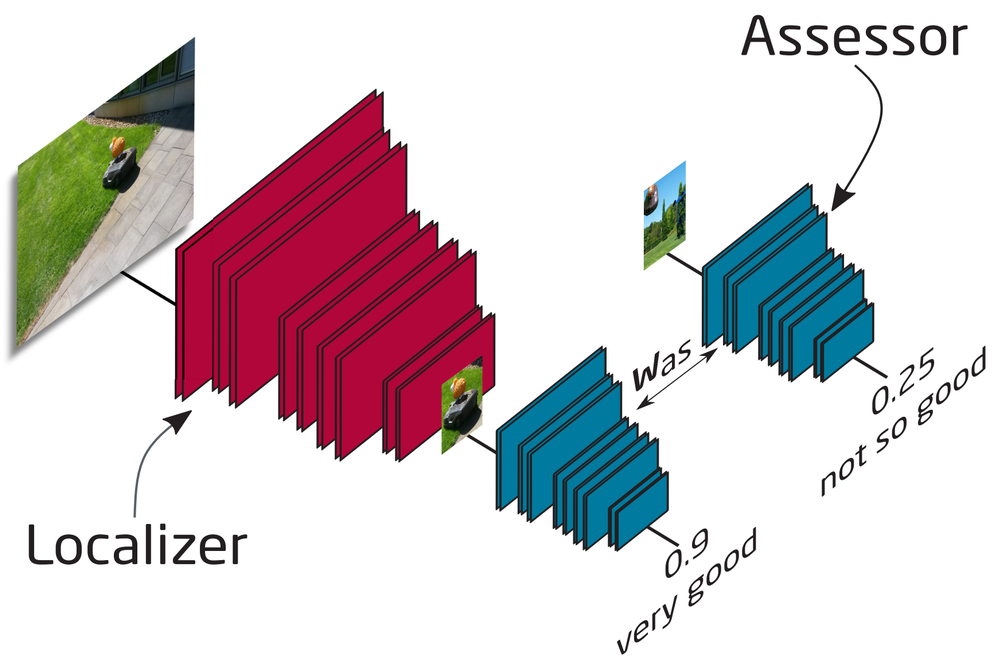
LoANs: Weakly Supervised Object Detection with Localizer Assessor Networks

Problem

- creating bounding box labels is a costly process
- current methods for weakly supervised object detection only use low-level or implicit cues for localizing objects

Proposed System



Idea

- use knowledge of a teacher network to train a student network in knowledge transfer fashion
- no bounding box annotation for training of student necessary
- teacher can be trained on artificial data

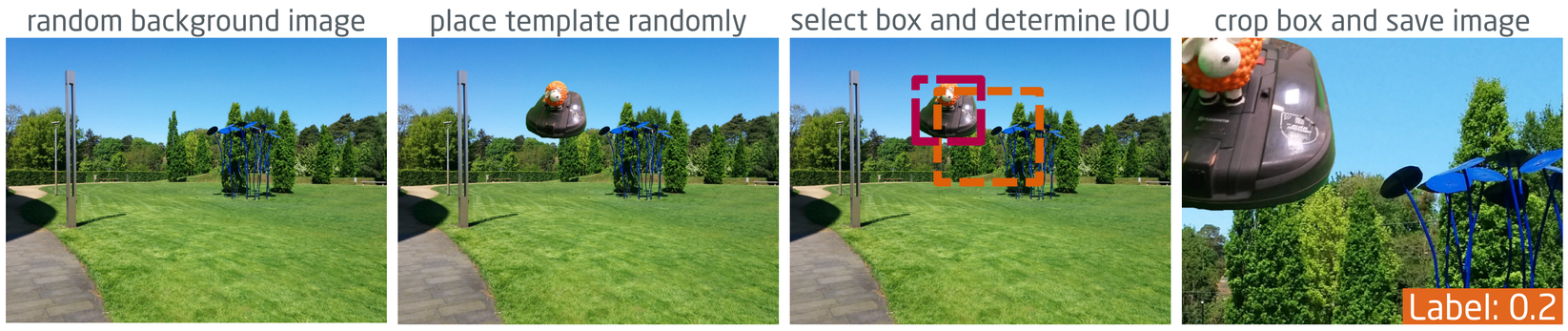
Assessor ("Teacher")

- predicts intersection over union (IOU) of image crop and bounding box of object
- trained in fully supervised setting on artificial data

Localizer ("Student")

- predicts region of interest that is likely to contain content that maximizes the output of the assessor
- trained under supervision of assessor and does not need any labeled images
- individual frames of a video can be used as training samples

Data Generation



Template Images

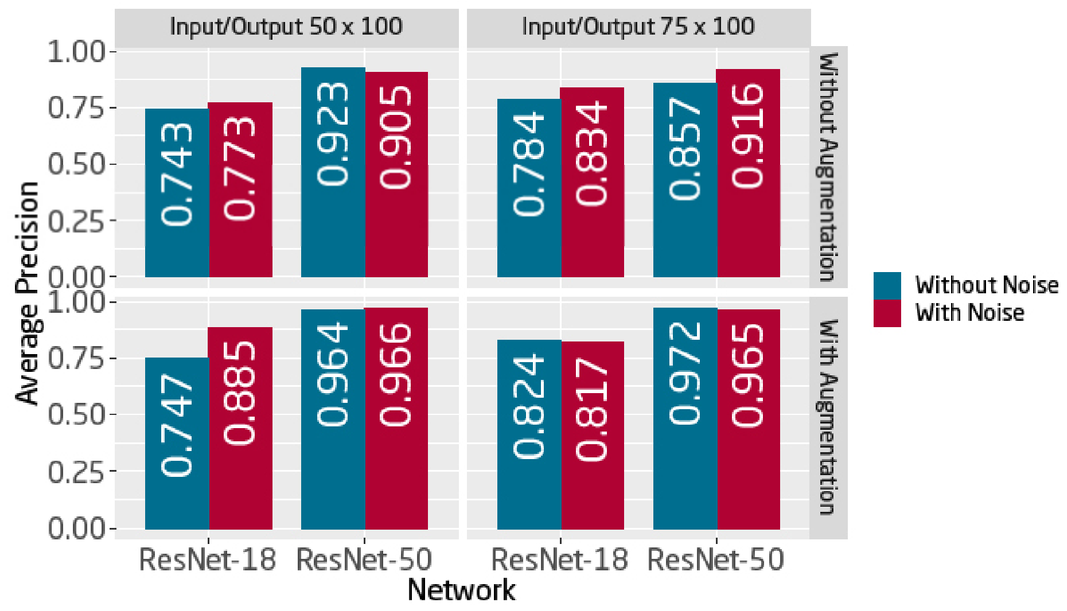
Experiments

- experimental results show on par performance to fully supervised system and robustness of the models, even if more than 50% of training data is noisy data

Results on Sheep Dataset

Method	224 x 244	300 x 300	512 x 512
SSD	-	0.887	0.969
Resnet-18	0.887	0.937	0.967
Resnet-50	**0.959**	**0.958**	**0.976**

Results on Figure Skating Dataset



Code, Models and Datasets

Get the code and all information you need to work with our models and datasets at:

https://github.com/Bartzi/loans

Authors

Christian Bartz, Haojin Yang, Joseph Bethge, Christoph Meinel
Multimedia Analysis and Deep Learning Team

Prof.-Dr.-Helmert-Str. 2-3 I D-14482 Potsdam | Germany
Tel.: +49331-5509228
E-Mail: christian.bartz@hpi.de
Web: https://bartzi.de/research/loans

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam