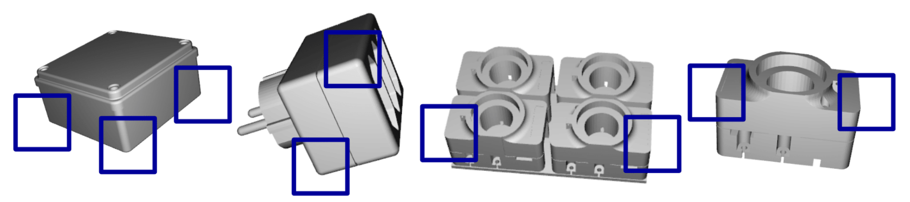
CorNet: Generic 3D Corners for 6D Pose Estimation of New Objects without Retraining

Giorgia Pitteri[1] Slobodan Ilic[3,4] Vincent Lepetit[2]

[1] Laboratoire Bordelais de Recherche Informatique, Université de Bordeaux , [2] LIGM (UMR 8049), École des Ponts, UPE, [3] Technische Universität München, [4] Siemens AG Corporate Technology

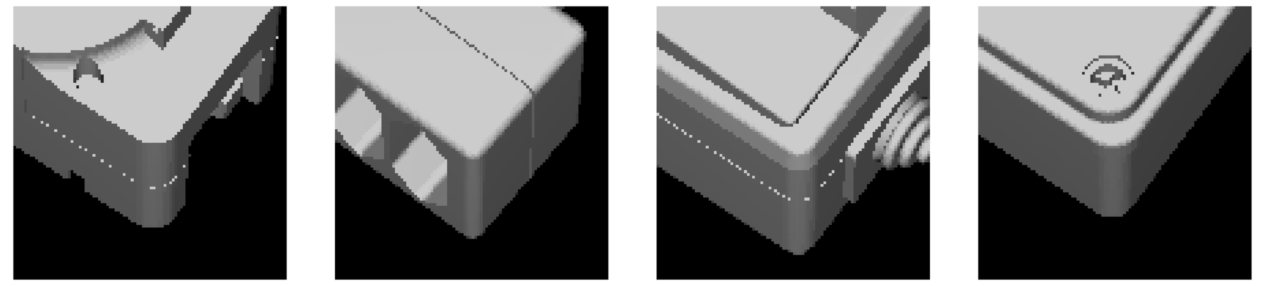
Introduction

- **Motivation:** 3D pose estimation from color images state-of-the-art methods rely on the power of Deep Learning however:
  - these methods have to be retrained for each new object
  - many different images of these new objects need to be available
  - domain transfer methods for training with synthetic data can be used but they take time
- **Goal & Contribution:**
  - 3D pose estimation of new objects **without additional learning** nor training images
- **Idea:**
  - Industrial objects often share common parts ⇒ Estimate the object's 3D pose by learning to estimate the pose of its 3D parts
  - Large number of industrial objects have prominent corners. Let's focus on corners!

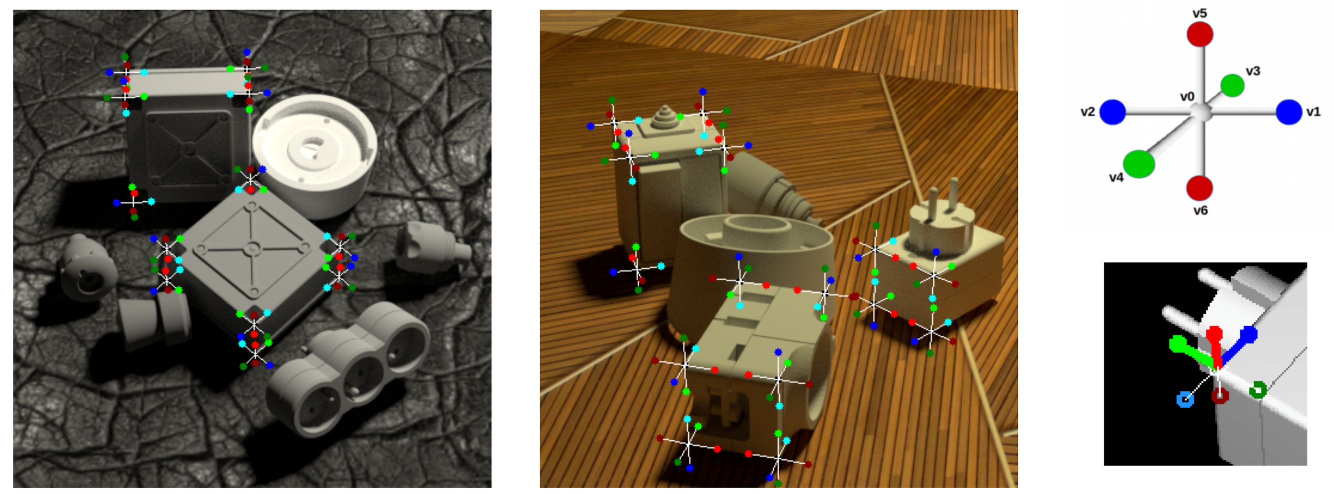


Training data for corner detection

- Considering generic corners of various shapes and appearances to be able to deal with new objects without retraining

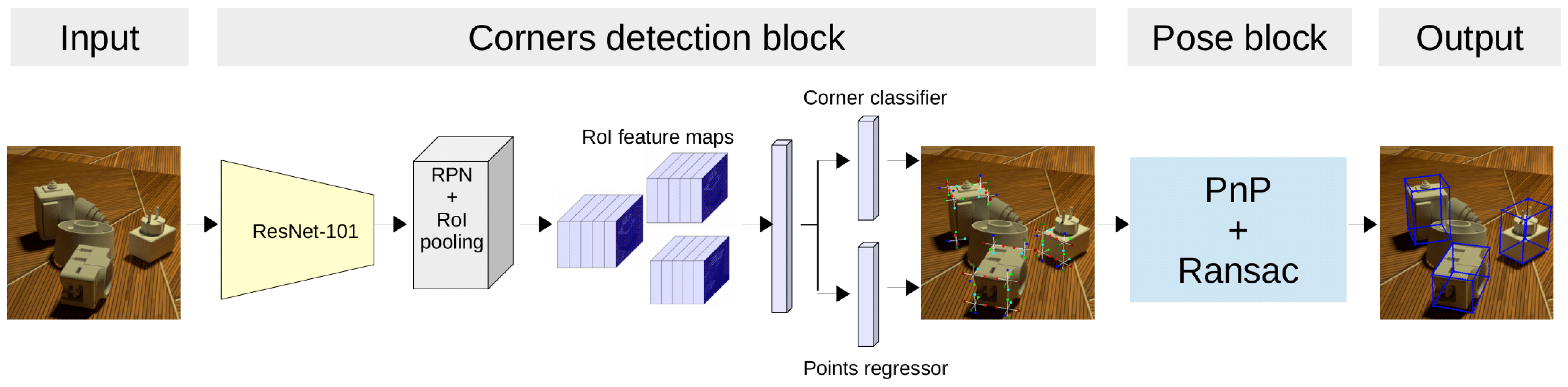


- Representing the 3D pose of a corner in terms of the 2D reprojections of a set of 3D virtual points. It makes it easy to compute the 3D pose of an object from several corners
- Examples of synthetic training images generated by rendering objects from the T-LESS dataset with annotated 2D reprojections of virtual points.



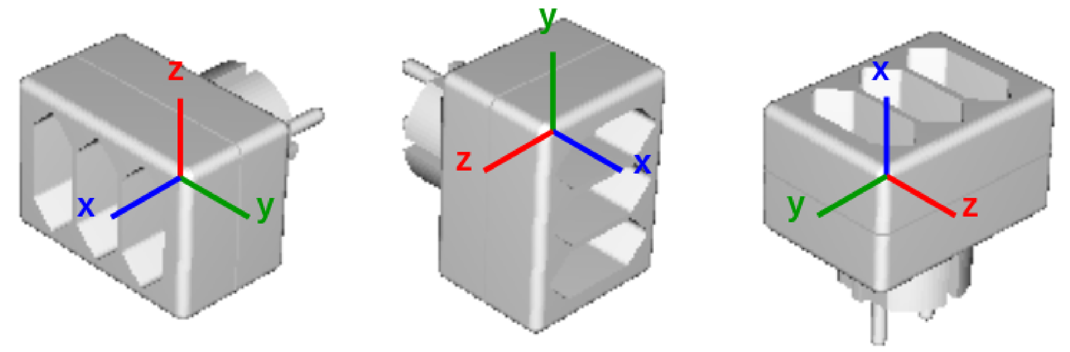
Method

- Corners detection and 2D reprojections of 3D virtual points prediction by adding specific branches to Faster R-CNN
- RANSAC-like pose estimation computation by solving a PnP agorithm from 2D-3D correspondences



Dealing with Corner Pose Ambiguities

- The same corner can look identical under different 3D poses. This implies that it is possible to predict the 3D pose of a corner up to some rigid motion ⇒ three possible arrangements of 3D virtual points

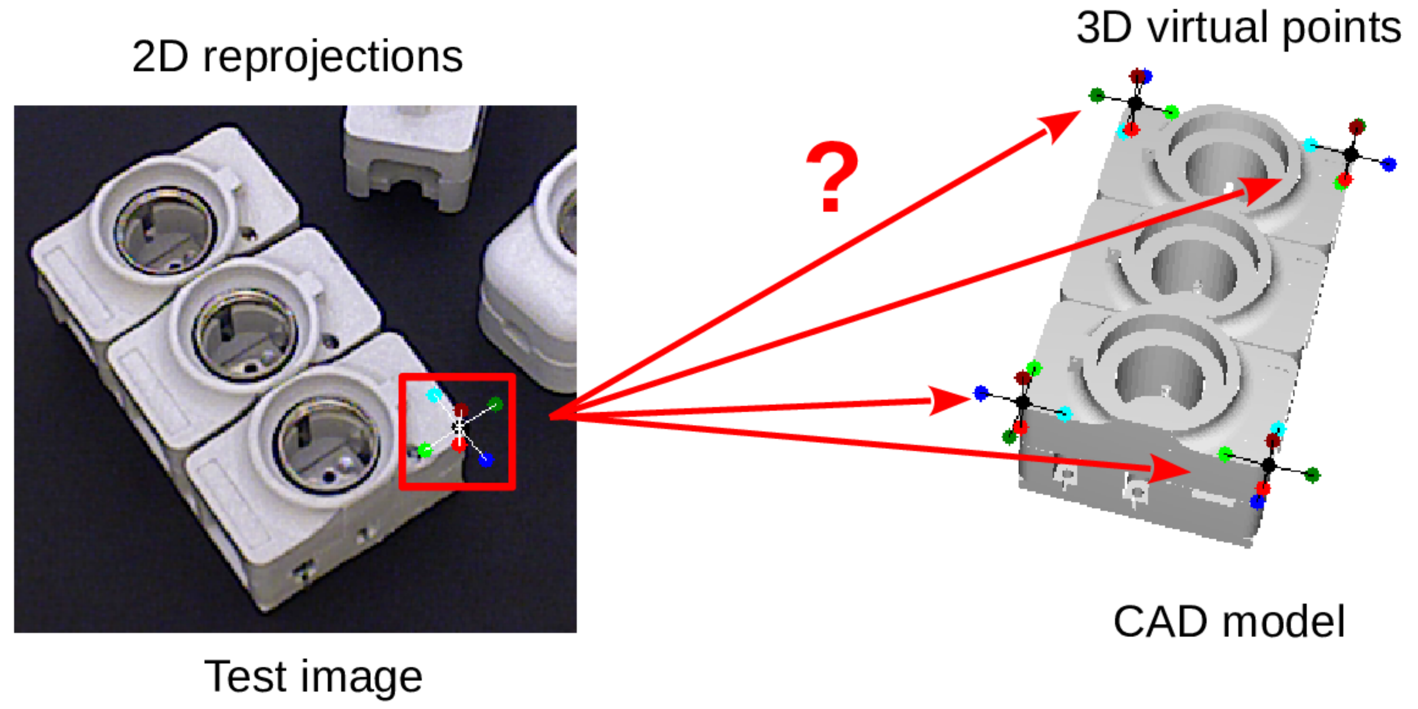


- Given one possible 3D pose $\mathbf{p}$ it is possible to generate the other poses by permuting properly the 2D reprojections of the virtual points.
- Introduction of two permutations Σ_1 and Σ_2 to permute the 2D reprojections of the 3D virtual points and to generate the other two possible poses from the estimated pose $\mathbf{p}$.
  - **Inference**: application of Σ_1 and Σ_2 to the estimated pose
  - **Training**: optimization done on the following cost function

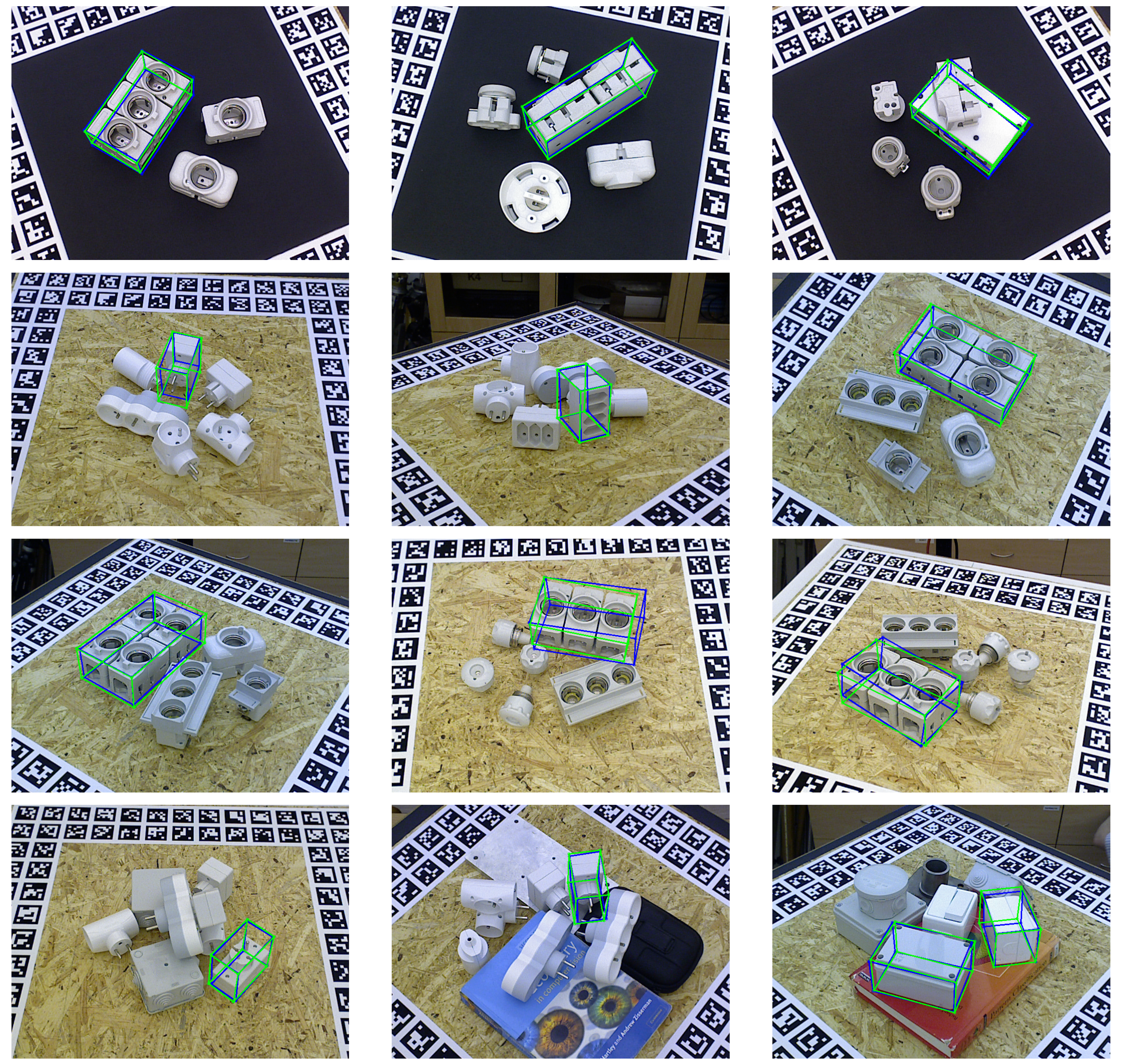
$$\boxed{loss = \min_{\Sigma \in \{I_d, \Sigma_1, \Sigma_2\}} \| P^{gt} - \Sigma(f_\theta(I)) \|}$$

Pose estimation algorithm

- 3D object pose estimation by matching the corners on the CAD model with their detected counterparts in the image by solving a PnP algorithm
- Application of the two permutations Σ_1 and Σ_2 to generate the other two possible poses to handle ambiguities
- RANSAC-like algorithm based on a **similarity score** to find the best pose among all the possible esimated poses
- **Similarity score**: cross correlation between the gradients of the input image and the image gradients of the CAD model rendered under the 3D pose hypothesys.



Results



Scene: Obj	$AD\{D\|I\}_{10\%}$	$AD\{D\|I\}_{20\%}$	$AD\{D\|I\}_{30\%}$	detection [%]
02: 7	68.3	80.1	83.7	67.3
03: 8	57.9	72.5	78.7	76.3
04: 26	28.1	47.2	56.2	48.3
04: 8	21.2	53.0	68.2	35.7
06: 7	36.8	61.7	78.7	73.7
08: 20	10.0	40.4	56.1	34.1
10: 20	27.8	47.2	58.3	30.0
11: 8	58.8	74.9	85.3	74.3
12: 7	23.1	44.6	47.7	54.6
13: 20	26.6	57.3	69.0	52.9
15: 29	48.0	59.1	76.7	38.3
14: 20	10.0	24.6	31.6	44.0
Average	34.7(±18.5)	55.2(±15.2)	65.9(±15.6)	52.5(±16.2)

References

[1] A. Crivellaro, M. Rad, Y. Verdie, K. Moo Yi, P. Fua and V. Lepetit: *Robust 3D Object Tracking from Monocular Images Using Stable Parts*, PAMI, 2017

[2] M. Rad and V. Lepetit: *BB8: A Scalable, Accurate, Robust to Partial Occlusion Method for Predicting the 3D Poses of Challenging Objects Without Using Depth*, ICCV, 2017

[3] T. Hodaň, P. Haluza, S. Obdržaĺek, J. Matas, M. Lourakis and X. Zabulis: *T-LESS: An RGB-D Dataset for 6D Pose Estimation of Texture-less Objects*, IEEE Winter Conference on Applications of Computer Vision (WACV), 2017

Acknowledgements

This research work was funded by Siemens AG, Corporate Technology Research & Technology Center, Munich.