Discovering places of hideout from a single RGB indoor view

Computer Vision (CS6350)
TPA - 11

1.   Problem Statement

Given an input image, the goal is to produce 2D bounding boxes for CGLs (potential hideouts) in that image. CGLs usually lie around the edges of items like furniture, pillars, doors, windows, staircase, etc. We refer to these items as obscuring items. We define image regions around edges of obscuring items as CGLs.

Since it is not feasible to classify any region of an image as a hideout (CGL) without looking at its surroundings, CGL detection requires context-aware detection and understanding of the complex 3D spatial relationships between neighboring regions within an image.

CGL detection finds applications in military counter-insurgency operations and intelligent scene surveillance with path planning for a robot.

2.   Input

Single RGB image which may or may not contain CGL(s).

3.   Output

- Bounding boxes (coordinates) for CGLs present in the input image.
- Demo should run on any test image.
- Team/Individual selecting this TPA will have to annotate at least 100 more samples from indoor and a few outdoor scenes.
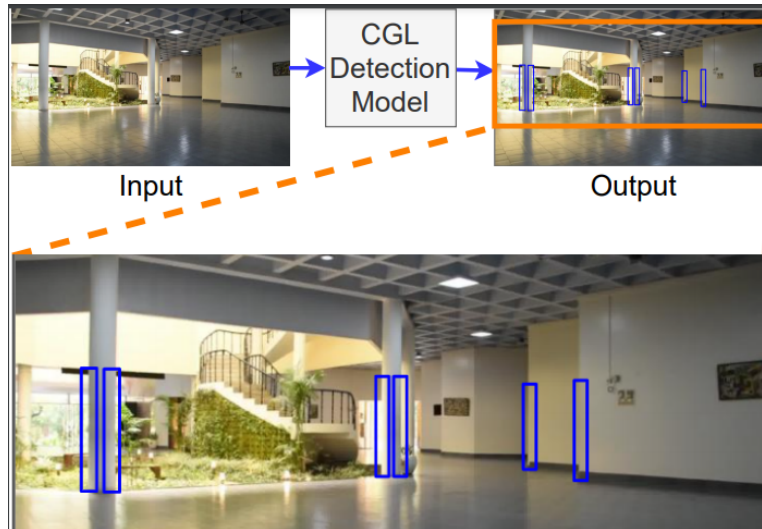
4. Dataset

In-house VPLab CGL detection dataset:
1400 images having around 13K annotated CGL instances.

5. References

[1] Fang, Yuxin, et al. "You Only Look at One Sequence: Rethinking Transformer in Vision through Object Detection." arXiv preprint arXiv:2106.00666 (2021).

[2] Carion, Nicolas, et al. "End-to-end object detection with transformers." European Conference on Computer Vision. Springer, Cham, 2020.

[3] Beal, Josh, et al. "Toward transformer-based object detection." arXiv preprint arXiv:2012.09958 (2020).

[4] Zhou, Xingyi, Vladlen Koltun, and Philipp Krähenbühl. "Probabilistic two-stage detection." arXiv preprint arXiv:2103.07461 (2021).

[5] Liang, Tingting, et al. "CBNetV2: A Composite Backbone Network Architecture for Object Detection." arXiv preprint arXiv:2107.00420 (2021).

[6] Dai, Xiyang, et al. "Dynamic Head: Unifying Object Detection Heads with Attentions." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

[7] Farhadi, Ali, and Joseph Redmon. "Yolov3: An incremental improvement." Computer Vision and Pattern Recognition, cite as (2018).

[8] Liu, Wei, et al. "Ssd: Single shot multibox detector." European Conference on Computer Vision. Springer, Cham, 2016.

[9] Ren, Shaoqing, et al. "Faster R-CNN: towards real-time object detection with region proposal networks." IEEE transactions on Pattern Analysis and Machine Intelligence 39.6 (2016): 1137-1149.