International Conference in Central Europe on Computer Graphics,
Visualization and Computer Vision (WSCG 2023)

Designing a Lightweight Edge-Guided Convolutional Neural Network for Segmenting Mirrors and Reflective Surfaces

Mark Edward M. Gonzales
mark_gonzales@dlsu.edu.ph

Joel P. Ilao
joel.ilao@dlsu.edu.ph

Lorene C. Uy
lorene_c_uy@dlsu.edu.ph

De La Salle University
Manila, Philippines

github.com/memgonzales/mirror-segmentation

Problem and Motivation

- Existing computer vision systems have difficulty detecting mirrors and reflective surfaces (Park & Park, 2021)



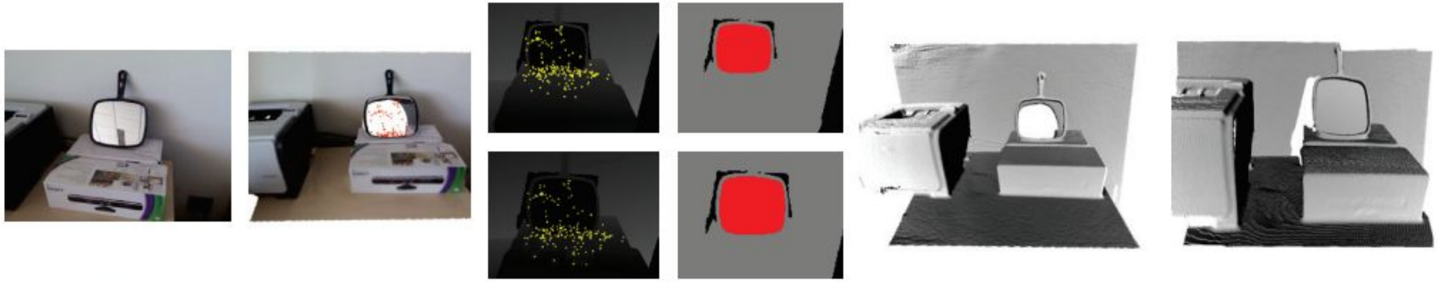


Park, D., and Park, Y.H. Identifying reflected images from object detector in indoor environment utilizing depth information. IEEE Robotics and Automation Letters, 6, 2, pp. 635-642, 2021.

GitHub

Problem and Motivation

- Presence of mirrors and reflective surfaces complicate tasks such as robot navigation (Anderson *et al.*, 2018) and 3D scene reconstruction (Zhang *et al.*, 2018)



Anderson, P. et al. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3674-3683, 2018.

Zhang, Y., Ye, M., Manocha, D., and Yang, R. 3D reconstruction in the presence of glass and mirrors by acoustic and visual fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40, 8, pp. 1785-1798, 2018.

GitHub

Problem and Motivation

- Mirrors pose potential hazards to autonomous driving and driving assistance systems (Zendel *et al.*, 2017)

Glare Spots

Distorted Reflections

Infinite Reflections

Zendel, O., Honauer, K., Murschitz, M., Humenberger, M., and Domínguez, G. Analyzing computer vision data - the good, the bad and the ugly. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6670-6680, 2017.

GitHub

A Challenging CV Task

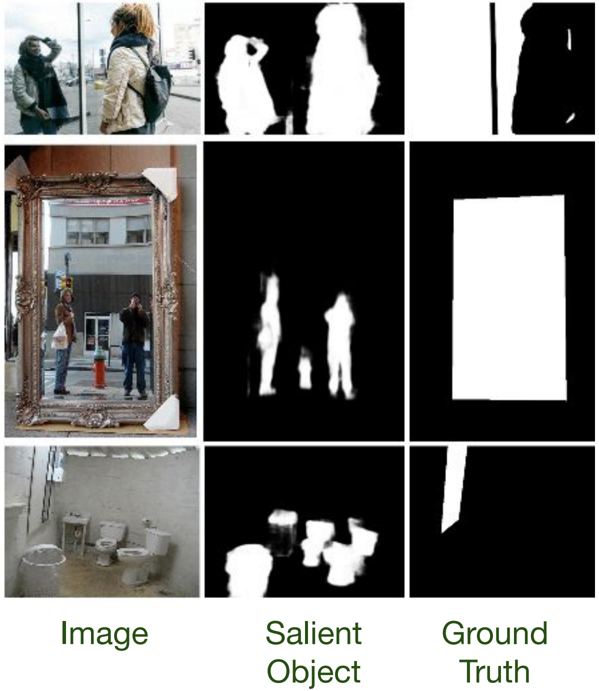
- **Semantic and instance segmentation?**
  - General object segmentation frameworks are unable to distinguish the reflection from the actual object (He *et al.*, 2017)
- **Salient object detection?**
  - Mirrors are not necessarily salient (Yang *et al.*, 2019)

He, K., Gkioxari, G., Dollár, P., and Girshick, R. Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980- 2988, 2017.
Yang, X. et al. Where is my mirror? Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8809-8818, 2019.

GitHub

A Challenging CV Task

- **Semantic and instance segmentation?**
    - General object segmentation frameworks are unable to distinguish the reflection from the actual object (He *et al.*, 2017)
- **Salient object detection?**
    - Mirrors are not necessarily salient (Yang *et al.*, 2019)



He, K., Gkioxari, G., Dollár, P., and Girshick, R. Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980- 2988, 2017.
Yang, X. et al. Where is my mirror? Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8809-8818, 2019.

GitHub

Mirror Detection Models

Model	Cues	Backbone	# Parameters
MirrorNet (Yang *et al.*, 2019)	Contextual contrasted features	ResNeXt-101	⚠️ 121.77M
PMDNet (Lin *et al.*, 2020)	Relational contextual contrasted features, edge features	ResNeXt-101	⚠️ 147.66M
SANet (Guan *et al.*, 2022)	Semantic associations	ResNeXt-101	⚠️ 105.84M

Yang, X. et al. Where is my mirror? Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8809-8818, 2019.

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

Guan, H., Lin, J., and Lau, R.W.H. Learning semantic associations for mirror detection. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5931-5940, 2022.

GitHub

Our Contributions

Gaps / Limitations	Our Contributions
Datasets are mostly limited to clear mirrors found in indoor scenes	Dataset consisting of outdoor mirrors and reflective surfaces (e.g., tinted car windows and building façades)
Designing lightweight mirror segmentation models remains an unexplored direction	Lightweight edge-guided CNN for segmenting mirrors and reflective surfaces

GitHub

Our Contributions

Gaps / Limitations	Our Contributions
Datasets are mostly limited to clear mirrors found in indoor scenes	Dataset consisting of outdoor mirrors and reflective surfaces (e.g., tinted car windows and building façades)
Designing lightweight mirror segmentation models remains an unexplored direction	Lightweight edge-guided CNN for segmenting mirrors and reflective surfaces

GitHub

Our Contributions

- **Dataset of outdoor mirrors and reflective surfaces** with 454 images and their corresponding ground-truth masks

- **Modified the architecture of PMDNet** (Lin *et al.*, 2020) and extensively tested different backbones and edge-related modules to guide segmentation

- **Pruned best-performing edge-guided CNN**, resulting in a lightweight model that performs competitively with PMDNet but with 78.20× fewer FLOPS and 238.16× fewer parameters

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

GitHub

Our Contributions

- **Dataset of outdoor mirrors and reflective surfaces** with 454 images and their corresponding ground-truth masks

- **Modified the architecture of PMDNet (Lin *et al.*, 2020)** and extensively tested different backbones and edge-related modules to guide segmentation

- **Pruned best-performing edge-guided CNN**, resulting in a lightweight model that performs competitively with PMDNet but with 78.20× fewer FLOPS and 238.16× fewer parameters

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

GitHub

Our Contributions

- **Dataset of outdoor mirrors and reflective surfaces** with 454 images and their corresponding ground-truth masks

- **Modified the architecture of PMDNet (Lin *et al.*, 2020)** and extensively tested different backbones and edge-related modules to guide segmentation

- **Pruned best-performing edge-guided CNN**, resulting in a lightweight model that performs competitively with PMDNet but with 78.20× fewer FLOPS and 238.16× fewer parameters

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

GitHub

Datasets

Publicly Available Datasets

Mirror Segmentation Dataset (MSD)

4018 images

Yang, X. et al. Where is my mirror? Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8809-8818, 2019.

Progressive Mirror Detection (PMD) Dataset

6016 images

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

GitHub

Our Proposed Dataset

De La Salle University – Outdoor Mirrors & Reflective Surfaces (DLSU-OMRS)

- **454 images** scraped from Shutterstock using the key phrases *outdoor mirror* and *street mirror*
- Ground-truth masks produced via **manual segmentation**
- **Average structural similarity index:** 28.67%

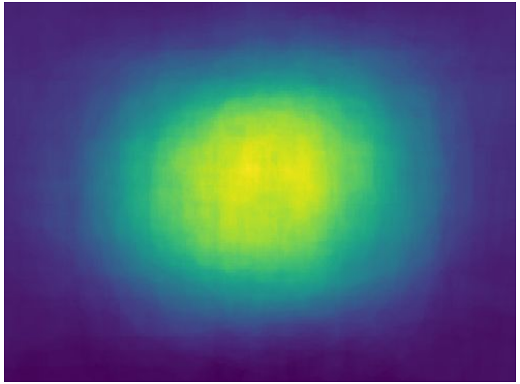




GitHub

Our Proposed Dataset

De La Salle University – Outdoor Mirrors & Reflective Surfaces (DLSU-OMRS)



Mirror Location Distribution

Mirror-to-Image Area Ratio

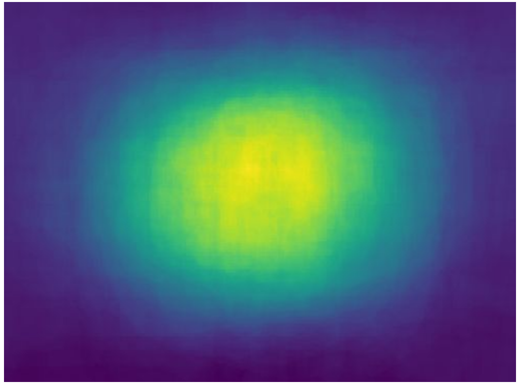
Color Contrast

Most mirrors are located near the center

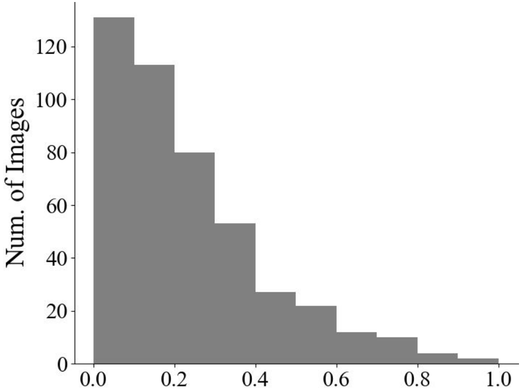
GitHub

Our Proposed Dataset

De La Salle University – Outdoor Mirrors & Reflective Surfaces (DLSU-OMRS)



Mirror Location Distribution



Mirror-to-Image Area Ratio

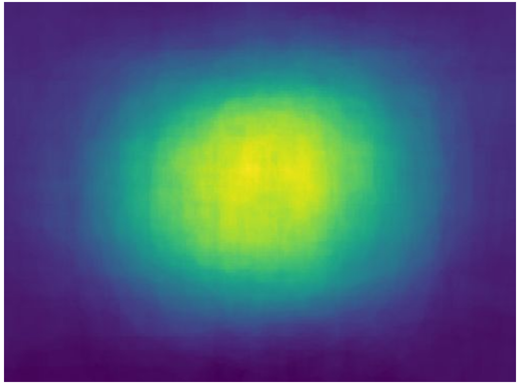
Color Contrast

Most mirrors occupy up to 20% of the image

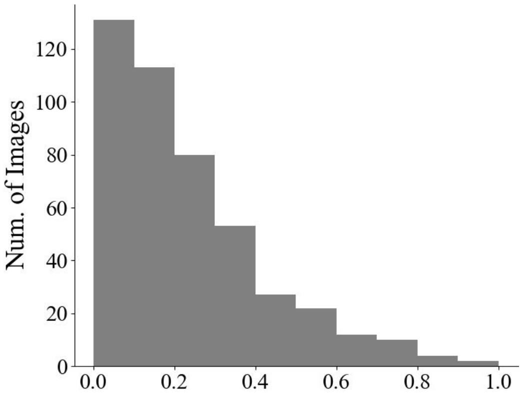
GitHub

Our Proposed Dataset

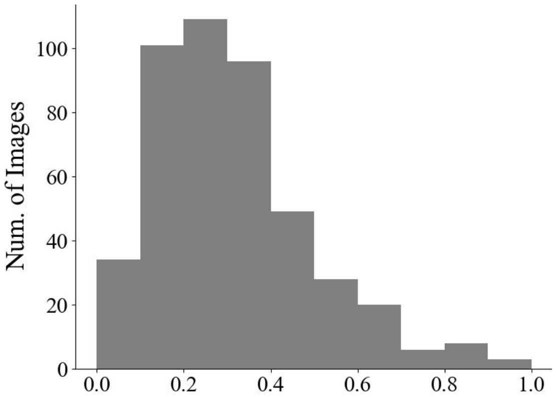
De La Salle University – Outdoor Mirrors & Reflective Surfaces (DLSU-OMRS)



Mirror Location Distribution



Mirror-to-Image Area Ratio



Color Contrast

The color contrast of most images is below 40%

GitHub

Our Proposed Dataset

De La Salle University – Outdoor Mirrors & Reflective Surfaces (DLSU-OMRS)

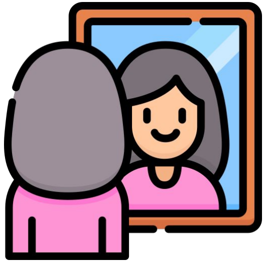
	# Images
One Mirror	338
Multiple Mirrors	116

By Shape	# Mirrors
Triangle	4
Quadrilateral	258
Polygonal	9
Round/Elliptical	160
Irregular	355

By Presence of Occlusions	# Mirrors
Present	192
Not Present	594

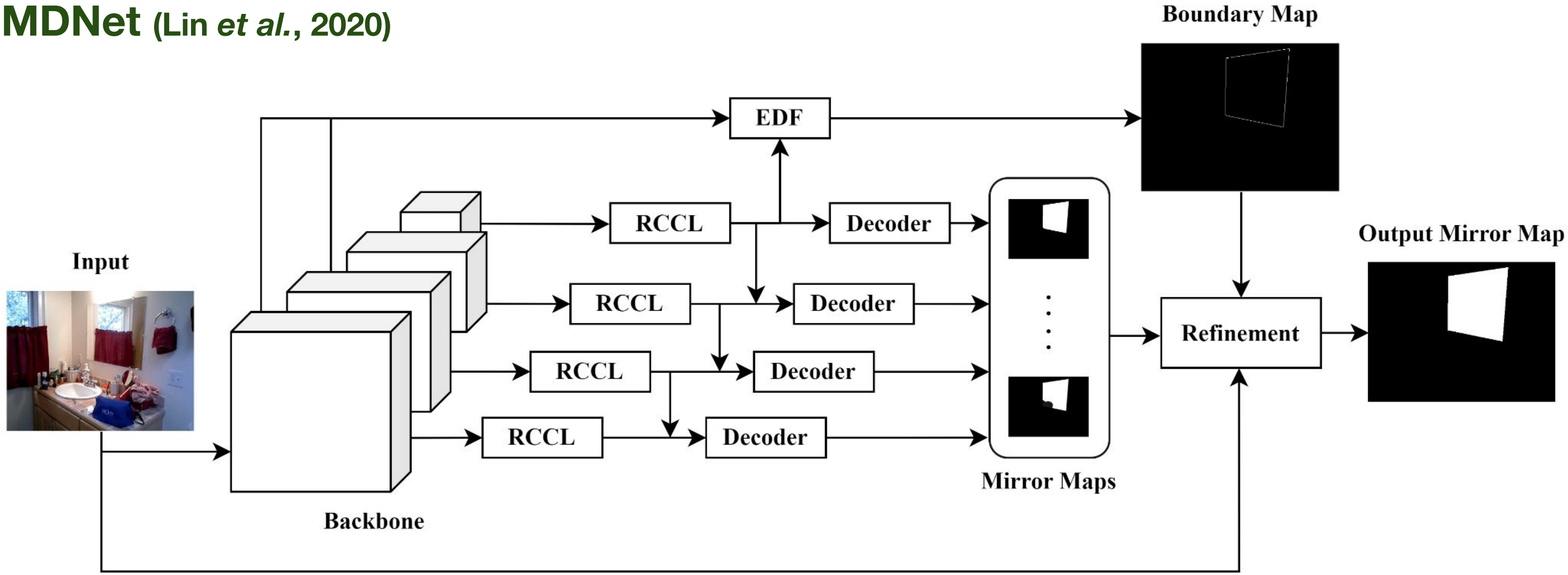
GitHub

Model Architecture



Base Model

PMDNet (Lin *et al.*, 2020)

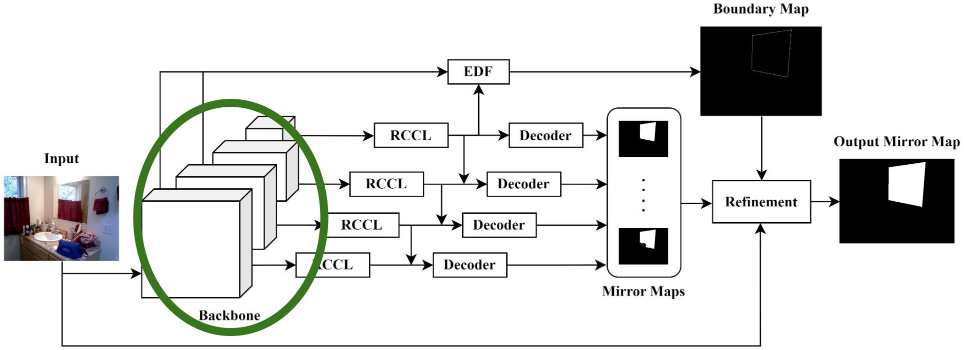


Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

GitHub

Changing the Feature Extraction Backbone

1. ResNet-50 (He et al., 2016)
2. Xception-65 (Chollet, 2017)
3. VoVNet-39 (Lee *et al.*, 2019)
4. MobileNetV3 (Howard *et al.*, 2019)
5. EfficientNetLite4 (Tan & Le, 2019)
6. EfficientNet-Edge-Large (Pruned) (Tan & Le, 2019)
7. EfficientNetV2-Medium (Tan & Le, 2019)



He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

Chollet, F. Xception: Deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1800-1807, 2017.

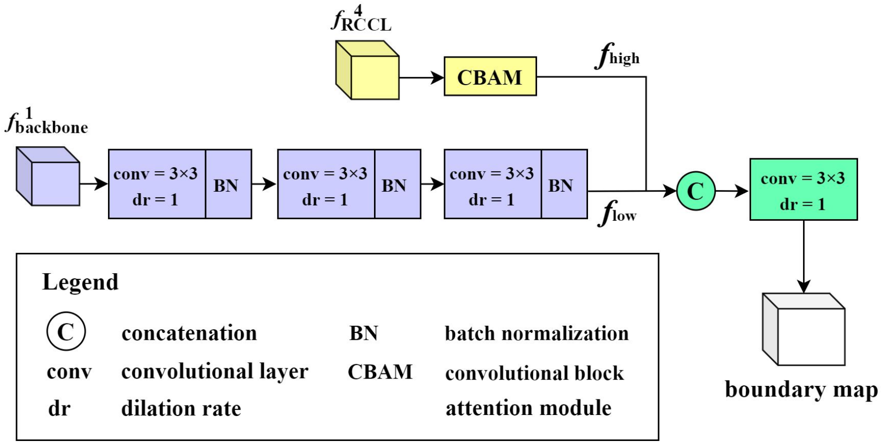
Lee, Y., Hwang, J.W., Lee, S., Bae, Y., and Park, J. An energy and GPU-computation efficient backbone network for real-time object detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 752-760, 2019.

Howard, A. et al. Searching for MobileNetV3. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 1314-1324, 2019.

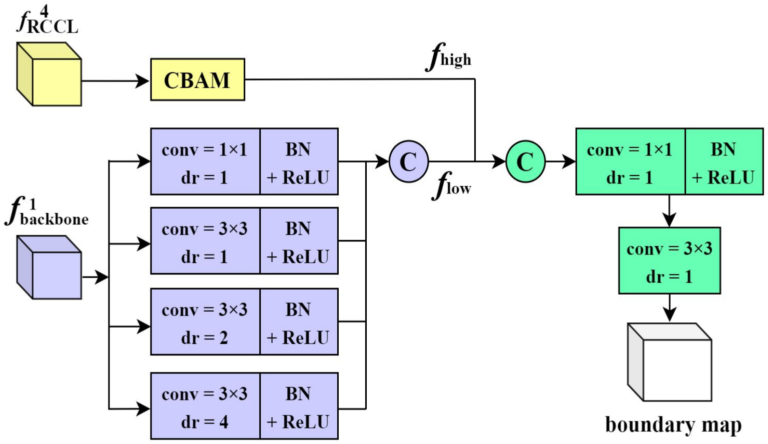
Tan, M., and Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, 97, pp. 6105-6114, 2019.

GitHub

Modifying the Edge Detection and Fusion Module



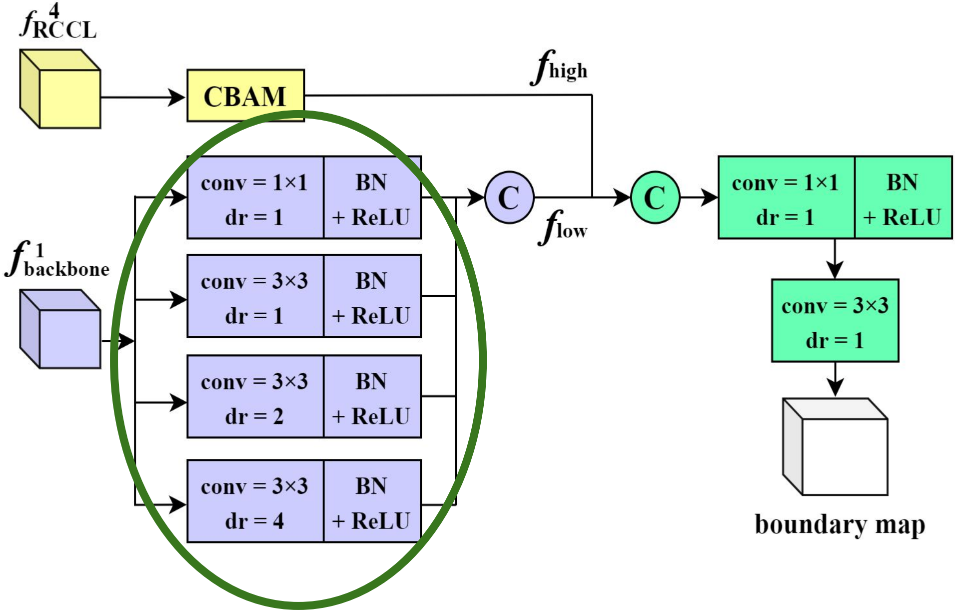
PMDNet



Ours

GitHub

Modifying the Edge Detection and Fusion Module



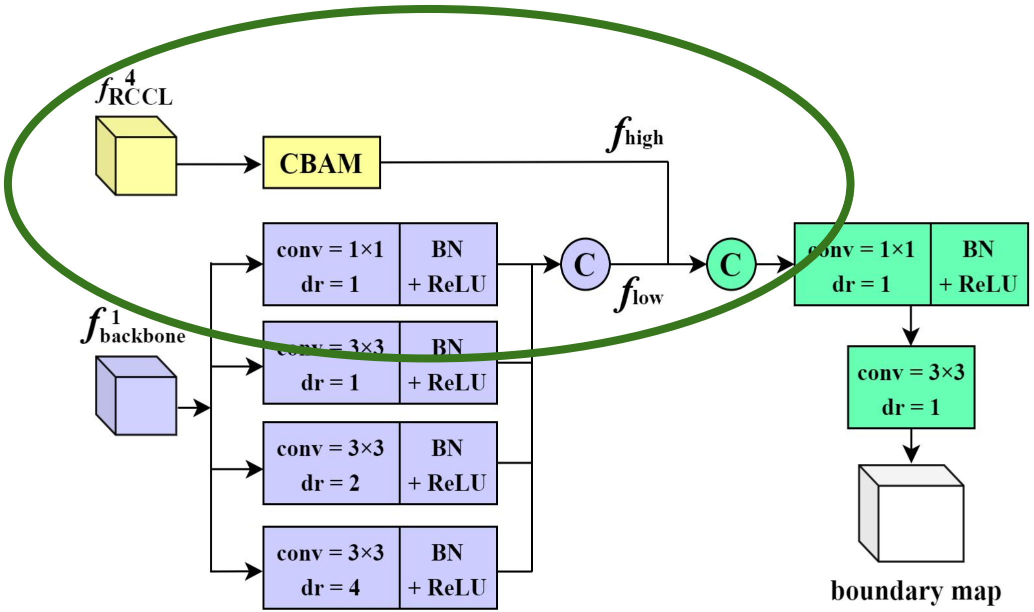
Low-Level Edge Features

We connect the side-output of the lowest-level backbone to a boundary extraction module with **four parallel conv. layers**, adapted from GDNet (Mei et al., 2022)

Mei, H. et al. Large-field contextual feature learning for glass detection. IEEE Transactions on Pattern Analysis & Machine Intelligence, 01, pp. 1-17, 2022

GitHub

Modifying the Edge Detection and Fusion Module



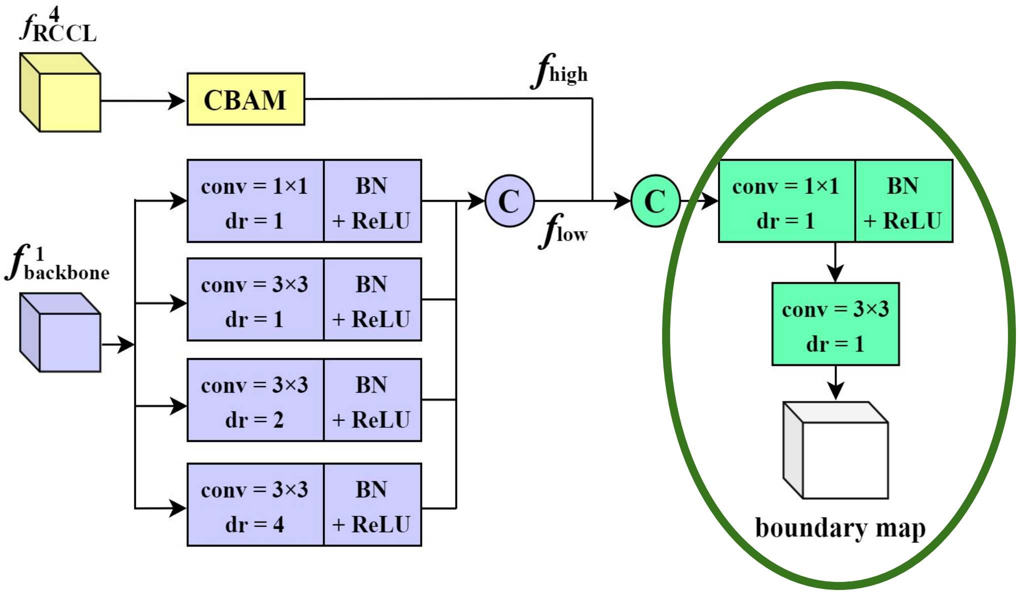
High-Level Edge Features

We share PMDNet's approach of using a **convolutional block attention module** (Woo et al., 2018), a lightweight module that infers spatial and channel attention maps

Woo, S., Park, J., Lee, J., Lee, J., and Kweon, I.S. CBAM: Convolutional block attention module. Proceedings of the European Conference on Computer Vision (ECCV), pp. 3-19, 2018.

GitHub

Modifying the Edge Detection and Fusion Module



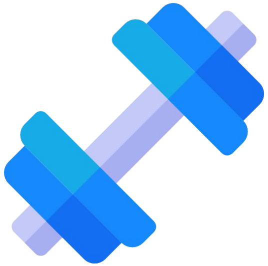
Edge Prediction Block

Original: 3×3 conv. layer

Ours: 1×1 conv. layer with BN and ReLU connected to a 3×3 conv. layer

GitHub

Model Training



Model Training

- **Library**: PyTorch

- **Training Dataset**
  - Training partition of split PMD dataset (5096 images)

- **Data Preprocessing and Augmentation**
  - Resized to 352 x 352
  - Random horizontal flipping
  - Jittering the brightness, contrast, saturation, and hue

GitHub

Model Training

- **Initial Learning Rate:** 1×10^{-3}
  - **Update:** Polynomial (power = 0.9)

- **Optimizer:** SGD
  - **Weight Decay:** 5×10^{-4}
  - **Momentum:** 0.9

GitHub

Loss Functions

Our Proposed Compound Loss Function

$$L = \boxed{\sum_{i=1}^{4} w_{mirror} \cdot L_{mirror}\left(\hat{M_i}, M\right)} + \boxed{w_{edge} \cdot L_{edge}\left(\hat{E}, E\right)} + \boxed{w_{output} \cdot L_{output}\left(\hat{M}, M\right)}$$

Mirror Maps

- IoU loss

Boundary Map

- Laplacian-based
- For emphasizing boundaries

Final Mirror Map

- Weighted IoU + BCE
- Draws the model to a larger receptive field

Zhao, T., and Wu, X. Pyramid feature attention network for saliency detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3080-3089, 2019.

Wei, J., Wang, S., and Huang, Q. F^3Net: Fusion, feedback and focus for salient object detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp.12321-12328, 2020.

GitHub

Our Proposed Compound Loss Function

$$L = \boxed{\sum_{i=1}^{4} w_{mirror} \cdot L_{mirror}\left(\hat{M}_i, M\right)} + \boxed{w_{edge} \cdot L_{edge}\left(\hat{E}, E\right)} + \boxed{w_{output} \cdot L_{output}\left(\hat{M}, M\right)}$$

Mirror Maps

- IoU loss

Boundary Map

- Laplacian-based
- For emphasizing boundaries

Final Mirror Map

- Weighted IoU + BCE
- Draws the model to a larger receptive field

Zhao, T., and Wu, X. Pyramid feature attention network for saliency detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3080-3089, 2019.

Wei, J., Wang, S., and Huang, Q. F^3Net: Fusion, feedback and focus for salient object detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp.12321-12328, 2020.

GitHub

Our Proposed Compound Loss Function

$$L = \boxed{\sum_{i=1}^{4} w_{mirror} \cdot L_{mirror}\left(\hat{M_i}, M\right)} + \boxed{w_{edge} \cdot L_{edge}\left(\hat{E}, E\right)} + \boxed{w_{output} \cdot L_{output}\left(\hat{M}, M\right)}$$

Mirror Maps

- IoU loss

Boundary Map

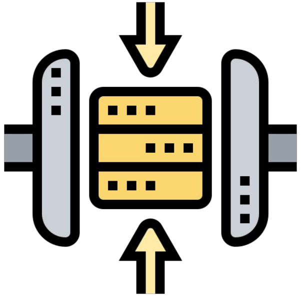
- Laplacian-based
- For emphasizing boundaries

Final Mirror Map

- Weighted IoU + BCE
- Draws the model to a larger receptive field

Zhao, T., and Wu, X. Pyramid feature attention network for saliency detection. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3080-3089, 2019.
Wei, J., Wang, S., and Huang, Q. F^3Net: Fusion, feedback and focus for salient object detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34, pp.12321-12328, 2020.

Model Compression



Filter Pruning via Geometric Median (FPGM)

- One-shot structured pruning technique (He et al., 2019)
- We applied FPGM on the convolutional and linear layers
- **Sparsity level:** 10%
- **Learning rate rewinding** (Renda et al., 2020) for 20 epochs was done to retrain unpruned weights from their final values

He, Y., Liu, P., Wang, Z., Hu, Z., and Yang, Y. Filter pruning via geometric median for deep convolutional neural networks acceleration. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4335-4344, 2019.

Renda, A., Frankle, J., and Carbin, M. Comparing rewinding and fine-tuning in neural network pruning. International Conference on Learning Representations, 2020.

GitHub

Model Evaluation

Test Datasets

- MSD (955 images)
- Test partition of split PMD dataset (571 images)
- DLSU-OMRS (454 images)

Metrics

- Maximum F-measure

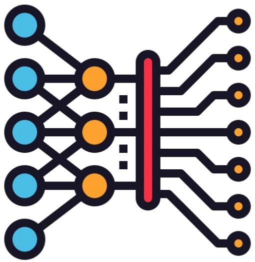
$$F_\beta = \frac{(1+\beta^2)\cdot \text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

- Mean absolute error

$$MAE = \frac{1}{w \cdot h} \sum_{x=1}^{w} \sum_{y=1}^{h} |\hat{Y}(x,y) - Y(x,y)|$$

GitHub

Results and Analysis



Model	Computational Complexity		MSD		PMD		DLSU-OMRS	
	GFLOPS ↓	# Params ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓
VST (Liu *et al.*, 2021)	46.36	44.48M	0.4290	0.2739	0.1317	0.261	0.5730	0.2274
PMDNet (Lin *et al.*, 2020)	118.86	147.66M	0.8350	0.0816	<u>0.8011</u>	<u>0.0324</u>	<u>0.8423</u>	**0.0878**
Ours – Compound Loss (CL) + Edge Extraction (EE)								
ResNet + CL + EE	116.46	130.12M	0.7695	0.1098	0.7524	0.0409	0.8042	0.1025
MobileNet + CL + EE	<u>6.61</u>	20.76M	0.7515	0.1153	0.7508	0.0427	0.8256	0.1006
EfficientNetLite + CL + EE	6.99	15.54M	0.7909	0.1027	0.7769	0.0387	0.8178	0.1048
EfficientNet + CL + EE	24.79	53.35M	<u>0.8483</u>	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
EfficientNet + CL + EE + Pruned	**1.52**	**0.62M**	**0.8498**	<u>0.0813</u>	0.7902	0.0364	**0.8456**	<u>0.0955</u>

Liu, N., Zhang, N., and Wan, K., Shao, L., and Han, J. Visual saliency transformer. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 4722-4732, 2021.

Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.

Model	Computational Complexity		MSD		PMD		DLSU-OMRS	
	GFLOPS ↓	# Params ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓
VST (Liu *et al.*, 2021)	46.36	44.48M	0.4290	0.2739	0.1317	0.261	0.5730	0.2274
PMDNet (Lin *et al.*, 2020)	118.86	147.66M	0.8350	0.0816	<u>0.8011</u>	<u>0.0324</u>	<u>0.8423</u>	**0.0878**
Ours – Compound Loss (CL) + Edge Extraction (EE)								
ResNet + CL + EE	116.46	130.12M	0.7695	0.1098	0.7524	0.0409	0.8042	0.1025
MobileNet + CL + EE	<u>6.61</u>	20.76M	0.7515	0.1153	0.7508	0.0427	0.8256	0.1006
EfficientNetLite + CL + EE	6.99	15.54M	0.7909	0.1027	0.7769	0.0387	0.8178	0.1048
EfficientNet + CL + EE	24.79	53.35M	<u>0.8483</u>	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
EfficientNet + CL + EE + Pruned	**1.52**	**0.62M**	**0.8498**	<u>0.0813</u>	0.7902	0.0364	**0.8456**	<u>0.0955</u>

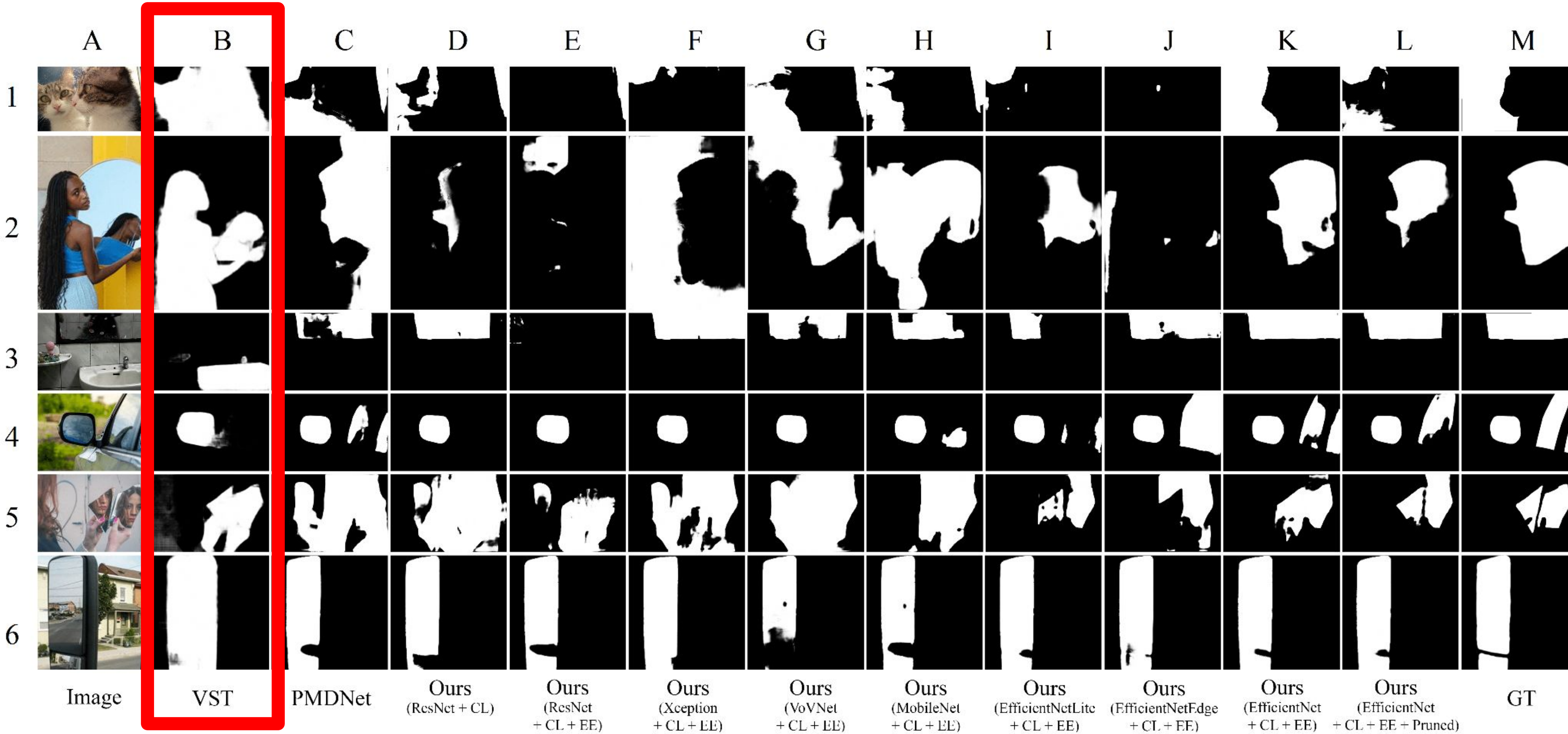
Our model that uses **EfficientNet** as a backbone and employs our compound loss function and edge extraction and prediction module **performed competitively with PMDNet**. It also has **4.79× fewer FLOPS** and **2.77× fewer parameters**

Model	Computational Complexity		MSD		PMD		DLSU-OMRS	
	GFLOPS ↓	# Params ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓
VST (Liu *et al.*, 2021)	46.36	44.48M	0.4290	0.2739	0.1317	0.261	0.5730	0.2274
PMDNet (Lin *et al.*, 2020)	118.86	147.66M	0.8350	0.0816	<u>0.8011</u>	<u>0.0324</u>	<u>0.8423</u>	**0.0878**
Ours – Compound Loss (CL) + Edge Extraction (EE)								
ResNet + CL + EE	116.46	130.12M	0.7695	0.1098	0.7524	0.0409	0.8042	0.1025
MobileNet + CL + EE	<u>6.61</u>	20.76M	0.7515	0.1153	0.7508	0.0427	0.8256	0.1006
EfficientNetLite + CL + EE	6.99	15.54M	0.7909	0.1027	0.7769	0.0387	0.8178	0.1048
EfficientNet + CL + EE	24.79	53.35M	<u>0.8483</u>	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
EfficientNet + CL + EE + Pruned	**1.52**	**0.62M**	**0.8498**	<u>0.0813</u>	0.7902	0.0364	**0.8456**	<u>0.0955</u>

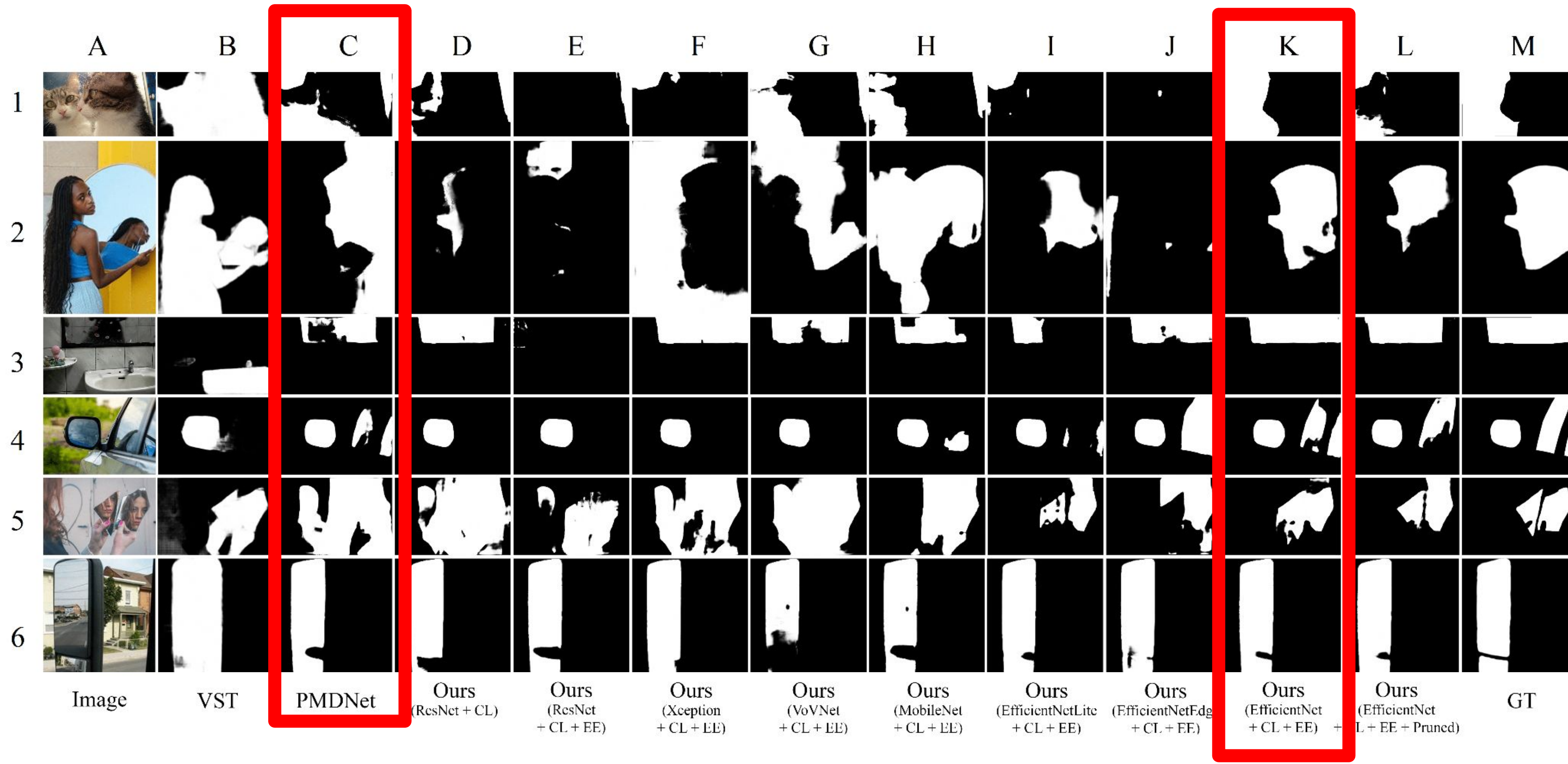
The **pruned version** of this model also **performed competitively with PMDNet** and slightly outperformed it on MSD and DLSU-OMRS in terms of F_β score. It also has **78.20× fewer FLOPS** and **238.16× fewer parameters**

Model	Computational Complexity		MSD		PMD		DLSU-OMRS	
	GFLOPS ↓	**# Params ↓**	F_β ↑	**MAE ↓**	F_β ↑	**MAE ↓**	F_β ↑	**MAE ↓**
VST (Liu *et al.*, 2021)	46.36	44.48M	0.4290	0.2739	0.1317	0.261	0.5730	0.2274
PMDNet (Lin *et al.*, 2020)	118.86	147.66M	0.8350	0.0816	<u>0.8011</u>	<u>0.0324</u>	<u>0.8423</u>	**0.0878**
Ours – Compound Loss (CL) + Edge Extraction (EE)								
ResNet + CL + EE	116.46	130.12M	0.7695	0.1098	0.7524	0.0409	0.8042	0.1025
MobileNet + CL + EE	<u>6.61</u>	20.76M	0.7515	0.1153	0.7508	0.0427	0.8256	0.1006
EfficientNetLite + CL + EE	6.99	15.54M	0.7909	0.1027	0.7769	0.0387	0.8178	0.1048
EfficientNet + CL + EE	24.79	53.35M	<u>0.8483</u>	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
EfficientNet + CL + EE + Pruned	**1.52**	**0.62M**	**0.8498**	<u>0.0813</u>	0.7902	0.0364	**0.8456**	<u>0.0955</u>

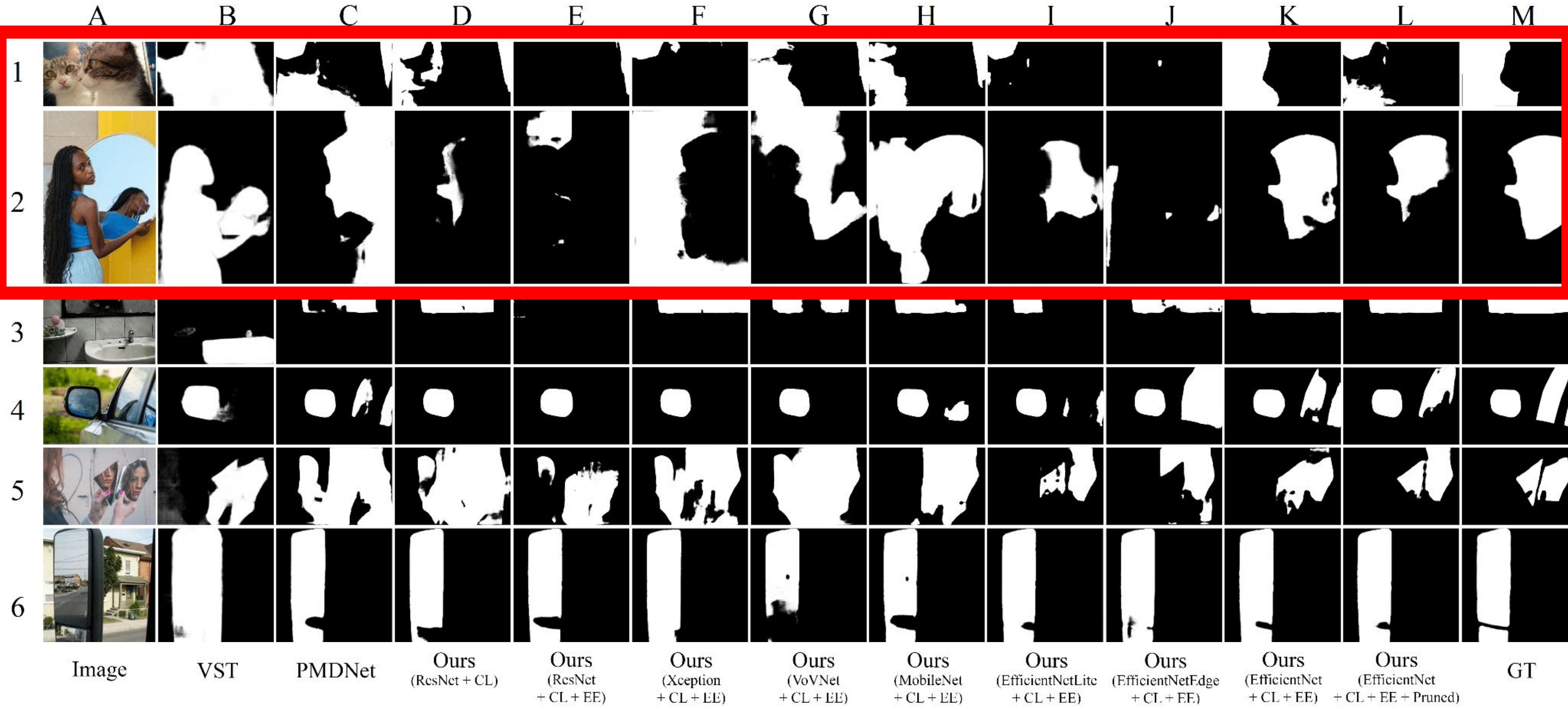
Although our model with an **EfficientNet-Lite** backbone was not able to outperform PMDNet, its F_β scores across all three benchmark datasets were **consistently within 0.06 points of the highest scores**



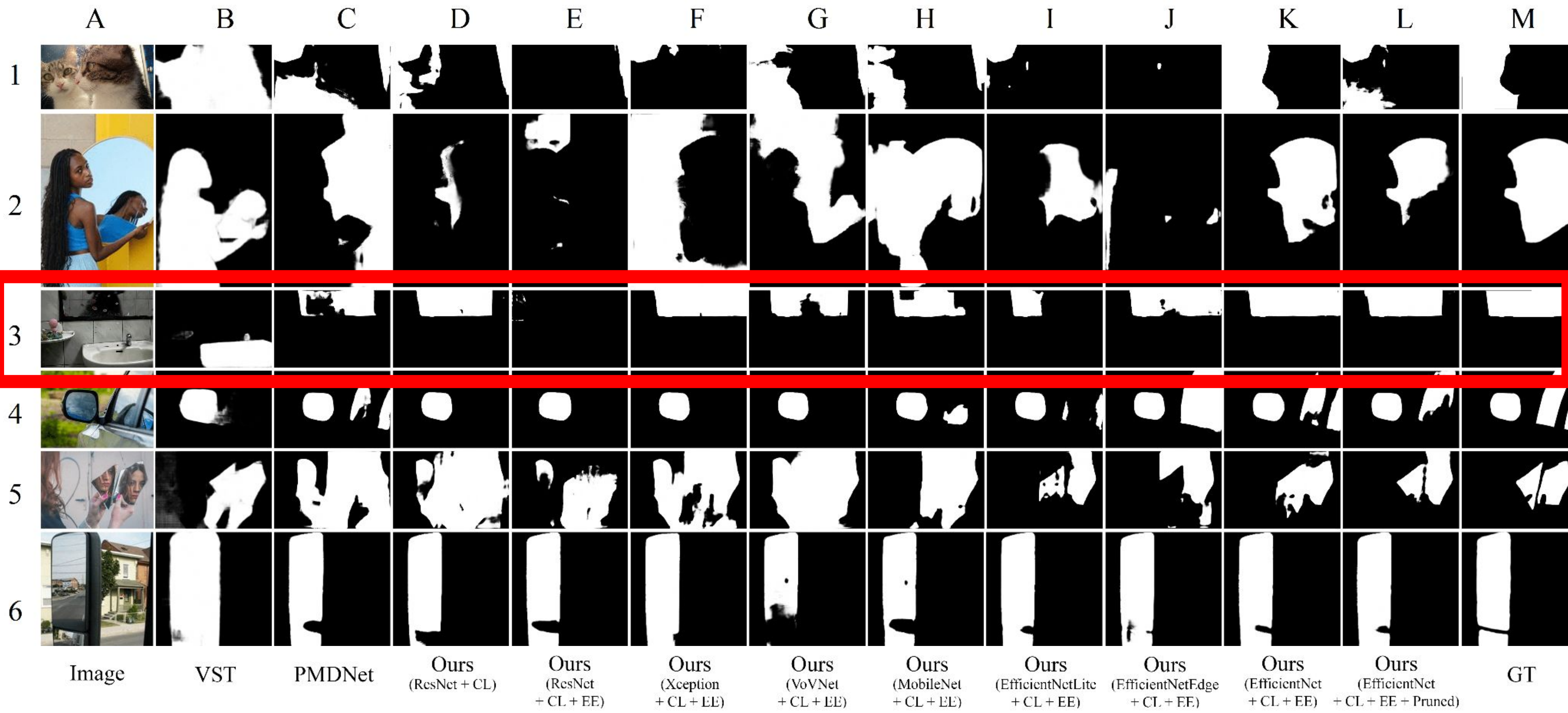
Salient object detection models may not necessarily tag mirrors as salient



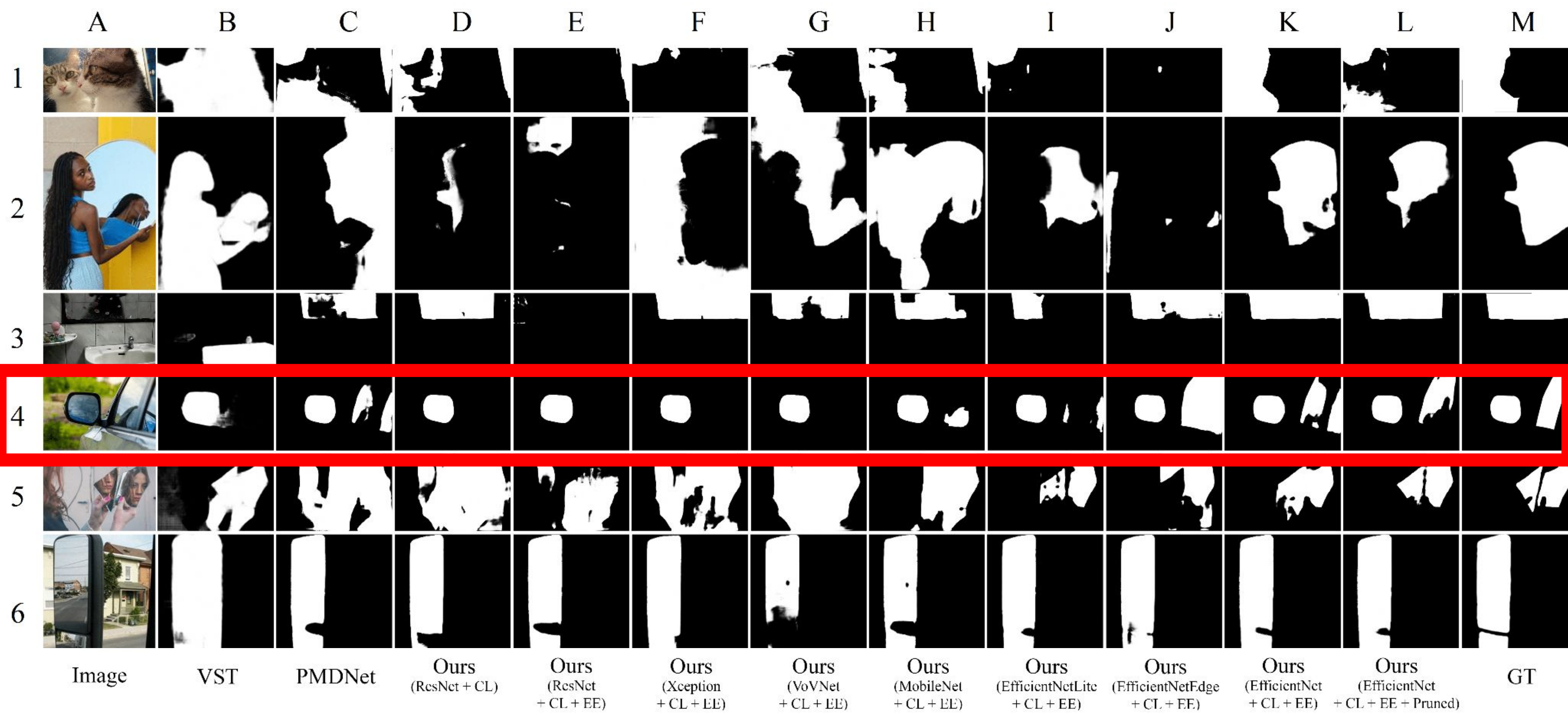
Our **best-performing model (column K)** can handle
some cases that may be challenging even for a state-of-the-art model (column C)



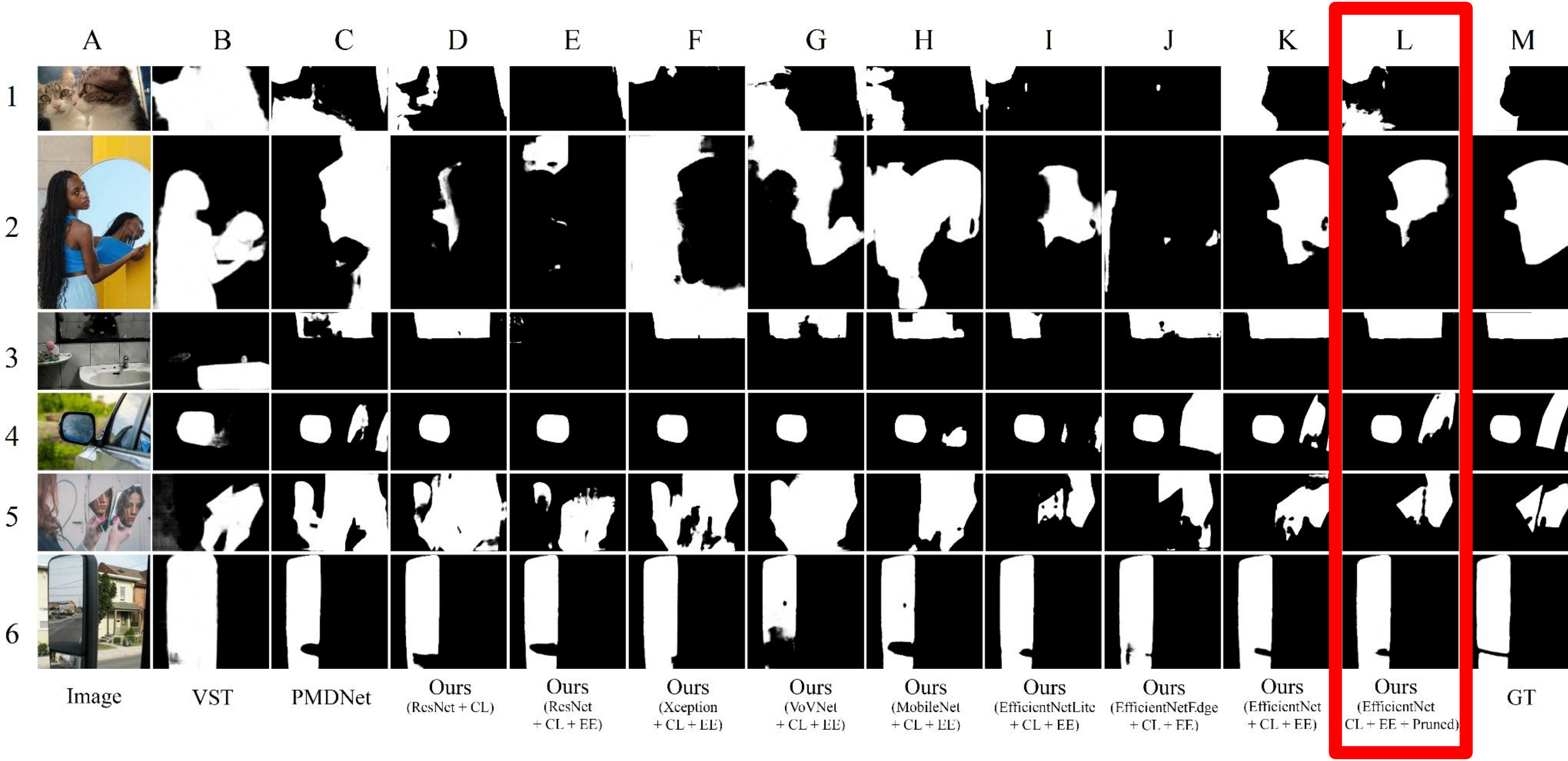
These success cases (column K) include images where the **object occludes the mirror** and, alongside its reflection, occupies a **large portion of the image**



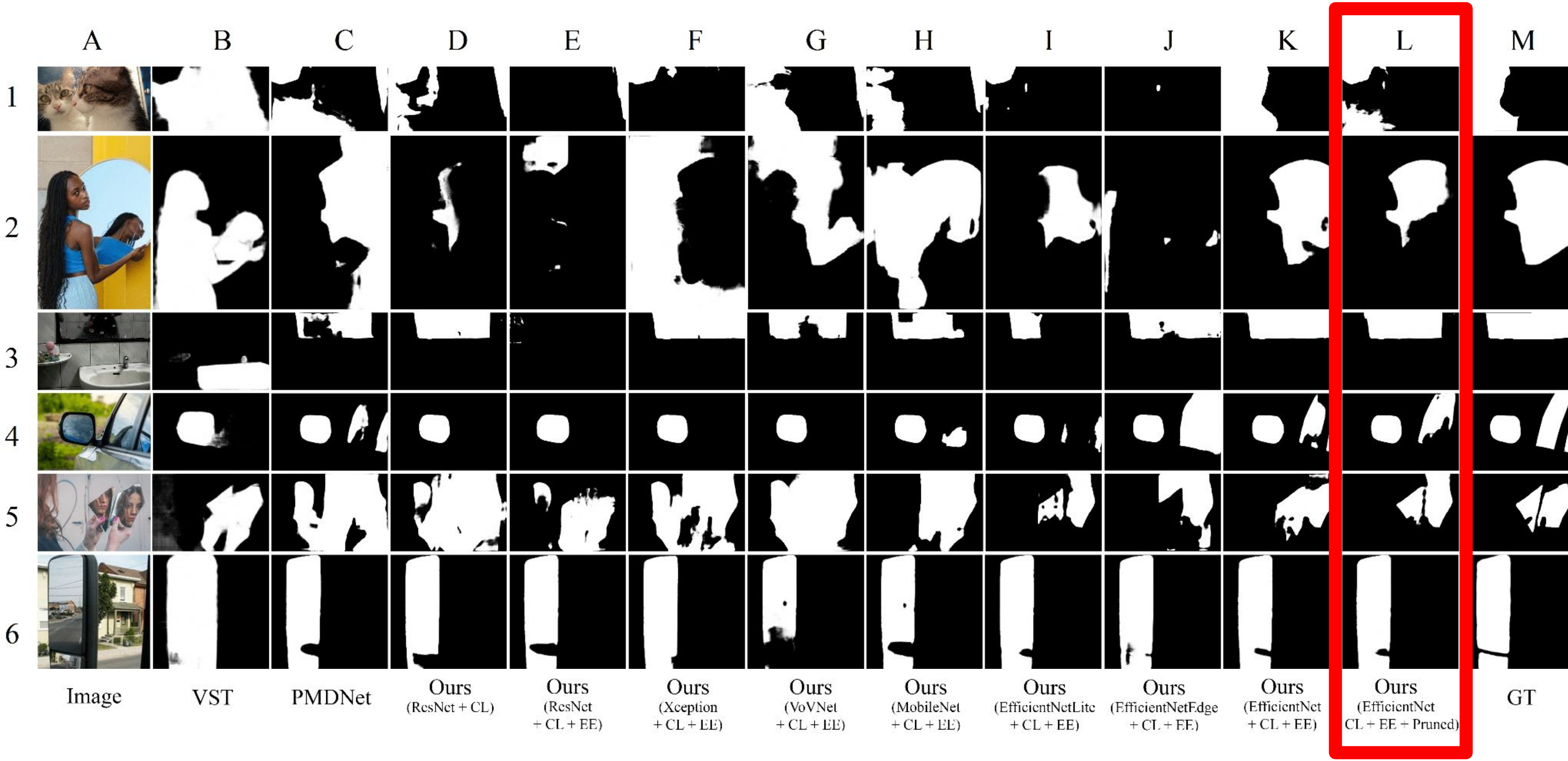
These success cases (column K) include images where the **reflection has a similar color to the mirror's frame**



These success cases (column K) include images where **multiple mirrors and reflective surfaces** are present

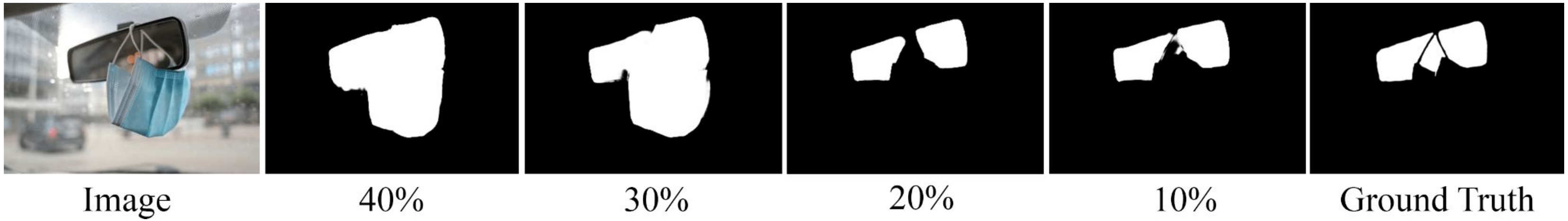


Our **pruned version** was able to segment **irregularly shaped mirror shards** (row 5)



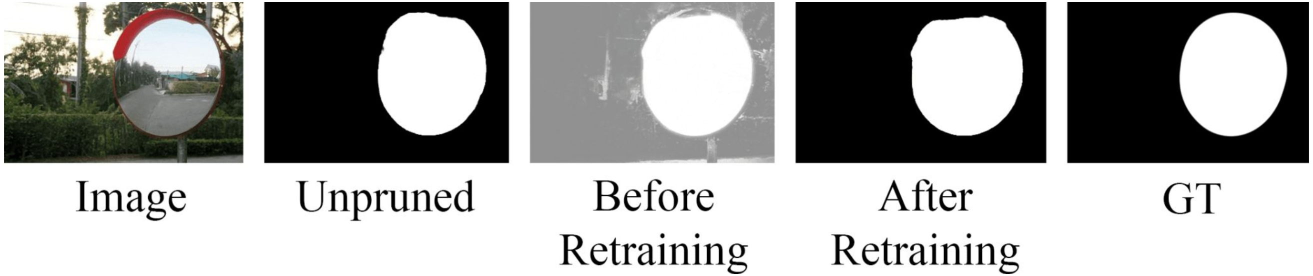
However, it has some **difficulty** handling cases
where the **object and reflection occupy the majority of the image** (rows 1 and 2)

Sparsity	MSD		PMD		DLSU-OMRS	
	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓
40%	0.6267	0.4633	0.6006	0.4790	0.6876	0.1485
30%	0.7695	0.0970	0.7566	0.0410	0.7963	0.1039
20%	0.8073	0.0905	0.7795	0.0352	0.8211	**0.0940**
10%	**0.8498**	0.0813	0.7902	0.0364	**0.8456**	0.0955
0%	0.8483	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032



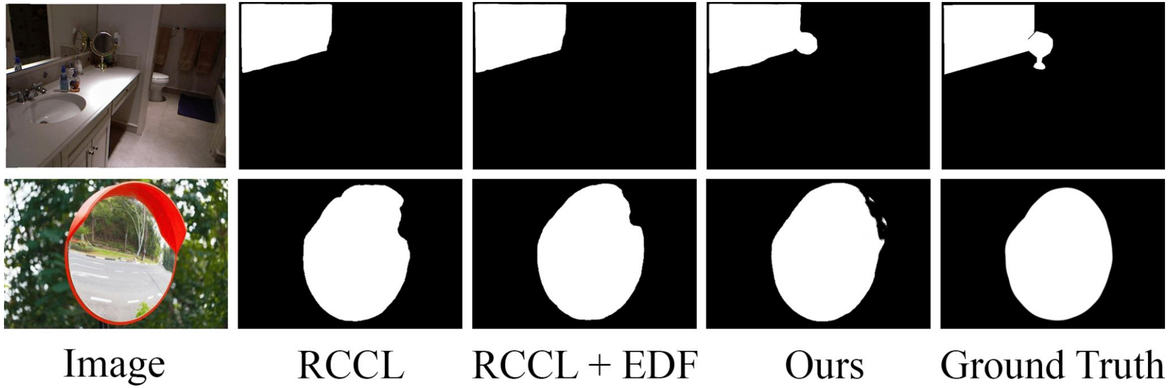
Performance of the Pruned Model

	MSD		PMD		DLSU-OMRS	
	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓
Unpruned	0.8483	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
Before Retraining	**0.8505**	0.4185	0.7858	0.4585	0.8432	0.4407
After Retraining	0.8498	0.0813	0.7902	0.0364	**0.8456**	**0.0955**



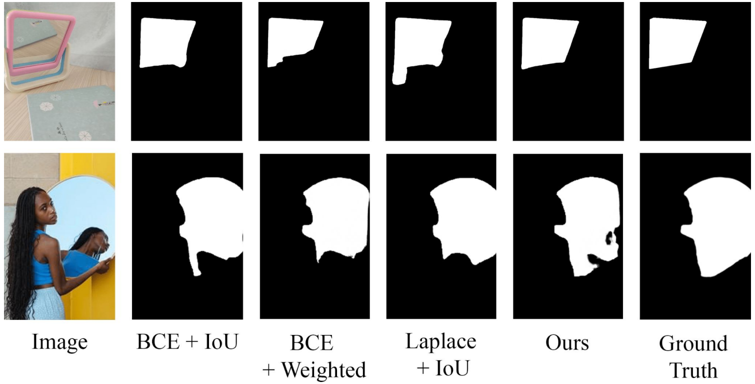
Performance of the Pruned Model

	MSD		PMD		DLSU-OMRS	
	F_β ↑	MAE ↓	F_β ↑	MAE ↓	F_β ↑	MAE ↓
RCCL	0.8052	0.0957	0.7957	0.0332	0.8300	0.0956
RCCL + EDF	0.8224	0.0949	0.8001	0.0335	**0.8389**	**0.0918**
Ours	**0.8483**	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032

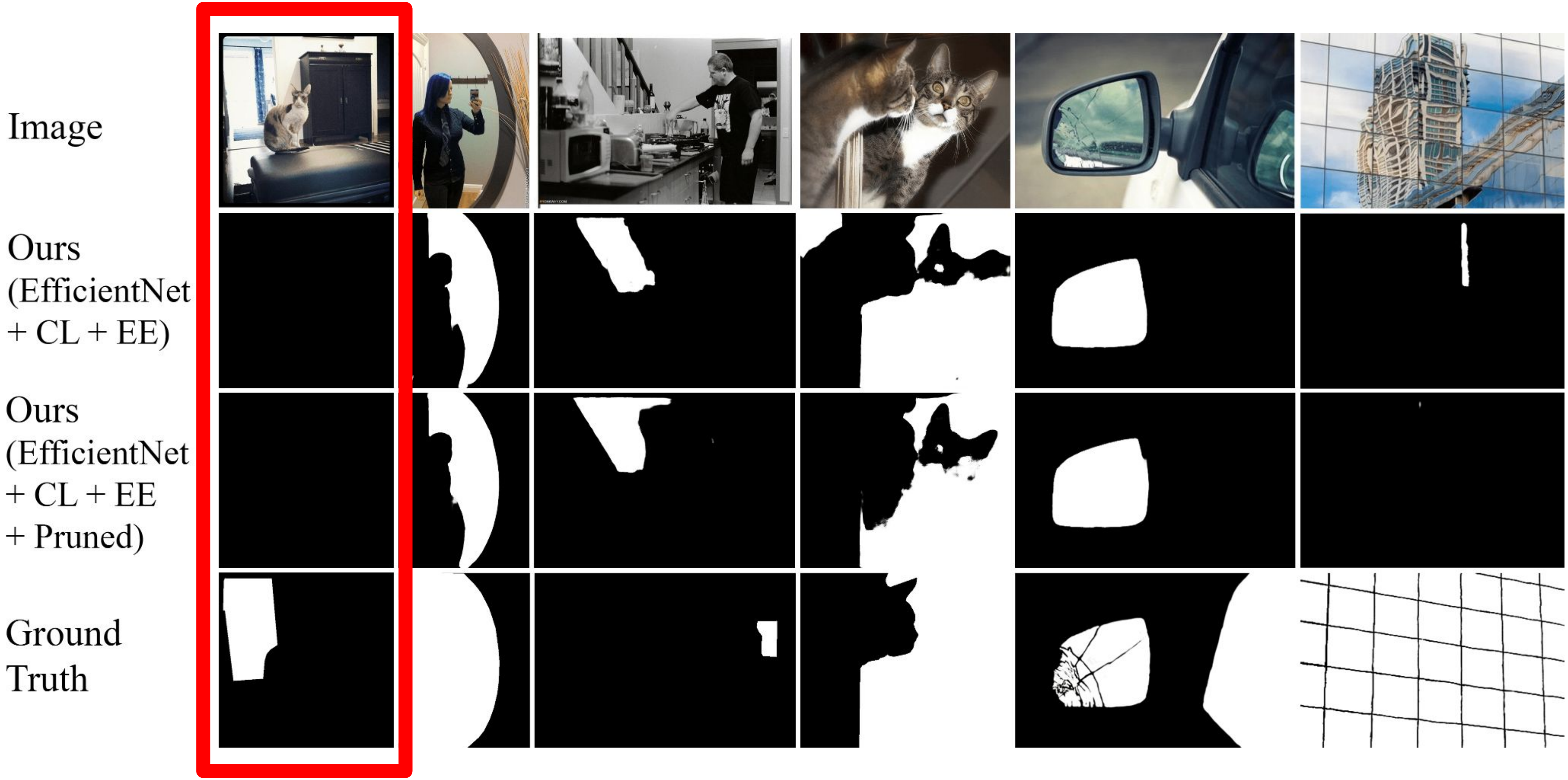


Performance of Ablated Models

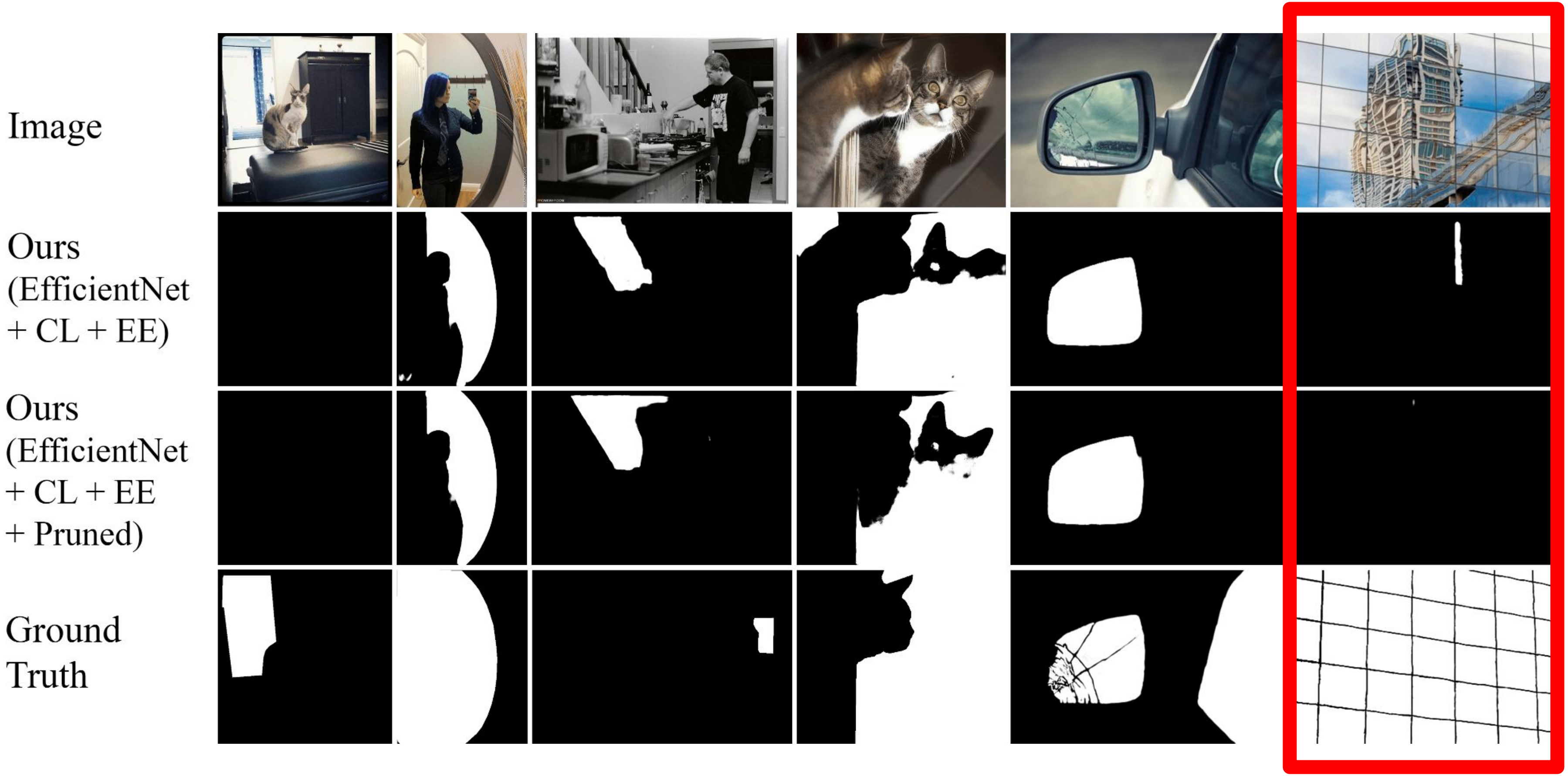
Loss	MSD		PMD		DLSU-OMRS	
	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓	$F_\beta \uparrow$	MAE ↓
BCE + IoU	0.8352	0.0949	0.8038	0.0320	0.8314	0.0969
BCE + Weighted	0.8163	0.0967	0.8073	0.0319	0.8470	0.0995
Laplace + IoU	0.8148	0.0950	0.7989	**0.0302**	**0.8553**	**0.0881**
Ours	**0.8483**	**0.0800**	**0.8117**	0.0313	0.8388	0.1032



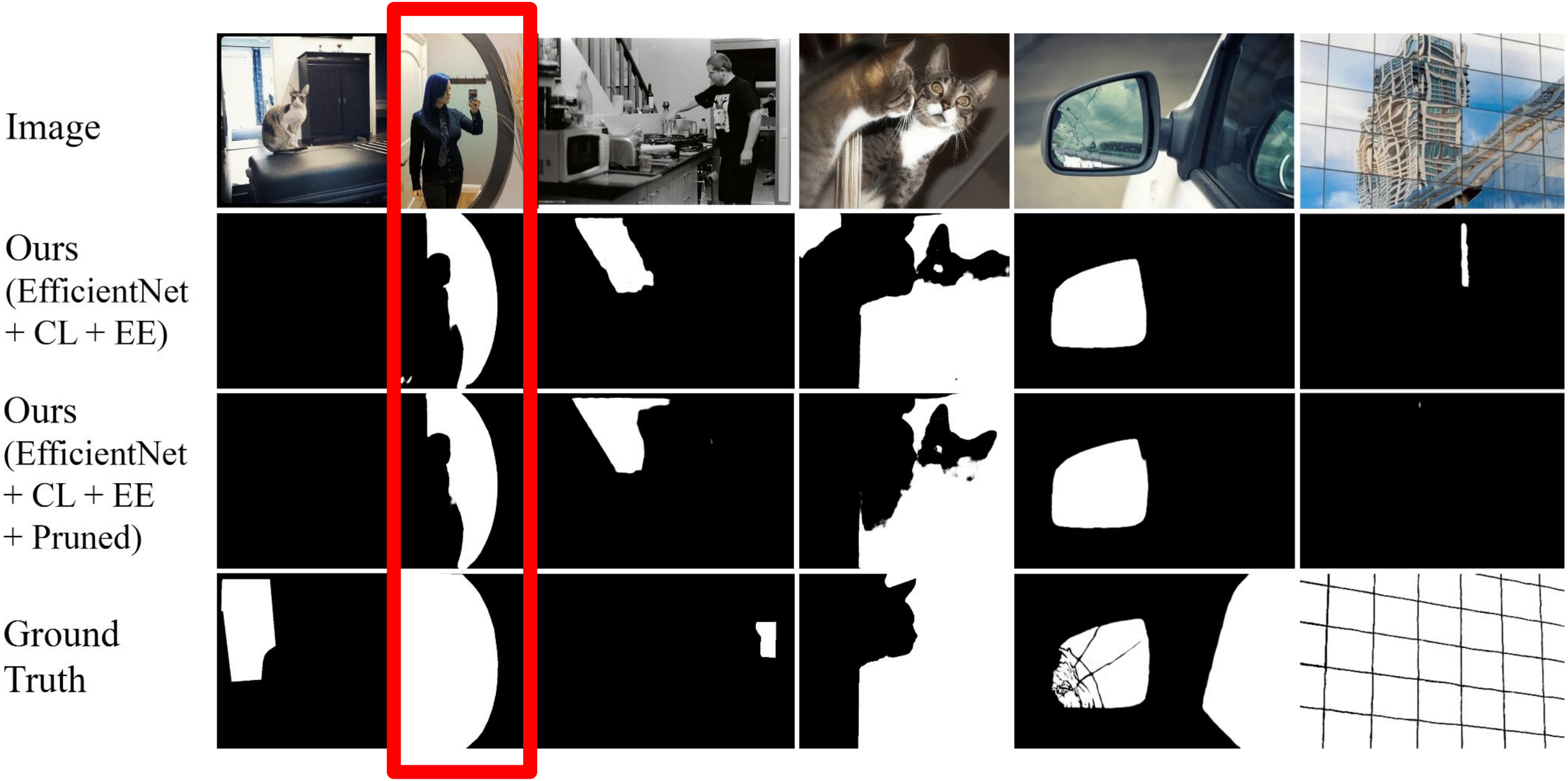
Performance Under Different Loss Functions



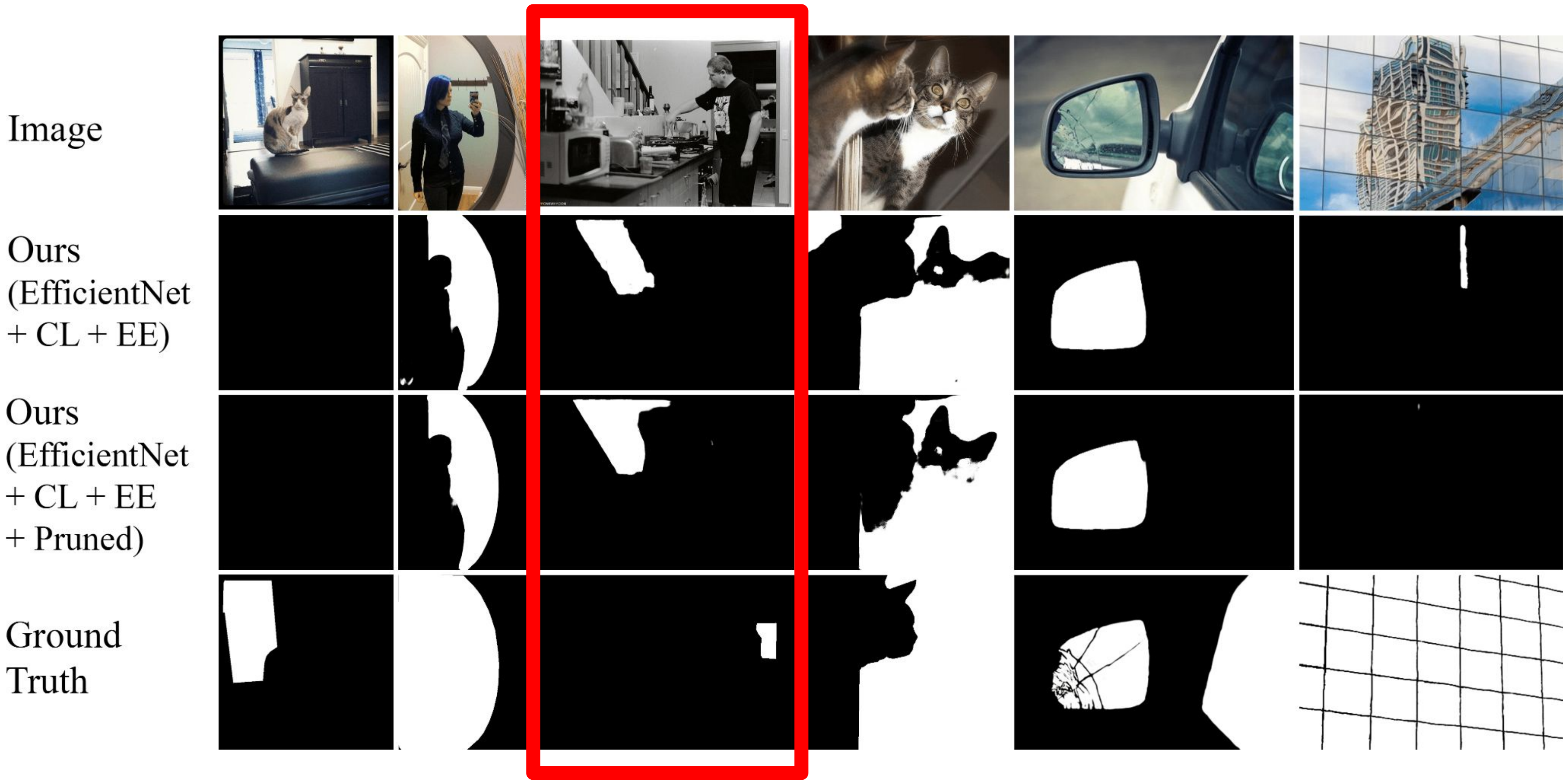
Failure Case: Contextual features inside and outside the mirror appear continuous



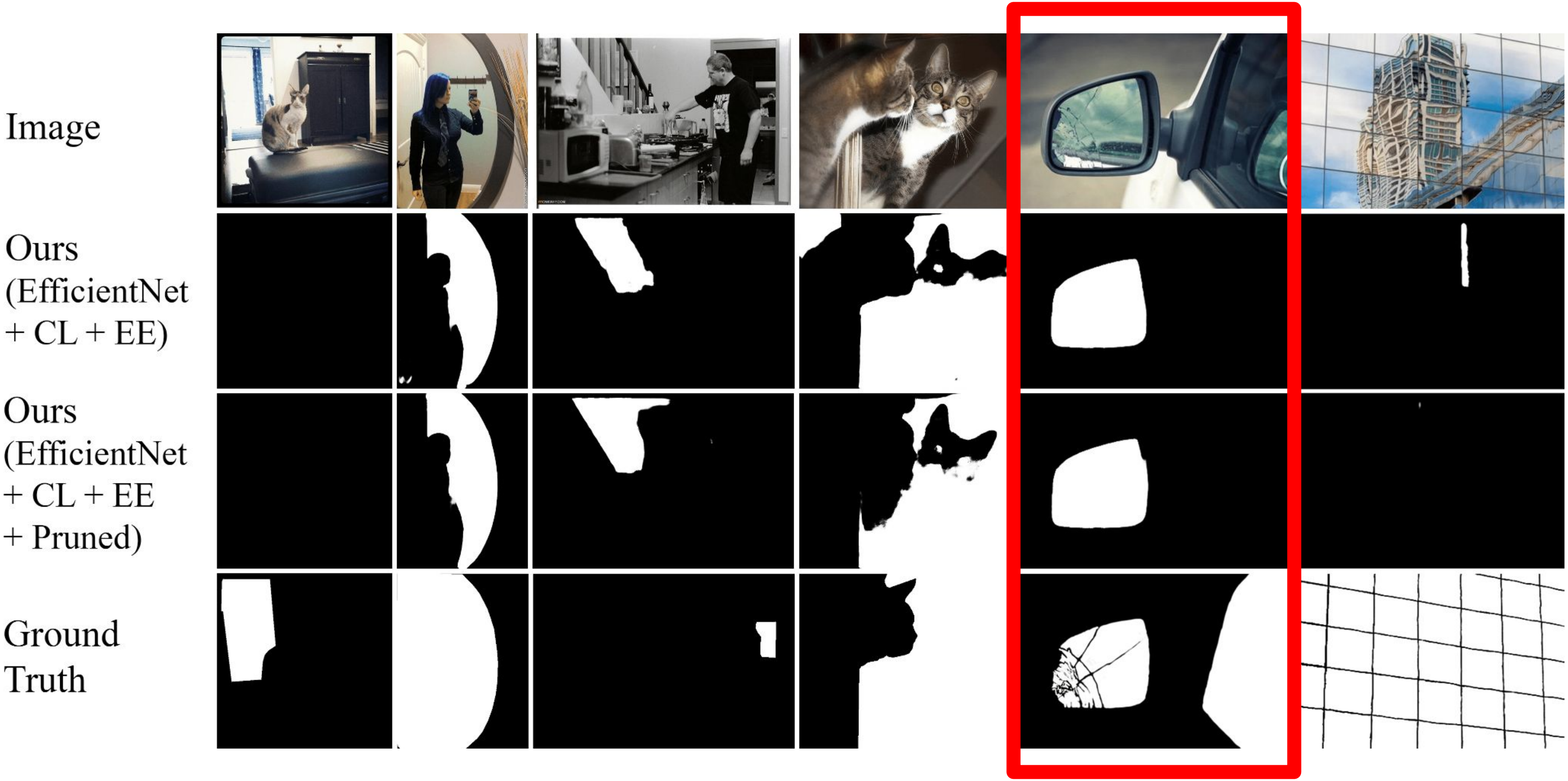
Failure Case: Available contextual features are inadequate



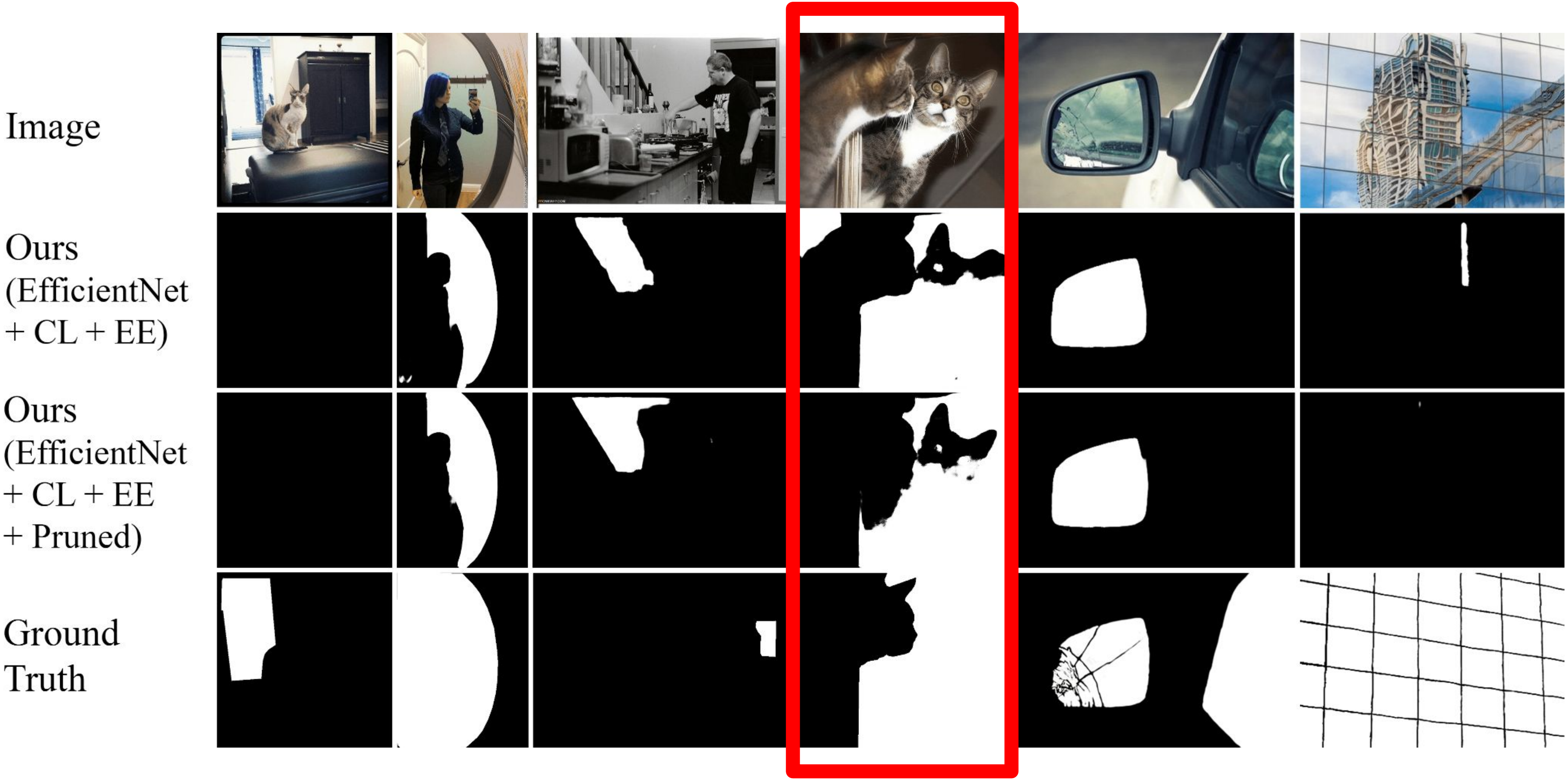
Failure Case: Sharp discontinuities are present within the mirror



Failure Case: Small mirrors in the background may be challenging to recognize



Failure Case: Heavily tinted reflective surfaces may be challenging to recognize



Failure Case: Some cases are challenging even for human observers

Conclusion

Conclusion

- We propose DLSU-OMRS, a **dataset of 454 outdoor mirrors and reflective surfaces** not well represented in existing mirror datasets

- We modified **PMDNet** architecture with different **feature extraction backbones and edge-related modules** to guide segmentation

GitHub

Conclusion

- Our best-performing model uses **EfficientNetV2-Medium** as its backbone
- Low-level edge features are captured via **parallel convolutional layers**
- High-level edge features are captured via a lightweight convolutional block attention module

GitHub

Conclusion

- **EfficientNetV2-Medium + Compound Loss + Edge Extraction**
    - F_β: 0.8483 (MSD), 0.8117 (PMD), 0.8388 (DLSU-OMRS)

- **Compressed model by filter pruning via geometric median**
    - F_β: 0.8498 (MSD), 0.7902 (PMD), 0.8456 (DLSU-OMRS)
    - 78.20× fewer FLOPS
    - 238.16× fewer parameters

GitHub

International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2023)

Designing a Lightweight Edge-Guided Convolutional Neural Network for Segmenting Mirrors and Reflective Surfaces

Mark Edward M. Gonzales
mark_gonzales@dlsu.edu.ph

Joel P. Ilao
joel.ilao@dlsu.edu.ph

Lorene C. Uy
lorene_c_uy@dlsu.edu.ph

De La Salle University
Manila, Philippines

github.com/memgonzales/mirror-segmentation

Model	Cues	Backbone	Limitations
MirrorNet (Yang *et al.*, 2019)	Contextual contrasted features	ResNeXt-101	● Blurry boundaries for some masks ● Difficulty in handling cases with insufficient contextual contrast between objects and reflection
PMDNet (Lin *et al.*, 2020)	Relational contextual contrasted features, multi-scale edge features	ResNeXt-101	● Blurry boundaries for some masks ● Difficulty in handling cases with insufficient correlational features inside and outside the mirror
PDNet (Mei *et al.*, 2021)	Color and depth discontinuities and correlations	ResNet-50	● False flagging of doorways as mirrors ● Difficulty in handling cases where discontinuities are hard to discern

Yang, X. et al. Where is my mirror? Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 8809-8818, 2019.
Lin, J., Wang, G., and Lau, R.H. Progressive mirror detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3694-3702, 2020.
Mei, H. et al. Depth-aware mirror segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3044-3053, 2021.

Model	Cues	Backbone	Limitations
SANet (Guan *et al.*, 2022)	Semantic associations	ResNeXt-101	• Inability to detect mirrors if semantic association labels are inadequate • Reliant on semantic annotations, which may not be readily available all the time
VCNet (Tan *et al.*, 2022)	Visual chirality	ResNeXt-101	• Difficulty in identifying boundaries of occluding objects with complex structures or shapes • Difficulty in excluding small occluding objects

Guan, H., Lin, J., and Lau, R.W.H. Learning semantic associations for mirror detection. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5931-5940, 2022.
Tan, X. et al. Mirror detection with the visual chirality cue. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1-13, 2022.

Model	Computational Complexity GFLOPS ↓	# of Params ↓	MSD F_β ↑	MSD MAE ↓	PMD F_β ↑	PMD MAE ↓	DLSU-OMRS F_β ↑	DLSU-OMRS MAE ↓
VST [Liu21]	46.36	44.48M	0.4290	0.2739	0.1317	0.261	0.5730	0.2274
PMDNet [Lin20a]	118.86	147.66M	0.8350	0.0816	<u>0.8011</u>	<u>0.0324</u>	<u>0.8423</u>	**0.0878**
Ours (Compound Loss)								
ResNet-50	105.47	129.04M	0.7548	0.1119	0.7650	0.0403	0.7874	0.1011
Ours (Compound Loss + Edge Extraction)								
ResNet-50	116.46	130.12M	0.7695	0.1098	0.7524	0.0409	0.8042	0.1025
Xception-65	75.28	129.12M	0.7800	0.0973	0.7566	0.0401	0.7643	0.1164
VoVNet-39	98.25	61.90M	0.7014	0.1196	0.7578	0.0412	0.7868	0.1088
MobileNetV3	<u>6.61</u>	20.76M	0.7515	0.1153	0.7508	0.0427	0.8256	0.1006
EfficientNet-Lite	6.99	15.54M	0.7909	0.1027	0.7769	0.0387	0.8178	0.1048
EfficientNet-Edge-Large (Pruned)	17.02	<u>10.42M</u>	0.7682	0.1082	0.7831	0.0349	0.8035	0.1044
EfficientNetV2-Medium	24.79	53.35M	<u>0.8483</u>	**0.0800**	**0.8117**	**0.0313**	0.8388	0.1032
Ours (Compound Loss + Edge Extraction + FPGM Pruning)								
EfficientNetV2-Medium	**1.52**	**0.62M**	**0.8498**	<u>0.0813</u>	0.7902	0.0364	**0.8456**	<u>0.0955</u>