Background
When open-set unlabeled data contain outliers from unseen classes, mainstream SSL methods experience significant performance drops, as it is impossible to generate correct closed-set pseudo-labels for outliers.

Motivation
A common strategy employed in prior research is to first detect and then filter outliers out. However, it is quite challenging to obtain a reliable outlier detector at the outset, especially when labels are extremely scarce. We observed that an unreliable detector can be more harmful than the outliers themselves, since it may wrongly exclude numerous inliers.

Core Approach of IOMatch
Can we jointly utilize open-set unlabeled data without the need for precise differentiation between inliers and outliers?
We achieve this by leveraging unified open-set targets as pseudo-labels:
• A standard closed-set classifier is used to predict the most likely seen class for an unlabeled sample, with the probability \( p = (p_1, ..., p_K) \).
• An additional multi-binary classifier is incorporated. Each binary classifier is designed to determine whether an unlabeled sample truly belongs to each seen class or not, with the probability \( o_k = (o_{k1}, o_{k2}) \).
• By combining these two predictions, we can estimate the likelihood of an unlabeled sample being an inlier \( p_k \times o_k \) of each seen class or an outlier \( \sigma p_k \times o_k \).
• We optimize an open-set classifier with these unified targets, via the consistency-regularized pseudo-labeling scheme.

Experiments

Conclusion
We proposed a simple yet effective open-set SSL framework, IOMatch, and we found:
➢ It is challenging, but not mandatory, to identify outliers before performing pseudo-labeling.
➢ What truly matters is the idea of joint inliers and outliers utilization. Producing unified open-set targets is just one way, and we can explore stronger techniques for this.