CORSMAL Challenge:
Audio-visual object classification for human-robot collaboration

Additional and related references

Multi-modal estimation of the properties of containers and their content: survey and evaluation

Details of the challenge as presented last year. Baselines with only audio as input are also provided for the classification of filling level and filling type.

[paper][code]

Improving filling level classification with adversarial training

[paper][webpage]

Audio Classification of the Content of Food Containers and Drinking Glasses

[paper][code][webpage]

See the glass half full: Reasoning about liquid containers, their volume and content.

[paper][webpage]

Seeing glass-ware: from edge detection to pose estimation and shape recovery

[paper]

ClearGrasp: 3D shape estimation of transparent objects for manipulation

[paper][code][webpage]

Learning to estimate pose and shape of hand-held objects from RGB images

[paper]

HOnnotate: A method for 3D Annotation of Hand and Object Poses

[paper][code][webpage]

Making sense of audio vibration for liquid height estimation in robotic pouring (PouringNet)

[paper][code][webpage]

Normalized object coordinate space for category-level 6d object pose and size estimation (NOCS)

[paper][code][webpage]

Mask R-CNN

Instance segmentation network

[paper][code]

SiamMask

Single object tracking using an object mask as reference

[paper][code][webpage]

ResNet-18 *(network available in PyTorch)*

[paper][code]

CORSMAL
Collaborative object recognition,
shared manipulation and learning