Beyond Reflections: Proposing Datasets and Labels to Achieve Accurate Discrimination between Glass and Mirror Objects

Kihyun Kim
Seoul National University
ki5477@snu.ac.kr

Soo Yong Kim
Seoul National University
ksyint1111@snu.ac.kr

Sung-eun Jang
Seoul National University
jse9512@snu.ac.kr

Abstract

Accurate segmentation of glass and mirror objects in image poses a significant challenge for data-driven models. The visual characteristics of glass and mirrors lack visual form, leading to misidentify objects that appear on these surfaces. Consequently, detecting glass and mirrors within images has remained an unresolved problem for quite some time. Previous approaches have attempted to tackle this issue by using features of mirror and glass. However, these methods and datasets are only studied for mirrors or glasses each, falling short of providing a more general segmentation model. Therefore, we made new dataset with 850 images with mirror and glass together. We tested our dataset with previous models and found that models confuse mirror and glass when they are together in the image. By this result, we argue that new model is needed to distinguish mirror and glass.

1. Introduction

The development of deep learning models is showing growth in various fields, and performance in the field of vision is affecting various domains. A lot of progress has also been made in image segmentation. The high accuracy and performance of deep learning-based methodologies have significantly outperformed traditional computer vision technologies. However, there are still various limitations to overcome, one of which is the segmentation of mirror and glass objects.

Both mirror and glass do not have a clear visual shape and usually contain shape of other objects. These reflections are caused by reflected light and projected light. This prevents the model from clearly analyzing the objects contained within the image. In addition, glass sometimes has clear reflection as a mirror due to differences in illumination. Because of these similarities, the mirror or glass segmentation models have been studied using similar approaches.

However, most of the studies so far have targeted either mirror or glass, and distinguishing the two has not been noted. In addition, most of the datasets mainly used in these studies are also designed focusing on only one of these two targets. However, mirrors and glass have very high visual similarities, and incorrect classification of these two objects can prevent the model from clearly understanding the space within the image.

To understand this problem situation, we propose a new dataset consisting of 850 images in which glass and mirrors appear simultaneously and add novel segmentation class, glass in mirror. In addition, we tested our dataset with six mirror or glass segmentation models. By this result, we found that current models fail to segment when mirror and glass are together in images. Our dataset and new label "glass in mirror" can help further researchers to develop mirror and glass segmenting models.

2. Recent study

Both mirror and glass segmentation models are similar, they use single image as input, and output as masked image where glass or mirror is white, else black. Figure 1 show example of model input, output mask and ground truth mask of recent models.

Figure 1. Example of mirror or glass detecting models. Right : Input image from GSD-S dataset including glass(or mirror) object. Middle : Ground truth mask of input image. Left : Prediction from GlassSemNet[10]. white pixels are predicted as glass

2.1. Mirror segmentation models

First mirror segmentation model MirrorNet[21] came out in 2019. The model learns discontinuity in both low level patterns and high level semantics. The model used ResNet[12] as backbone and extracted features at every level of ResNet. Low level features are used to extract pattern, pixel discontinuity. High level features are used to extract semantic discontinuity at the edge of mirrors. The idea of using ResNet as backbone and extracting every level feature became basic idea for mirror or glass segmentation in image after this model. Next model is PMDNet[9] in 2020. The model added mirror boundary detection and learns contextual contrast between inside and outside of mirror. This increased model performance but couldn't understand the image with large mirror, unable to find contrast. MirrorSemNet[5] learns semantic association between mirror and other objects in the image. They built the model in two part, first part is segmenting and classifying objects in image. The second part is learning semantic association by using semantic segmentation. By learning semantic association, the model can predict most possible place for mirror in the image. After this, researchers focused more on mirror properties and model efficiency.

Model [20] added visual chirality to PMDNet[9] model. Visual chirality is concept that model can understand difference between horizontal flipped image and not. They convolved visual chirality filter in pixel level to find out if some region in image is flipped or not. If it is flipped, the model thinks the region is inside mirror. Model [3] used attention and YOLO(You Look Only Once)[11] to detect mirror, focused on detecting correct bounding polygon. Lastly, model [15] uses different feature extracting CONVnet at high and low level from backbone ResNet. The model only use CNN and pooling, batch norm. The model is light weight but still performs quite well, better accuracy than [21,9,5]. They used input augmentation of rotating 90 degrees, by comparing normal and rotated image, the model can learn useful features.

2.2. Glass segmentation models

Glass segmentation is much harder than mirror because reflection in the glass is transparent, overlapped with background. Also, if the glass window is very clean, there is no reflection. In this case, even human can't detect glass in the image. So glass segmentation models were developed after mirror segmentation model. GDNet[6] is the first glass segmentation model. This model is similar to MirrorNet[21]. The model added concatenation layer, merging high level and low level feature after feature extraction. They also tested GDNet with MSD dataset, miror dataset proposed by MirrorNet[21]. The accuracy score outperformed MirrorNet[21]. Models [13,8] detect light scattering and reflection in glasses. This improved model accuracy but didn't fit well in clear glass images. GlassSemNet[10] is similar to MirrorSemNet [5], learning semantic association of glass and other objects in images. Model [22] use ensemble method to outperform and focus on reflection in the glass.

3. Benchmark datasets

We use two benchmark datasets, PMD[9]and GSD-S[10]. PMD dataset contains 6,461 mirror images with ground truth masks. Images are gathered from 6 major image datasets, ADE20K[2], NYUD-V2[14], MINC[18], Pascal-Context[16], SUNRGBD[19] and COCO-Stuff[4]. They filtered images with mirror label in it and relabelled mirror mask by hand.

GSD-S dataset[10] contains 4,519 images with glass object in it. Images are gathered from 4 big datasets, SUNRGBD[19], 2D-3D Semantics[7], MatterPlot3D[1], and COCO-Stuff[4]. It is relabelled by human like PMD dataset and also has ground truth semantic segmentation. There are total 43 object classes in ground truth.

4. Proposed dataset

We downloaded about 1,000 glass and mirror related images from Google using Chrome extension "Download All Images" [A]. This extension helps download all images in current web page to zip file. We searched images in Google by keywords like "Glass and mirror", "Mirror like Glass", "Glass like Mirror", "Mirror in bathroom", "Glass in bathroom" and so on. Also, we used negative filters like "only mirror" or "only glass". By using negative filters we tried to get images with mirror and glass together as many as possible.

After downloading, we filtered out some unwanted images from our dataset. We built our own standard for choosing the appropriate images. First we filtered out images without glass or mirror, since we need at least one of these objects. Next we filtered out images with large text, images with large white background, and same image with different size. After filtering we got 850 glass or mirror image dataset.

We considered three labels, mirror, glass and glass in mirror. We colored mirror with green, glass with red, and glass in mirror with blue for GT mask. We thought glass in mirror label important because it looks like glass region but it is not a real glass, just reflection inside mirror. If the model can really distinguish mirror and glass, it shouldn't segment glass in mirror region as glass since it is not real. We didn't consider "mirror in glass" label because it is okay for model to recognize mirror in glass as mirror. When human look at glass, we can see recognize objects inside the glass is real and it's behind the glass. Also, there were not many cases that has mirror inside glass region.

In annotation process we used "CVAT" [B], which is open platform for making mask for segmentation by hand. Since our benchmark datasets PMD[9] and GSD-S[10] also have hand-made mask, we tried to make masks similar to

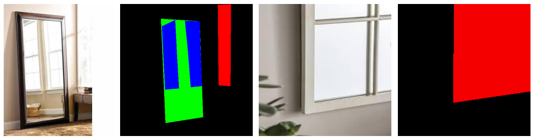

Figure 2. Examples of our dataset. First two images from left shows image with GT mask having all three labels. Other two images shows window shaped mirror example from our dataset. We can see window frame is reflected by mirror.

benchmark datasets by having three standards. First, we only used straight lines to make mask regions. Circular or oval shape were also made by many straight lines like benchmark datasets. Second, we couldn't split masks by all window lattices in image because Window size and image size are all different. So we made our own rule, we split mask only when inner lattice is thicker than half of outer lattice. Third, in low quality image, it is hard to recognize mirror and glass. We used mirror and glass reflection feature to find them. If reflection is clear, we label it as mirror. When reflection is transparent, it is glass. When we can't find reflection properties, we searched for light scattering or edge of mirror or glass to label masks.

In total we had 447 glass labeled images, 489 mirror labeled images, and 267 mirror and glass together. Mirror and glass together includes 181 images with glass in mirror label. Figure2 shows two examples of our 850 datasets. Our 850 glass and mirror dataset not only has glass and mirror together images, but also has glass like mirror or mirror like glass, which might be harder for models to recognize.

5. Experiment

We selected six benchmark models[5,6,8,9,10,21] which use similar features to segment mirror or glass. In glass segmentation models[6,8,10], the authors said they take similar approach like mirror segmentation models because mirror and glass are similar. Thanks to this idea, we classified models using similar methods. Models[6,21] use discontinuity at edges of mirror or glass, models[8,9] use object reflection in mirror and light scattering in glass, which are properties by reflection. Models[5,10] use semantic association of objects in image. Since mirror and glass are similar and segmentation models use similar features, we assumed that these models can't distinguish mirror and glass when they are together. So we tested these six models by our dataset with mirror and glass together. We evaluated glass and mirror models pairwise to compare similar models. Three datasets are used for testing, PMD[9], GSD-S[10] and our 850 dataset. We show qualitative evaluation by output images and then show quantitative results all together in section 6 for fair comparison.

5.1. Semantic association models

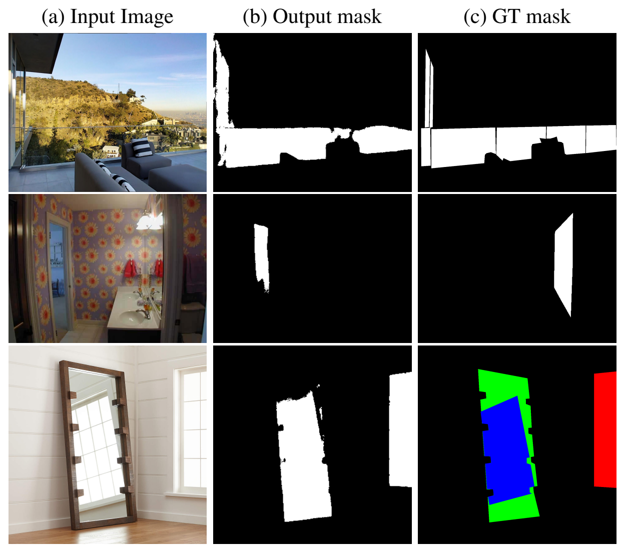

Figure 3. Sample outputs by model in [10], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

Samples of glass model[10] output images are shown in figure 3. The model is trained with GSD-S dataset and tested with GSD-S, PMD, and our dataset. In Figure 3 row 1, image is very similar to ground truth since the model is trained by same dataset. All other outputs had high similarity to ground truth. Second row of figure 3 is tested with PMD dataset, which is mirror image. The model didn't think mirror as glass but masked opened door as glass. There were many other examples in PMD dataset like row 2, where model thought open window or bright part as glass. We think the reason is because the model only learns semantic relation, doesn't think of glass property like reflection.

Lastly, we tested the model with our dataset. Our dataset has many images with glass and mirror together. In Figure 3 row 3, the model classified window(red) and glass in mirror(blue) as glass quite well, but confused edge of mirror(green) and glass in mirror(blue) region. However, when there is no glass in image, model tends to recognize mirror as glass.

In table 1, we can see that IOU value is lower than GlassNet[8], but lower(better) MAE score. It is because GlassSemNet[10] often outputs gray colored output with slight brighter on glass or mirror. IOU metric only concerns 0 or 1, so IOU value of gray scale output is 0. But in MAE it is better than wrong segmentation of other object as glass.

You can find the result of model[5] in figure 4. The model is trained with PMD dataset and tested with GSD-S dataset, PMD dataset, and the 850 dataset we suggest. As we can easily expect, we can confirm from the second row

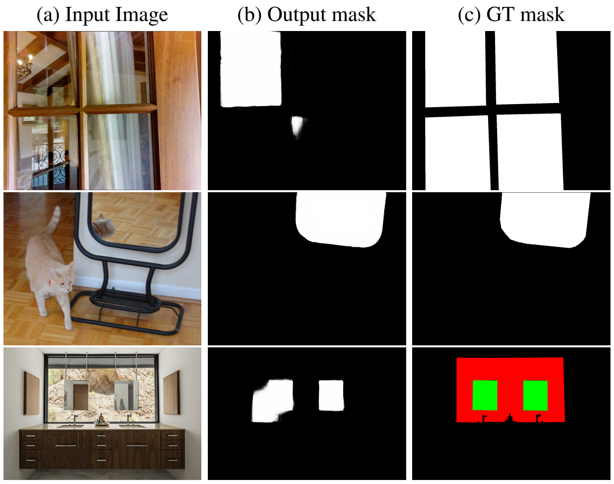

Figure 4. Sample outputs by model in [5], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

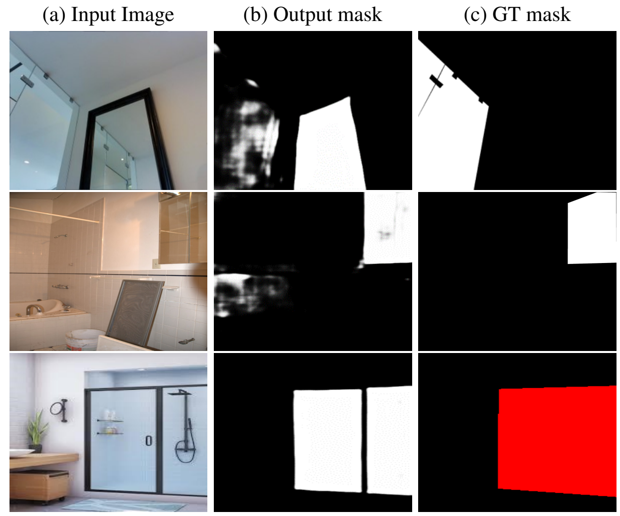

Figure 5. Sample outputs by model in [6], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

of figure 4 that the prediction results of the model for PMD dataset are very similar to ground truth. These results were the same for all image data used in the experiment. However, the results of the GSD-S data, which can be found in the first row, show that the model is disturbed by the glass object in the image. We can interpret that the model did not understand the difference between glass and mirror, and it was also confirmed in other image data used in the experiment. The limitations of this model can be confirmed more clearly in the third row of Figure 4. The model clearly predicted the location of the target mirror object. However, it was disturbed by glass objects around the mirror, causing an error in predicting some of the glass region as a mirror. This decline in precision was observed equally in images in which mirror and glass objects existed at the same time. In particular, the closer the mirror and glass were placed on the image, this became severe.

This trend can also be confirmed through the figures in Table 1. The IOU of the model[5] in the PMD targeting the mirror are higher than the other five comparison models, and the MAE is also relatively low. And the output for GSD-S is the opposite, which can be interpreted that the model shows good performance for simple forms of input images. However, it shows relatively low performance for image sets that include glass-in-mirror and glass near mirror cases like our proposed dataset.

5.2. Learning discontinuity models

The predicted outputs from model[6] show in Figure 5. As can be seen from Figure 5, the model effectively detected objects with discontinuity in the image. However, we confirm that there are still some limitations in the model. It can be seen in the first row of Figure 5 that the model is also difficult to clearly distinguish between glass and mirror. The discontinuity in the image is more pronounced in the mirror object than in the glass object, which makes the model seem to have confused the mirror contained in the image with the glass object. Also, in second row of Figure 5, the model tended to recognize some mirror objects contained in the PMD dataset as glass. Interestingly, it seems like the model confused the lines caused by the black tiles in the input image as a discontinuity. Nevertheless, for most of the given input images, the model provided excellent prediction output. As shown in the last row of Figure 5, the model showed high performance by accurately segmenting the frame of the window.

You can additionally confirm this through the differences in the model's IOU, MAE, and F1 score for GSD-S targeting glass objects and the dataset we propose, and PMD containing mirror objects from Table 1. The model produced a significantly different output from the ground truth mask in the mirror dataset, which shows that the model is working correctly.

The MirrorNet[21] was trained by mirror dataset named MSD, proposed by MirrorNet[21]. MSD dataset is previous version of PMD dataset which has less images in it. In Figure 6, we can see the output of MirrorNet[21]. In first row, we can see the model recognizes glass as mirror. Even in PMD dataset the model often predicted wrong. Digging deeper into the flaws of the model, when the glass is at the center and the mirror is at the side or further spot from the focus of camera, the model often masked the glass. Finally in row 3 we can see the model got wrong output even though mirror is big and centered in the image. In table 1 we can see the model got low IOU and MAE value in all three datasets.

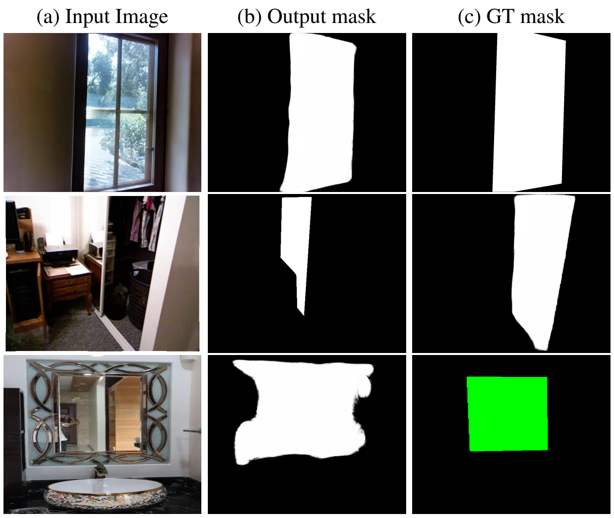

Figure 6. Sample outputs by MirrorNet[21], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

Since the model only find edge of mirror, it often misunderstands mirror and glass.

5.3. Mirror reflection and Glass scattering models

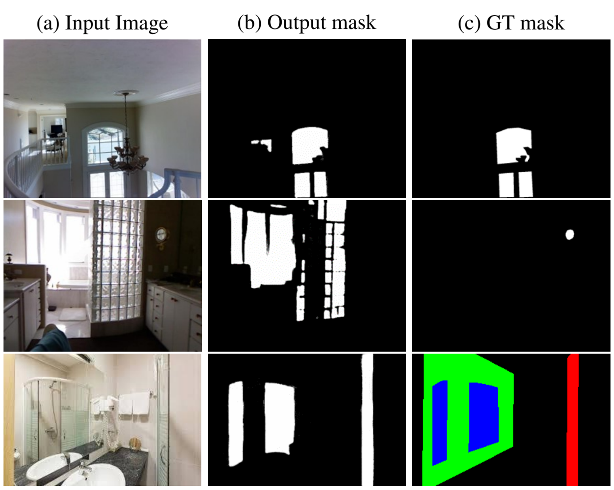

Figure 7. Sample outputs by model in [8], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

The experimental result image of the glass segmentation model[8] can be found in Figure 7. The model accurately segmented glass objects in both PMD and GSD-S. In particular, we can see that the model output for GSD-S in the first row detects even a small glass not shown in the groud truth.The IOU, F1 score, and MAE for GSD-S identified in Table 1 can also be confirmed through good figures of the

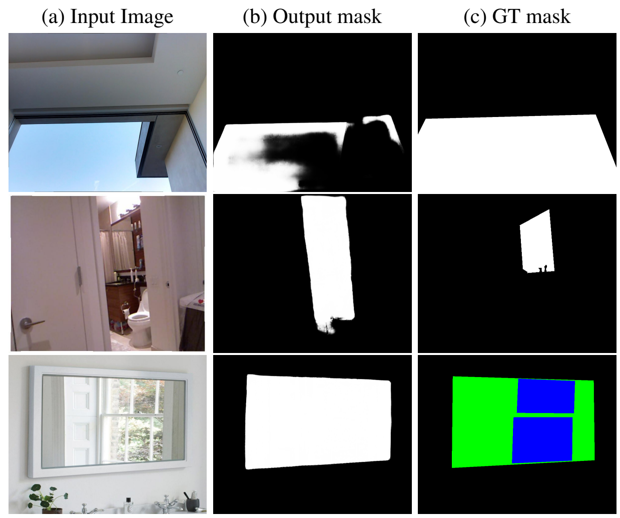

Figure 8. Sample outputs by PMDNet[9], First row : GSD-S[10] dataset, Second row : PMD[9] dataset, Third row : our glass and mirror dataset

model.The values for the PMD also indicate that the model can distinguish the mirror from the glass to some extent.

However, this model, like other glass segmentation models, can also be seen to make errors in the glass-in-mirror case, which is included in our proposed datasets. This can be confirmed once again in the figures in Table 1. The model shows relatively low performance on the dataset proposed by GSD-S.

In Figure 8 we can see the result of PMDNet[9]. PMDNet also have similar disadvantage as MirrorNet[21]. In first row of Figure 8, PMDNet segmented part of glass as mirror. This happened when background of glass was bright. There were many incorrect predictions like this in GSD-S dataset. Also in second row tested with PMD dataset, we found that when the mirror is surrounded by lots of bright lights the model hardly does masking although this model is trained with PMD dataset. Lastly, when there are many objects inside mirror reflection, The model doesn't catch the mirror region and only catch other objects inside the mirror.

In Figure 8 row 3 however, the model segmented mirror quite well compaired to other models. The model had same flaw in bright images as other datasets. In table 1 we can see PMD[9] model got highest IOU score. The model also got high score in GSD-S dataset, which means the model thought glass as mirror when there was only glass dataset. We think it is because of the flaw of model, segmenting wrong in bright images.

Dataset		PMD			GSD-S			Ours		
Models	Venue	IOU	F_β	MAE	IOU	F_β	MAE	IOU	F_β	MAE
MirrorNet[21]	ICCV 2019	0.3580	0.4148	0.1594	0.3430	0.5225	0.1591	0.2312	0.3522	0.1933
PMDNet[9]	CVPR 2020	0.4002	0.8980	0.0452	0.3810	0.5342	0.1391	0.4532	0.3642	0.1462
MirrorSemNet[5]	CVPR 2022	0.6684	0.8437	0.0493	0.1686	0.4795	0.1948	0.1090	0.3266	0.1791
GDNet[6]	CVPR 2020	0.1794	0.2721	0.3007	0.3745	0.5154	0.3347	0.2498	0.3799	0.3635
GlassNet[8]	CVPR 2021	0.1137	0.2531	0.5462	0.721	0.821	0.061	0.3512	0.4720	0.3001
GlassSemNet[10]	NeurIPS 2022	0.2469	0.4152	0.1072	0.7575	0.8556	0.0346	0.2935	0.4514	0.1318

Table 1. Top three rows are mirror models, evaluated with mirror mask of our dataset. Down three are glass model, evaluated with glass mask of our dataset. Red colored scores are scores by opposite mask(ex. glass models evaluated with mirror mask). If red scores are high, it means the model confuse mirror and glass.

(a) Input Image (b) GTmask (c) MirrorNet [21] (d) PMDNet [9] (e) MirrorSemNet [5] (f) GDNet [6] (g) GlassNet [8] (h) GlassSemNet [10]

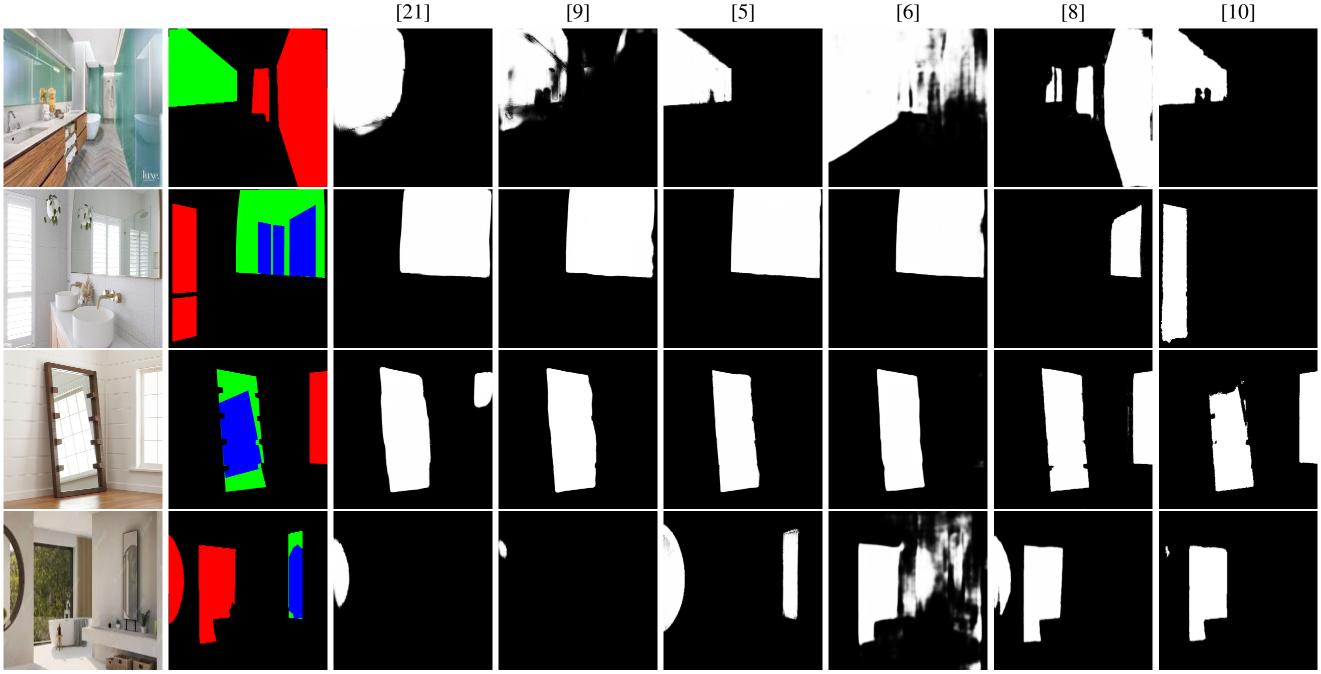

Figure 9. Sample outputs from six models with images from our dataset

6. Result

For evaluation, we used IOU(Intersection of Uinion), MAE(Mean Absolute Error) and F-beta(Maximum beta score) score with beta = 0.3. We changed our GTMask images to Gray scale images for evaluation. For mirror models[5,9,21], we colored only green mask(mirror) as white and everything else black. Mask for Glass models[6,8,10] vice versa.

In table 1 , we can see evaluation score of six models by three datasets. Both edge detection models [21,6] didn't segment well as we see in section 5.2. MirrorNet[21] got lowest IOU in PMD dataset compared to other mirror segment models and GDNet[6] vice versa. The reflection and light scattering detection models[9,8] got highest IOU value in our dataset. In section 5.3, we saw that these models works pretty well in PMD and GSD-S dataset. But PMDNet[9] also got high IOU value in GSD-S dataset, which means the model thinks glass as mirror when there is only glass in image. We think it is because of brightness that confuse model that we mentioned in section 5.3 PMDNet. GSD-S dataset consists of many glass window, which is bright. Lastly, semantic association understanding models [5,10] got highest performance in their own dataset PMD and GSD-S, but didn't work well in our datset. We think it is because these models only understand nearby object relation and no understanding of mirror or glass features.

To sum it up, models detecting reflection or light scattering in image got highest performance in our dataset. Edge detection models and semantic association models work poorly when mirror and glass are together in one image.

Even human often use light scattering in glass and reflection in mirror to distinguish them. So it is easy to understand that these two models found way to distinguish mirror and glass. However, their performance decreased compared to glass or mirror only images. Also, there are many cases when mirror models detect glass or glass models detecting mirror. In Figure 9 (row 1 col h) and (row 3 col f), we can see glass models detect mirror as glass. So current models confuse mirror and glass in image when they are together.

7. Limitation and further research

We expect our proposed new label ”Glass in the Mirror” to be of great help to future research. Our dataset not only allows glass segmentation models to make decisions by considering the re-section in the mirror, but can also be used to evaluate existing models. In addition, the proposed dataset can also help the mirror segmentation model. The model will be able to receive information through the dataset that can more clearly distinguish the similar characteristics of glass objects and mirror objects.

Nevertheless, there are several distinct limitations in our dataset in terms of the number and quality of data. We initially planned a dataset containing 1000 images, but we didn’t achieve this. It takes a lot of time and human resources to detect mirrors and glass objects to create a ground truth mask. In addition, when acquiring images, Google uses the results of searching for specific keywords, which took a lot of time to select the right image for the data set. In addition, these images often appear in duplicate and have different sizes, which cost a lot of effort to preprocess.

The second problem is that most of the images with mirror and glass are taken in bathroom. This problem was also in benchmark dataset PMD[9] and GSD-S[10]. PMD images were usually taken in bathroom or darker place and GSD-S were taken with outside, bright light. It is obvious that most common place where mirrors and glass exist at the same time is the bathroom. The bias of dataset can be an obstacle on the training of deep learning-based models in the future.

Third problem is that in some cases, distinction between glass and mirror is challenging for humans, too. There were some cases like very clear glass window or very small glass or mirror that we couldn’t label correctly. Models are affected by image size too. In Figure 9 last row, we can see most models didn’t perform well compare to other images. It is because the image size was small, making hard for model to detect objects. Also in row 1 of Figure 9, mirror and glass is distorted by position of the camera, which also decreased accuracy of models.

These limitation in dataset are open to improvement in the future, and we expect them to be used in various computer vision task studies if they are successfully addressed.

8. Conclusion

We propose new dataset consisting of 850 images with glass and mirror objects. Also we made new label glass in mirror that can help further researchers make model that can distinguish mirror and glass. We evaluated our dataset with six recent models, which use similar approach in segmenting glass and mirror. Models using reflection and light scattering properties worked best on our dataset, understanding difference between mirror and glass. But all six models performed lower in our dataset compared to mirror or glass only dataset.

References

[1] A. Chang, et al. Matterport3d: Learning from rgb-d data in indoor environments. In International Conference on 3D Vision (3DV),arXiv, arXiv:1709.06158. 2017.

[2] B. Zhou, et al. Scene parsing through ade20k dataset. In Conference on Computer Vision and Pattern Recognition (CVPR), 2017. p. 5122-5130.

[3] F. Li, et al. Mirror-Yolo: An attention-based instance segmentation and detection model for mirrors. arXiv preprint arXiv:2202.08498, 2022.

[4] H. Caesar, J. Uijlings, and V. Ferrari. Coco stuff: Thing and stuff classes in context. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018. p. 1209-1218.

[5] H. Guan, J. Lin and W.H Lau. Learning Semantic Associations for Mirror Detection. In Conference on Computer Vision and Pattern Recognition(CVPR). 2022. p. 5941-5950.

[6] H. Mei, et al. Don’t hit me! Glass Detection in Real-world Scenes. In Computer Vision and Pattern Recognition (CVPR), 2020, p. 3687-3696.

[7] I. Armeni, et al. Joint 2D-3D-Semantic Data for Indoor Scene Understanding. arXiv preprint, arXiv:1702.01105, 2017.

[8] J. Lin, Z. He and R. W. H. Lau. Rich Context Aggregation with Reflection Prior for Glass Surface Detection, In Conference on Computer Vision and Pattern Recognition (CVPR),2021, p. 13410-13419.

[9] J. Lin, Z. He and R. W. H. Lau. Progressive Mirror Detection, In Conference on Computer Vision and Pattern Recognition (CVPR), 2020, p.3697-3705.

[10] J. Lin, et al. Exploiting Semantic Relations for Glass Surface Detection. In Advances in Neural Information Processing Systems 35, 2022, p.22490-22504.

[11] J. Redmon, et al. 2016. You Only Look Once: Unified, Real-Time Object Detection. In Computer Vision and Pattern Recognition (CVPR). p. 779–788.

[12] K. He, X. Zhang, S. Ren and J. Sun. Deep Residual Learning for Image Recognition, 2016 In Conference on

Computer Vision and Pattern Recognition (CVPR), 2016, p. 770-778.

[13] L. Yu, et al. Progressive glass segmentation. In Transactions on Image Processing, 2022, p. 31: 2920-2933.

[14] N. Sliberman, et al. Indoor segmentation and support inference from rgbd images. In European Conference on Computer Vision(ECCV), 2012. p.746-760.

[15] R. He, J. Lin, J and W.H Lau. Efficient Mirror Detection via Multi-level Heterogeneous Learning. arXiv preprint arXiv:2211.15644, 2022.

[16] R. Mottaghi, et al. The Role of Context for Object Detection and Semantic Segmentation in the Wild, In Conference on Computer Vision and Pattern Recognition(CVPR), 2014, p. 891-898.

[17] R. Zhang, et al. The unreasonable effectiveness of deep features as a perceptual metric. In conference on computer vision and pattern recognition(CVPR). 2018. p. 586-595.

[18] S. Bell, et al. Material recognition in the wild with the materials in context database. In Conference on Computer Vision and Pattern Recognition(CVPR), 2015. p. 3479-3487.

[19] S. Song, S. P. Lichtenberg, and J. Xiao. Sun rgb-d: A rgb-d scene understanding benchmark suite. In Conference on Computer Vision and Pattern Recognition(CVPR), 2015. p. 567-576.

[20] X. Tan, et al. Mirror Detection With the Visual Chirality Cue. In Transactions on Pattern Analysis and Machine Intelligence, 2022. p. 3492-3504.

[21] X. Yang, et al. Where Is My Mirror?. In International Conference on Computer Vision(ICCV).2019.p.8808-8817.

[22] Z. Mao, et al. A dataset and ensemble model for glass façade segmentation in oblique aerial images. In Geoscience and Remote Sensing Letters, 2022, 19: 1-5.

[A] Chrom web store ”Download All Images”, URL: https://chrome.google.com/webstore/detail/download-all-images/ifipmflagepipjokmbdecpmjbibjnakm

[B] Open data annotation platfrom CVAT, URL: https://www.cvat.ai/