# Novelty-based Generalization Evaluation for Traffic Light Detection: Supplementary material

July 8, 2021

## 1 Implementation detail: Normalization

The density scores from a KDE are inversely proportional to the novelty. That is, a lower density score means higher novelty. Secondly, these density values may have also negative values. However, its not possible to use negative weights to perform weighted average for our $\mathcal{G}$ score. As the smaller density values imply higher novelty, we take the absolute of the negative values as the novelty score. On the other hand, the larger density values imply lower novelty. To reduce the impact of these samples we take the inverse of the positive density values (c.f. Figure 1). The zero density values are replaced by its nearest negative value prior to aforementioned operation.

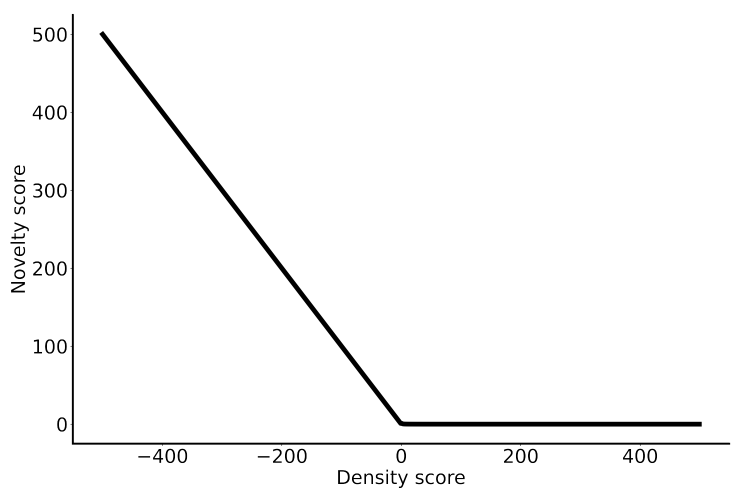


Figure 1: KDE density score normalization and transformation as novelty score

## 2 False positive analysis with KDE and LOF

In Section 4-A, we performed a comparative study of the various novelty scorers. Although the ROC scores show that the LOF was better than KDE we use KDE as the scorer because it had much steady behaviour with increasing novel samples in the dataset. As shown in Figure 3, the LOF scorer shows a higher FP rate in comparison to the KDE. For this reason we chose KDE, the second best as our novelty scoring algorithm.

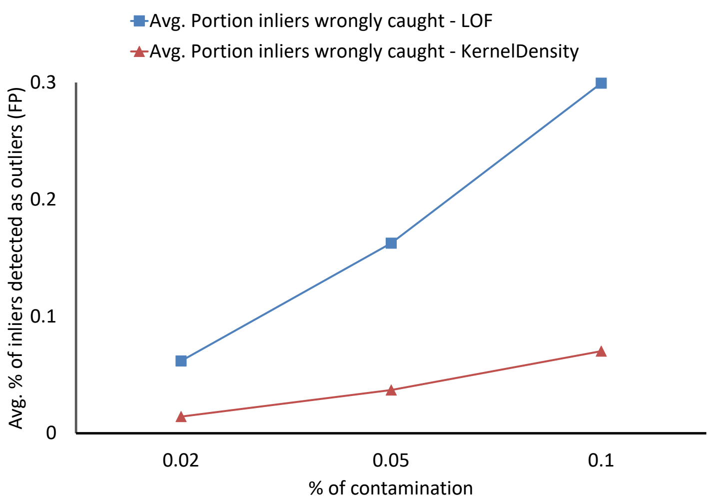


Figure 2: False positive increase with increasing novel objects in dataset for LOF

## 3 Other figures

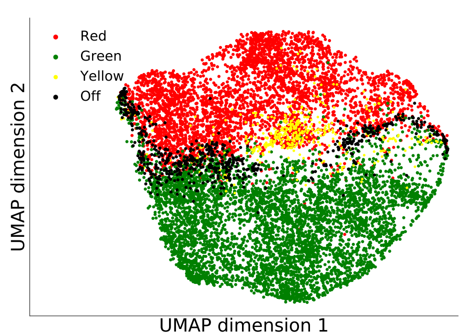


Figure 3: BSTLD data on 2-dimensional space using UMAP