PerSense: Training-Free Personalized Instance Segmentation in Dense Images

Muhammad Ibraheem Siddiqui, Muhammad Umer Sheikh, Hassan Abid, Muhammad Haris Khan
Mohamed Bin Zayed University of Artificial Intelligence
[muhammad.siddiqui, muhammad.umer, hassan.abid, muhammad.haris]@mbzuai.ac.ae

BMVC 2025

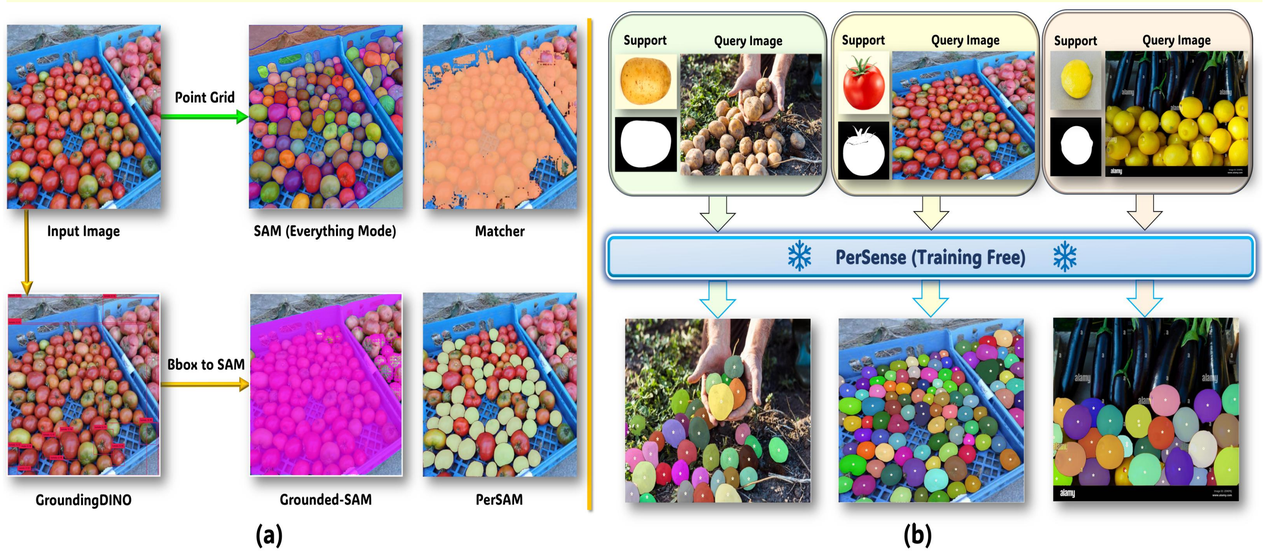
MOHAMED BIN ZAYED UNIVERSITY OF ARTIFICIAL INTELLIGENCE

Paper

Project Page

Motivation

🔎 The emergence of foundational models has significantly advanced **segmentation approaches**. However, challenges remain **in dense scenarios**, where occlusions, scale variations, and clutter impede **instance delineation**.



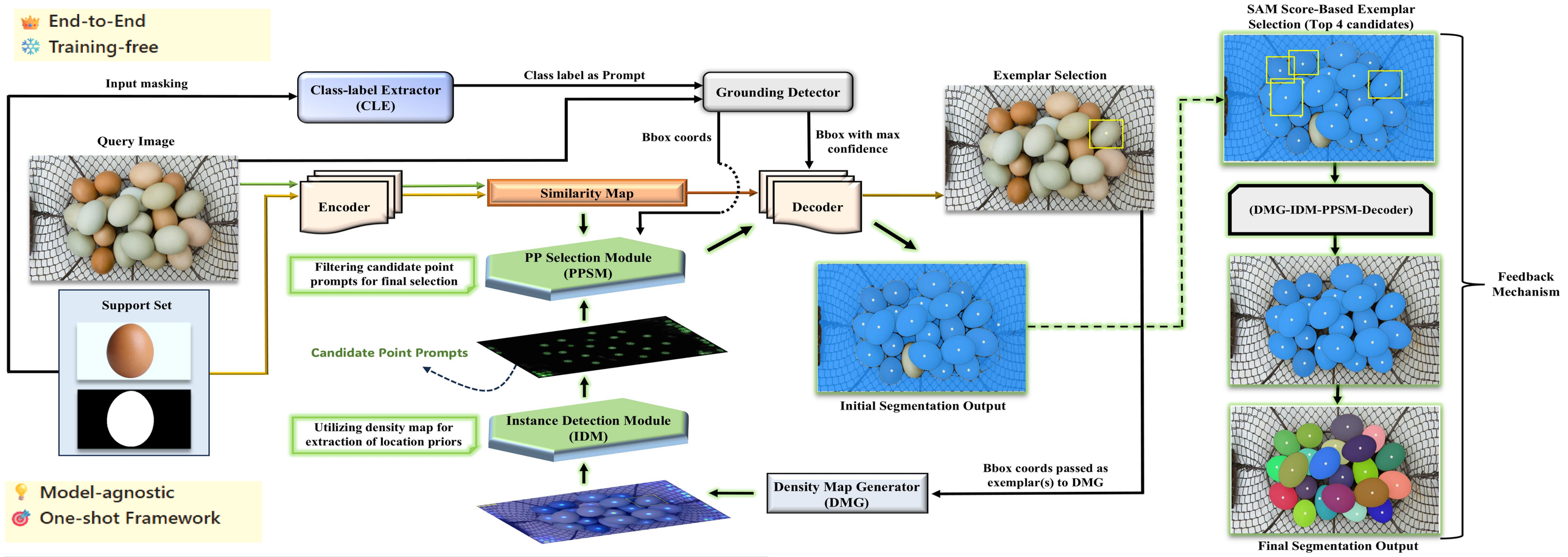
(a) Recent methods struggle in dense scenes, while SAM's "everything mode" segments all regions indiscriminately. **(b)** Introducing **PerSense**, a **training-free** and **model-agnostic one-shot framework** offering an **end-to-end automated pipeline** for personalized instance segmentation in dense images.

Contributions

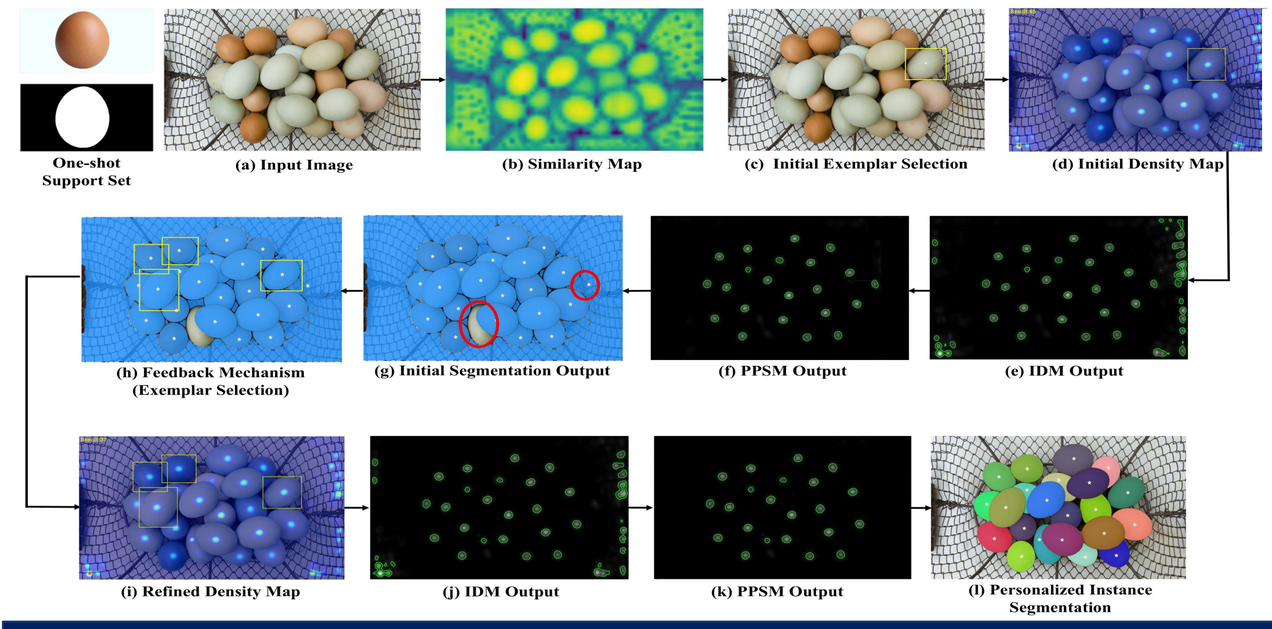
Towards developing PerSense framework, we make following core contributions:

- ❑ We develop a new baseline capable of automatically generating instance-level point prompts via proposing a novel **Instance Detection Module (IDM)** that leverages density maps, encapsulating the spatial distribution of objects.
- ❑ To mitigate false positives within generated point prompts, we design **Point Prompt Selection Module (PPSM)**. Both IDM and PPSM transform density maps into personalized point prompts for instance-level segmentation, offering seamless integration within our model-agnostic framework.
- ❑ We introduce a **feedback mechanism** that enables PerSense to improve the accuracy of density maps by automating the exemplar selection process for density map generation (DMG).
- ❑ We demonstrate the potential of PerSense for vision-based industrial automation (quality control, cargo monitoring) and medical applications such as cellular segmentation.
- ❑ To promote algorithmic advances and effective tools for this relatively underexplored task, **we introduce PerSense-D, a diverse dataset** exclusive to **personalized instance segmentation in dense images**.

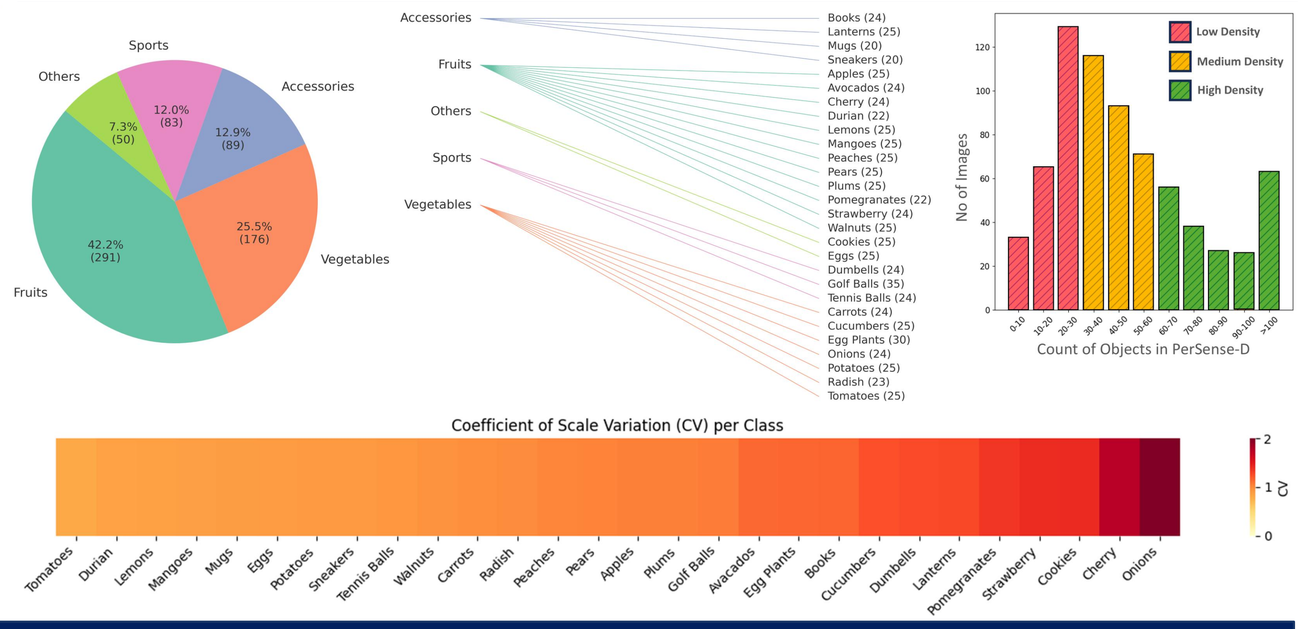
PerSense Framework



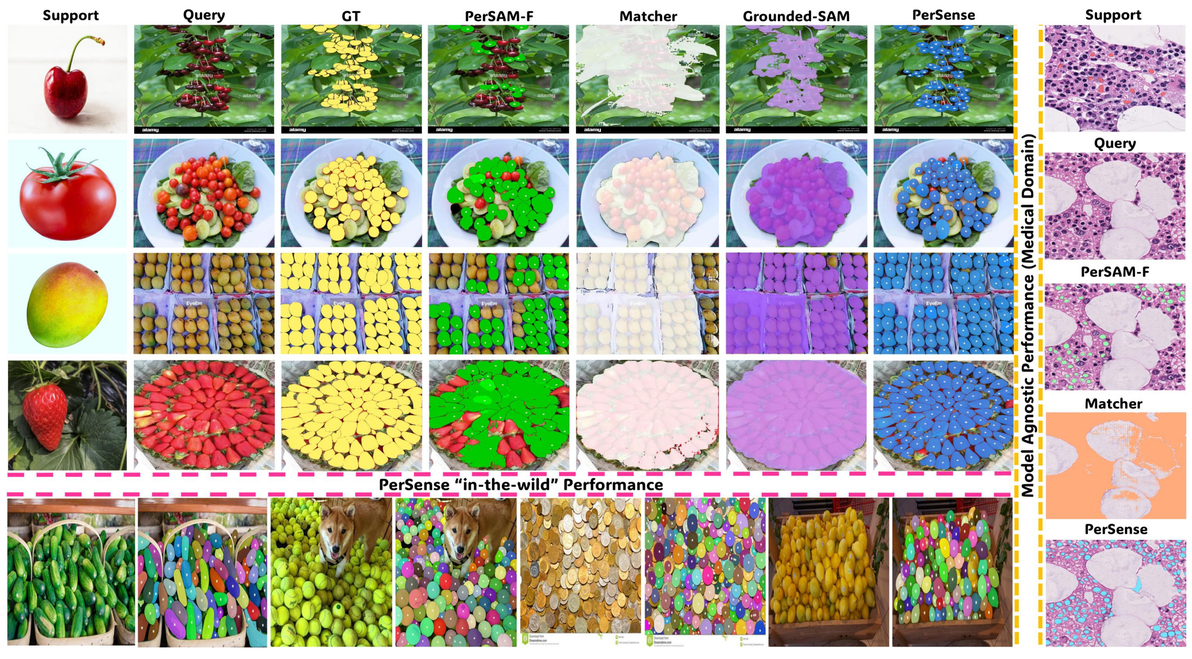
Step-by-step Analysis



New PerSense-D Benchmark



Results



Method	Venue	Density-based Low	Density-based Med	Density-based High	Class-wise Overall
End-to-end trained / fine-tuned					
C3Det [11]	CVPR 2022	52.70	46.64	39.11	48.60
SegGPT [33]	ICCV 2023	59.81	53.34	52.05	55.50
PerSAM-F† [37]	ICLR 2024	38.18	34.84	26.73	29.30
PseCo† [8]	CVPR 2024	53.99	65.23	68.55	61.83
GeCo† [24]	NIPS 2024	63.92	63.40	74.49	65.95
Training-free († indicates methods using SAM)					
PerSAM† [37]	ICLR 2024	32.27	28.75	20.25	24.45
TFOC† [30]	WACV 2024	62.78	65.38	65.69	62.63
Matcher† [18]	ICLR 2024	58.62	58.30	68.00	62.80
GroundedSAM† [28]	arXiv 2024	58.36	66.24	64.97	65.92
PerSense† (DMG1)	(ours)	**66.36**	67.27	74.78	70.96
PerSense† (DMG2)	(ours)	59.84	**73.51**	**77.57**	**71.61**

Table 1: Comparison of PerSense with SOTA on PerSense-D Benchmark

Method	Venue	COCO-20^i F0	F1	F2	F3	Mean	LVIS-92^i Mean
In-domain training							
HSNet [21]	CVPR 21	37.2	44.1	42.4	41.3	41.2	17.4
VAT [6]	ECCV 22	39.0	43.8	42.6	39.7	41.3	18.5
FPTrans [36]	NIPS 22	44.4	48.9	50.6	44.0	47.0	-
MIANet [34]	CVPR 23	42.4	52.9	47.7	47.4	47.6	-
LLaFS [38]	CVPR 24	47.5	58.8	56.2	53.0	53.9	-
COCO as training data							
Painter [32]	CVPR 23	31.2	35.3	33.5	32.4	33.1	10.5
SegGPT [33]	ICCV 23	56.3	57.4	58.9	51.7	56.1	18.6
Training-free (excluding PerSAM-F [37])							
PerSAM† [37]	ICLR 24	23.1	23.6	22.0	23.4	23.0	11.5
PerSAM-F† [37]	ICLR 24	22.3	24.0	23.4	24.1	23.5	12.3
Matcher† [18]	ICLR 24	**52.7**	**53.5**	**52.6**	**52.1**	**52.7**	**33.0**
PerSense†	(ours)	47.8	49.3	48.9	50.1	49.0	25.7

Table 2: Comparison of PerSense with SOTA on COCO and LVIS