SEGMENT ANY ANOMALY (SAA+)

OCT 2024

UNSUPERVISED AND SUPERVISED LEARNING

- Learning a representation of normal samples AND calculate discrepancy between Test and train image
- auto-encoder-based reconstruction
- memory-based normal distribution

ZERO-SHOT ANOMALY SEGMENTATION (ZSAS)

- Goal: Anomaly segmentation without requiring training data
- ZSAS create anomaly map

$$\mathbf{A} \in [0, 1]^{h \times w \times 1}$$

SEGMENT ANY ANOMALY FOR ZSAS

- Use prompt to roughly retrieve coarse anomaly region
  - Use Anomaly Region Generator(GroundingDINO)
- anomaly region refined into pixel-wise segmentation masks
  - Anomaly Region Refiner(Like SAM)

ANOMALY REGION GENERATOR

- Use GroundingDINO which pre-trained on large-scale language-vision datasets
  - pre-trained on large-scale language-vision
- Input:
  - language prompts
  - image I
- Output:
  - bounding-box-level region
  - confidence score set S

$$\mathcal{R}^B, \mathcal{S} := \text{Generator}(\mathbf{I}, \mathcal{T})$$

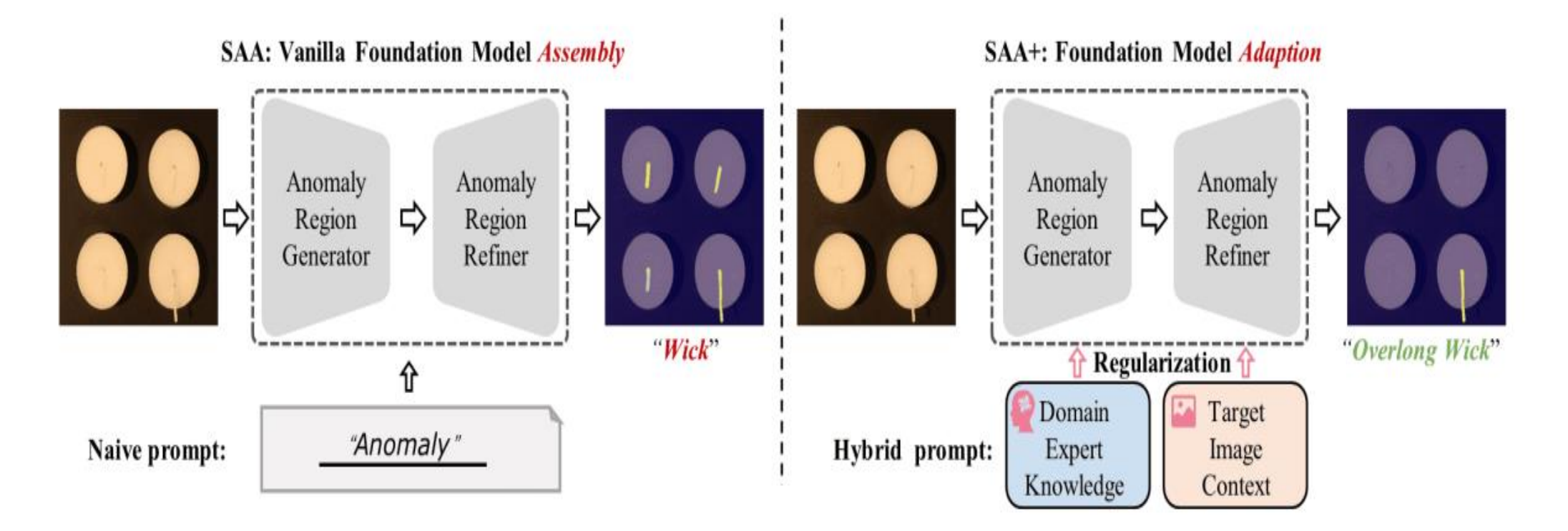
ANOMALY REGION REFINER

- Refine bounding-box-level anomaly region
- Use Segmentation Model(SAM)
  - assign a label to each individual pixel in an image
  - accepts various types of prompts as input
- Overall:

$$\mathcal{R} := \text{Refiner}(\mathbf{I}, \mathcal{R}^B)$$

$$\mathcal{R}, \mathcal{S} := \text{SAA}(\mathbf{I}, \mathcal{T}_n)$$

SAA STRUCTURE



SAA+

- Prompt Generated from Domain Expert Knowledge
- Type of Prompt:
  - Class-agnostic prompts
    - General prompts like "anomaly" OR "Defected"
  - Class-specific prompts:
    - designed based on expert knowledge
    - Like "black hole" and "white bubble"

$$\mathcal{P}^L = \{\mathcal{T}_\mathrm{a}, \mathcal{T}_\mathrm{s}\}$$

ANOMALY OBJECT PROPERTY AS PROMPT

1. Anomaly Location
   - anomalies may appear outside the inspected objects
   - calculate the intersection over union (IoU) between the potential anomaly regions and inspected object
   - Applying IoU threshold
2. Anomaly Area
   - anomalies should be smaller than inspected object

$$\mathcal{R}^P, \mathcal{S}^P := \text{Filter}(\mathcal{R}, \mathcal{P}^P)$$

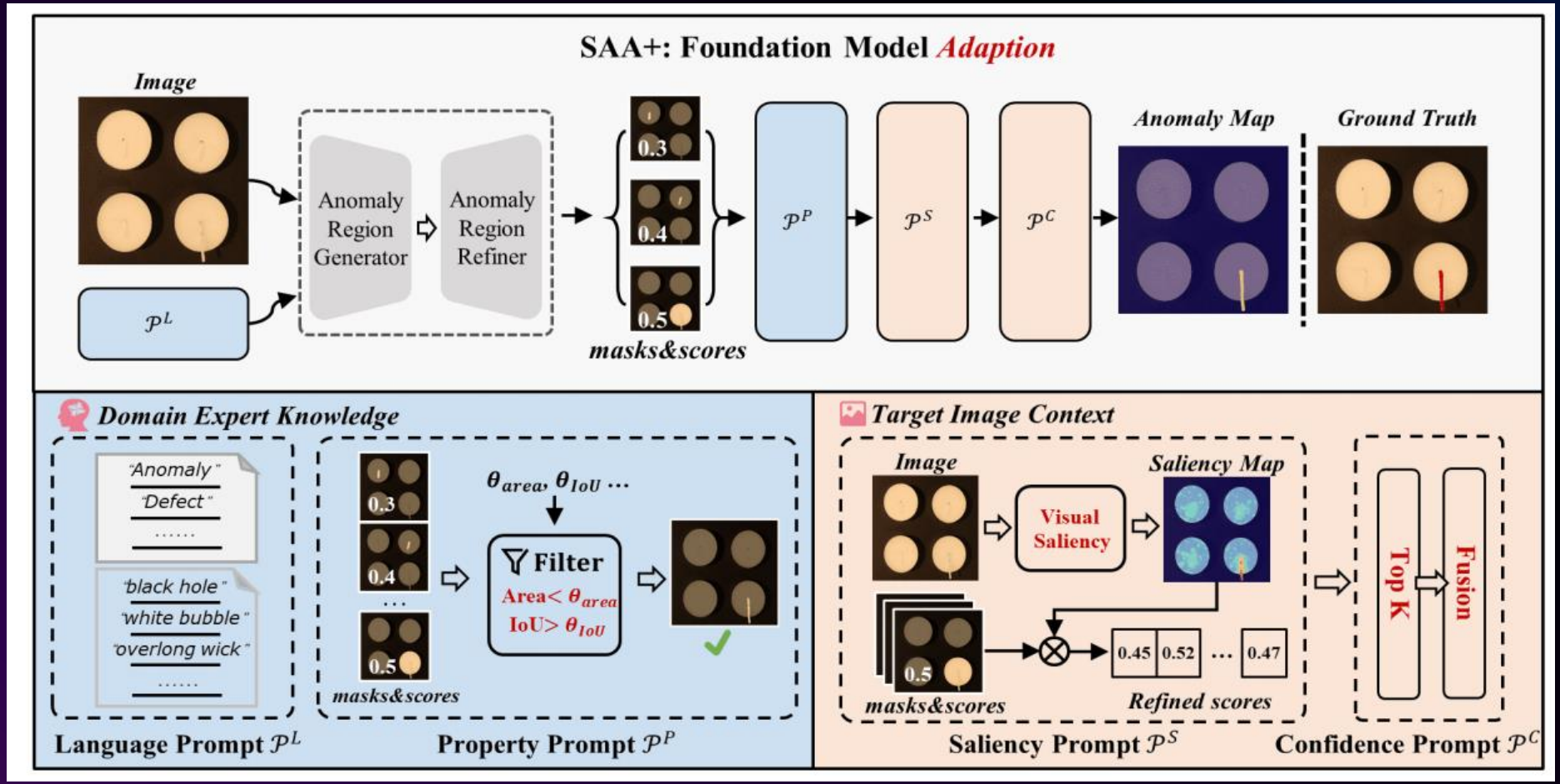
PROMPTS DERIVED FROM TARGET IMAGE CONTEXT

1. Anomaly Saliency as Prompt
   - calculate a saliency map for the input image

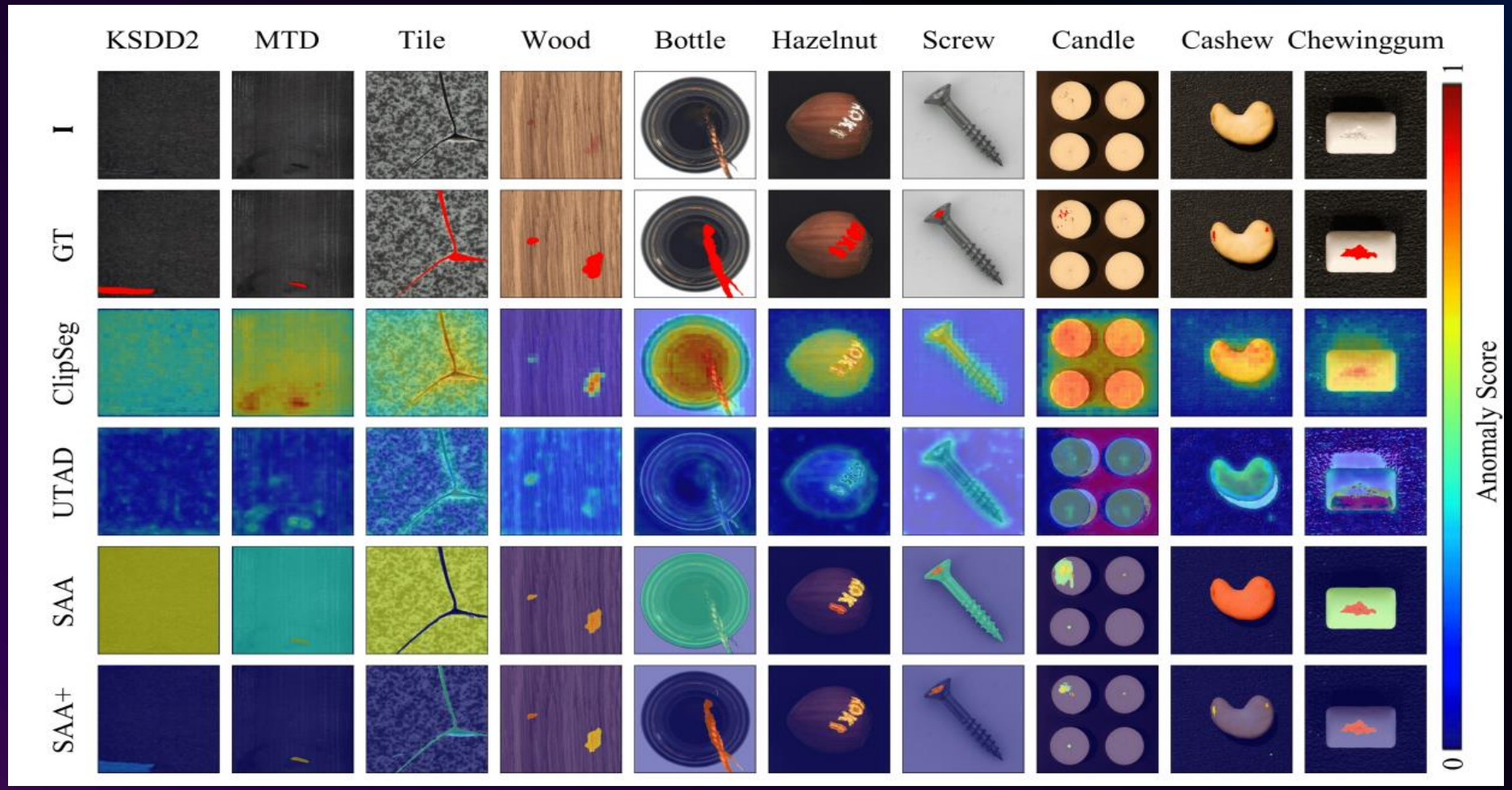
$$\mathbf{s}_{ij} := \frac{1}{N} \sum_{\mathbf{f} \in N_p(\mathbf{f}_{ij})} (1 - \langle \mathbf{f}_{ij}, \mathbf{f} \rangle)$$

2. Anomaly Confidence as Prompt
   - K candidates with the highest confidence scores
   - Usthis K candidates to estimate the final anomaly map

OVERVIEW OF SAA+



RESULTS


KSDD2
MTD
Tile
Wood
Bottle
Hazelnut
Screw
Candle
Cashew
Chewinggum
I
GT
ClipSeg
UTAD
SAA
SAA+
Anomaly Score
1
0