PEEKABOO: Hiding Parts of an Image for Unsupervised Object Localization

Paper ID: 305

Hasib Zunair and A. Ben Hamza
Concordia University, Montreal, Canada

British Machine Vision Conference 2024

We aim to localize unfamiliar and salient objects without the need for class specific training

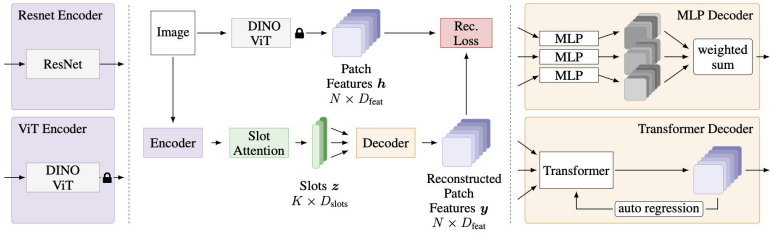
- Supervised learning is time-consuming and fails in cases of novel objects due to the finite nature of object classes.
- Unsupervised learning has challenges due to absence of visual information like appearance, type and number of objects

Image credit: Link

Existing unsupervised methods are complex and do not model visual context; fails when objects are small, reflective or under poor illumination

Model

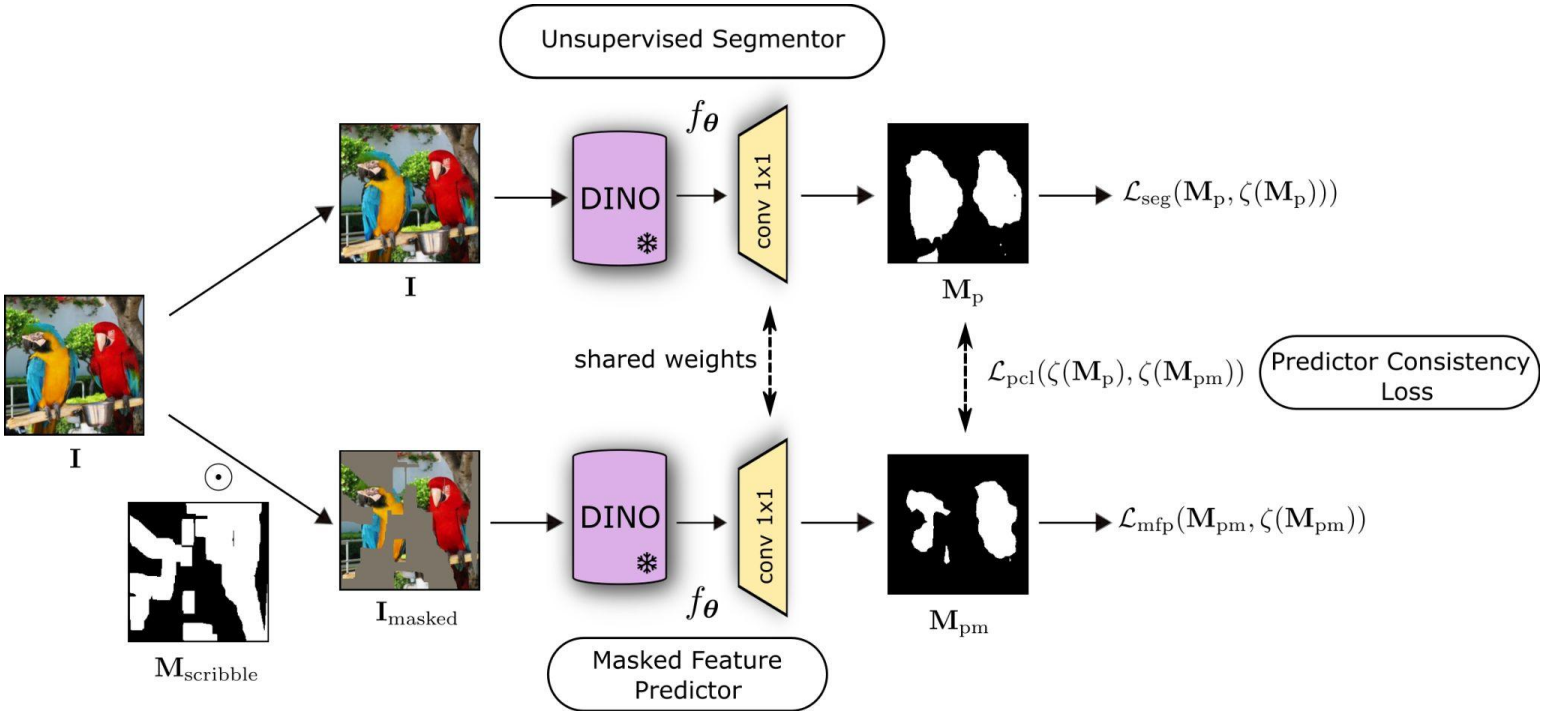
- Multiple stages of training
- Millions of learnable parameters
- Ensembles



Data:

- Uses both large-scale real world and synthetic data

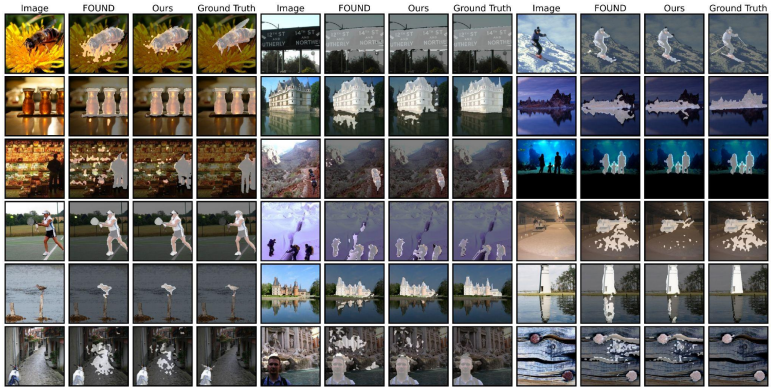
We introduce Peekaboo, a self-supervised single stage approach to localize novel and salient objects in an unsupervised manner



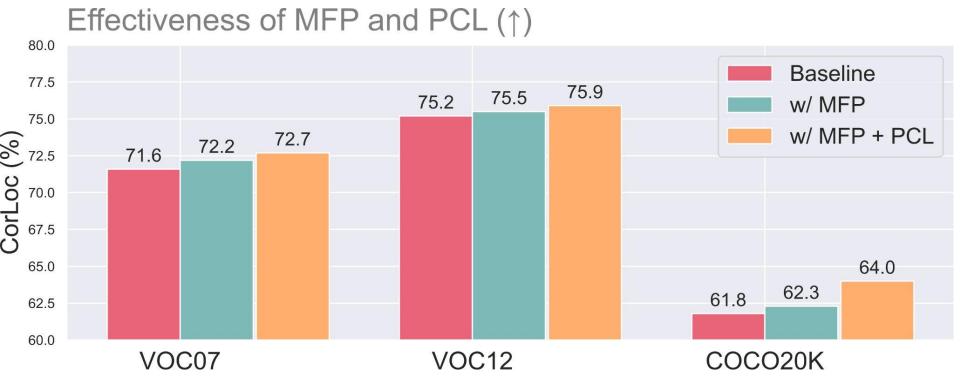
Peekaboo outperforms SOTA method on Single Object Discovery and Unsupervised Saliency Detection tasks

Method	Learning	DUT-OMRON			DUTS-TE			ECSSD		
		Acc	IoU	max F_β	Acc	IoU	max F_β	Acc	IoU	max F_β
HS [42]		84.3	43.3	56.1	82.6	36.9	50.4	84.7	50.8	67.3
wCtr [67]		83.8	41.6	54.1	83.5	39.2	52.2	86.2	51.7	68.4
WSC [18]		86.5	38.7	52.3	86.2	38.4	52.8	85.2	49.8	68.3
DeepUSPS [25]		77.9	30.5	41.4	77.3	30.5	42.5	79.5	44.0	58.4
BigBiGAN [37]		85.6	45.3	54.9	87.8	49.8	60.8	89.9	67.2	78.2
E-BigBiGAN[37]		86.0	46.4	56.3	88.2	51.1	62.4	90.6	68.4	79.7
Melas-Kyriazi et al. [23]		88.3	50.9	-	89.3	52.8	-	91.5	71.3	-
LOST [30]		79.7	41.0	47.3	87.1	51.8	61.1	89.5	65.4	75.8
DSM [22]		80.8	42.8	55.3	84.1	47.1	62.1	86.4	64.5	78.5
TokenCut [30]		88.0	53.3	60.0	90.3	57.6	67.2	91.8	71.2	80.3
SelfMask [29]	✓	90.1	**58.2**	-	92.3	62.6	-	94.4	78.1	-
FOUND† [30]	✓	90.7	57.1	79.9	93.5	63.7	85.2	**94.9**	**80.6**	95.1
DeepCut [0]	✓	-	-	-	-	59.5	-	-	74.6	-
WSCUOD [22]	✓	89.7	53.6	64.4	91.7	59.9	73.1	92.2	72.7	85.4
PEEKABOO (Ours)	✓	**91.5**	57.5	**80.4**	**93.9**	**64.3**	**86.0**	94.6	79.8	**95.3**
LOST + BS [30]	✓	81.8	48.9	57.8	88.7	57.2	69.7	91.6	72.3	83.7
DSM + CRF [22]	✓	87.1	56.7	64.4	83.8	51.4	56.7	89.1	73.3	80.5
WSCUOD + BS [22]	✓	90.9	58.5	68.3	92.5	63.0	76.4	92.8	74.2	89.6
TokenCut + BS [30]	✓	89.7	61.8	69.7	91.4	62.4	75.5	93.4	77.2	87.4
SelfMask + BS [29]	✓	91.9	**65.5**	-	93.3	66.0	-	**95.5**	**81.8**	-
FOUND + BS† [30]	✓	91.7	60.9	69.1	94.0	66.1	75.0	95.2	81.7	93.0
PEEKABOO + BS (Ours)	✓	**92.4**	61.2	**71.4**	**94.4**	**66.3**	**77.4**	94.9	80.6	**93.7**

Method	Learning	VOC07	VOC12	COCO20K
Zhang et al. [60]		46.2	50.5	34.8
DDT+ [51]		50.2	53.1	38.2
rOSD [55]		54.5	55.3	48.5
LOD [56]		53.6	55.1	48.5
DINO [0]		45.8	46.2	42.1
LOST [30] (ViT-S/16)		61.9	64.0	50.7
LOST + CAD [30]		65.7	70.4	57.5
DSM [22] (ViT-S/16)		62.7	66.4	52.2
TokenCut [30] (ViT-S/16)		68.8	72.1	58.8
TokenCut + CAD [30]		71.4	75.3	62.6
SelfMask [29]	✓	72.3	75.3	62.7
FOUND† [30]	✓	71.7	75.6	61.1
FreeSOLO [55]	✓	56.1	56.7	52.8
DeepCut [0]	✓	69.8	72.2	61.6
WSCUOD [22]	✓	70.6	72.1	63.5
DINOSAUR [22]	✓	-	70.4	**67.2**
PEEKABOO (ViT-S/8) (Ours)	✓	**72.7**	**75.9**	64.0



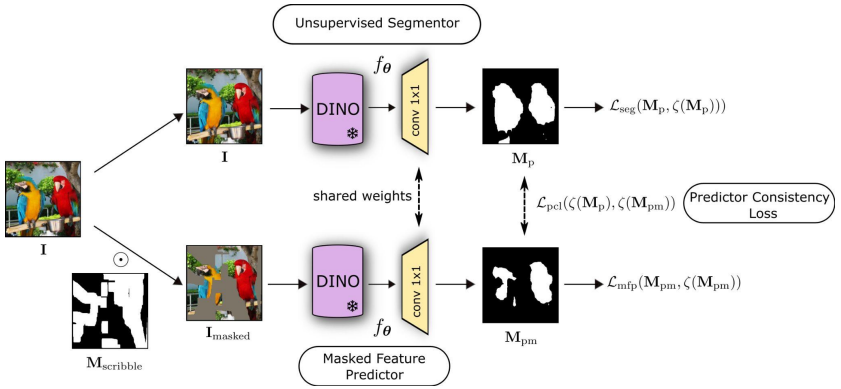
Peekaboo components complement each other; is better with high masking of images



Peekaboo can detect multiple unfamiliar objects of different shapes and scales; basically which are not background

Consider using Peekaboo in your research!

- A self-supervised segmentation model with zero-shot generalization to unfamiliar images and objects that are small, reflective or under poor illumination without the need for additional training.
- Compared to existing methods it does not require:
  - Multiple stages of training
  - Millions of learnable parameters
  - Combinations of multiple networks
  - Large scale real world and synthetic data



Project Page: https://hasibzunair.github.io/peekaboo/