Bauhaus-Universität Weimar

BASIC CRACK DETECTOR USING SUPERVISED MACHINE LEARNING CLASSIFIERS

(SUPPORT VECTOR MACHINES AND DECISION TREES)

FINAL PROJECT

IMAGE ANALYSIS AND OBJECT RECOGNITION

SUMMER SEMESTER 2024

GROUP 16

Aditya Kumar Gupta

Olubunmi Emmanuel Ogunleye

Ajay Satish Patil

PROGRAM: DIGITAL ENGINEERING

FACULTY OF CIVIL ENGINEERING / MEDIA

Under The Guidance of

Prof. Dr. Volker Rodehorst

M.Sc. Mariya Kaisheva

Task 1: Data Acquisition

The Crack Images were collected at different locations around the city of Weimar.

Task 2: Data Annotations and Augmentations

The annotations are carried out using GIMP software. In the annotations, the cracks are labelled with white (255) as foreground and non-crack regions are labelled with black background (0). We obtained a total of 12 Images after annotation and augmentation.

Task 3: Data Split

In order to evaluate the performance of the algorithm, the dataset was splitted into a subset. Training dataset which represents 80% of the entire data and testing dataset which represents 20% of the entire dataset.

Task 4: Crack Segmentation

a. The images are converted to gray scale and thresholding is performed in order to

S/N	Statistics	Value
1	Number of Images Collected	12
2	Dataset size after augmentation	12
3	Software used for Annotation	GIMP
4	Annotated Images Number of Crack Region	9
5	Annotated Images Number of Non-Crack Region	9

   obtain a binarized form of the image.
b. Morphological operators such as Opening, Closing and Thinning are performed in order to clean up the binarized image.
c. Connected Component Analysis is performed in order to extract discrete regions in the images.
d. Features extracted from the regions obtained from connected component analysis are.
   - Number of Pixels (Area)
   - Perimeter
   - Eccentricity

The numbers of pixels and perimeter are selected specifically as the cracks will have less pixels in relation to the perimeter, therefore the classifier would be capable of training and distinguishing the crack and non-crack regions.

e. The classifier used to filter out non-crack regions are Support Vector Machines (SVM) and Decision Trees. They are supervised machine learning algorithms for semantic segmentation of features as learnt in the course.

Task 5: Crack Analysis

a. Performance Metrics:

The result from the Support vector Machine Classifier yielded 50% accuracy, while the Decision Tree classifier yielded 100% accuracy. This implies that the Decision Tree Classifier performed better.

b. The length of crack for test images is reported in the command window of MATLAB.

```
Number of Crack region: 9
Number of Non Crack region: 9
Warning: Colon operands must be real scalars. This warning will become an error in a future release.
> In CrackDetectionProject>CalculateCrackLength (line 225)
In CrackDetectionProject (line 111)

Crack length of the region is : 4483 pixels
Accuracy of SVM Classifier: 50%
Accuracy of Decision Tree Classifier: 100%
fx >>
```

c. Comments and Observations:
- Some Cracks were not detected in few of the data samples, this could be because of the size of the crack in those images.
- A sufficient dataset could possibly uncover scenarios of overfitting or underfitting if such exists within the model.