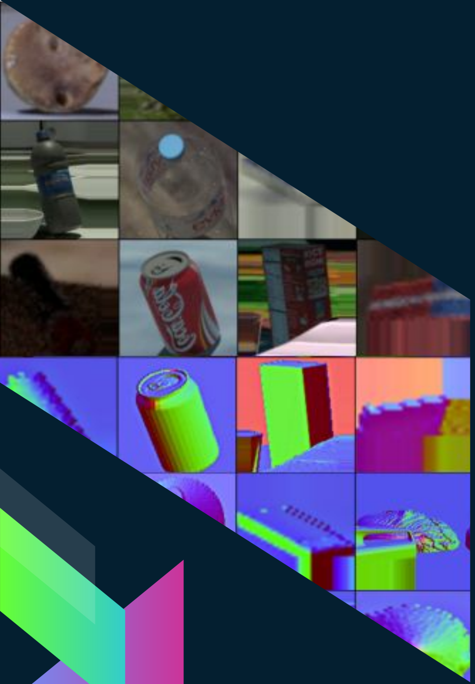
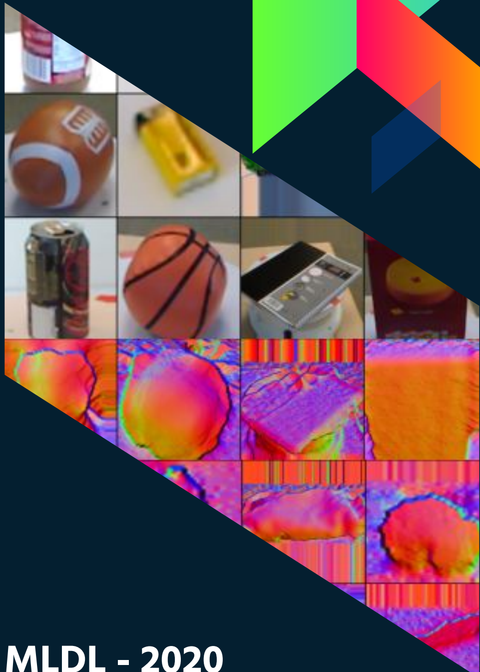


POLITECNICO
DI TORINO

Unsupervised Domain Adaptation through rotation as regression for RGB-D Object Recognition



Gullotto Marco
Giammarinaro Silvia

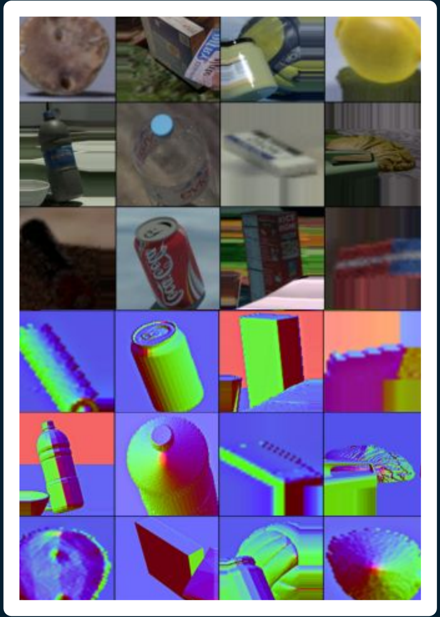
MLDL - 2020

ABOUT THE PROJECT

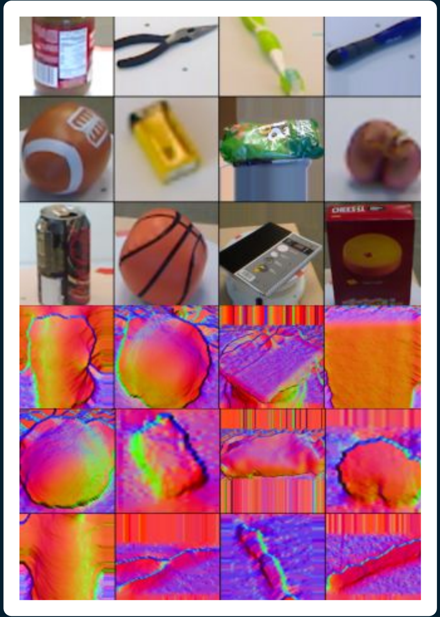
- What is Domain adaptation?
- Why are we using RGB-D images?
- What are the tasks?
- What is our goal?

Gullotto Marco
Giammarinaro Silvia

DATASET



SynROD as source domain

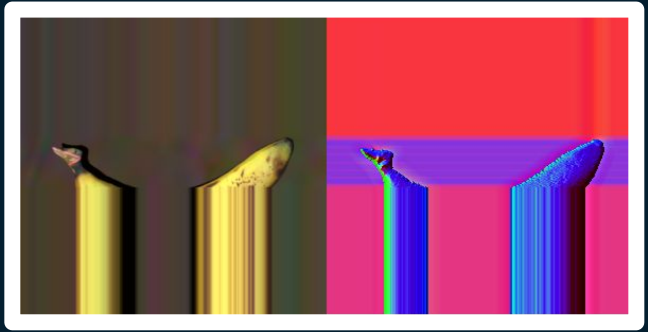


ROD as target domain

Gullotto Marco
Giammarinaro Silvia

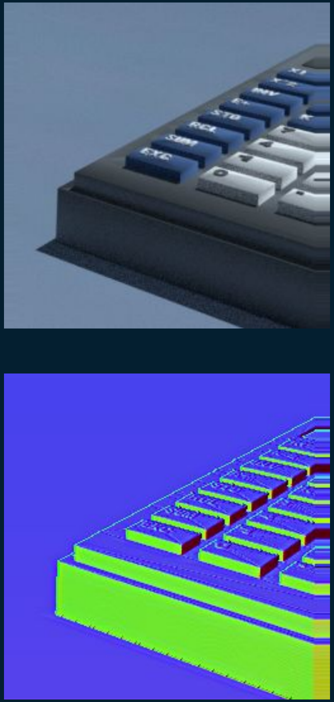
RELATED WORK

- Multy-modality CNN
- Image's stretch
- Color encoding for depth images
- Pretext task as artifact
- Loss entropy

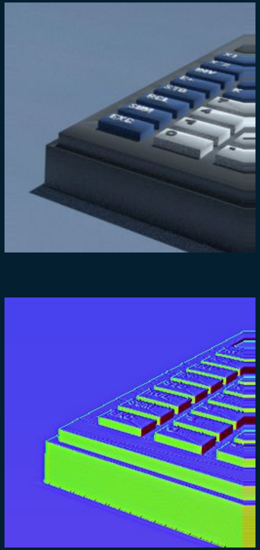


Gullotto Marco
Giammarinaro Silvia

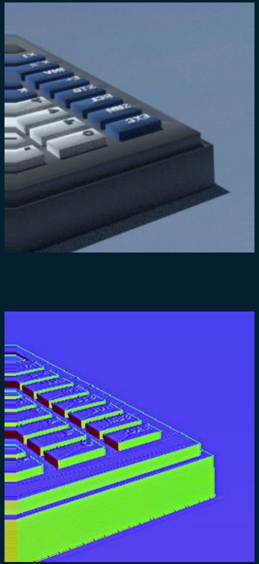
PREPROCESSING



Raw images



Resize to
256x256 px

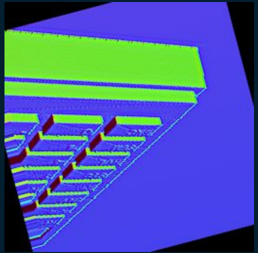


Horizontal flip with
probability p=0.5

Gullotto Marco
Giammarinaro Silvia

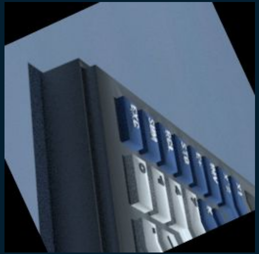
PREPROCESSING

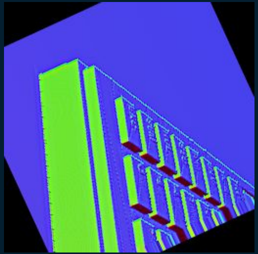
31.45°



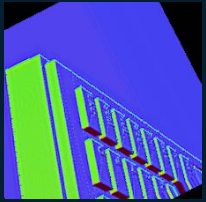
165.21°

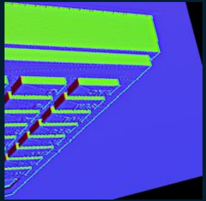
Relative rotation with 133.76°





Absolute rotation with 65.30°

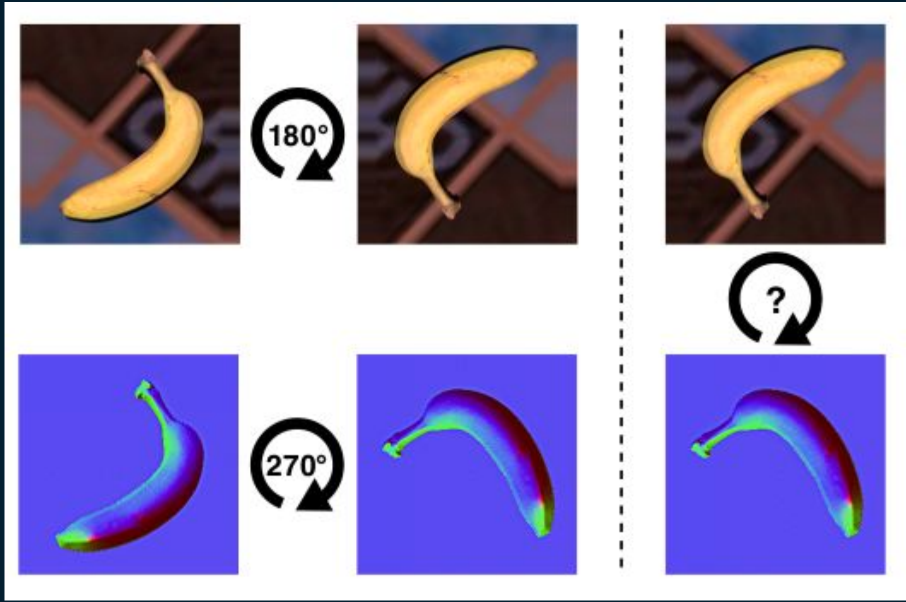




Center Crop to
224x224 px

Gullotto Marco
Giammarinaro Silvia

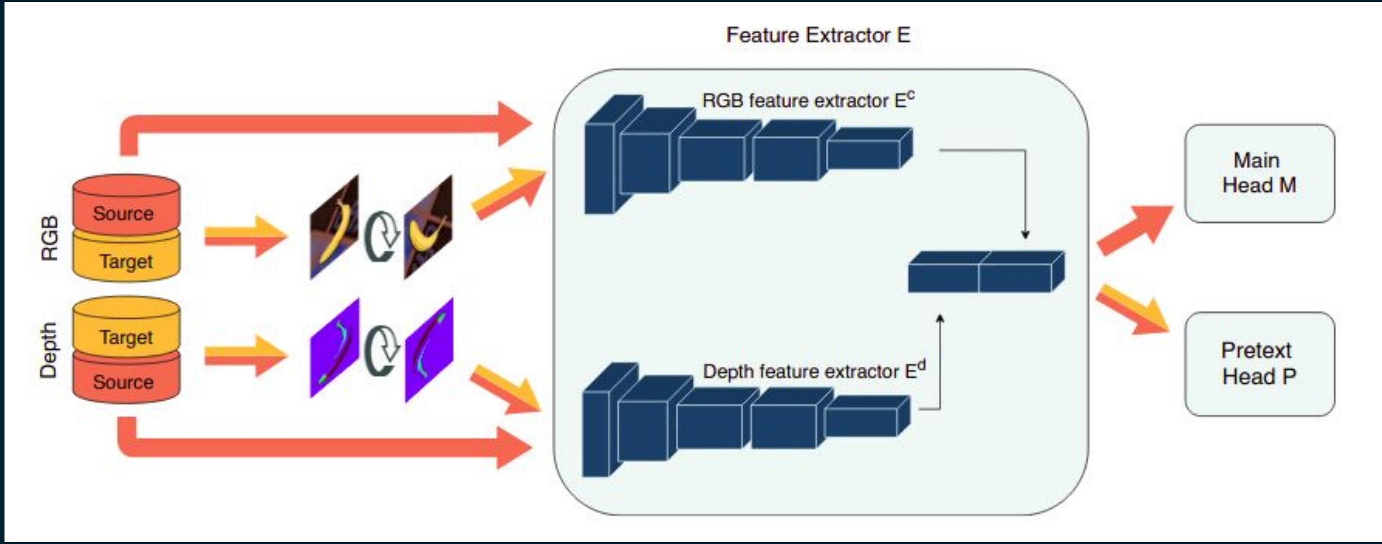
PRETEXT TASK



By how much should the RGB image be rotated to align with the depth image?

Gullotto Marco
Giammarinaro Silvia

NETWORK ARCHITECTURE



Gullotto Marco
Giammarinaro Silvia

OPTIMIZATION PROBLEM

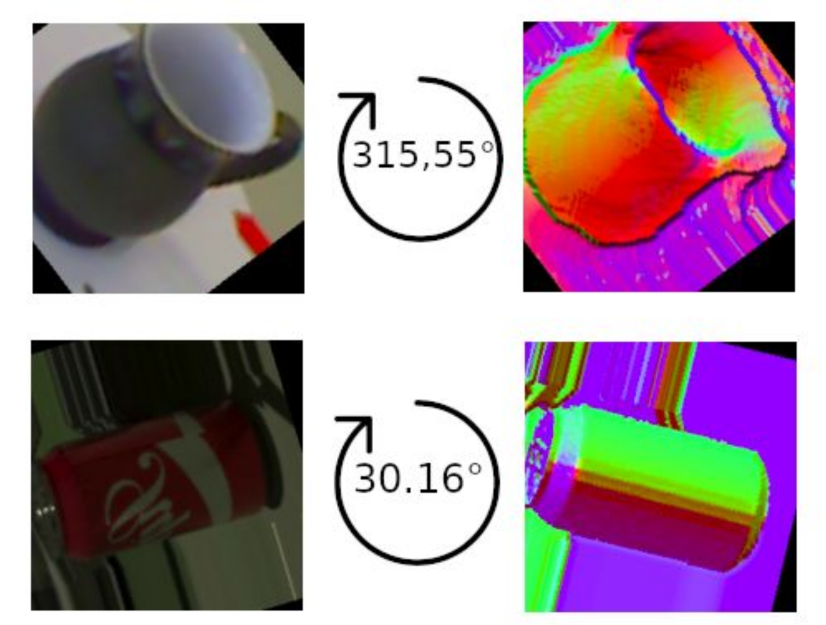
$$\mathcal{L} = \lambda_p \cdot \mathcal{L}_p + \mathcal{L}_M + \alpha \cdot \mathcal{L}_{en}$$

Cross-entropy pretext task

Cross-entropy main task

Loss entropy

OUR VARIATIONS



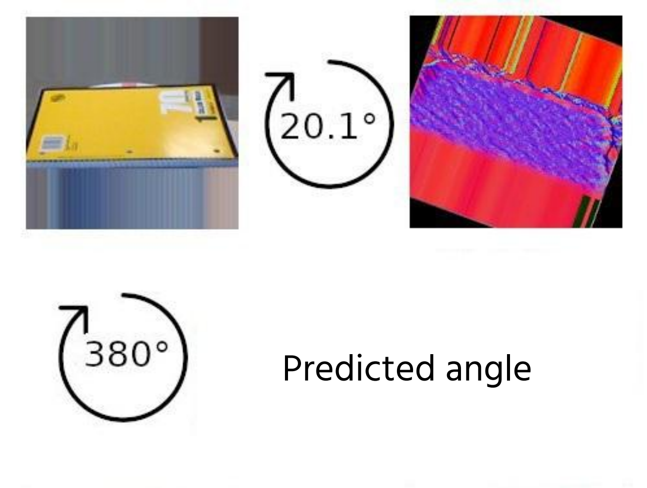
Absolute Rotation

Relative Rotation

Gullotto Marco
Giammarinaro Silvia

OUR VARIATIONS

Loss = (True angle - Predicted angle)^2



Loss = $(20.1° - 380°)^2 >> 0$

Gullotto Marco
Giammarinaro Silvia

ROTATION AS REGRESSION

- Using sine and cosine we can rewrite the problem as:

$$\vartheta \rightarrow (sin(\vartheta), cos(\vartheta))$$

- We can obtain the value of the angle in degrees using:

$$output \leftarrow atan2(sin_{predicted}, cos_{predicted})$$

- The new loss becomes:

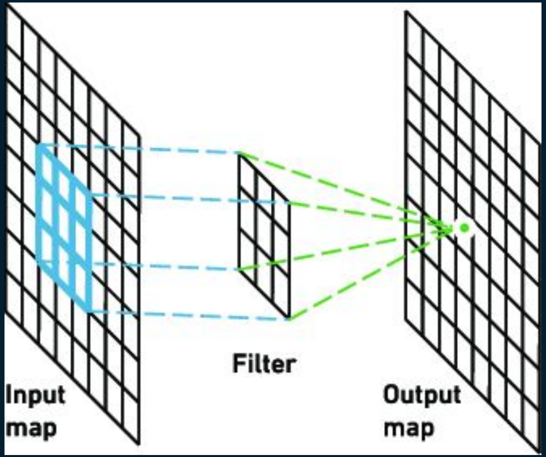
$$\mathcal{L}_{regr} = \frac{1}{2}((sin_{pred} - sin(y))^2 + (cos_{pred} - cos(y))^2)$$

Gullotto Marco
Giammarinaro Silvia

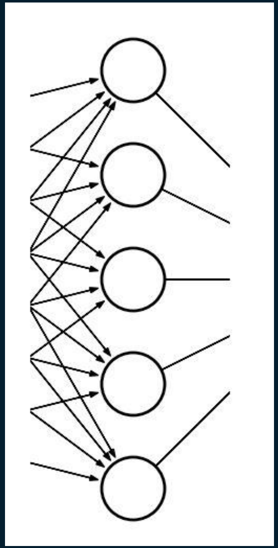
ROTATION AS REGRESSION

$$\mathcal{L} = \lambda_p(\underbrace{\mathcal{L}^S_{sin} + \mathcal{L}^S_{cos}}_{\text{MSE losses for the samples taken from the source domain}} + \underbrace{\mathcal{L}^T_{sin} + \mathcal{L}^T_{cos}}_{\text{MSE losses for the samples taken from the target domain}}) + \underbrace{\mathcal{L}_M}_{\text{Cross-entropy main task}} + \alpha \cdot \underbrace{\mathcal{L}_{en}}_{\text{Loss entropy}}$$

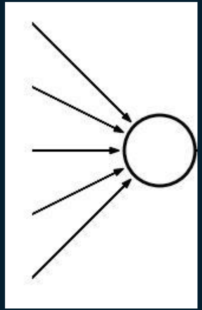
PRETEXT HEAD



2 convolutional layers



Fully connected with
100 neurons



Output neuron

Gullotto Marco
Giammarinaro Silvia

EXPERIMENTS

RGB
DEPTH

RGB-D e2e

Loghmani et al.
cross-modality
algorithm

Source only

Absolute and relative
rotation as regression

Gullotto Marco
Giammarinaro Silvia

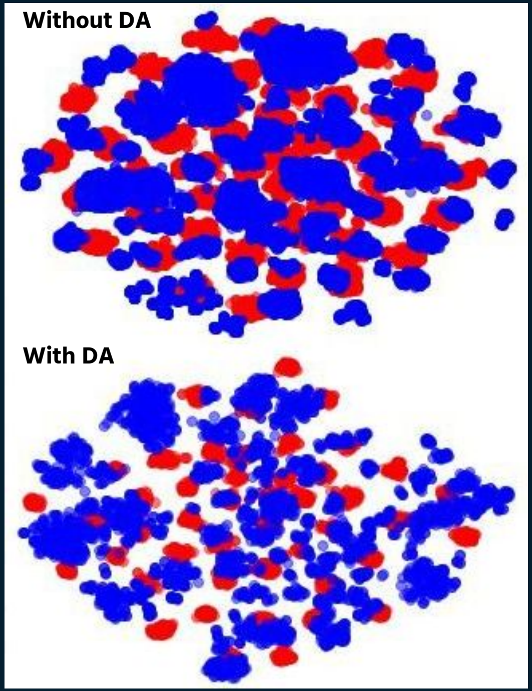
RESULTS

POLITECNICO
DI TORINO

Method	Modality	synROD $\rightarrow$ ROD
Source only	RGB Depth RGB-D e2e	51.4% 13.1% 47.9%
Loghmani et al. impl.	RGB-D	57.6% ± 1.2%
Abs. rotation as regression	RGB-D	55% ± 0.9%
Rel. rotation as regression	RGB-D	59% ± 0.8%

CONCLUSIONS

- t-SNE algorithm
- Other trials
  - Different color encoding
  - Noise patterns
  - ARID dataset



Gullotto Marco
Giammarinaro Silvia

THANKS FOR YOUR ATTENTION!

Gullotto Marco
Giammarinaro Silvia

Template by SlidesCarnival

MLDL - 2020