

Data distillation: towards omni-supervised learning

Summary

1. Propose data distillation, which distills knowledge from unlabeled data without the requirement of training a large set of models.
2. Data distillation involves 4 steps:
 - training a model on manually labeled data.
 - applying the trained model to multiple transformations of unlabeled data
 - converting the predictions on the unlabeled data into labels by ensembling the multiple predictions.
 - retraining the model on the union of the manually labeled data and automatically labeled data.
3. Show up to 3 points AP improvement over a strong Mask R-CNN baseline on the human keypoint detection task of the COCO dataset.