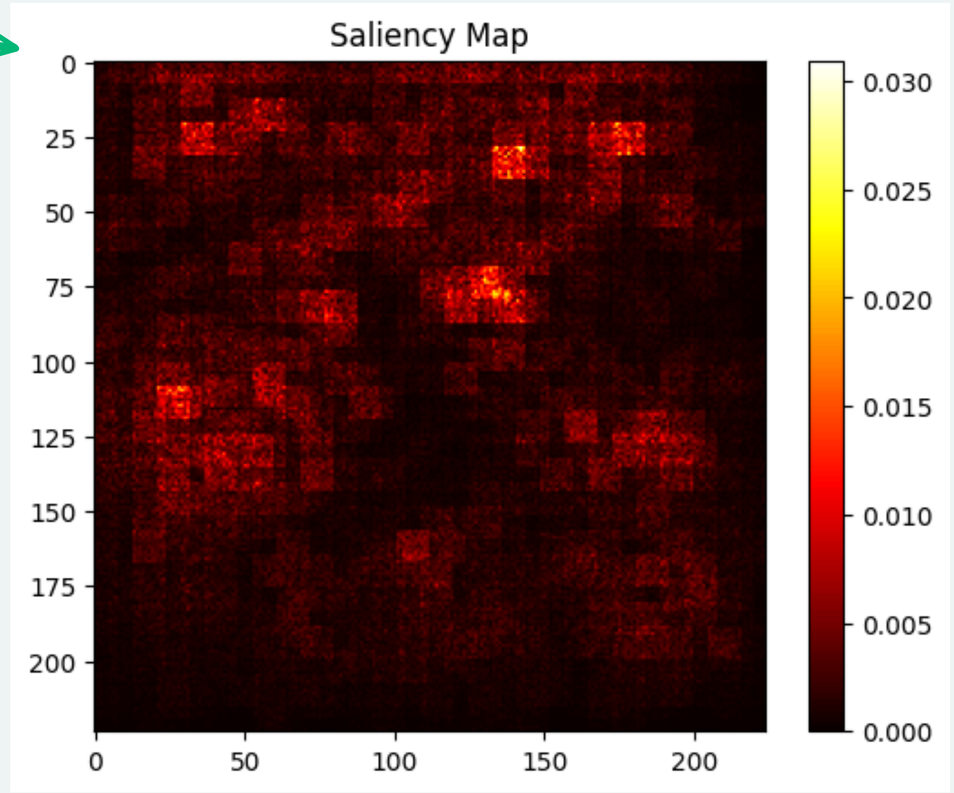
LIME Explanation

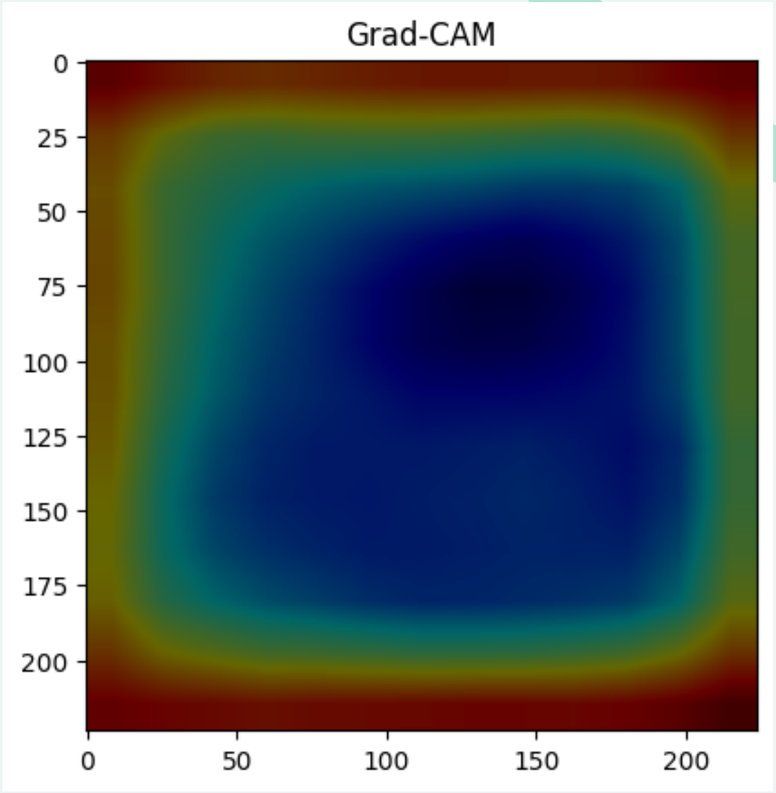


RESULTS (CONTINUED)

XAI METHODS

Saliency Map





Grad-CAM

1
2
3
4
5
6
7
8
9

Thank you