Introduction
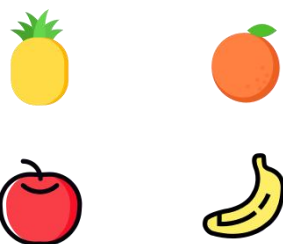
What is Soft Exemplar?

We call manually annotated exemplar bounding box is termed as the hard exemplar. Dynamically selecting exemplars in image features without annotation is called soft exemplar.
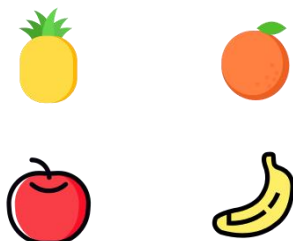
What is Open-set Object Counting?

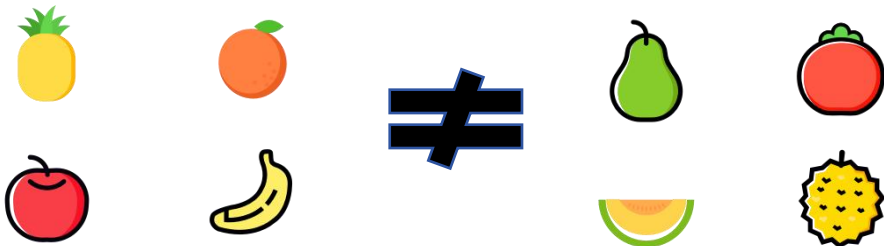Query the number of target objects in an image through text input.

Train set　　　**Test set**

Close-set

Open-set

Limitation of previous methods

- **The text-based zero shot method has the lowest annotation cost but poor performance.**

- **The exemplar based few shot method performs better but requires expensive annotation of example boxes.**

Objective

- **Balance the limitations mentioned above**

  🚀 Generate multi-scale soft exemplars based on semantics without the need for annotation costs.

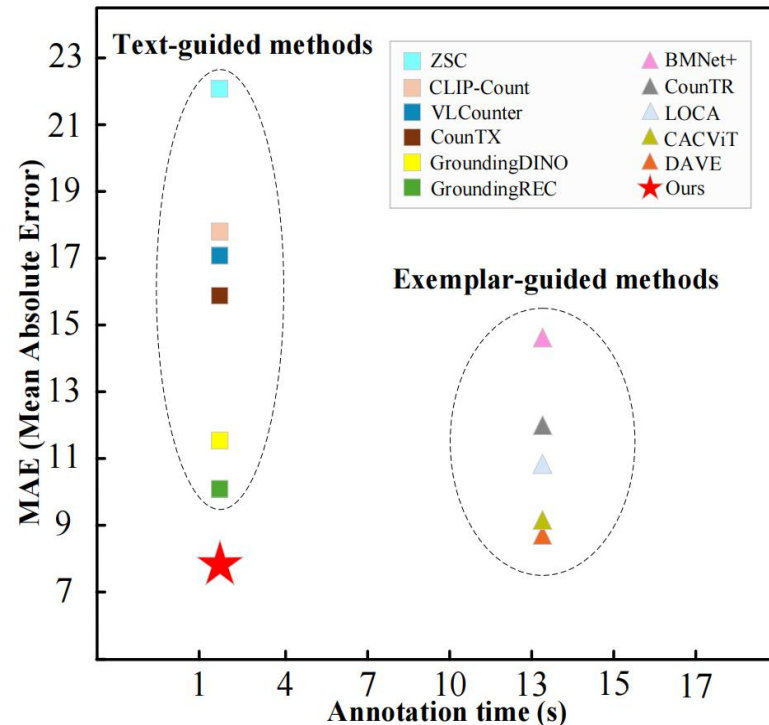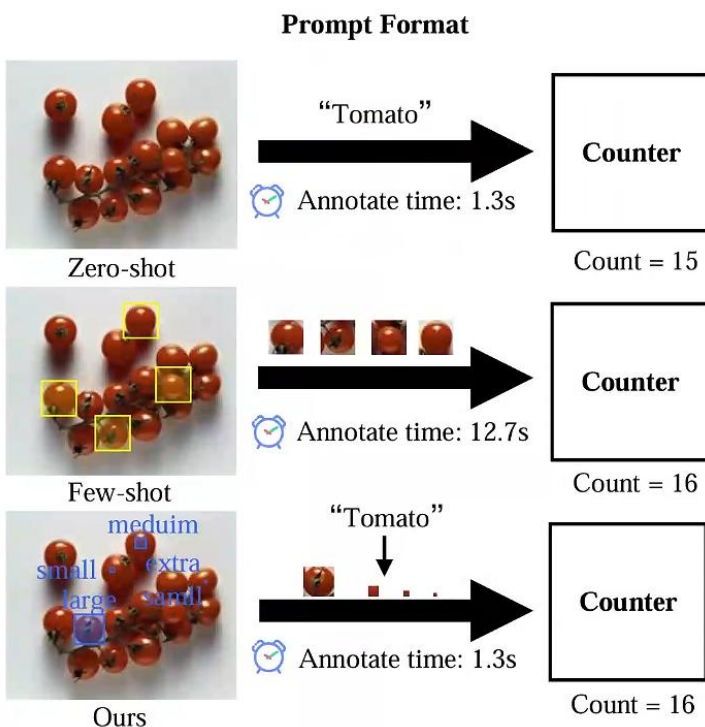 🚀 Improve counting accuracy by filtering error soft exemplars through clustering.

Fig 1. The advantages of our method

Method

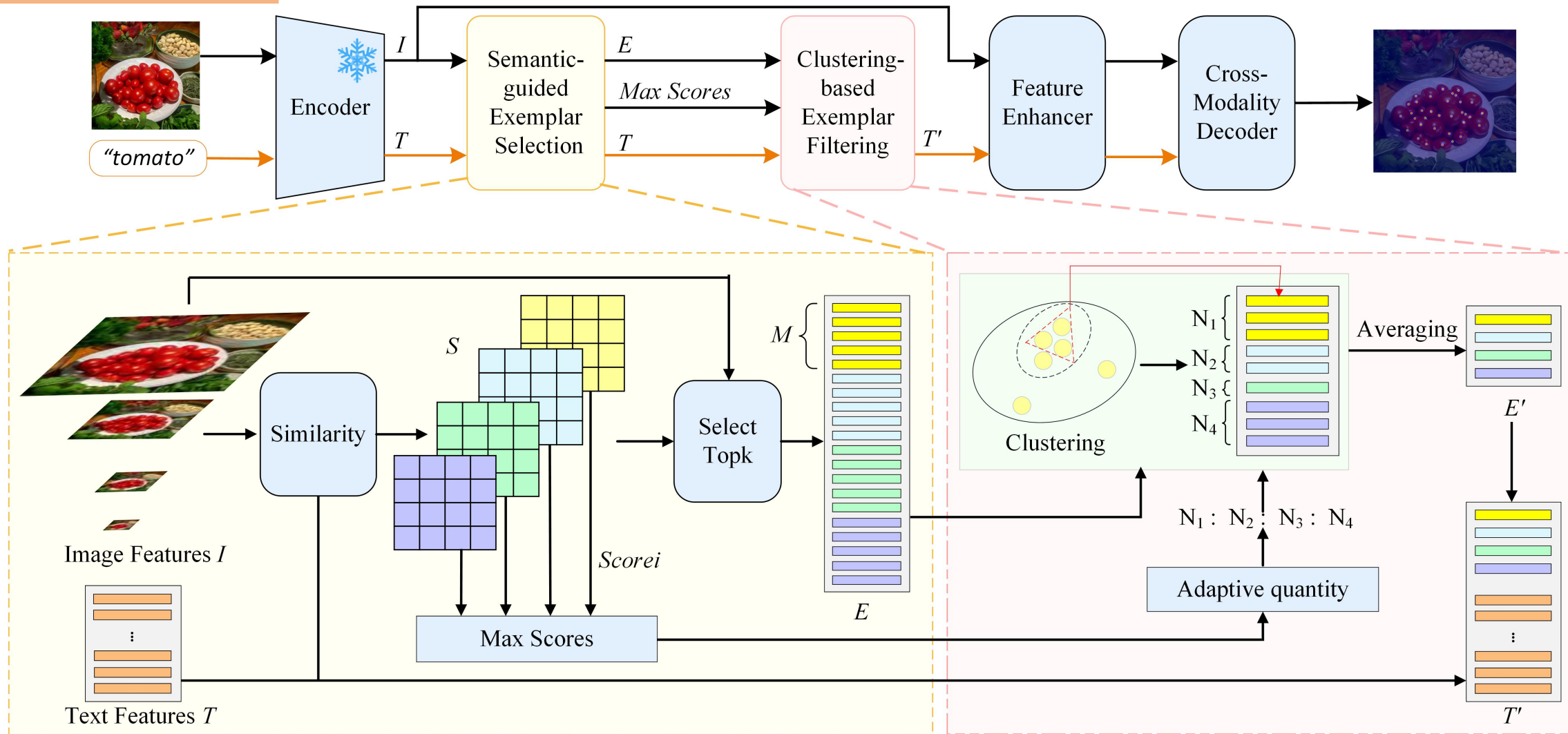**Fig 2. Overall architecture and detailed diagram of proposed methods**

Experiments

Table 1. Comparison with state-of-the-art methods

Method	Prompt Format	Publication Venue	Year	Val set		Test set	
				MAE	RMSE	MAE	RMSE
ZSC [48]	Text	CVPR	2023	26.93	88.63	22.09	115.17
CLIP-Count [19]	Text	ACM MM	2023	18.79	61.18	17.78	106.62
VLCounter [20]	Text	AAAI	2024	18.06	65.13	17.05	106.16
CountTX [1]	Text	BMVC	2023	17.10	65.61	15.88	106.29
GroundingDINO [29]	Text	ECCV	2024	12.07	57.02	11.52	97.37
COUNTGD [2]	Text	NeurIPS	2024	12.14	47.51	12.98	98.35
GroundingREC [5]	Text	CVPR	2024	10.06	58.62	10.12	107.19
FamNet [35]	Visual Exemplars	CVPR	2021	23.75	69.07	22.08	99.54
BMNet+ [37]	Visual Exemplars	CVPR	2022	15.74	58.53	14.62	91.83
CountTR [28]	Visual Exemplars	BMVC	2022	13.13	49.83	11.95	91.23
LOCA [42]	Visual Exemplars	ICCV	2023	10.24	32.56	10.79	56.97
CACViT [44]	Visual Exemplars	AAAI	2024	10.63	37.95	9.13	48.96
DAVE [32]	Visual Exemplars	CVPR	2024	8.91	28.08	8.66	32.36
COUNTGD [2]	Visual Exemplars & Text	NeurIPS	2024	7.10	26.08	5.74	24.09
CountSE (ours)	Text	–	–	**8.51**	**54.93**	**7.84**	**82.99**

Fig 3. Soft Exemplar Visualization

(a) Accurate soft exemplars.

"seagull" GT: 53 Pred:52
"polka dot" GT:32 Pred:32
"apple" GT:548 Pred:612
"stamp" GT:65 Pred:66

(b) Inaccurate soft exemplars.

"chair" GT:32 Pred:29
"shirt" GT:27 Pred:20
"nail polish" GT:146 Pred:134
"toilet paper roll" GT:32 Pred:41