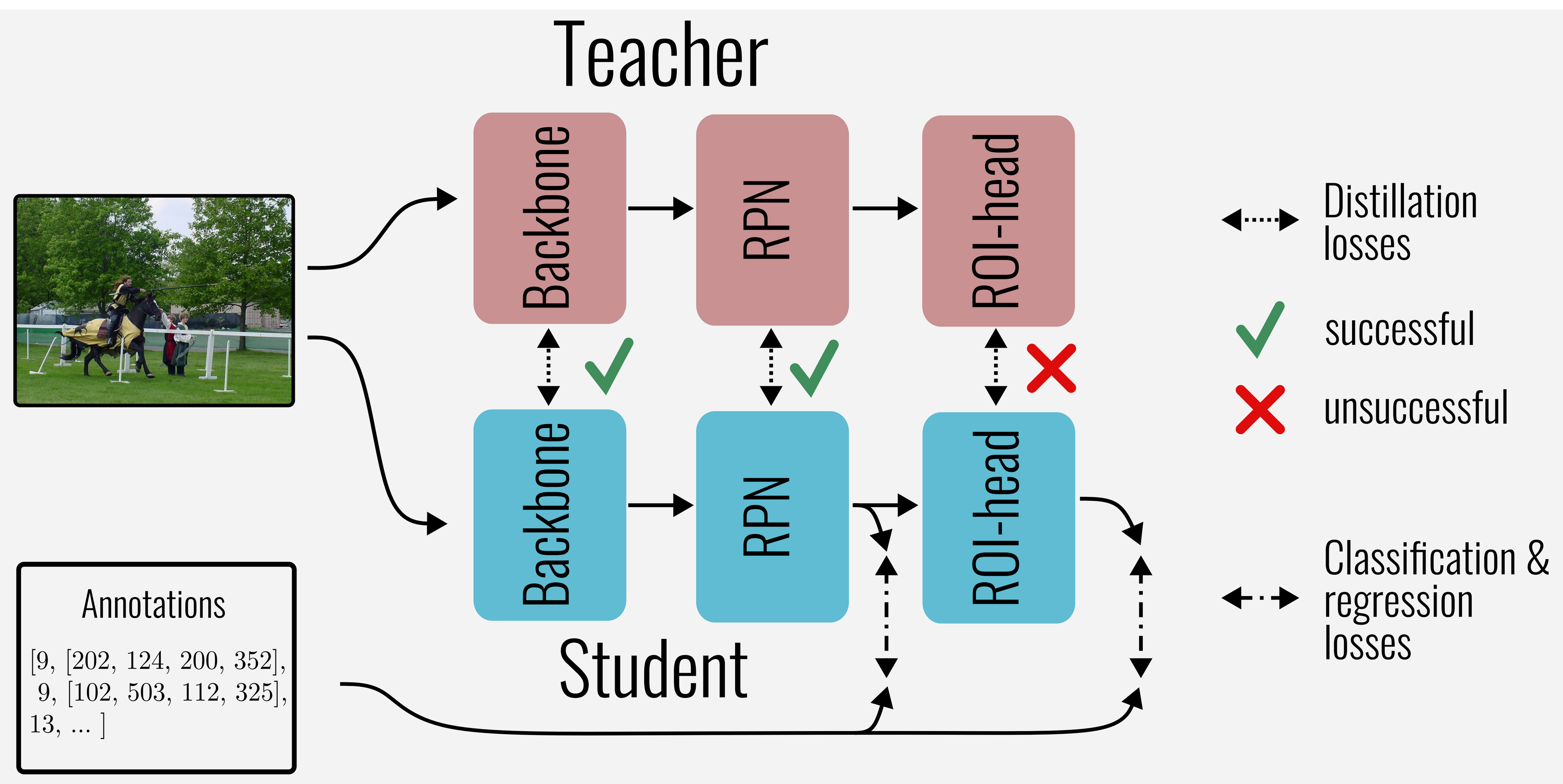
KU LEUVEN

RE-EXAMINING DISTILLATION FOR CONTINUAL OBJECT DETECTION

Eli Verwimp[1] - Kuo Yang[2] - Sarah Parisot[2] - Lanqing Hong[2] - Steven McDonagh[2] - Eduardo Pérez Pellitero[2] - Matthias De Lange[1] - Tinne Tuytelaars[1]

[1] KU Leuven - [2] Huawei Noah's Ark Lab

Introduction



2 GOALS

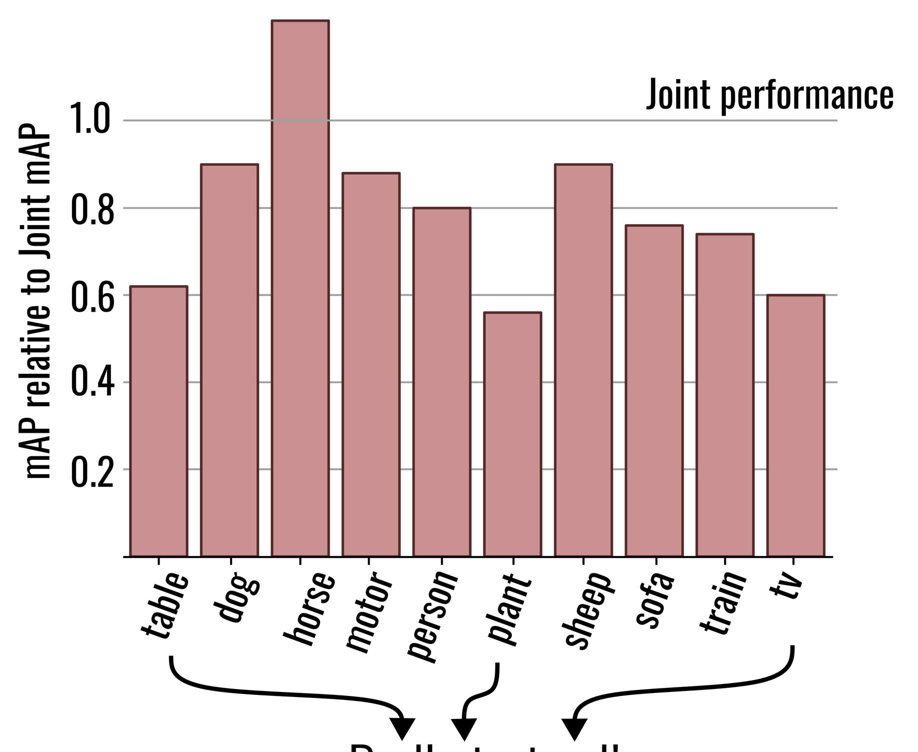
Retain old knowledge from former tasks present in the teacher model, by using knowledge distillation from the teacher to the student model

Learn new tasks from newly annotated data as well as possible. These can be new classes or new domains.

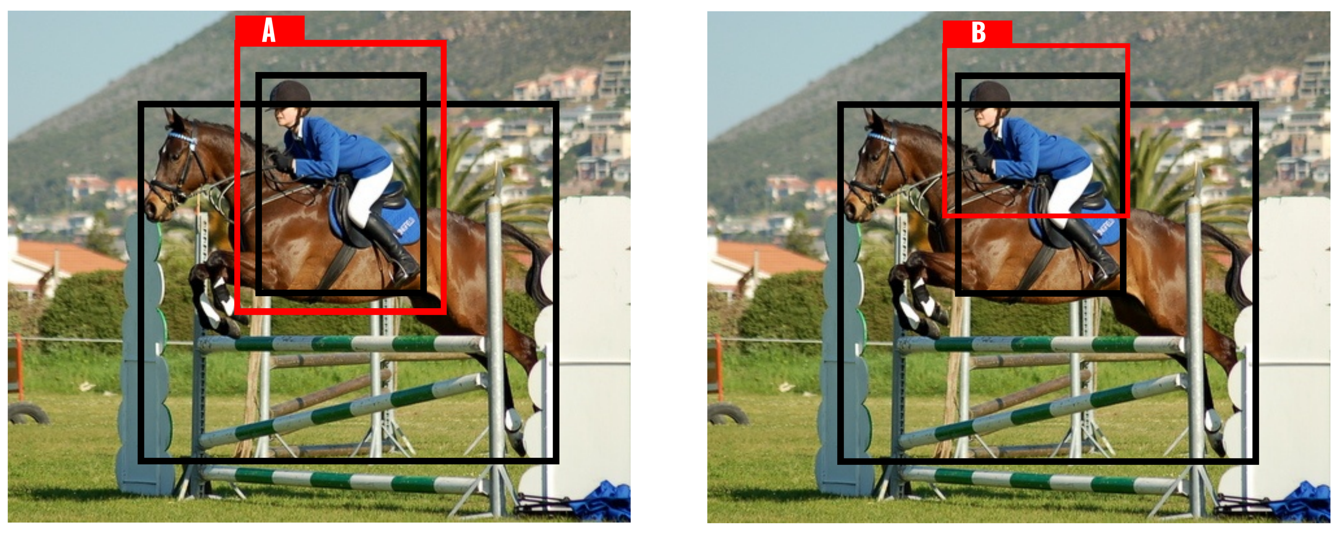
Problems

Illustrated on VOC10+10

Not all classes from the new task are learned well

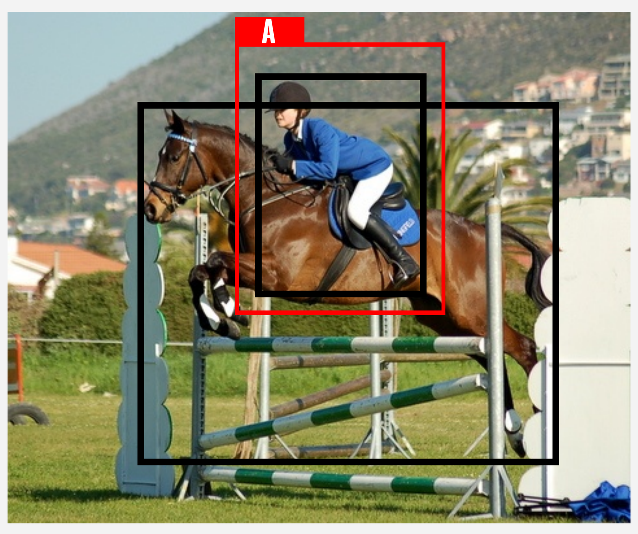


New objects (e.g. person) were present in old task's images (e.g. horse). The teacher has learned to classify these as background or another class and these predictions are used in knowledge distillation. This is a problem when overlap is high (A) and lower (B).

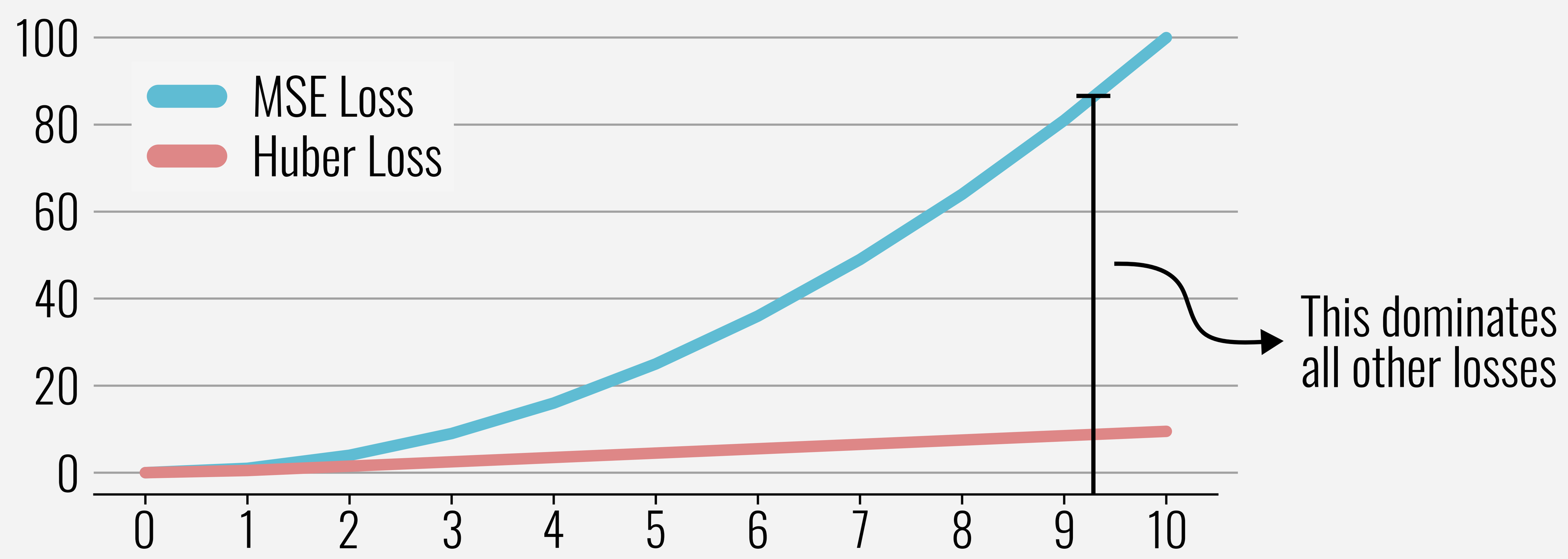


Solutions

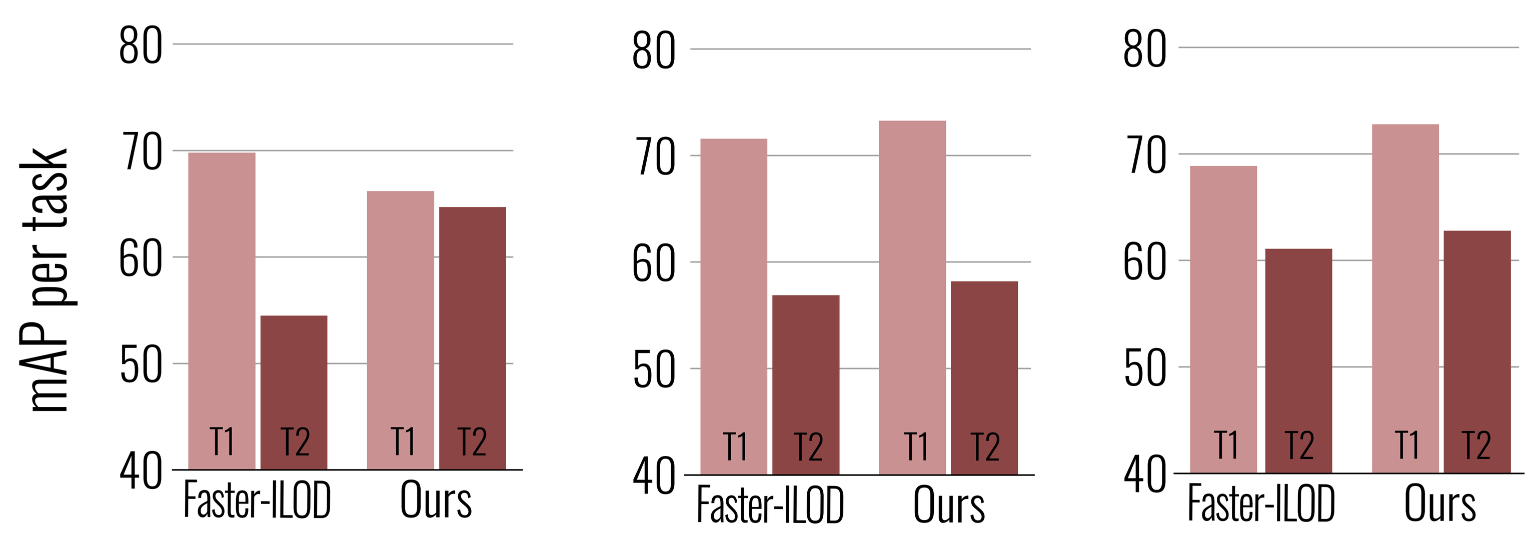
Selective Distillation (solves A)
Don't use proposals overlapping with new GT in distillation loss

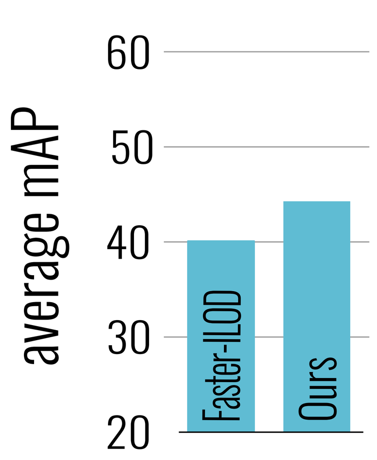


Huber Loss (solves B)
Penalizes large differences less than MSE



Results





Conclusion

1. Most continual learning issues are in the ROI-heads
2. Careful application of distillation improves effectiveness
3. Room for improvement by leveraging rehearsal techniques (see paper)

Scan for online paper!

BMVC 2022