ENAC / PROJET DE MASTER 2020-2021
SECTION DE GÉNIE CIVIL

Automation of the diagnosis process in the railway system : Detection of defects in concrete sleepers using vision-based machine learning models

Author : Linah CHARIF

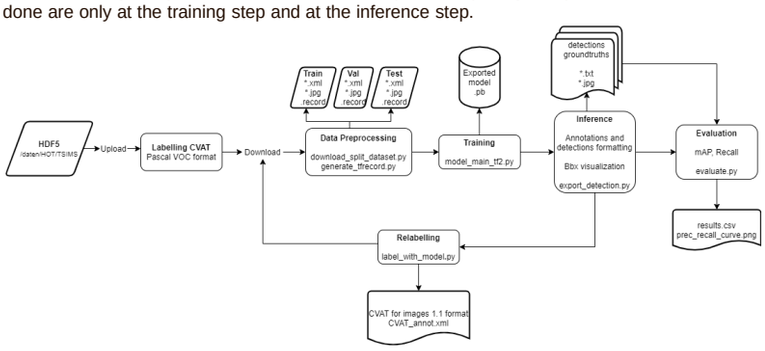
Supervision : Prof. Nikolas GEROLIMINIS [1] / Ilir FETAI [2] / Lucian Stefan ANCU [3]

[1] Urban Transport Systems Laboratory (LUTS) EPFL / [2] Competence Center for Machine Perception, SBB / [3] Measurement and Diagnosis Department (MUD), SBB

The diagnosis in the railway infrastructure has always been done with manual inspection, thanks to rail inspectors that walk over the track and check the defects on the rail surface, fasteners and sleepers. In the case of concrete sleepers, rail inspectors classify defects according to their size and occurrence over 20 sleepers. The manual inspection is error prone, it is time consuming and it can be sometimes risky. In an effort to overcome these issues, SBB has developed a project called AISI (Artificial Intelligence Streckeninspection) in order to automatize the rail inspection by collecting and analyzing images with diagnosis trains equipped with cameras. The objective of this thesis is integrated within the AISI Project, since its goal is the conception of a deep learning model to detect and classify concrete sleepers defects using computer vision and more specifically, object detection. The defects studied in this thesis are cracks and flakings (or spallings). In the literature, many researchers have tackled the subject of defect detection in concrete, especially in civil infrastructures. The most common methods a combination of image-processing and Machine Learning (image pre-processing and a classifier) whereas the latest methods use deep learning (CNN, FCN and RPN, YOLO). The proposed methodology uses a Fast-RCNN model developed with Tensorflow Object Detection API. The evaluation of the model is given on a total of 996 defects and its precision and recall are given. The model shows average results as the number of false positives is high. It has been noted that the labels provided needed to be improved. The proposed approach has a great potential of development since a large amount of training data can be added thanks to the produced relabelling process. Also, the designed pipeline establishes a base that can be adjustable for various algorithms and configurations.

Currently, the railway inspections are still greatly human dependent : usually the entire network is inspected visually and the defects are collected manually. In the conduct of a more efficient and sustainable diagnosis, the Swiss Federal Railways (SBB CFF FFS) is gathering track images captured with high resolution cameras placed on diagnosis trains travelling up to a speed of up to 160 km per hour. SBB aims to develop supervised algorithms to detect anomalies and classify defects on rails and sleepers that are then subject to verification by experts. As of today, SBB has already tackled the subject of rail surface defect detection using Machine Learning algorithms, in the scope of a project called Artificial Intelligence Streckeninspection (AISI). Within this same project, SBB aims to develop a vision-based defect detection system focused on the sleepers which is the main research topic of this thesis.

The main goal of the proposed approach is the detection and classification of defects on concrete sleepers using computer vision. The figure below shows the flowchart of the proposed approach. The structure of the code allows to run each one script for each process. This pipeline could be easily adapted for other frameworks (for example Pytorch), the modifications to be done are only at the training step and at the inference step.



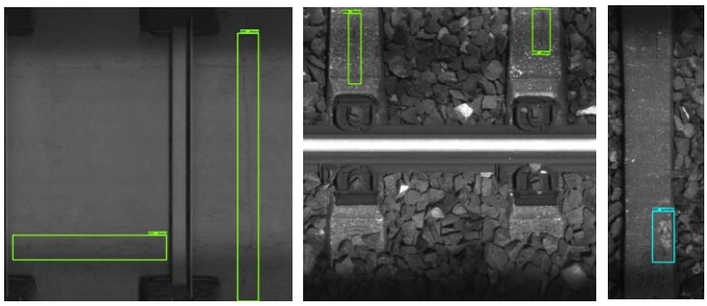
The first step to developing a computer vision model is to collect and label the images. The images used in the case of this thesis are extracted from different databases originally stored in HDF5 format, that gathers the measurements of diagnosis trains (DFZ and gDFZ) using the MERMEC technology V-Cube, Figure below (source: MERMEC official website).

Two different sets of images are used to construct the dataset, equally divided into three channels (left, central, right). The samples consist of a mix of concrete sleepers, wood sleepers, and tunnel platforms. At a later stage, the wood sleepers are discarded, and all the images are cropped to have a 1024x1024 size, to match the required input size for model training. The labelling process has been done using Computer Vision Annotation Tool (CVAT). The data is split into Train, Validation and Test as presented in the Table below :

	Train set	Validation set	Test set
Dataset 1	80%	20%	-
Dataset 2	-	50%	50%
Size (pixel)	1024x1024	(3000 or 1024)x1024	3000x1024
Total number of defects	3491 (64%)	1006 (18%)	996 (18%)

The chosen object detector is Faster-RCNN, it has been demonstrated that it is the model that achieves the best performance in both 2016 COCO object detection challenge in comparison with Single-Shot Detector (SSD) and Region-based Fully Convolutional Network (R-FCN). Faster-RCNN is a two stage detector, meaning that it is a combination of two sequential processes : object proposal and object classification. Prior to these two stages, a CNN extracts features from the images.

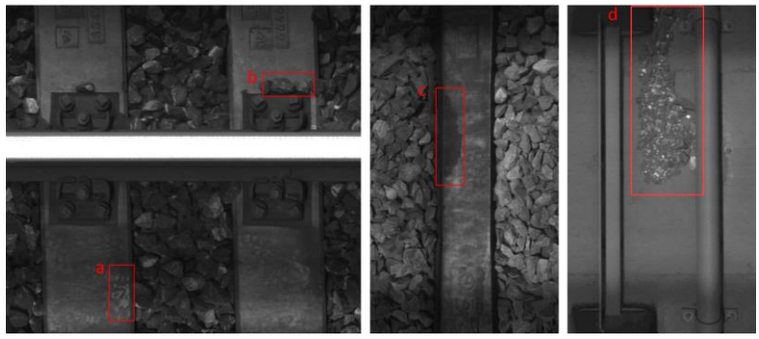
The main framework that has been used is TensorFlow object detection API. As a CNN base the Inception Resnet V2 architecture is applied. A pretrained model on COCO 2017 dataset can be found in TensorFlow 2 Detection Model Zoo. In this case, Transfer Learning is used to adapt of the pretrained model to our dataset. In order to challenge the Tensorflow Object Detection API model, another library has been used as an experiment. It is called MMDetection. It is an open source object detection toolbox based on PyTorch. The training and validation process of the Faster R-CNN model was performed on a computer with 4 NvidiaTesla V100-SXM2 GPUs, 32 GB memory (3 GPUs for training, 1 GPU for validation) working under CentOSLinux 8 operating system. The figure below shows the Tensorflow model outputs for cracks in green and flakings in blue.



The evaluation is done on the test set on a total of 2279 images, with a total of 275 cracks and 406 flakings. The metrics used to assess the performance for each class are precision and recall, using a score threshold of 0,5 and an IoU threshold of 0,2.

Model	Class	Precision	Recall	True Positives	False Positives
Tensorflow	Crack	55%	32%	88	72
	Flaking	11%	23%	92	767
MMDetection	Crack	40%	10%	27	41
	Flaking	10%	7%	29	273

It can be generally noted that the cracks are better detected than flakings which have a great diversity in shape and size. Also, the model outputs a lot of false positives in the case of flakings which can be due to a confusion between the flaking and the rest of the sleeper like in the images below (a) Footprints, b) Rocks, c) Water stain, d) Undefined defect).



In summary, in this thesis has been applied state-of-the-art object detection methods as a mean to automate the diagnosis of railway infrastructures. The results obtained are promising since this model reduces the number of false positives by 80% in comparison with the former system used by SBB. However, some improvements are still needed to meet the AISI project requirements for production. It is to note that this process is still dependent on human intervention, upstream for the labelling process and downstream for the validation of detected defects. However, it limits the field excursions and it increases the efficiency of the diagnosis. The developed approach satisfies the AISI project mission statement and will be further refined within SBB to become a common practice in the diagnosis of concrete sleepers.