JUNE 18-22, 2023
CVPR VANCOUVER, CANADA
VISION Workshop

Parcel3D: Shape Reconstruction from Single RGB Images for Applications in Transportation Logistics

Alexander Naumann, Felix Hertlein, Laura Dörr and Kai Furmans
FZI Research Center for Information Technology and Karlsruhe Institute of Technology (KIT)
anaumann@fzi.de - hertlein@fzi.de

Karlsruher Institut für Technologie

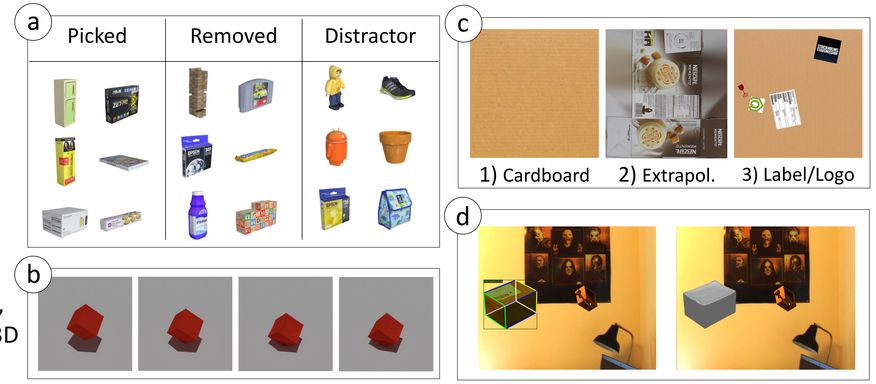
FZI

1) Introduction

- **Motivation**
  - Increasing amounts of valuable goods in postal supply chains, transported using parcels → need for automated parcel monitoring w.r.t. damage and tampering
  - Multisensory setups are not always feasible (e.g. last-mile delivery) → focus on single RGB images to enable maximum flexibility
- **Contributions**: Novel synthetic dataset and targeted neural network architecture
- **Related work**: see [1] and online at a-nau.github.io/cv-in-logistics

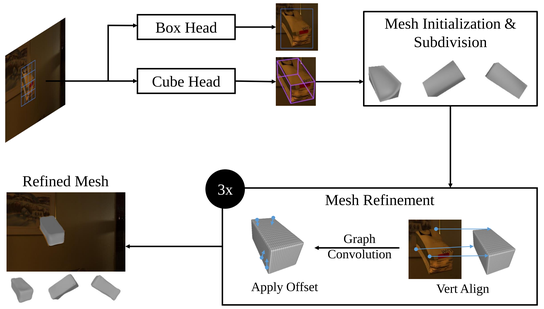
2) Synthetic Dataset Parcel3D

- >13,000 images of damaged and intact parcels
- **Model Selection**
  - (a) Automated filtering of GSO dataset [2] and
  - (b) physics-based simulation for deformations
- **Texture Generation**
  - (c.1) Generate synthetic parcel textures and
  - (c.2) extrapolate existing textures,
  - (c.3) both with additional labels and logos
- **Environment** (d): Realistic scenes & lighting conditions [3]
- **Annotations** (d): bounding box, class label, segmentation mask, parcel corner points, full 3D mesh with rotation & translation, 3D bounding box of pristine shape



3) CubeRefine R-CNN Architecture

- Extends Cube R-CNN [4] by an iterative mesh refinement [5]
- Estimates 3D bounding box of pristine parcel and potentially deformed current shape simultaneously
- Enables detailed damage assessment by comparing two 3D meshes
- 3D bounding box facilitates tampering detection



4) Evaluation & Conclusion

Model	Dataset	Box AP	Mesh AP_{50}	Mesh AP_{75}	Chamfer Distance (↓)	Normal Consistency
Pix2Mesh [5]	Intact	96.0 (0.5)	89.6 (1.0)	48.5 (1.6)	0.311 (0.086)	0.901 (0.001)
Mesh R-CNN [6]	Intact	95.9 (0.5)	**92.9 (1.6)**	67.0 (2.8)	0.225 (0.088)	0.914 (0.001)
Cube R-CNN [4]	Intact	97.1 (0.1)	92.0 (0.3)	74.4 (2.0)	0.159 (0.016)	0.925 (0.001)
CubeRefine R-CNN (ours)	Intact	**97.1 (0.0)**	92.8 (0.2)	**77.2 (1.2)**	**0.128 (0.002)**	**0.929 (0.001)**
Pix2Mesh [5]	Damaged	95.1 (0.6)	84.3 (1.2)	12.4 (1.4)	0.750 (0.553)	0.866 (0.002)
Mesh R-CNN [6]	Damaged	94.6 (0.5)	**91.1 (0.5)**	**26.1 (1.9)**	0.860 (0.436)	**0.880 (0.002)**
Cube R-CNN [4]	Damaged	95.0 (0.2)	32.6 (0.5)	0.1 (0.0)	0.494 (0.004)	0.806 (0.000)
CubeRefine R-CNN (ours)	Damaged	**95.2 (0.1)**	70.7 (0.7)	4.1 (0.2)	**0.293 (0.003)**	0.861 (0.000)
Pix2Mesh [5]	Real	**74.4 (1.9)**	27.8 (2.1)	2.3 (0.6)	2.112 (0.060)	0.744 (0.006)
Mesh R-CNN [6]	Real	70.6 (5.0)	29.4 (2.7)	4.9 (1.5)	2.153 (0.073)	0.742 (0.008)
Cube R-CNN [4]	Real	43.4 (6.9)	30.1 (5.8)	**13.3 (4.3)**	0.875 (0.041)	0.808 (0.003)
CubeRefine R-CNN (ours)	Real	41.5 (5.8)	**32.3 (4.2)**	13.1 (3.0)	**0.814 (0.062)**	**0.828 (0.006)**

Table: Quantitative performance analysis of mesh reconstruction when trained on Parcel3D and evaluated on different datasets. We repeated all trainings five times and report mean values with standard deviations in parentheses.

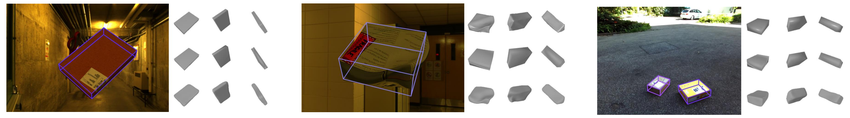


Figure: Exemplary qualitative results. We show the input image with the projected 3D bounding box on the left, and a 3×3 grid of mesh reconstructions on the right. Each column shows a different viewing angle, and the rows contain ground truth, 3D bounding box and refined mesh, respectively.

- **Limitations**: Only deformation damages; parcel damage != freight damage; diverse real-world deformations are challenging
- **Conclusion**: Parcel3D allows transfer to real-world data; CubeRefine R-CNN performs competitively and is the only method enabling 3D damage quantification

[1] Naumann et al., "Literature Review: Computer Vision Applications in Transportation Logistics and Warehousing", arXiv:2304.06009
[2] Downs et al., "Google Scanned Objects: A High-Quality Dataset of 3D Scanned Household Items", ICRA 2022
[3] Gardner et al., "Learning to predict indoor illumination from a single image", ACM Transactions on Graphics 2017
[4] Brazil et al., "Omni3D: A Large Benchmark and Model for 3D Object Detection in the Wild", CVPR 2023
[5] Wang et al., "Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images", ECCV 2018
[6] Gkioxari et al., "Mesh R-CNN", ICCV 2019

Physically present at CVPR