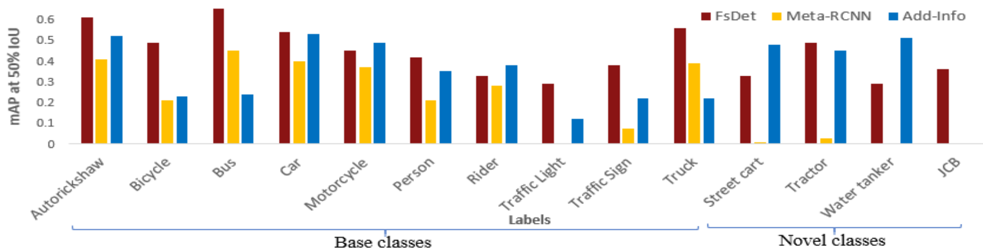

Figure 5: Classwise performance of metric-learning and meta-learning based techniques in detecting rare objects in IDD dataset. The last 4 classes represent the rare categories. FsDet, a metric learning based approach, performs better than meta-learning approaches in both base and novel classes.

4.3 Discussion

Our experiments uncovered several findings on the nature of road objects and the behaviour of few-shot object detection networks in the context of driving. Results from Tables **??** and **??** demonstrate that cosine similarity based TFA architectures (FsDet) outperforms meta-learning based architectures (Meta-RCNN, Meta-Reweight and Add-Info) on novel-class performance by 11.2 mAP points on IDD-10 (split 1) and 1.0 mAP point on IDD-OS split. We attribute lower inter-class distance between new object categories as the probable reason for the lower performance of meta-learning over similarity-based methods. This aspect of IDD makes it a unique dataset well suited for evaluation of few-shot object detection in a real-world, driving scenarios.

Comparing the base class performance in Table **??** against the roofline performance metrics in Table **??**, we demonstrate a lower degradation in base-class performance when adopting TFA architecture (FsDet) over its meta-learning counterparts, after the introduction of novel classes. Meta-learning techniques like Meta-RCNN and Add-Info suffer a significant reduction in base-class performance except when additional features were provided to the final prediction head of the object detector.

Figure **??** shows class-level confusion among all classes in IDD-OS split trained on 10-shot data samples using TFA architecture. In particular, the confusion between $truck$ vs. car, $bicycle$ vs. $motorcycle$ and $water\text{-}tanker$ vs. car classes are high, with the maximum being 40%. This observation can possibly be explained by the fact that road objects in context, share a large number of low-level features with other object classes, thus posing a challenge for few-shot algorithms to differentiate. This observation here is in line with that by the authors of MetaDet (**?**) that confusion between classes is the primary challenge in few-shot learning scenarios. This is further echoed by the authors of Meta-Reweight (**?**) that there exists a high confusion of 50% among classes in PASCAL VOC dataset.

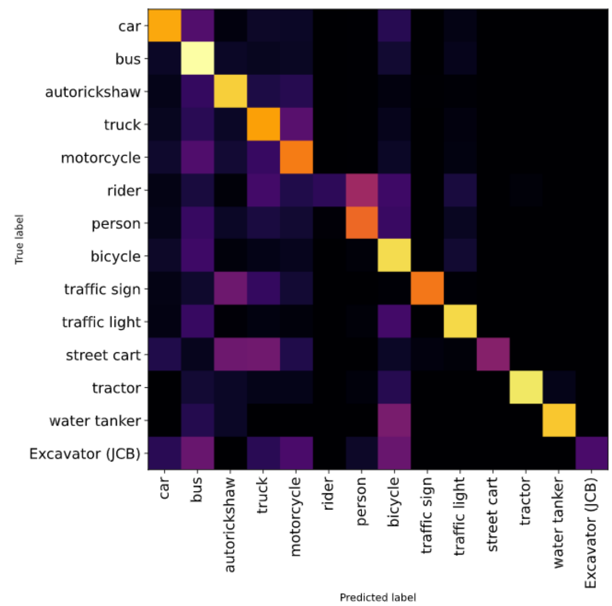

Figure 6: Confusion matrix plotted for class prediction results from IDD validation dataset showing confusion between classes when trained on IDD-OS 10-shot split on FsDet network.

5 Conclusion

We analyzed the performance of state-of-the-art methods for few-shot object detection, using a real-world dataset (IDD) which inherently contains class-imbalanced data from driving scenarios. Our evaluation of methods was for two tasks: same-domain and open-set representations. To evaluate these settings, we expanded a publicly available dataset with additional class labels in the open-set representation. By creating an extension of IDD, we hope to pave a way for many future works in few-shot learning with real-world datasets. Based on our experiments, we conclude that cosine similarity based TFA network (FsDet) outperforms meta-learning based networks in both the tasks by 11.2 and 1.0 mAP points in novel class performance respectively. We conclude that meta-learning networks while achieving great strides, tend to under perform even simpler baselines from metric-learning based methods. We also observe that class-confusions remains an open challenge in any few-shot learning paradigm and can be the focus of further improvements.

Modi aperiam quam deleniti unde, a eveniet iusto laborum eligendi minima ut et qui soluta, ducimus laboriosam iusto, deserunt quis ratione, impedit veniam blanditiis aperiam architecto est officia sed qui?Repellendus assumenda accusantium ullam quis culpa qui, expedita impedit mollitia tempore accusamus est?