

Agenda

Problem Setting

Copy & Paste Data Augmentation

Classification

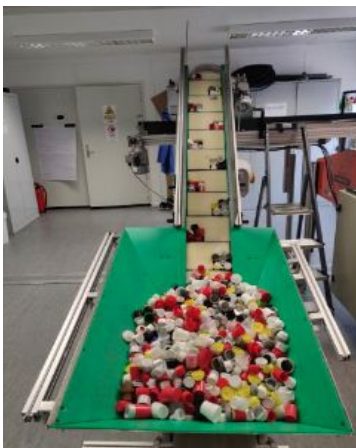
Instance Segmentation

Summary

Problem Setting

Task

- Setting: plastic lids for recycling
- Task: object counting - only one lid allowed
- Given: small unlabeled dataset
 - 2751 images, 0-5 lids in trays



Recycling Production Line

Dataset: Images



Label: 0



Label: 1



Label: 2



Label: 3



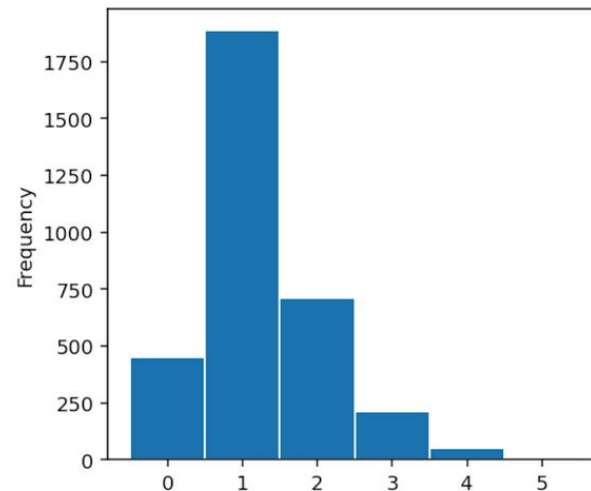
Label: 4



Label: 5

Dataset: Labeling

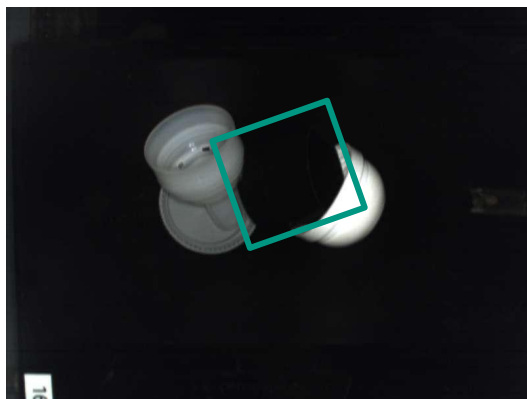
- Image-level labels for dataset
- Problems:
 - Small amount of data
 - Class Imbalance
 - Hard cases:
 - Different Colors - 25.7%
 - Overlapping - 22.5%
 - Transparent - 16.1%
 - Inside - 6.9%
 - Dark Color - 5.3%
 - One Color - 3.9%
 - Open Lid - 0.6%
 - Edge - 0.5%



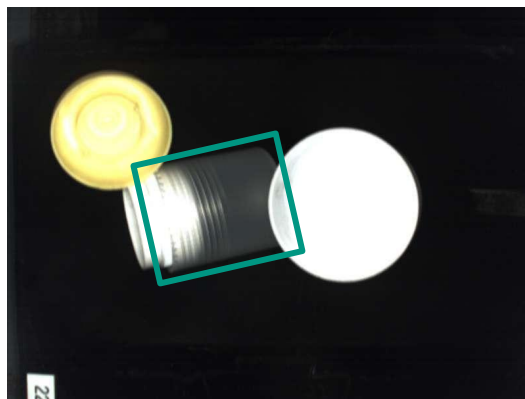
Frequency of Labels 0-5 in Original Dataset

Dataset: Analysis

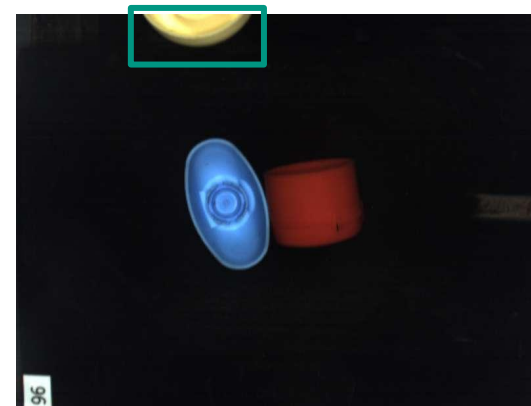
- Focus on: overlapping, edge, dark, transparent
- Target and tags stratified train/test split



Overlapping and Dark
Label: 4



Overlapping and Transparent
Label: 4



Edge
Label: 3

Copy & Paste Data Augmentation

Copy & Paste Data Augmentation - Approach

- Inspired by [Dwibedi17, Ghiasi20]
- Synthetic images of label 0: add gaussian noise
- Synthetic images of label 1-5:
 - Background: empty tray
 - Objects: from images of label 1
 - Positions of object based on heatmap
- Copy & Paste
 - Binary masks of lid
 - Lid with black background
 - Bitwise_or of background and mask
 - Add lid with black background



Binary Mask of Object

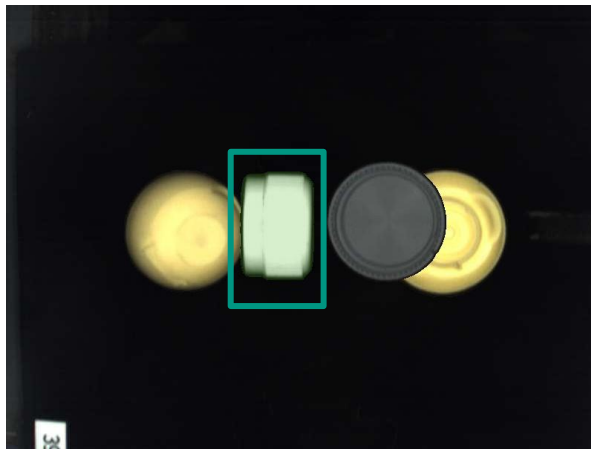


Object on black background

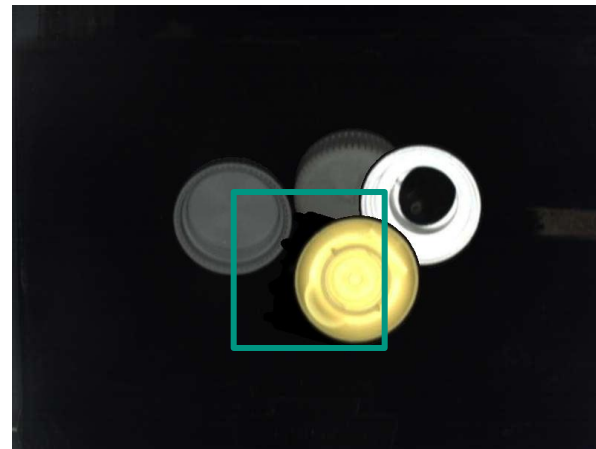
Copy & Paste Data Augmentation - Tags



Rotate

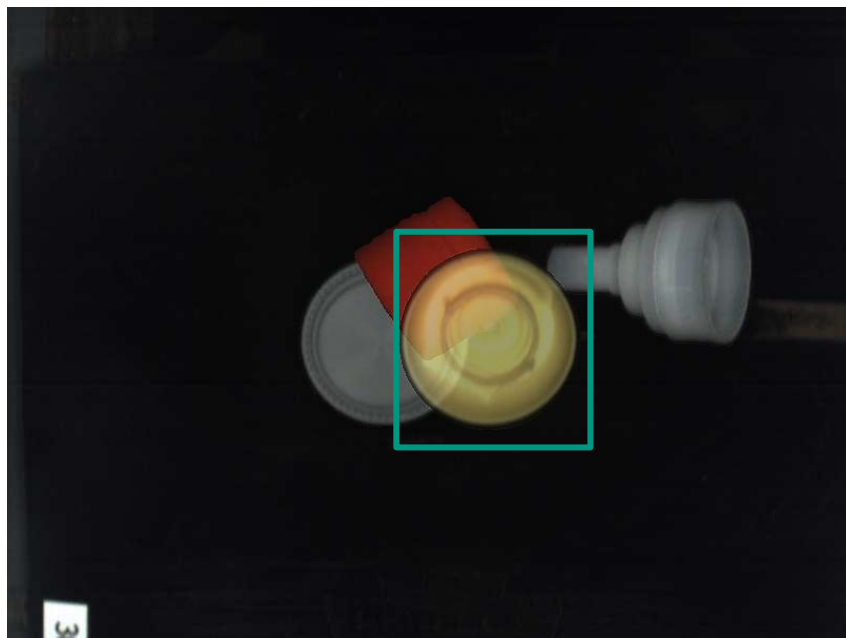


Color

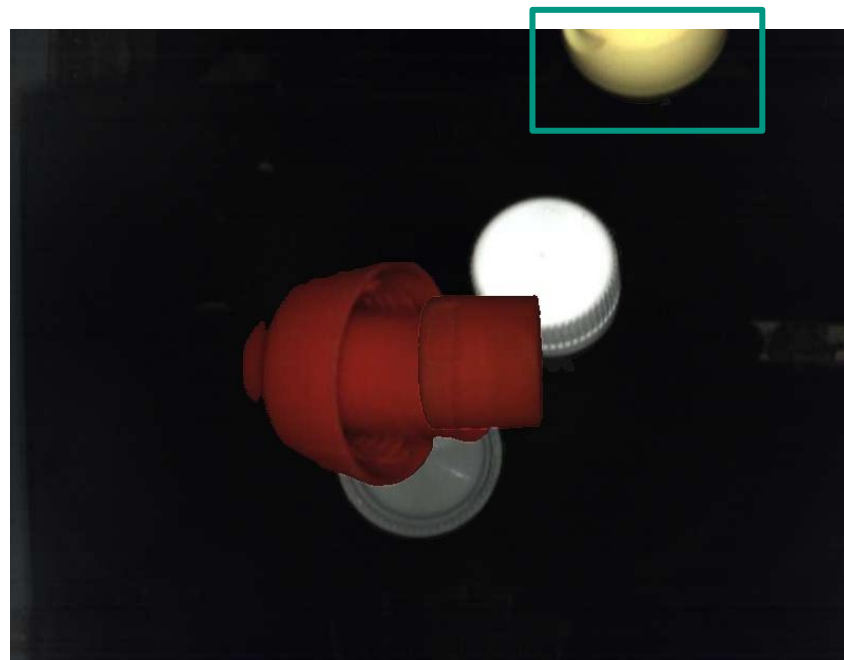


Dark

Copy & Paste Data Augmentation - Tags



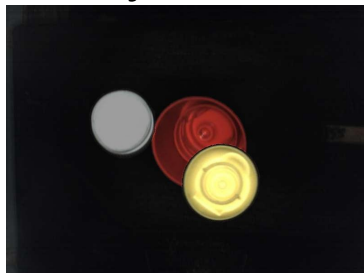
Transparent



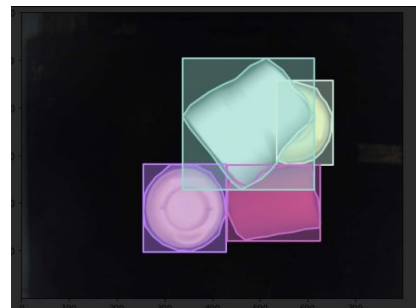
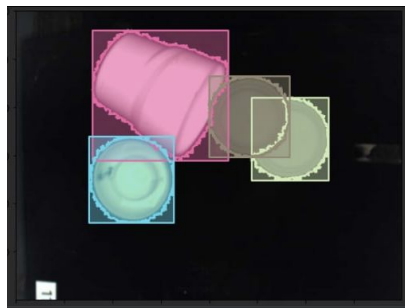
Edge

Copy & Paste Data Augmentation - Annotation

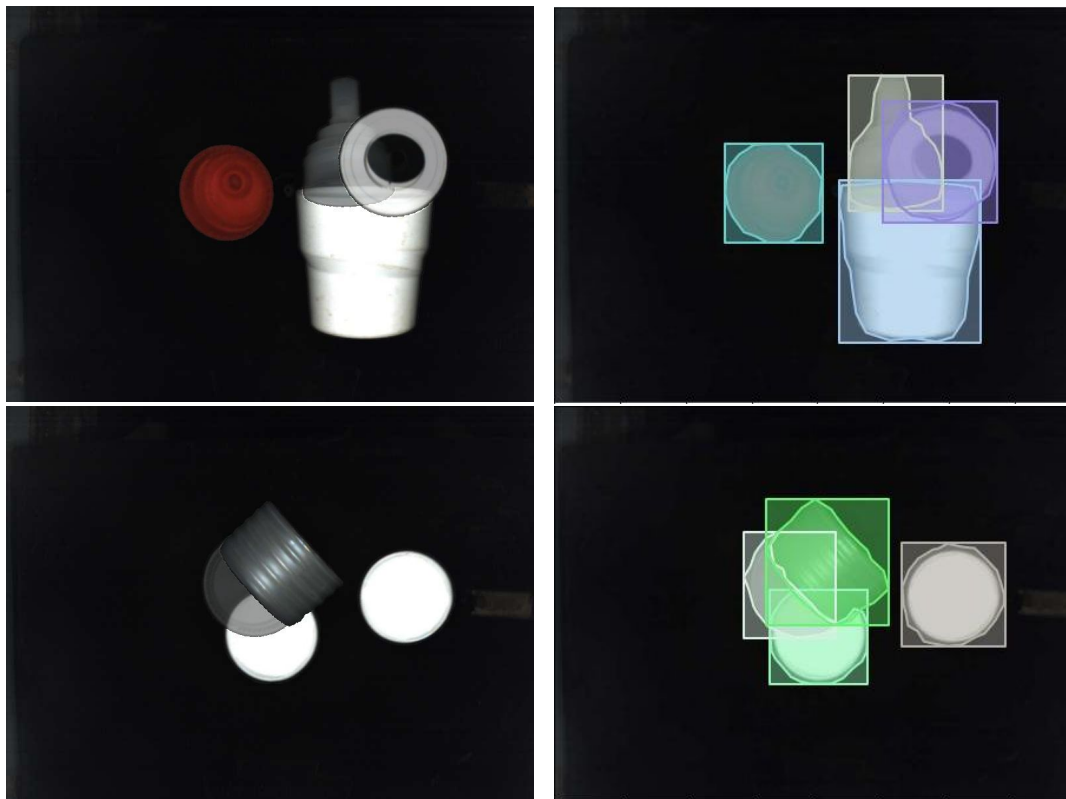
- Binary mask of each lid in synthetic image



- Smoothed contour of masks with spline interpolation



Copy & Paste Data Augmentation - Annotation

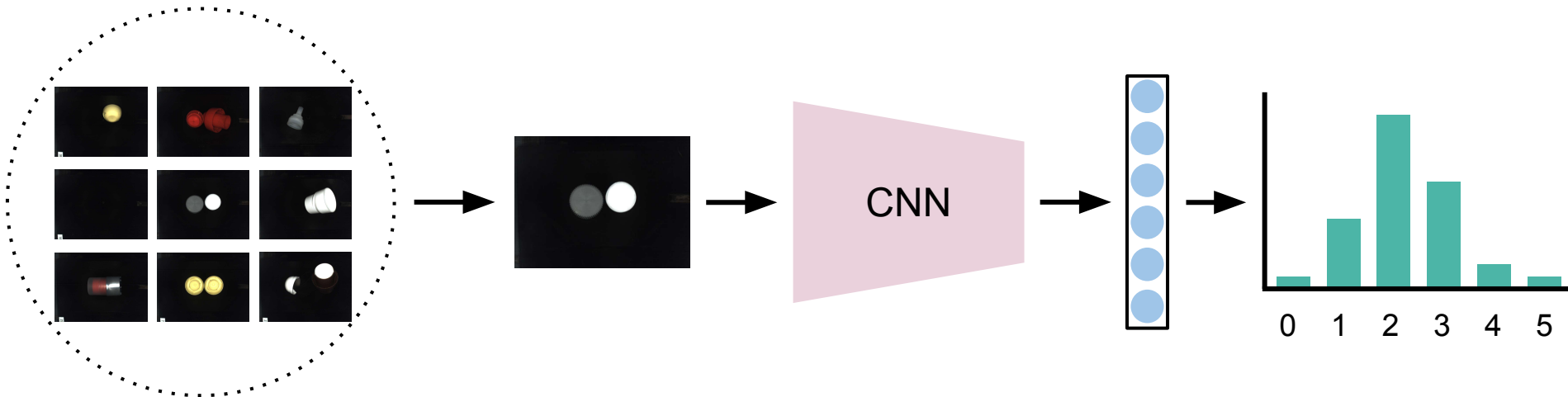


Synthesized Images and Their COCO Annotations

Classification

Classification: Idea

Approach: Salient Object Subitizing (SOS) [Zhang15, Zhang16]



Classification: Settings

Implementation Set-Up:

- ResNet18
- Cross Entropy Loss with Weights
- Stratified Batch Sampling
- Stochastic Gradient Descent with Learning Rate 0.001
- Augmentation – Random Horizontal Flip

Metric:

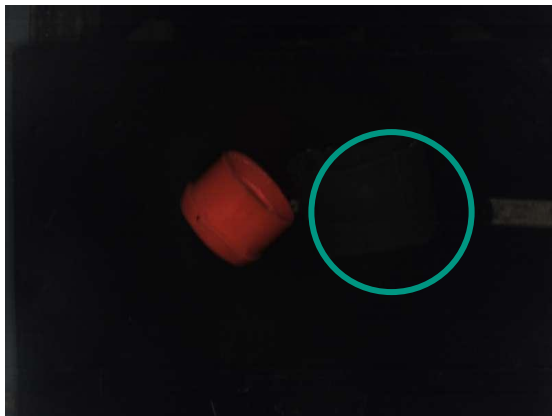
- Decision Accuracy =
$$\frac{\textit{Accepted Trays} + \textit{Rejected Trays}}{\textit{All Trays}}$$

Classification: Results

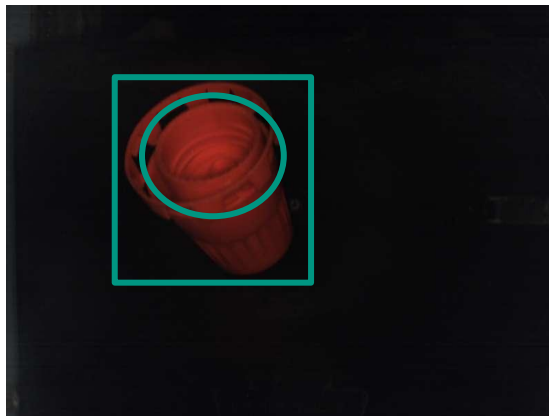
Name	Accuracy	Decision Accuracy
SOS	0.66 (± 0.02)	0.71 (± 0.02)
GoogLeNet	0.90 (± 0.01)	0.95 (± 0.00)
ResNet18	0.92 (± 0.01)	0.97 (± 0.00)
ResNet18_weights	0.93 (± 0.01)	0.97 (± 0.00)
ResNet18_synthetic	0.95 (± 0.01)	0.98 (± 0.00)

Evaluation of Classification Approaches

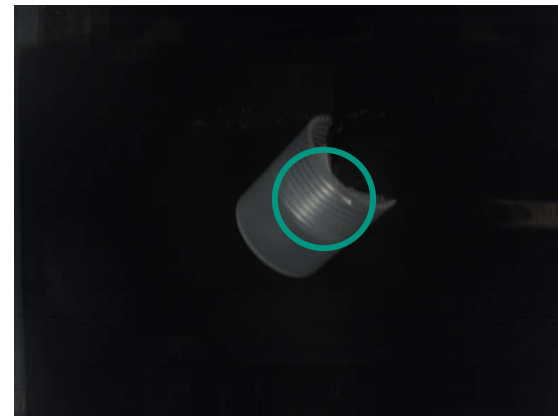
Classification: Misclassified Images



Predicted Label: 1
Correct Label: 2



Predicted Label: 1
Correct Label: 2



Predicted Label: 2
Correct Label: 3

Instance Segmentation

Instance Segmentation

- Train: fully synthesized data
- Test:
 - Separate fully synthesized test dataset for COCO evaluation
 - The same real-world test dataset used for classification
- Model:
 - Mask R-CNN (ResNet50 + FPN)
 - SGD, 5000 iterations, annealing LR 0.001 with steps (3000, 4000)

Name	AP	AP50	AP75
Seg_all	80.25	96.85	88.33

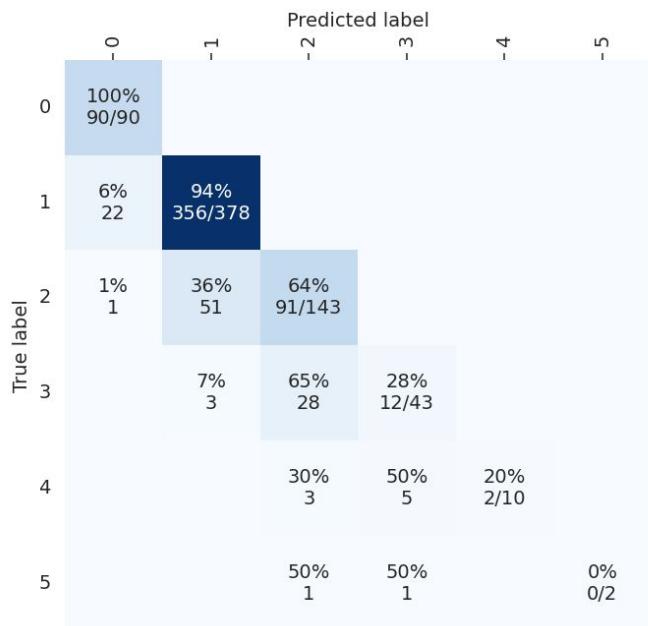
COCO Evaluation of Segmentation Model on Separate Synthesized Test Dataset

Segmentation: Ablation of Augmentation

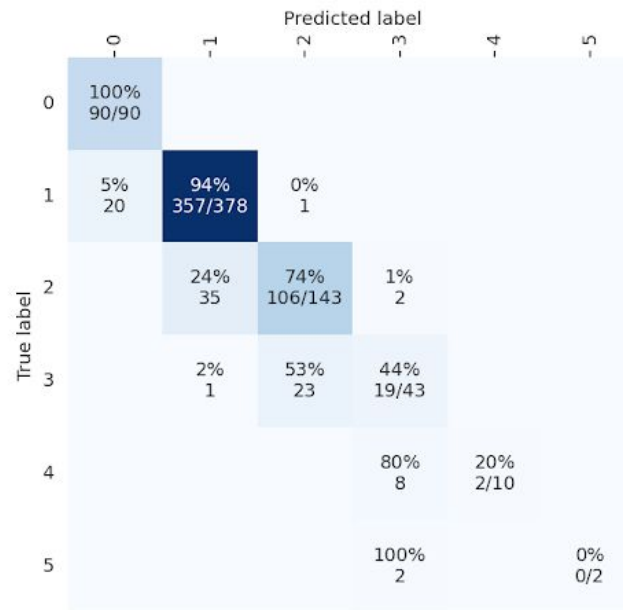
Name	Rotate	Color	Edge	Dark	Transparent	Accuracy	Decision Accuracy
Seg_no	x	x	x	x	x	0.83	0.89
Seg_edge	+	+	+	x	x	0.83	0.89
Seg_edge_dark	+	+	+	+	x	0.85	0.90
Seg_all	+	+	+	+	+	0.86	0.91

Ablation Study of Different Augmentations for Segmentation Model on Real-world Test Dataset

Segmentation: Ablation of Augmentation



Confusion Matrix of Seg_no Model
(without Any Transformation)



Confusion Matrix of Seg_all Model
(with All Transformations)

Segmentation vs. Classification

Name	Label 0	Label 1	Label 2	Label 3	Label 4	Label 5	Accuracy	Decision Accuracy
Seg_all	1.00	0.94	0.74	0.44	0.20	0.00	0.86	0.91
Classification_best	1.00	0.99	0.87	0.70	0.70	0.50	0.95	0.98

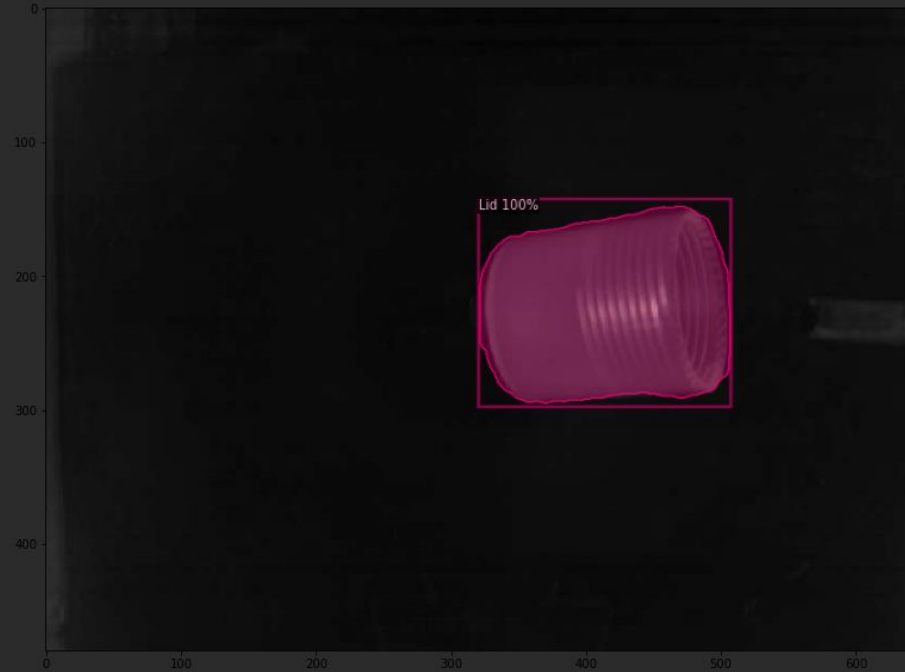
Comparison of the Best Segmentation (Seg_all) and Best Classification Approach (Classification_best)

Segmentation: Qualitative Results

LABEL = 0 PREDICTED = 0 FILENAME = presorted/trial (1531).jpg

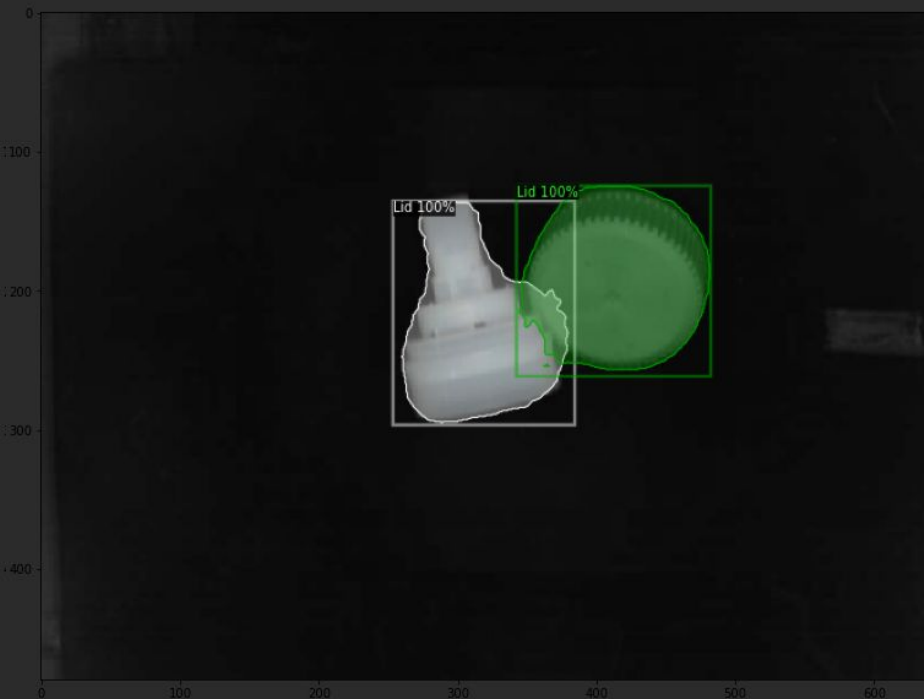


LABEL = 1 PREDICTED = 1 FILENAME = presorted/trial (1379).jpg



Segmentation: Qualitative Results

LABEL = 2 PREDICTED = 2 FILENAME = presorted/trial (973).jpg

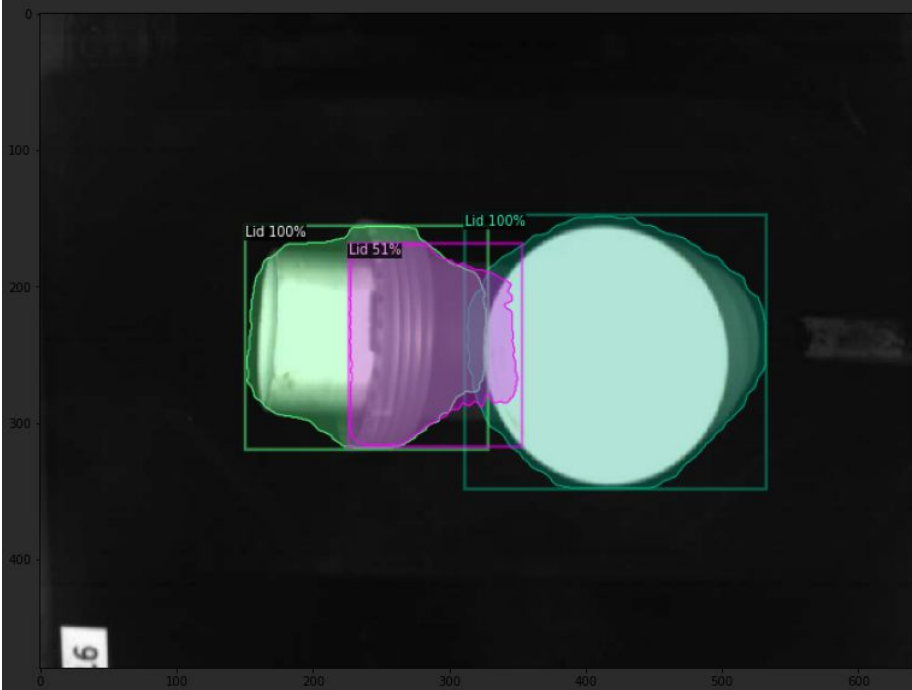


LABEL = 3 PREDICTED = 3 FILENAME = harder_samples/Anlernbilder (14).jpg

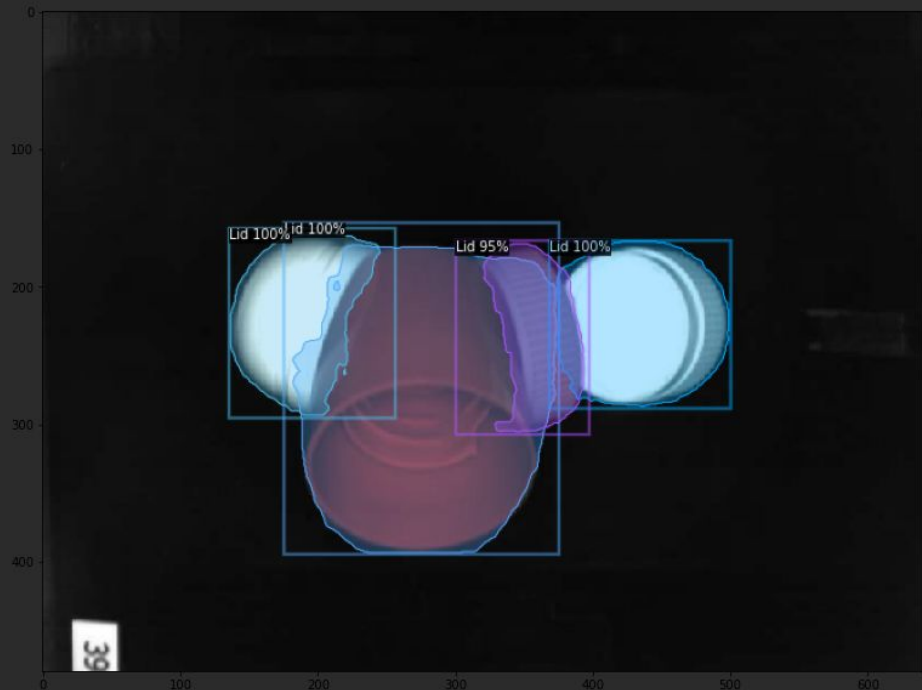


Segmentation: Qualitative Results

LABEL = 3 PREDICTED = 3 FILENAME = harder_samples/Anlernbilder (462).jpg

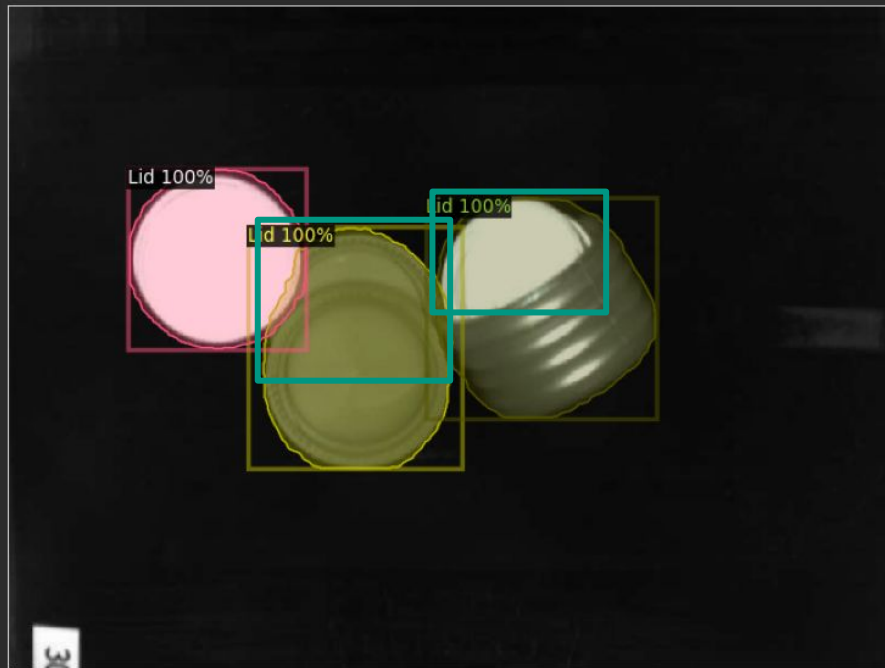


LABEL = 4 PREDICTED = 4 FILENAME = harder_samples/Anlernbilder (37).jpg

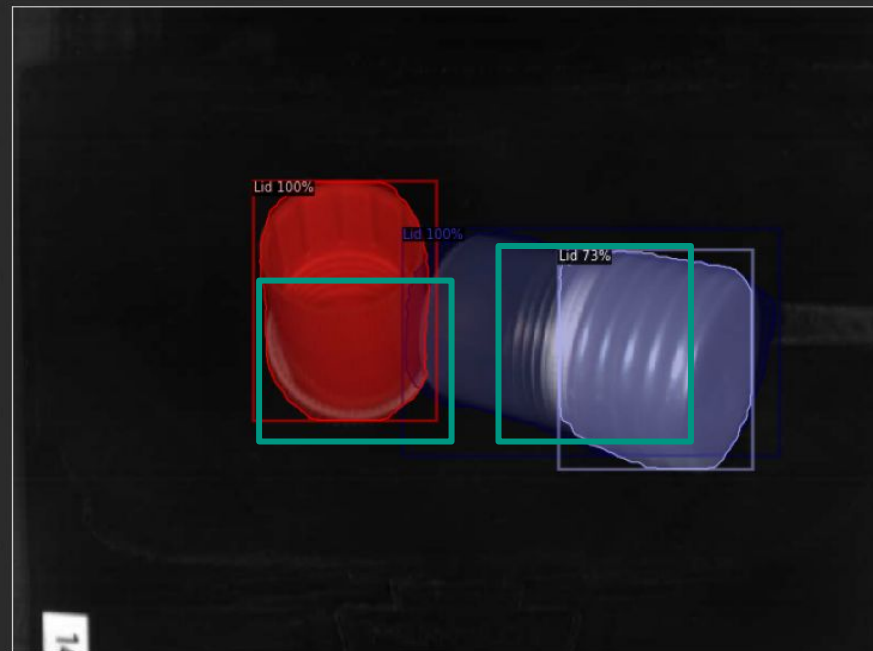


Segmentation: Qualitative Results

LABEL = 5 PREDICTED = 3 FILENAME = harder_samples/Anlernbilder (154).jpg



LABEL = 5 PREDICTED = 3 FILENAME = cluttered/Anlerndaten (697).jpg



Summary

- Unlabeled small dataset
- Copy & Paste Data Augmentation
- Classification: best performance
 - Model Tuning
 - Synthesized Data Improves Performance
- Instance Segmentation: poorer performance
 - But: interpretability
 - Synthesized data improves performance
- Future works:
 - Mixed approach: combine classification and segmentation
 - Improve for more complex object contours

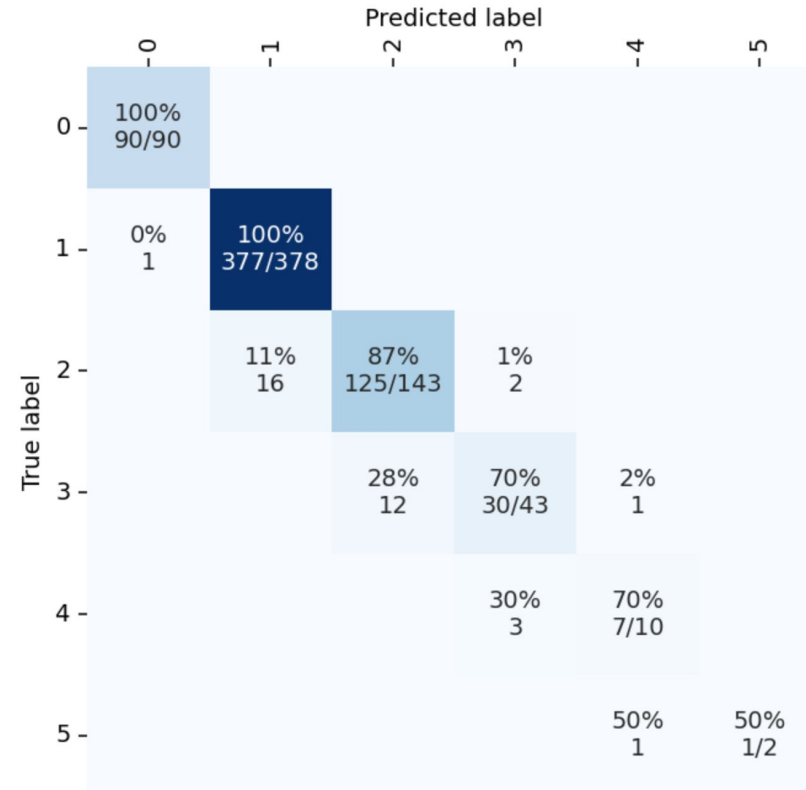
Thank you for your attention

References

- [Dwibedi17] Debidatta Dwibedi, Ishan Misra, and Martial Hebert. “Cut, Paste and Learn: Surprisingly Easy Synthesis for Instance Detection”. In: CoRR abs/1708.01642 (2017). arXiv: 1708.01642. url: <http://arxiv.org/abs/1708.01642>.
- [Ghiasi20] Golnaz Ghiasi et al. “Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation”. In: CoRR abs/2012.07177 (2020). arXiv: 2012.07177. Url: <https://arxiv.org/abs/2012.07177>.
- [Zhang15] Jianming Zhang, Shugao Ma, Mehrnoosh Sameki, Stan Sclaroff, Margrit Betke, Zhe Lin, Xiaohui Shen, Brian Price and Radomír Měch. “Salient Object Subitizing.” In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
- [Zhang16] Jianming Zhang, Shugao Ma, Mehrnoosh Sameki, Stan Sclaroff, Margrit Betke, Zhe Lin, Xiaohui Shen, Brian Price and Radomír Měch. “Salient Object Subitizing.” Journal version under review, 2016. Url: <https://arxiv.org/abs/1607.07525>.
- The code of our project: <https://github.com/yayapa/AnomaliesRecycling>

Appendix

Classification: Confusion Matrix



Tags



different colors



one color



transparent

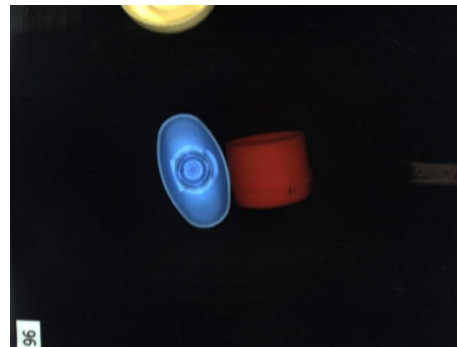


dark color

Tags



open lid



edge



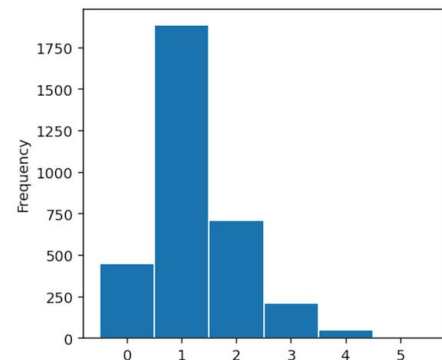
inside



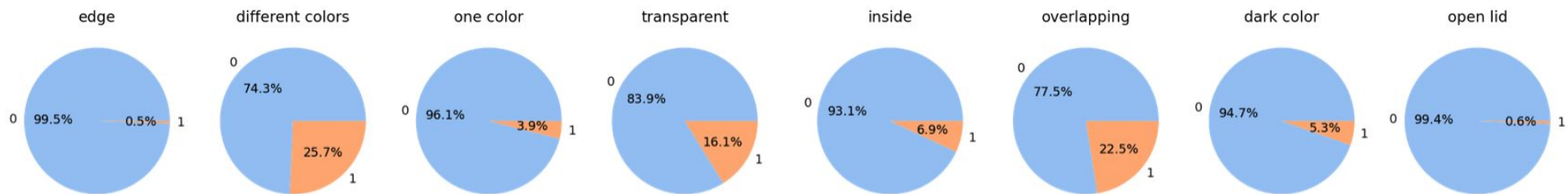
overlapping

Dataset: Labeling and Analysis

- Image-level labels for dataset
- Problems:
 - Small amount of data
 - Class Imbalance
 - Hard cases - analysis with tags



Frequency of Labels 0-5 in Original Dataset

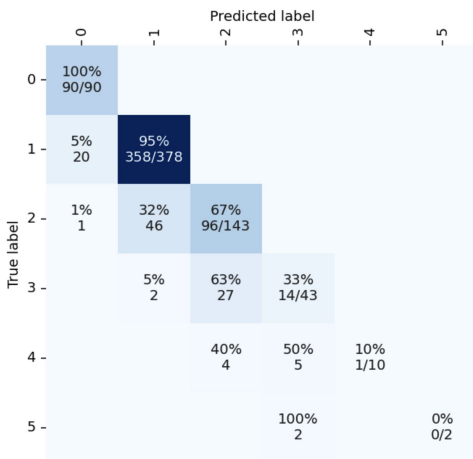


Amount of Data Corresponding to Hard Cases Marked with Tags

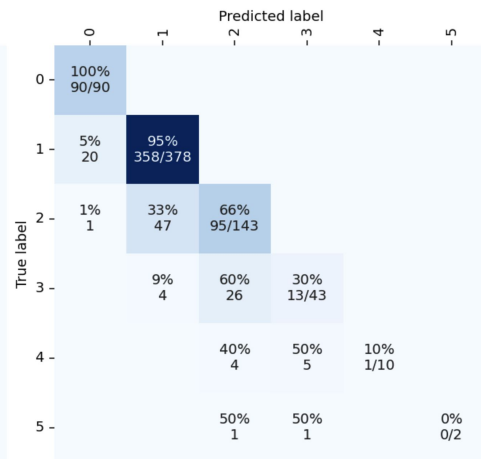
Segmentation: Ablation of Augmentation

Name	Rotate	Color	Edge	Dark	Trans parent	Label 0	Label 1	Label 2	Label 3	Label 4	Label 5	Accuracy	Decision Accuracy
Seg_no	x	x	x	x	x	1.00	0.94	0.64	0.28	0.20	0.00	0.83	0.89
Seg_edge	+	+	+	x	x	1.00	0.94	0.66	0.26	0.10	0.00	0.83	0.89
Seg_edge _dark	+	+	+	+	x	1.00	0.95	0.69	0.35	0.30	0.00	0.85	0.90
Seg_all	+	+	+	+	+	1.00	0.94	0.74	0.44	0.20	0.00	0.86	0.91

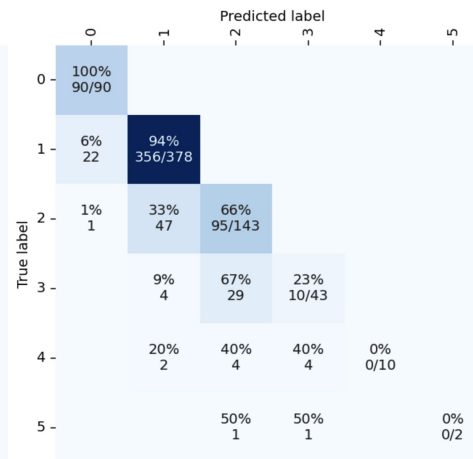
Segmentation: Ablation Of Test Threshold



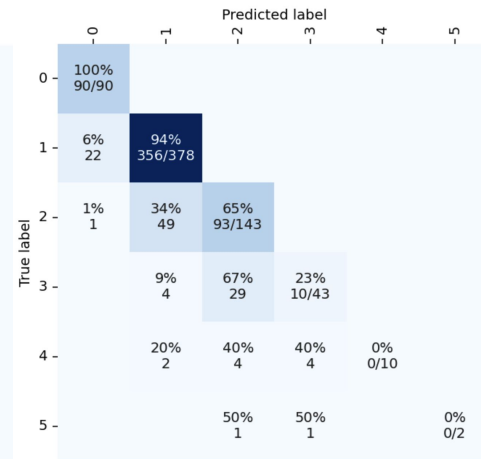
Threshold: 0.5
Accuracy: 0.84
Decision Accuracy: 0.9



Threshold: 0.7
Accuracy: 0.84
Decision Accuracy: 0.89



Threshold: 0.9
Accuracy: 0.83
Decision Accuracy: 0.89



Threshold: 0.95
Accuracy: 0.82
Decision Accuracy: 0.88

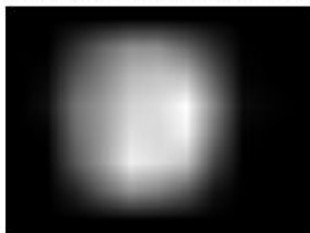
Class Activation Map: GradCAM

data/test/1/trial (704).jpg

input image



grad-cam backprop for label 1 (predicted 1)



data/test/0/trial (1307).jpg

input image



grad-cam backprop for label 0 (predicted 1)

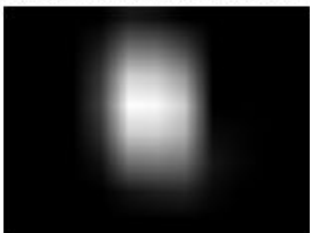


data/test/2/Anlerndaten (618).jpg

input image



grad-cam backprop for label 2 (predicted 1)



data/test/2/Anlerndaten (596).jpg

input image



grad-cam backprop for label 2 (predicted 1)

