



OAFORMER: LEARNING OCCLUSION DISTINGUISHABLE FEATURE FOR AMODAL INSTANCE SEGMENTATION

Zhixuan Li^{1,2}, Ruohua Shi², Tiejun Huang² and Tingting Jiang^{✉1,2}

¹ Advanced Institute of Information Technology, Peking University, Hangzhou, China

² National Engineering Research Center of Visual Technology, School of Computer Science, Peking University, Beijing, China



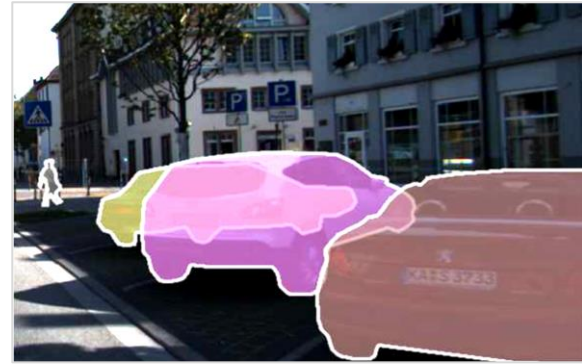
北京大学
PEKING UNIVERSITY

What is amodal instance segmentation?

Predict the complete mask of the occluded instance, including both **visible** and **invisible** regions.



Visible Instance Segmentation

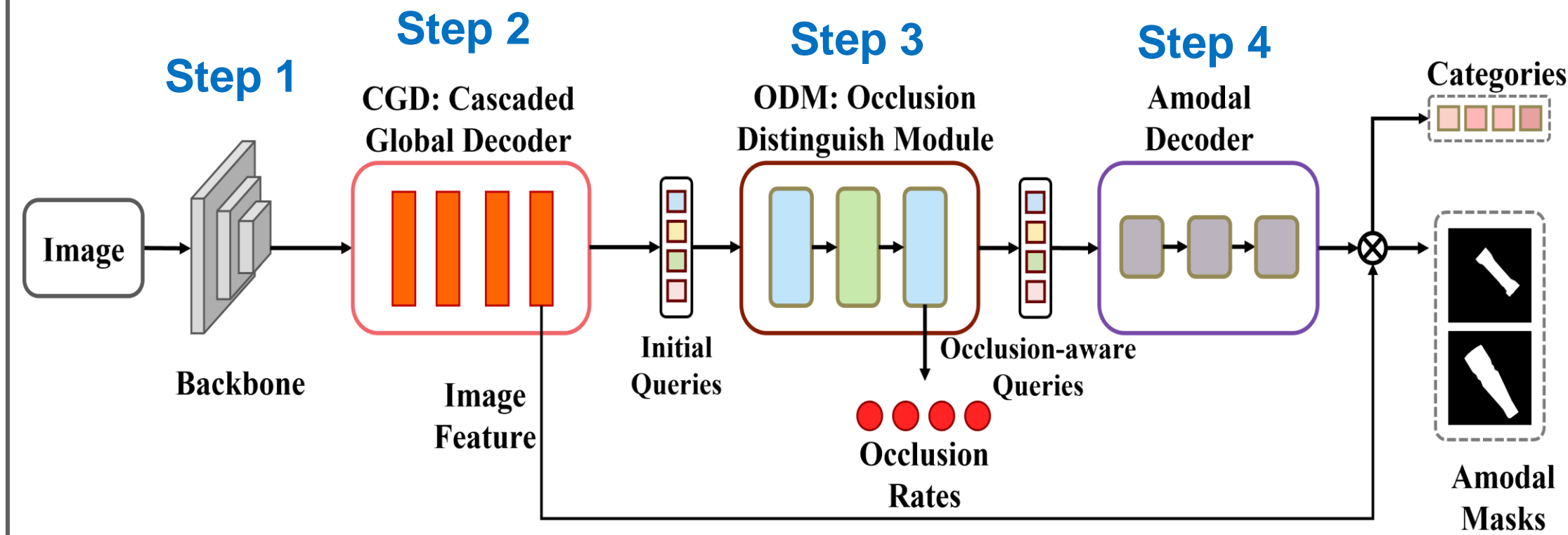


Amodal Instance Segmentation

Problems in existing AIS methods:

confused for distinguishing whether the object is occluded

Occlusion Confusing Problem



Our approach:

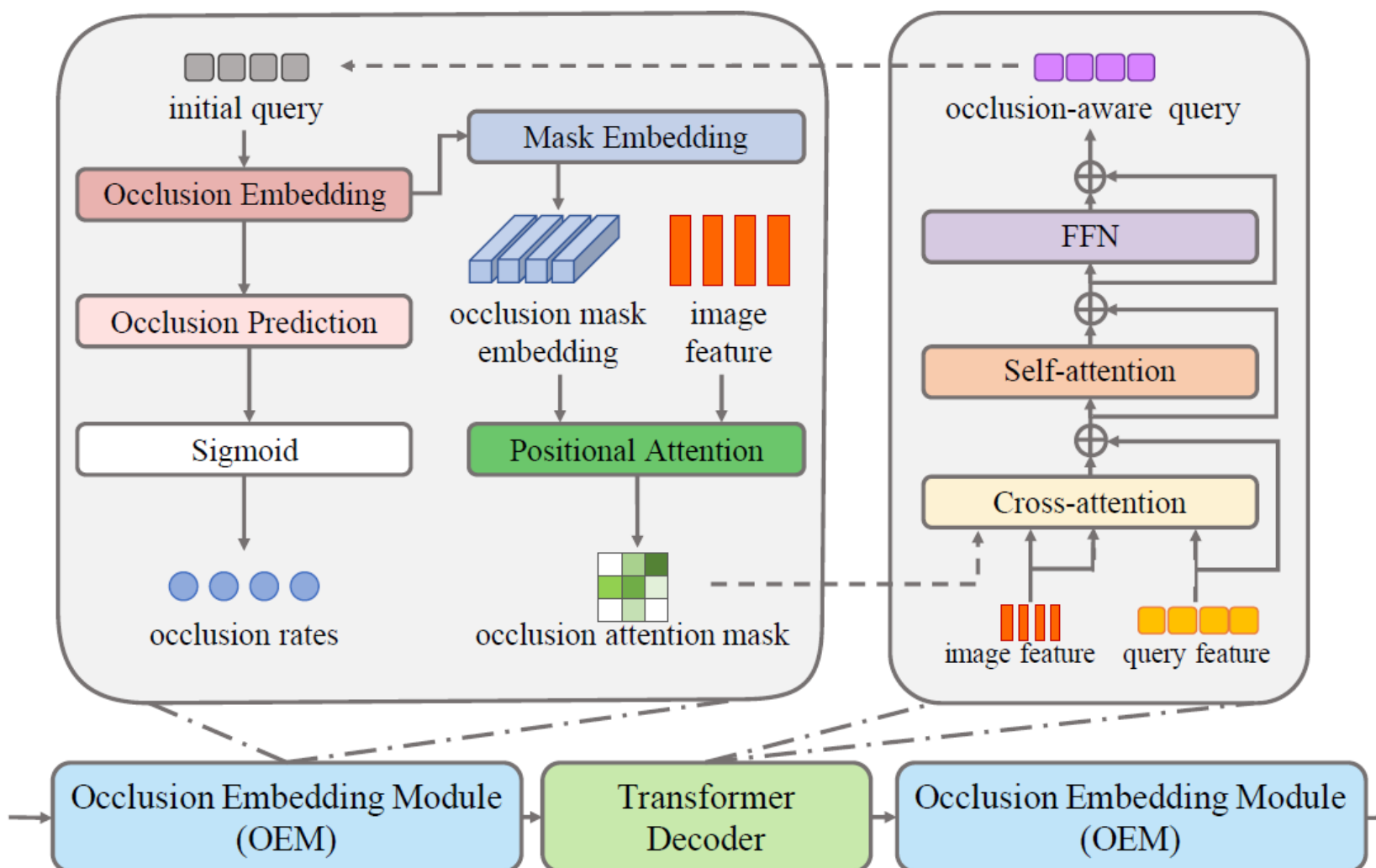
Step 1: Multi-level feature extraction

Step 2: Cascaded feature enhancement

Step 3: Learn and embed occlusion distinguishable information

Step 4: Amodal masks and categories prediction

Occlusion Distinguish Module



Experiments:

Comparison with SOTA

Method	Supervision	D2SA			COCOA-clS		
		AP _{avg}	AP ₅₀	AP ₇₅	AP _{avg}	AP ₅₀	AP ₇₅
Mask-RCNN [22]	bbox, mask, cls	63.6	83.9	68.0	33.7	56.5	35.8
ORCNN [7]	bbox, mask, cls	64.2	83.6	69.1	28.0	53.7	25.4
SLN [13]	bbox, mask, cls	25.1	30.8	29.4	14.4	23.6	15.8
BCNet [11]	bbox, mask, cls	50.9	66.9	57.2	22.1	32.3	24.5
ShapeDict [15]	bbox, mask, cls	70.3	85.1	75.8	35.4	56.0	38.7
A3D [16]	bbox, mask, cls	68.5	N/A	N/A	34.9	N/A	N/A
Ours (w/o ODM)	mask, cls	61.7	78.7	63.3	33.9	45.0	35.8
Ours (w/ ODM)	mask, cls	72.5	86.5	76.1	37.4	49.7	40.5

Summary:

1. This paper proposes an end-to-end transformer-based method named OAFormer to handle the occlusion confusing problem in the AIS task
2. OAFormer contains two novel components that learn and embed each instance's occlusion information from instance-level and spatial-level
3. Experiments show that OAFormer can achieve state-of-the-art performance on the D2SA and COCOA-clS datasets.

References:

[22] He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. *ICCV*. 2017

[7] Follmann, P., König, R., Hartinger, P., Klostermann, M., Böttger, T.: Learning to see the invisible: End-to-end trainable amodal instance segmentation. *WACV*. 2019

[13] Zhang Z., Chen A., Xie L., Yu J., and Gao S., "Learning semantics-aware distance map with semantics layering network for amodal instance segmentation," in *ACM MM*, 2019, pp. 2124–2132.

[11] Ke, L., Tai, Y.W., Tang, C.K.: Deep occlusion-aware instance segmentation with overlapping bilayers. *CVPR*. 2021

[15] Xiao, Y., Xu, Y., Zhong, Z., Luo, W., Li, J., Gao, S.: Amodal segmentation based on visible region segmentation and shape prior. *AAAI*. 2020

[16] Li Z., Ye W., Jiang T., and Huang T., "2D amodal instance segmentation guided by 3D shape prior," in *ECCV*, 2022.