

Metric Benchmarking Tool

Christian Hellert¹, Christian Brunner²,
Dominik Brüggemann³, Tom Thielo²,
Jonas Schneider²

Overview

The *Metric Benchmarking Tool (MBT)* is an application to perform standard benchmarks within the project KI Absicherung. It is designed to enable and simplify evaluations regarding the effectiveness of mechanisms implemented to improve pedestrian recognition. The tool uses the KI-Absicherung dataset with available enriched metadata and 2D bounding box predictions provided in the project specific output format to compute a user specified set of object detection metrics. Figure 1 shows the workflow of the tool.

Configuring the Tool

A global configuration file is used to adjust the MBT to the needs of the user. The configuration file is read in and settings are distributed to the respective processing modules displayed as blue boxes below.

Loading Data

First, annotations and predictions are loaded using the *Annotation Loading and Prediction Loading* modules. 2D bounding box annotations and predictions are used along with the available enriched meta information.

Correlating Bounding Boxes

Next, the *Correlation* module takes the annotations and predictions and performs a correlation between them. The matching is based on an Intersection over Union (IOU) threshold. By default a n-to-m matching is computed.

Applying Filters and Reducing Matching

After correlation, the *Filtering* module takes annotation, prediction and matching data frames and filters them according to user defined rules. The result is a filtered matching data frame, which is forwarded to a reduction step. Here the matching is reduced from the n-to-m to a 1-to-1 matching, as required for the computation of most metrics.

Computing Metrics

Based on the provided user configuration, the selected metrics are calculated. Results are stored into a data frame, which then contains the metric value along with the name and the identifier of the metric. Figure 2 shows example results for a set of commonly used metrics on validation data.

Metric	Value
Precision	23.78%
Recall	82.27%
F1 Score	0.37
VOC mAP	63.79%

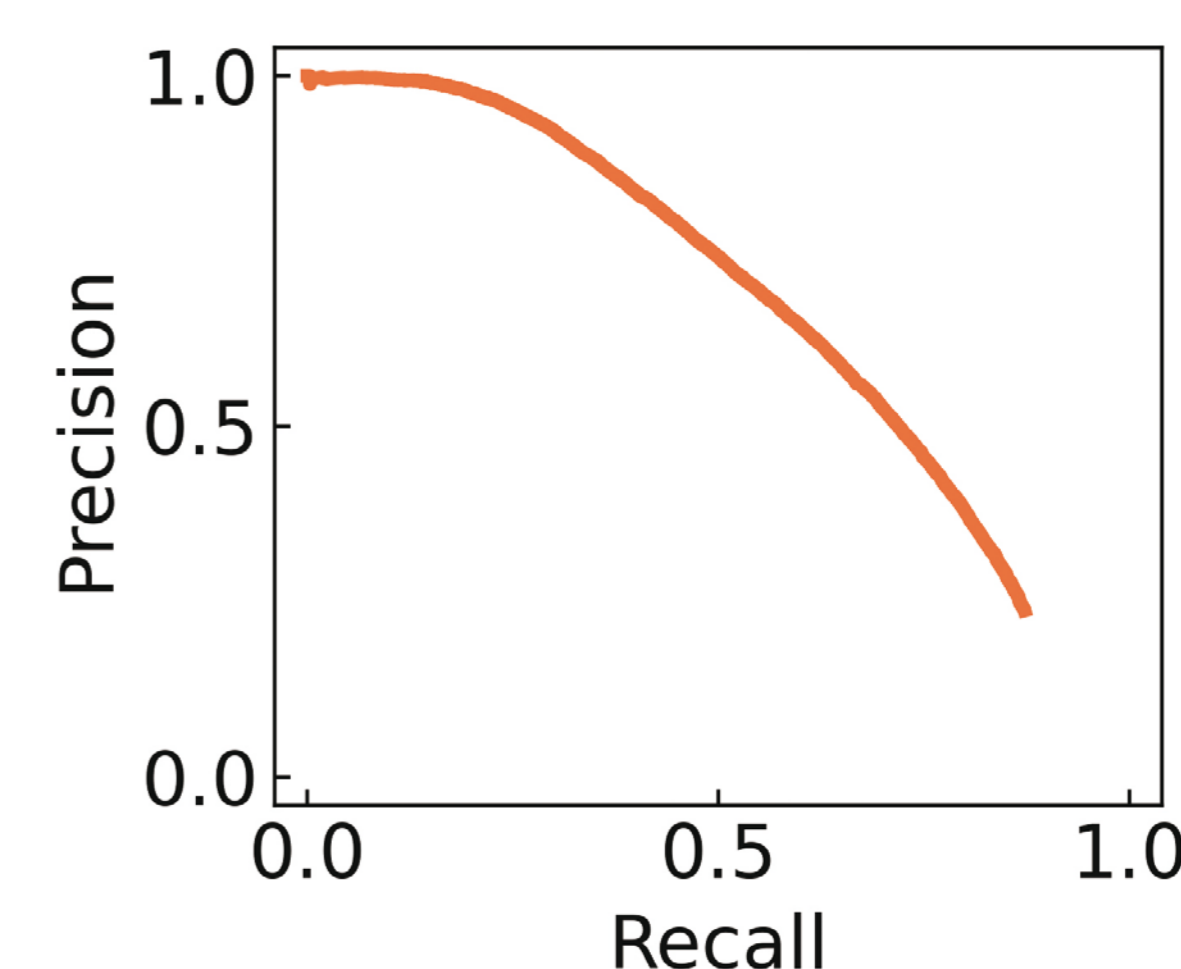


Figure 2: Metric results and Precision-Recall curve for critical pedestrians (category A) within the KI-A Tranche 3 to 5 validation data split.

Exporting Results

Finally, the *Exporting* module is invoked to write the results into JSON files for easy further processing according to the KI Absicherung exchange format.

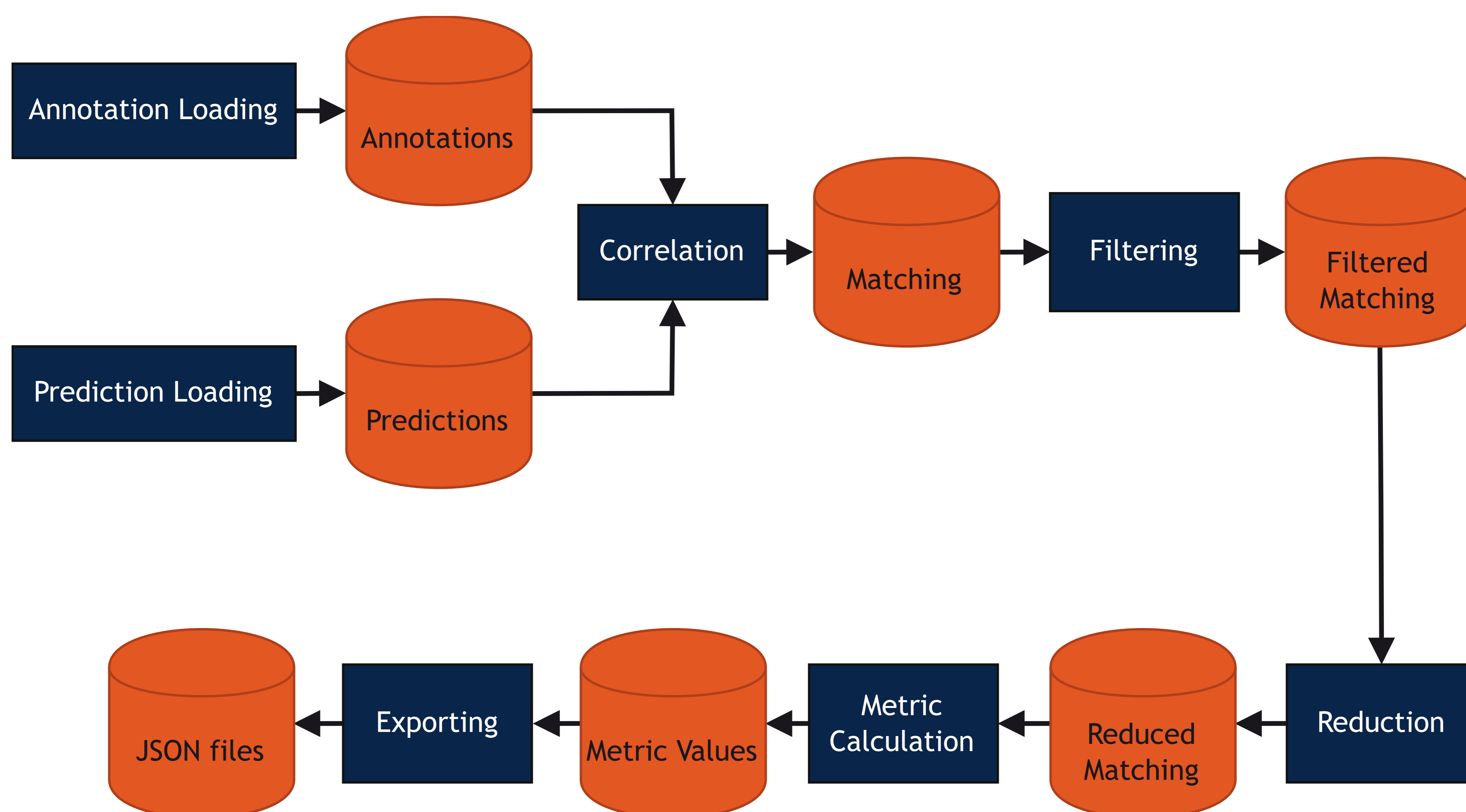


Figure 1: Workflow of the Metric Benchmarking Tool. Blue boxes represent processing modules and orange boxes represent data structures.



For more information contact:
christian.hellert@continental-corporation.com
christian.brunner@efs-auto.com
dbrueggemann@uni-wuppertal.de

KI Absicherung is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.



Supported by:



on the basis of a decision by the German Bundestag