**BASIC CRACK DETECTOR USING SUPERVISED MACHINE LEARNING CLASSIFIERS**

**(SUPPORT VECTOR MACHINES AND DECISION TREES)**

**FINAL PROJECT**

**IMAGE ANALYSIS AND OBJECT RECOGNITION**

**SUMMER SEMESTER 2024**

**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.

A graph and a diagram

Description automatically generated

A comparison of a graph and a diagram

Description automatically generated

A graph and a diagram

Description automatically generated

A close-up of a crack in a stone

Description automatically generated

A close-up of a graph

Description automatically generated

A graph and a diagram

Description automatically generated

A graph and a diagram

Description automatically generated

A close-up of a graph

Description automatically generated

A graph and a picture of a person

Description automatically generated

A graph and a diagram

Description automatically generated

|  |  |  |
| --- | --- | --- |
| S/N | Statistics | Value |
| 1 | Number of Images Collected | 11 |
| 2 | Dataset size after augmentation | 38 |
| 3 | Software used for Annotation | GIMP |
| 4 | Annotated Images Number of Crack Region | 22 |
| 5 | Annotated Images Number of Non-Crack Region | 280 |

A close-up of a graph

Description automatically generated

**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 38 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**

1. The images are converted to gray scale and thresholding is performed in order to obtain a binarized form of the image.
2. Morphological operators such as Opening, Closing and Thinning are performed in order to clean up the binarized image.
3. Connected Component Analysis is performed in order to extract discrete regions in the images.
4. 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.

1. 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**

1. Performance Metrics:

The result from the Support vector Machine Classifier yielded 86.7% accuracy, while the Decision Tree classifier yielded 85% accuracy. This implies that the SVM Classifier performed better.

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

A screenshot of a computer

Description automatically generated

1. 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.