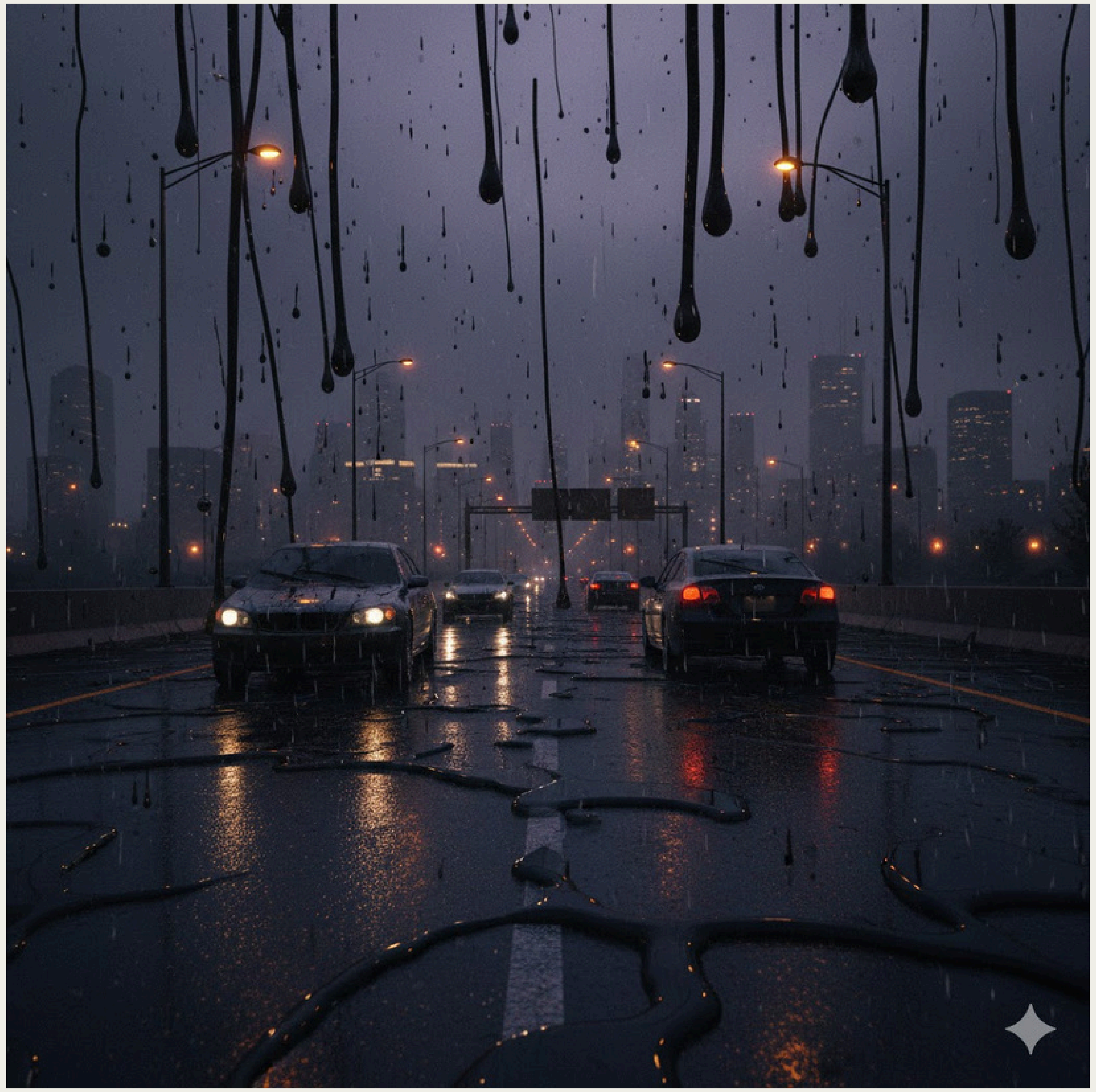
Advances in Post-Processors for Open Set Recognition

WALTER J. SCHEIRER
UNIVERSITY OF NOTRE DAME

COMPUTER SCIENCE AND ENGINEERING

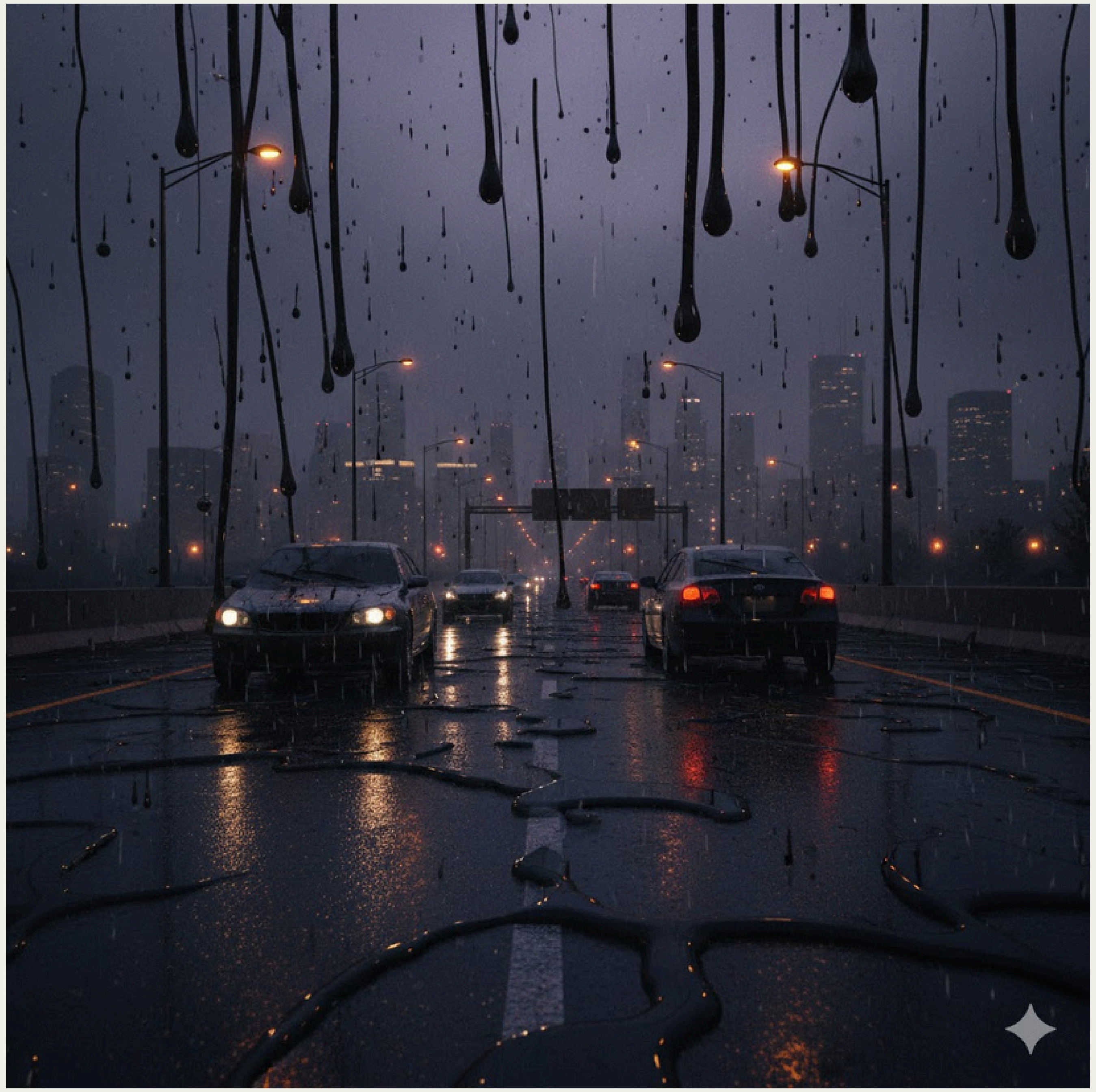
COMPUTER VISION @ND

THE NOVELTY PROBLEM IN AI: AN EXAMPLE



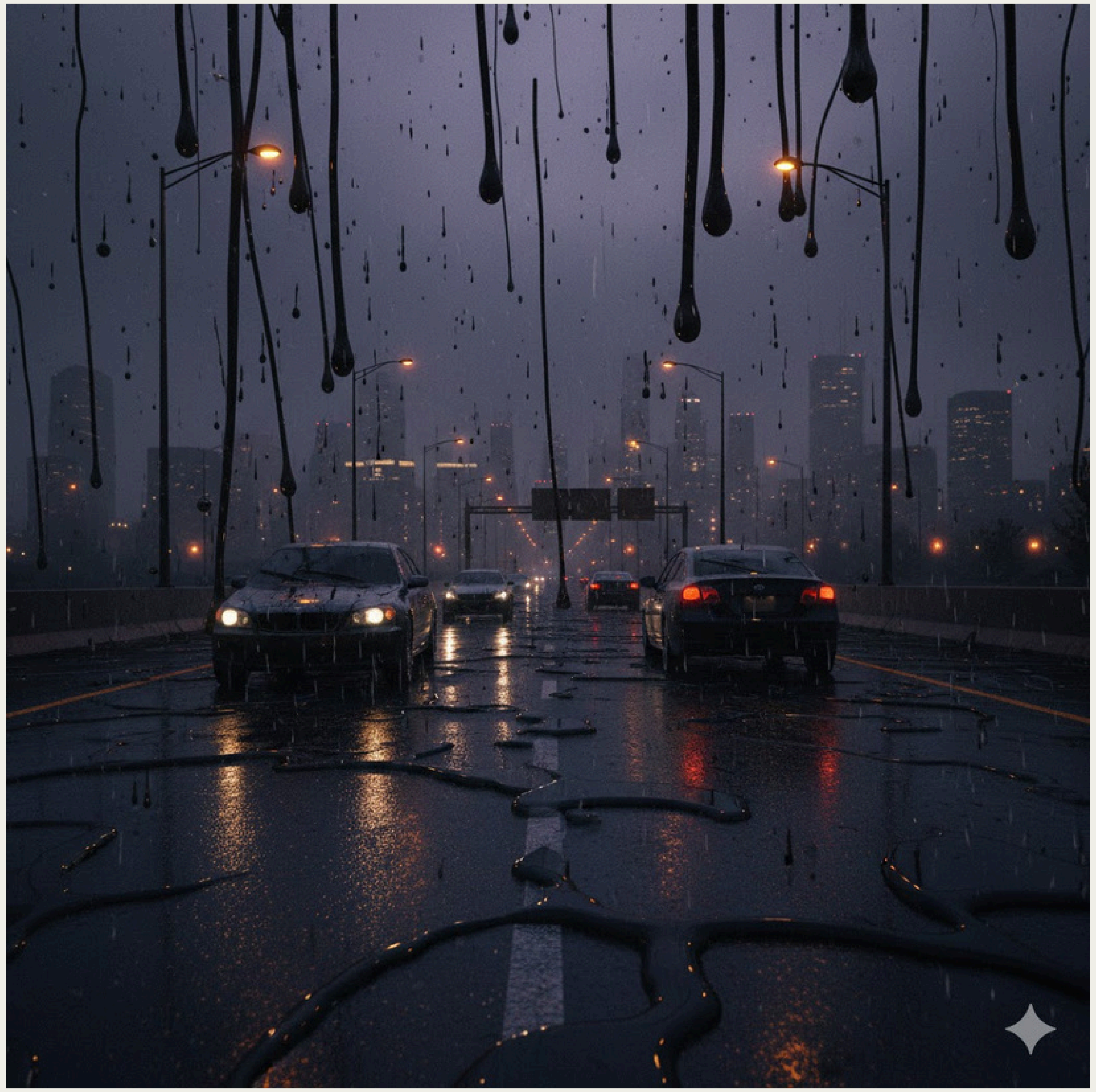
Nano Banana Prompt: Create a photorealistic image of a highway where it is raining **black droplets of oil**

THE NOVELTY PROBLEM IN AI: AN EXAMPLE



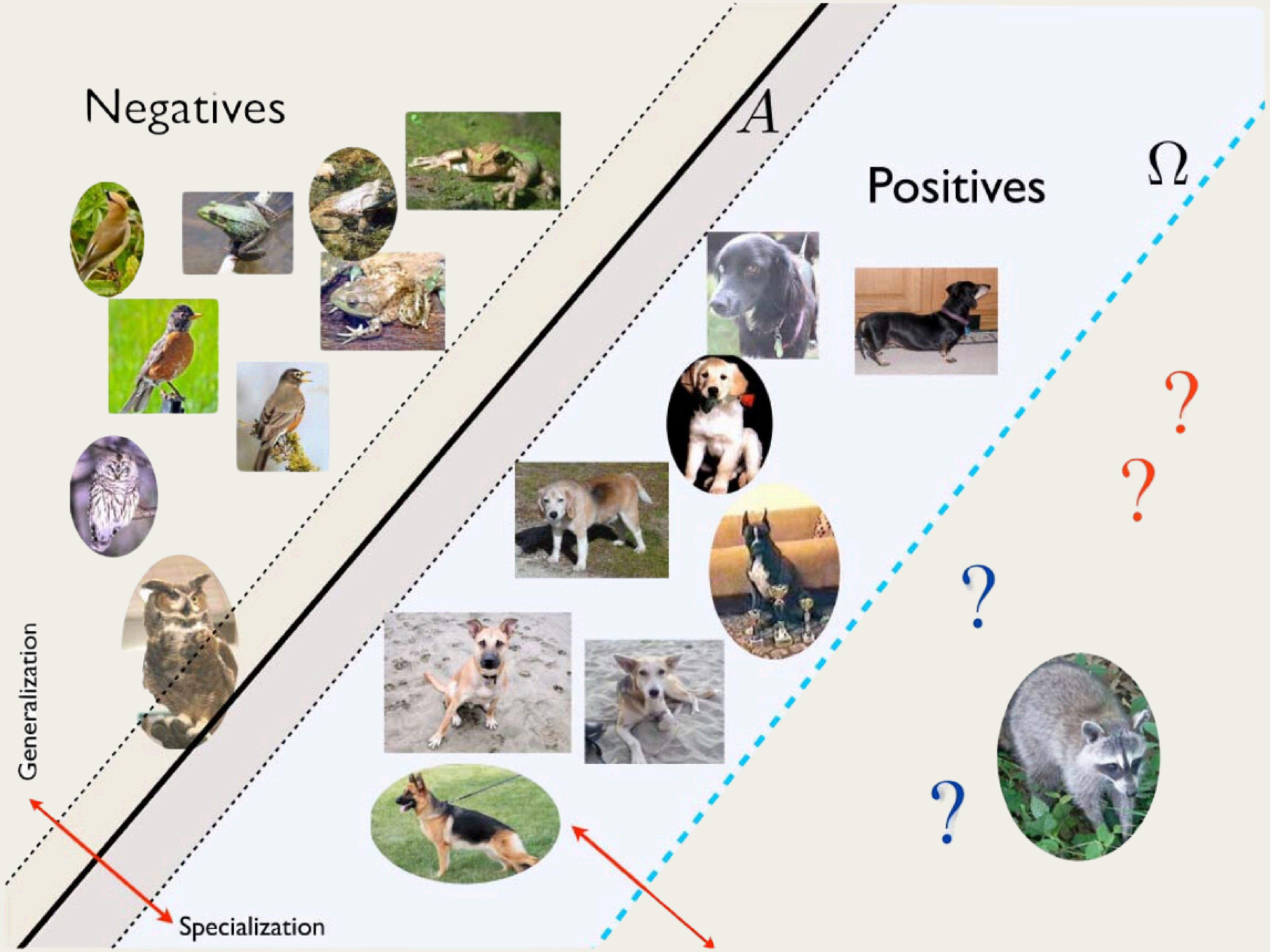
ChatGPT Prompt: Describe the driving conditions in this scene

THE NOVELTY PROBLEM IN AI: AN EXAMPLE



ChatGPT Response: The driving conditions look very hazardous. **It appears to be raining heavily, with poor visibility and wet, reflective road surfaces that can reduce traction.** There are dark, viscous-looking droplets or streaks falling and splattered across the road, which could further obscure visibility and make the surface slippery. Traffic is present, and the low light (likely dusk or nighttime) combined with city lighting glare adds to the difficulty. Overall, this would be a dangerous situation requiring very cautious driving — or avoiding the road altogether.

WHY IS NOVELTY SUCH A CONFOUND?



Training

We always have incomplete knowledge of the world at training time for open world domains.

Basis

There is no basis from which to generalize to novel data from known data, both of which are typically far away from each other in a feature space.

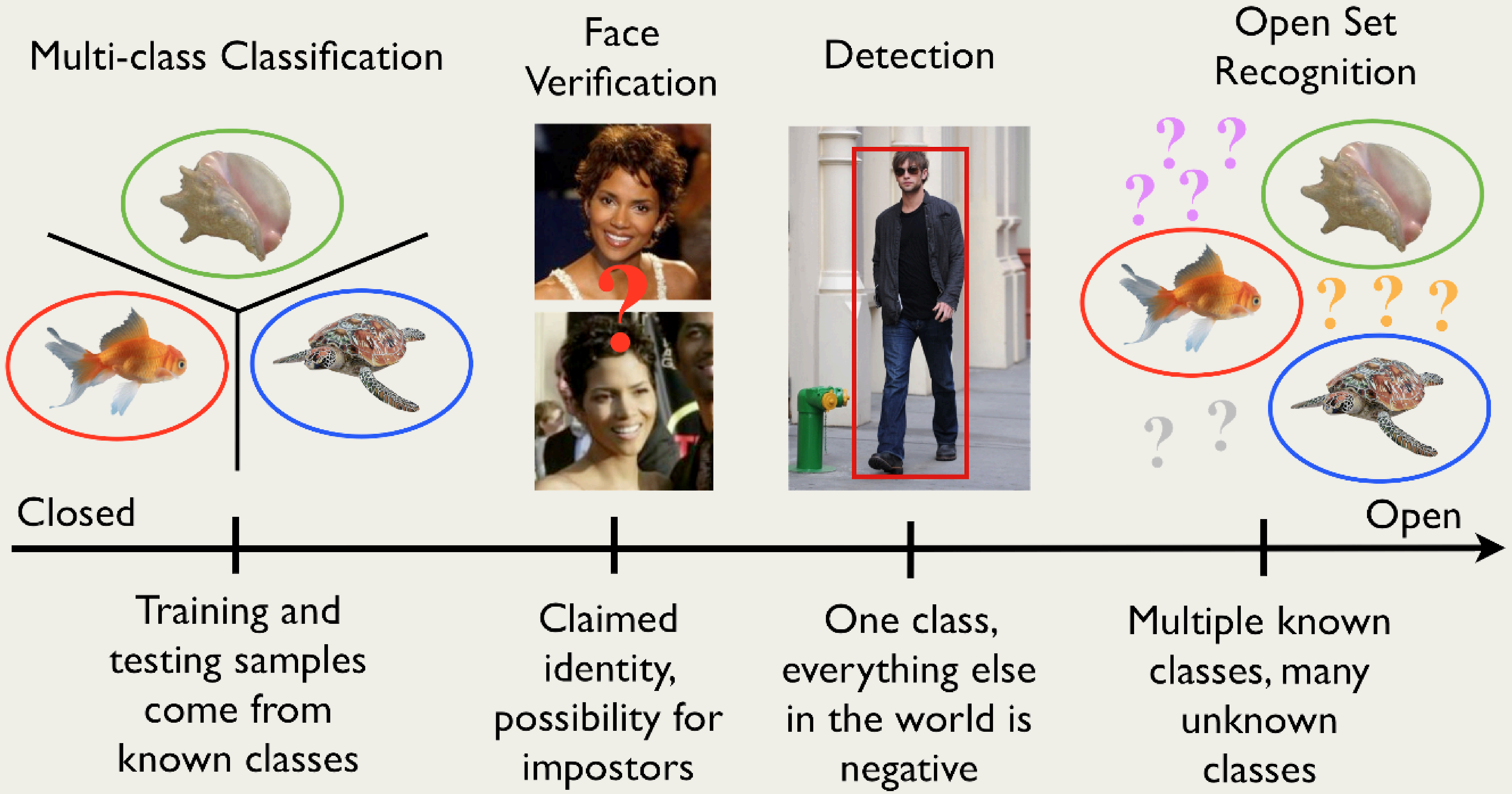
Generation

New things appear in an open environment all of the time. This isn't restricted to object classes. Novel activities and interactions are also significant confounds.

Scheirer et al., "Toward Open Set Recognition," *IEEE T-PAMI* 2012

OPEN SET RECOGNITION

There is a finite set of known objects in myriad unknown objects, combinations, and configurations — labeling something new, novel, or unknown should always be a valid outcome.



Scheierer et al., "Toward Open Set Recognition," *IEEE T-PAMI* 2012

OPEN SET RECOGNITION STRATEGIES

Feature Extractors: Training of networks that extract feature representations that enforce a better separation between known and novel classes when compared.

Examples: Objectosphere (Dhamija et al. NeurIPS 2018), Adversarial Reciprocal Points Learning (Chen et al. T-PAMI 2021), Class-Specific Semantic Reconstruction (Huang et al. T-PAMI 2023)

Post-Processors: Provide a novelty detection capability as a post-hoc add-on to any closed set network. Processes extracted features.

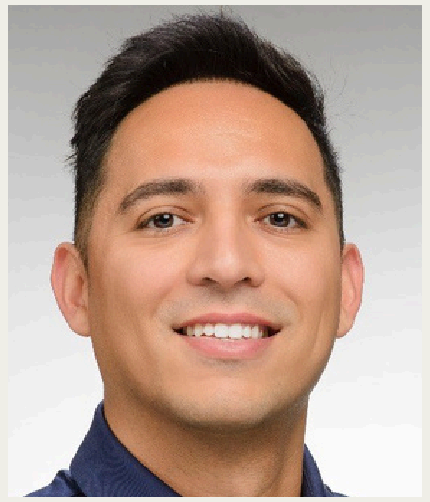
Examples: OpenMax (Bendale and Boult CVPR 2016), PostMax (Cruz et al. ECCV 2024), GHOST (Rabinowitz et al. AAAI 2025), COSTARR (Rabinowitz et al. ICCV 2025)

Related strategies: Anomaly Detection, Out-of-Distribution Detection, Open Vocabulary Object Detection

ADVANTAGES OF POST-PROCESSORS

- Practitioners encounter the open set recognition problem during deployment, where networks have already been trained
    - Post-processor methods allow pre-trained networks to accommodate open set recognition without re-training
- Architectures have broadly differing training recipes, incorporating OSR losses into training may yield wildly different results
    - Post-processors ensure consistency by not affecting base model performance

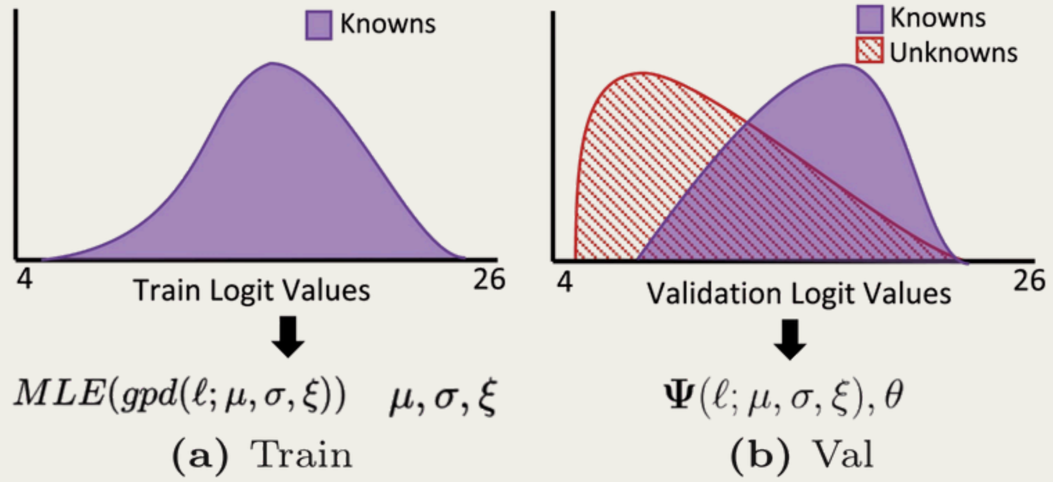
LET'S CONSIDER THREE RECENT APPROACHES



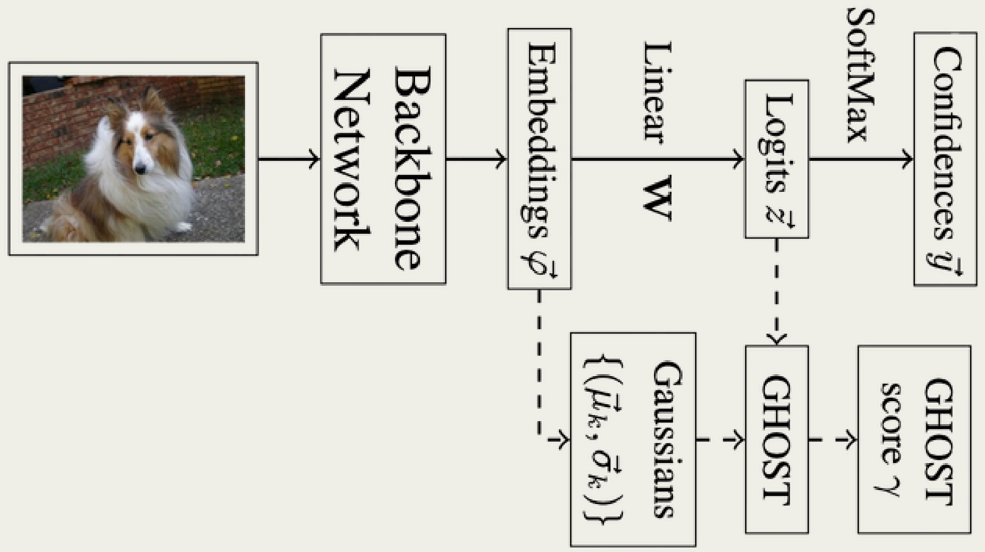
Steve Cruz

Ryan Rabinowitz

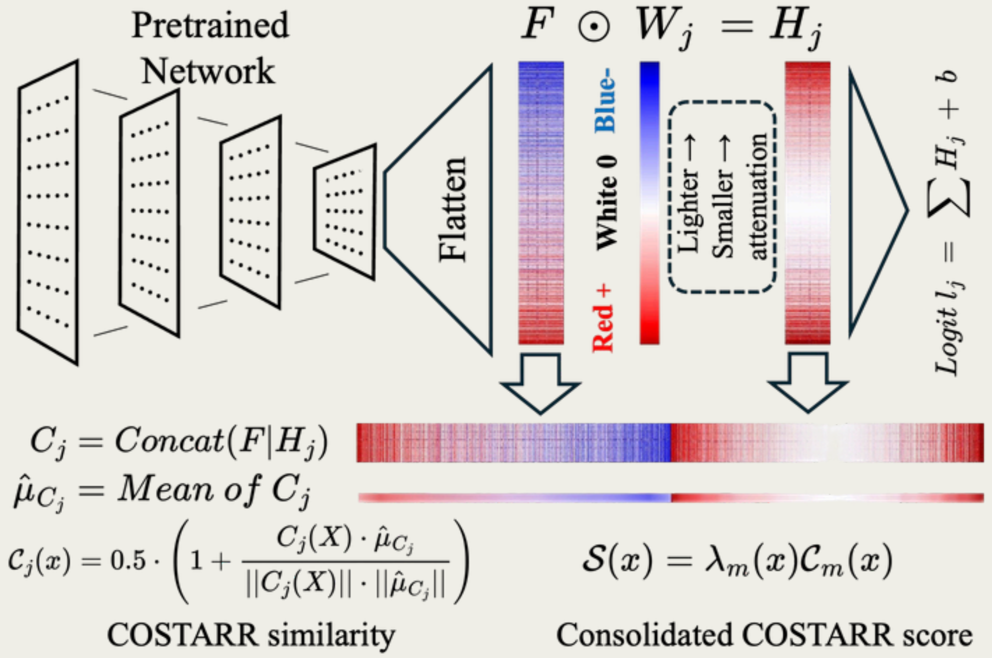
PostMax



GHOST



COSTARR



POSTMAX: WHAT DOES IT ADDRESS?

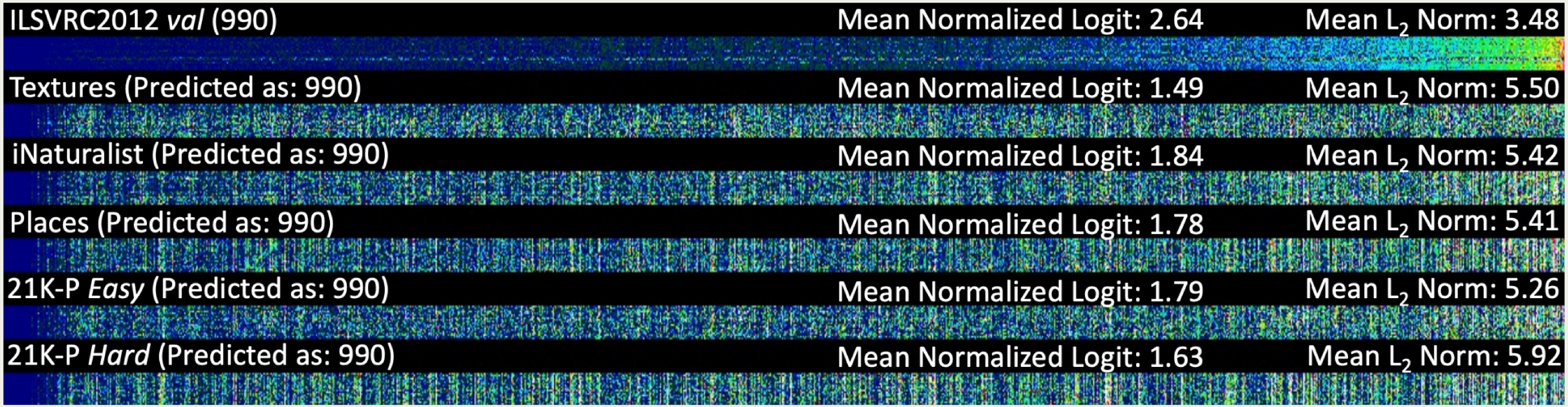
- **Operational thresholds** are common in open set recognition
  - The distribution of scores above the threshold leads to a Peak-over-Threshold formulation
  - Best modeled as an extreme value problem

- Post-processing refinement of the logit of the maximal class
  - Normalizes logits by deep feature magnitudes
  - Maps them to proper probabilities using the Extreme Value Theory-based generalized Pareto distribution

Cruz, S., Rabinowitz, R., Günther, M., & Boult, T. E. (2024, September). Operational Open-Set Recognition and PostMax Refinement. In European Conference on Computer Vision (pp. 475-492). Cham: Springer Nature Switzerland.

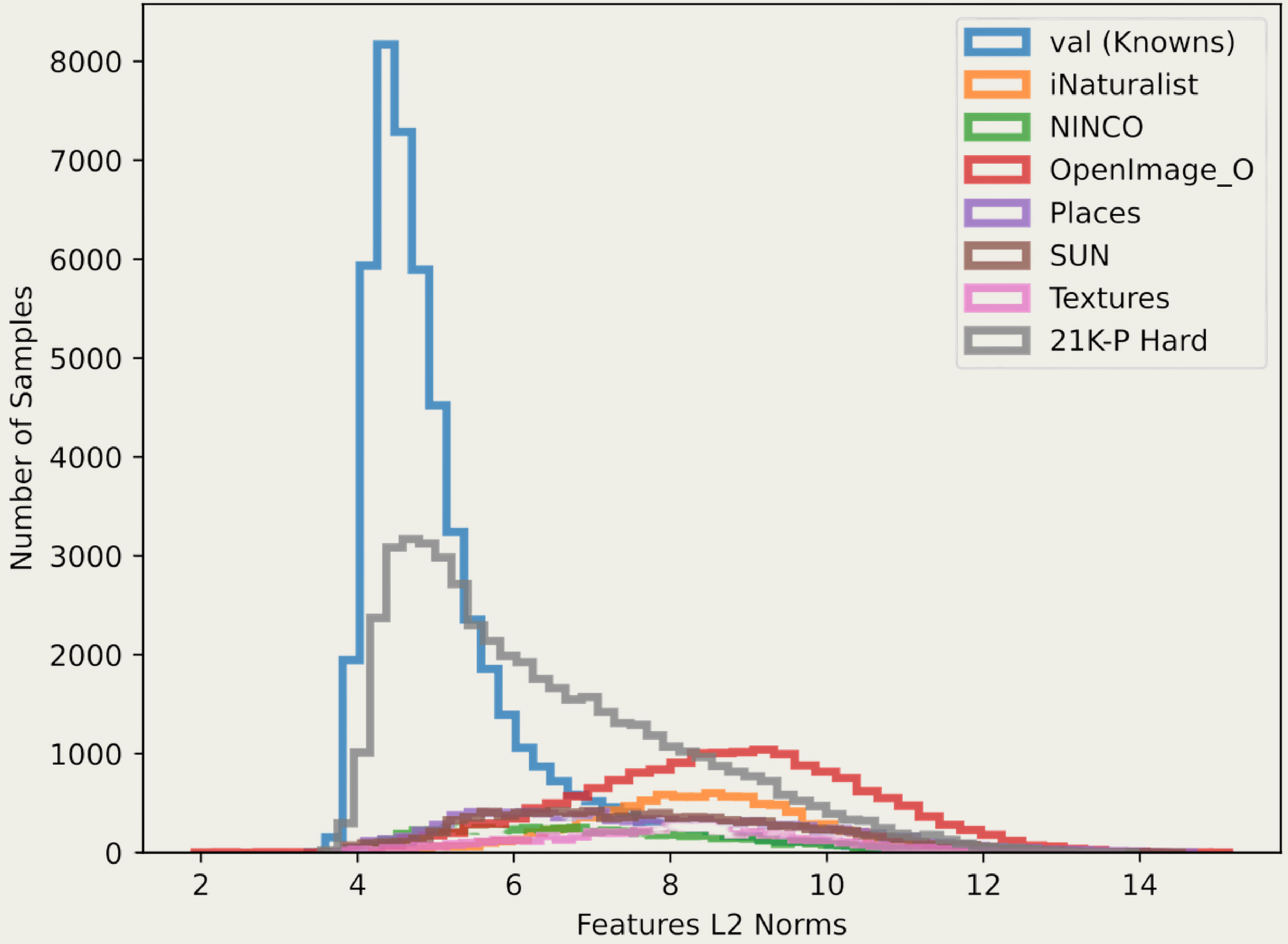
POSTMAX MOTIVATION

Squared deep feature values from a pre-trained ViT



Unknowns exhibit high feature values (bright colors) in regions where known sample activations are low.

POSTMAX MOTIVATION



Features from unknowns have, on average, larger L_2 norms than those of the correct class (knowns)

POSTMAX APPROACH

Algorithm 1 POSTMAX FITTING. This implements the probability distribution estimation of the proposed PostMax algorithm.

Require: D_{train}, $\mathrm{FV}(x)$, $\mathrm{L}(x)$

$M \leftarrow \{\}$

for each $(x, t) \in D_{\text{train}}$ **do**

  if $\arg\max \mathrm{L}(x) = t$ **then**

    $\ell \leftarrow \frac{\max \mathrm{L}(x)}{\|\mathrm{FV}(x)\|}$

    $M \leftarrow M \cup \{\ell\}$

  end if

end for

$\mu, \sigma, \xi \leftarrow \texttt{scipy.stats.genpareto.fit}(M)$

Use features and logits and divide by the feature magnitude to create a set for GPD fitting (Extreme Value Theory) and get probabilities.

We'll come back to thresholding for unknown detection in a little bit!

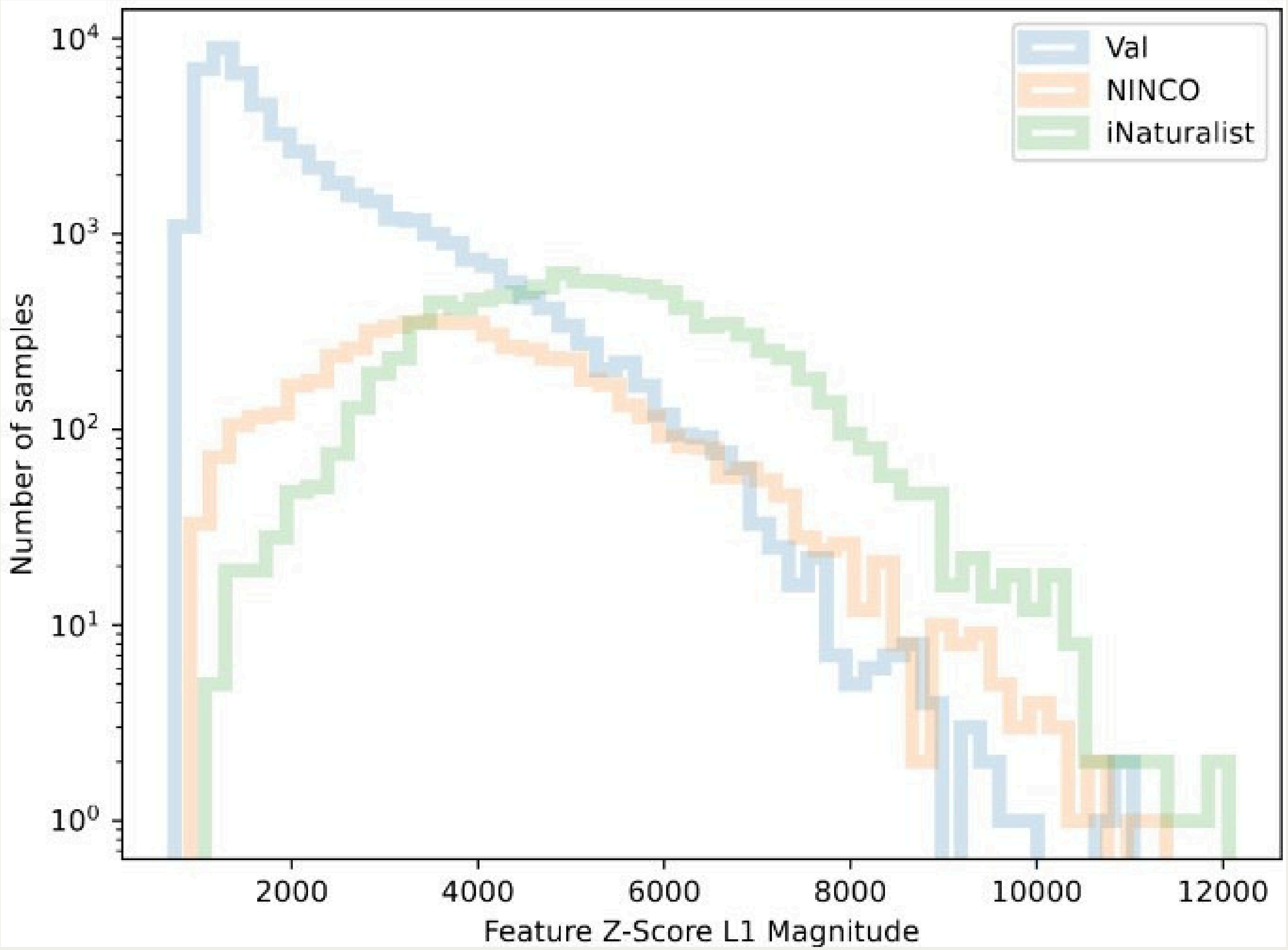
GHOST: WHAT DOES IT ADDRESS?

- Evaluations in open set recognition tend to focus on overall performance
  - **Fails to provide insights into performance across individual classes**
  - Affects handling of unseen or novel classes and ensuring a fair evaluation

- Gaussian Hypothesis Open Set Technique (GHOST)
  - Hyperparameter-free model of deep features using class-wise multivariate Gaussian distributions with diagonal covariance matrices

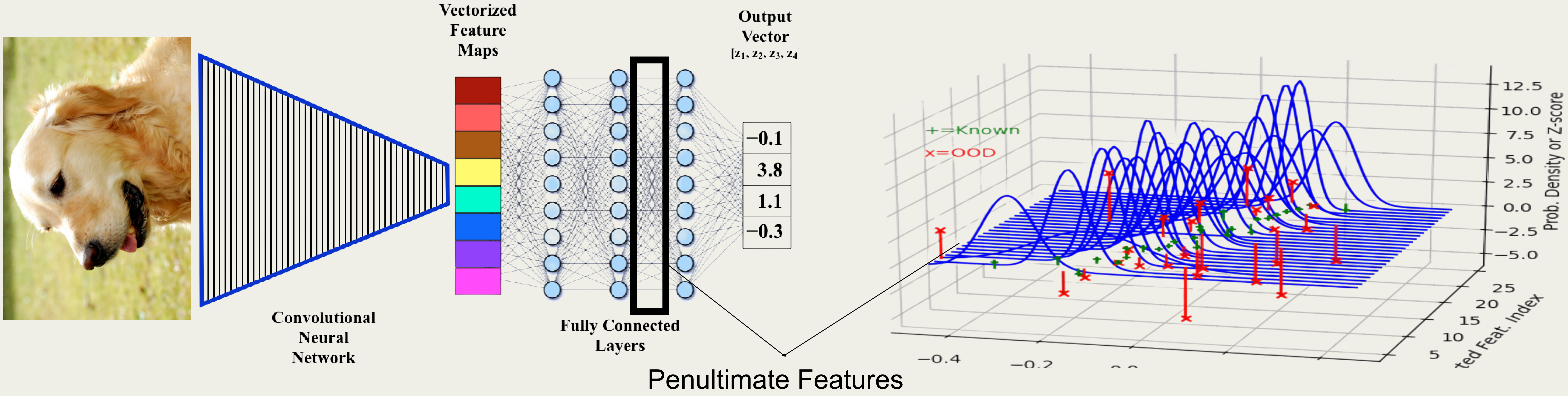
Rabinowitz, R., Cruz, S., Günther, M., & Boult, T. E. (2025, April). Ghost: Gaussian hypothesis open-set technique. In Proceedings of the AAAI Conference on Artificial Intelligence(Vol. 39, No. 6, pp. 6666-6674).

GHOST MOTIVATION



GHOST APPROACH: TRAINING

“When input is from a class seen in training, each value in the network can be reasonably approximated by a multivariate Gaussian and that, importantly, out-of-distribution samples would be more likely to be inconsistent with the resulting Gaussian model.” (Rabinowitz 2025)



There is evidence in favor of this trend using the Shapiro-Wilk test for normality and Holm’s step-down procedure for family-wise error control. For MAE-H and ConvNeXtV2-H, only 2.79% and 3.13% of distributions rejected the null hypothesis (normality).

GHOST APPROACH: PREDICTION

```python
def norm(self, logits, FVs):
    pred = torch.max(logits, dim=1).indices
    normalized_logits = torch.zeros(logits.shape)

    for c in tqdm(self.Gaus_dict.keys()):
        class_mask = pred == c
        mean_vector, std_vector = self.Gaus_dict[c]
        FV_Z_Score = (FVs[class_mask]-mean_vector)/std_vector
        FV_Z_Score = torch.abs(FV_Z_Score)
        diff_score = torch.sum(FV_Z_Score, dim=1)

        normalized_logits[class_mask,c] = logits[class_mask,c]/diff_score


    return normalized_logits
```

We'll come back to thresholding for unknown detection in a little bit!

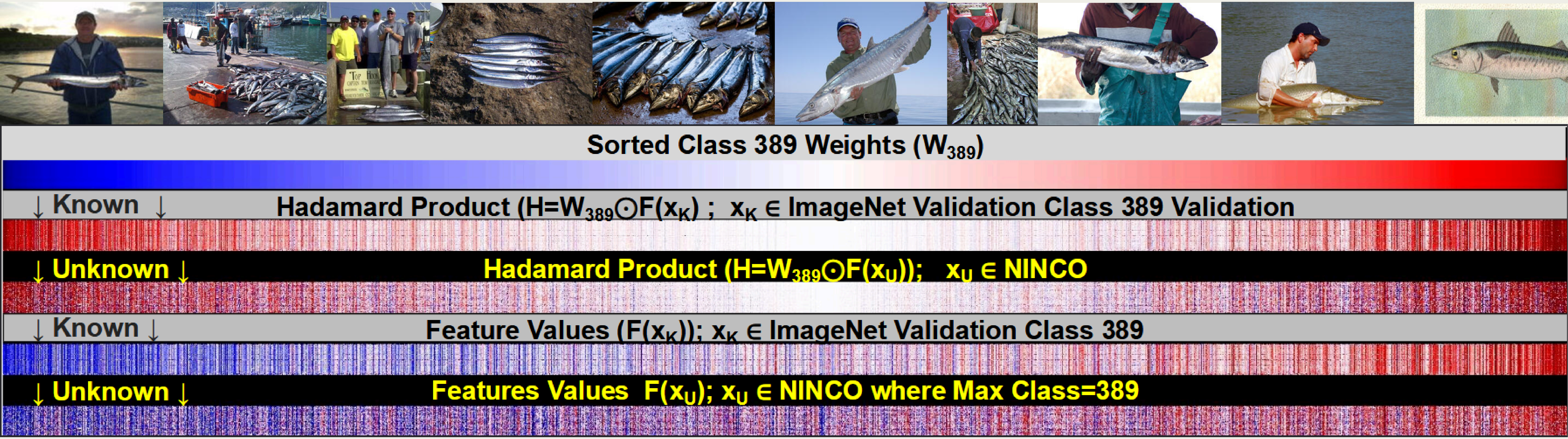
COSTARR: WHAT DOES IT ADDRESS?

- One novelty detection strategy: the **familiarity hypothesis** (Dietterich and Guyer 2022)
  - Detect novelty by the absence of familiar features
  - But why do approaches implicitly focus on only familiar features?

- New attenuation hypothesis to explain what is happening here
  - Small weights learned during training attenuate features and serve a dual role
    - They differentiate known classes while discarding information useful for distinguishing known from unknown classes

Rabinowitz, Ryan, Steve Cruz, Walter Scheirer, and Terrance E. Boult. "COSTARR: Consolidated Open Set Technique with Attenuation for Robust Recognition." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 4146-4155. 2025.

COSTARR MOTIVATION

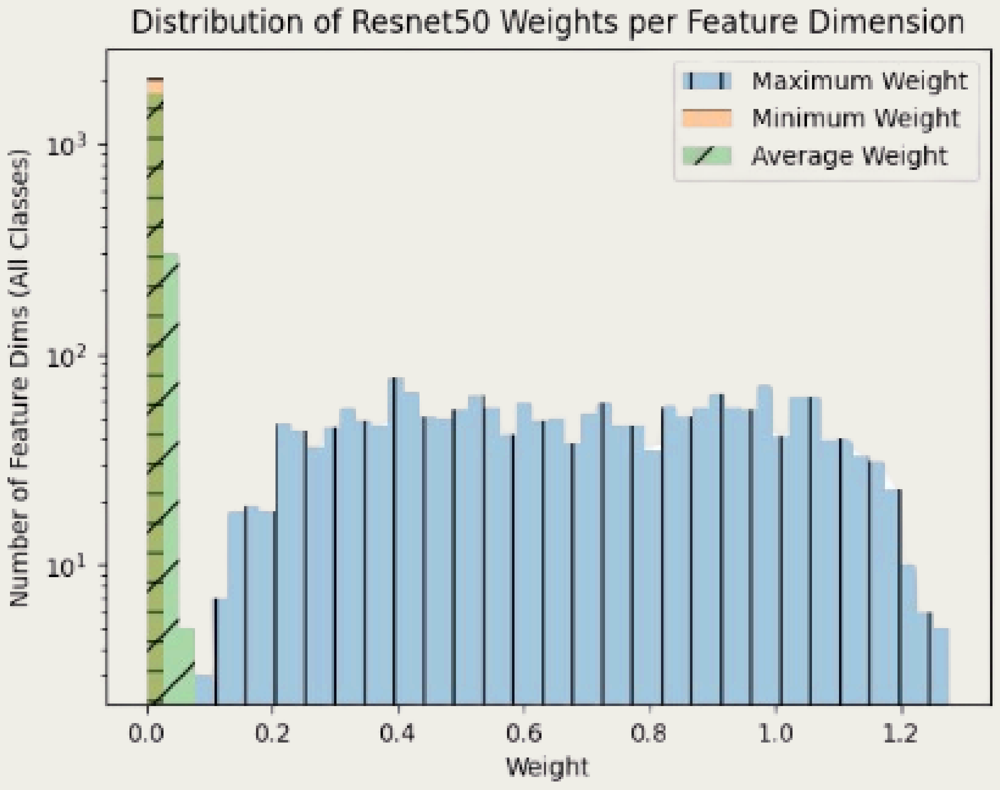


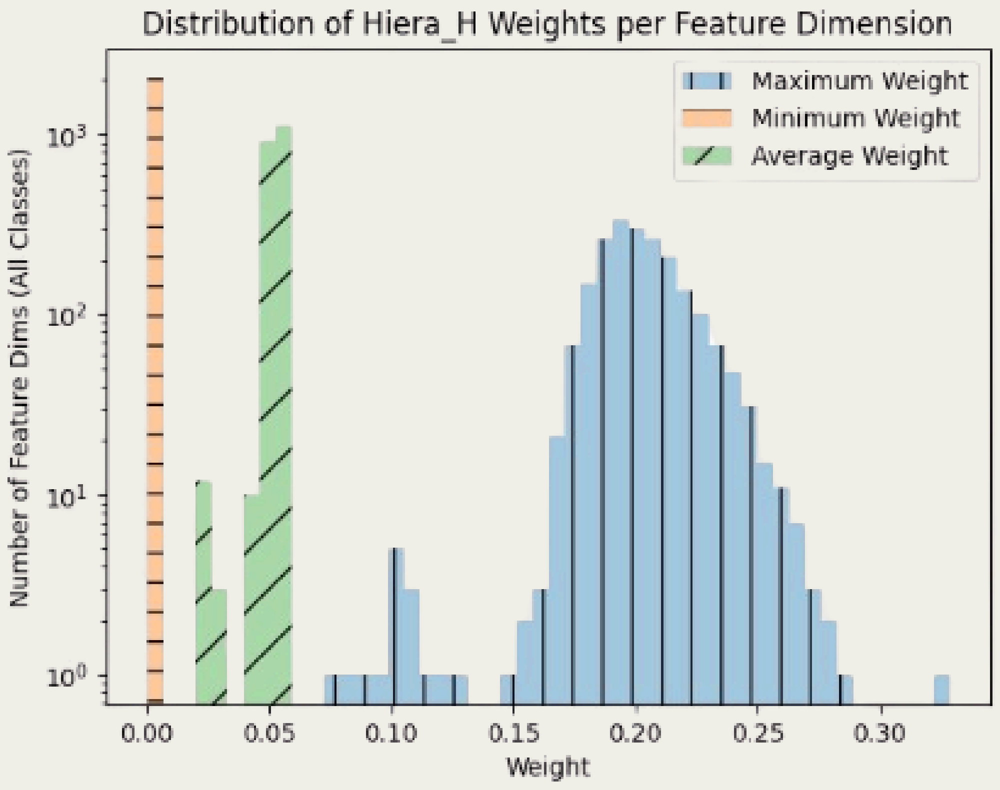
RED=Positive

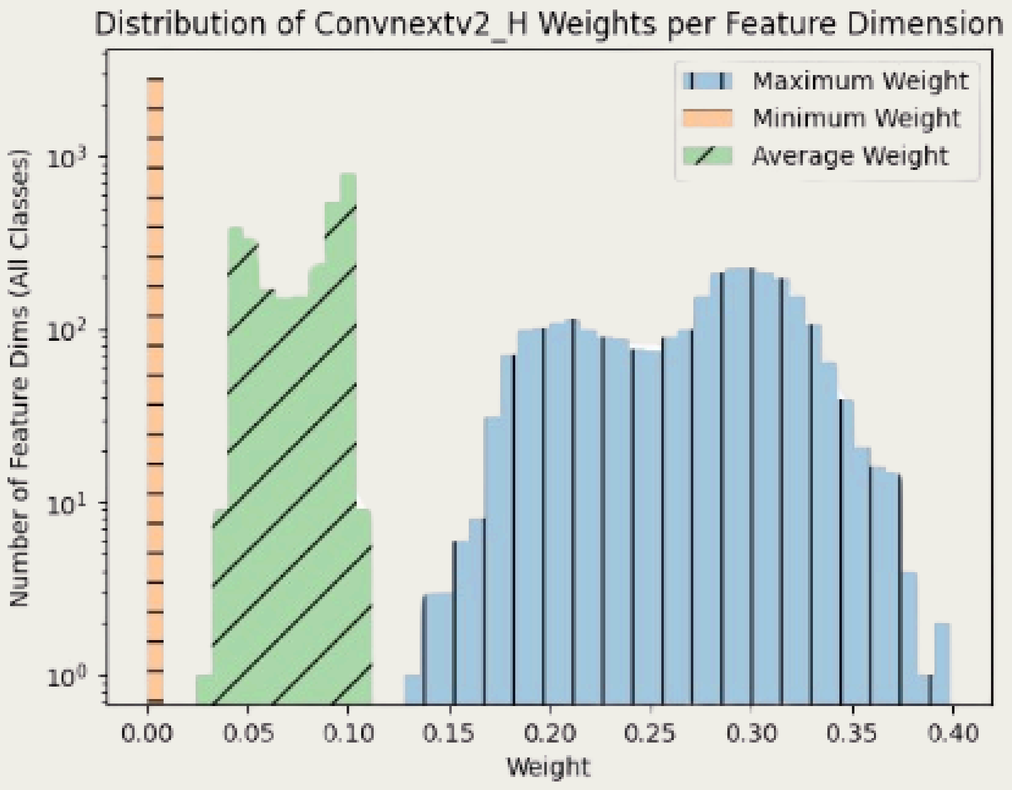
BLUE=Negative

Features which help distinguish known / unknown are being attenuated by classifier weights.

COSTARR MOTIVATION

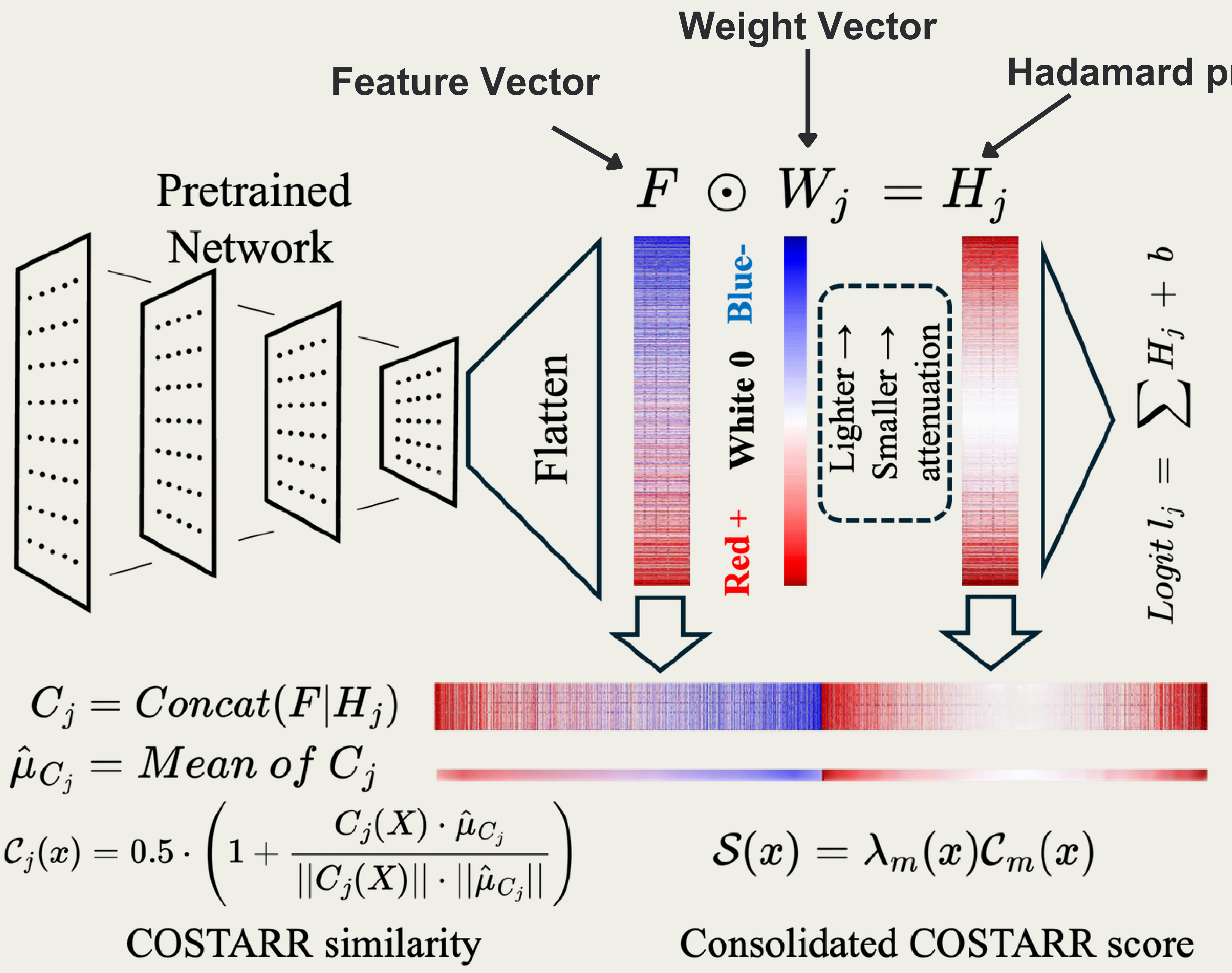






Every feature dimension is attenuated by some class.

COSTARR APPROACH



We'll come back to thresholding for unknown detection in a little bit!

EXPERIMENTS: HOW DO WE MEASURE?

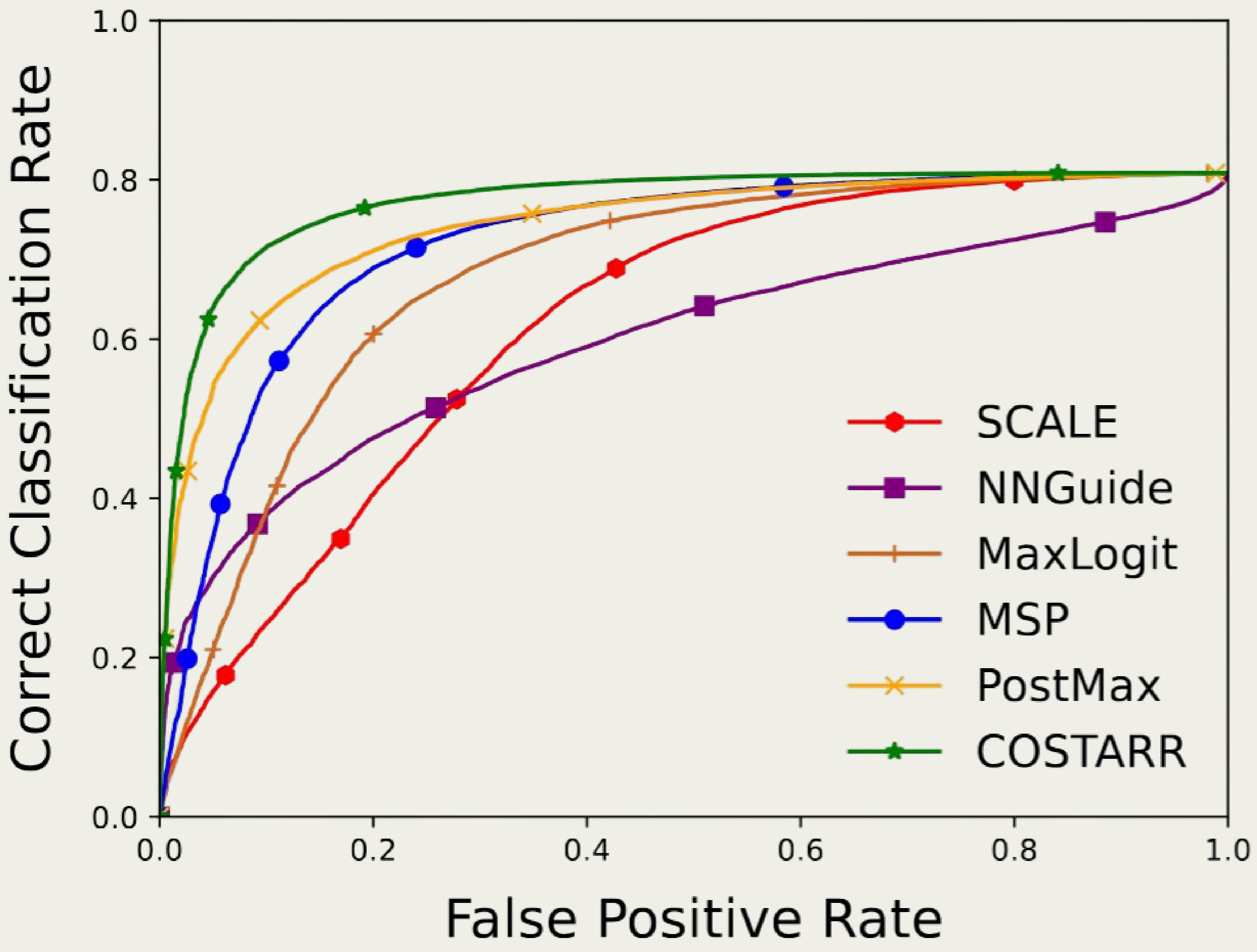
- Traditional Metrics 👎
  - AUROC / FPR95
    - Designed for binary problems
- Newer Metrics 👍
  - Open Set Classification Rate (OSCR)
  - Operational Open Set Accuracy (OOSA)
  - Coefficient of Variation (CoV) of Correct Classification Rate

OPEN SET CLASSIFICATION RATE CURVE

$$\mathrm{CCR}(\theta) = \frac{\left|\{(x_n, t_n) \in \mathcal{K} \wedge \hat{k} = t_n \wedge \gamma \geq \theta\}\right|}{|\mathcal{K}|},$$

$$\mathrm{FPR}(\theta) = \frac{\left|\{x_n \in \mathcal{U} \wedge \gamma \geq \theta\}\right|}{|\mathcal{U}|}.$$

The Open Set Classification Rate (OSCR) Curve plots the balance between correctly classified known samples and falsely accepted unknown samples.



Dhamija, A. R., Günther, M., & Boult, T. (2018). Reducing network agnostophobia. Advances in Neural Information Processing Systems, 31.

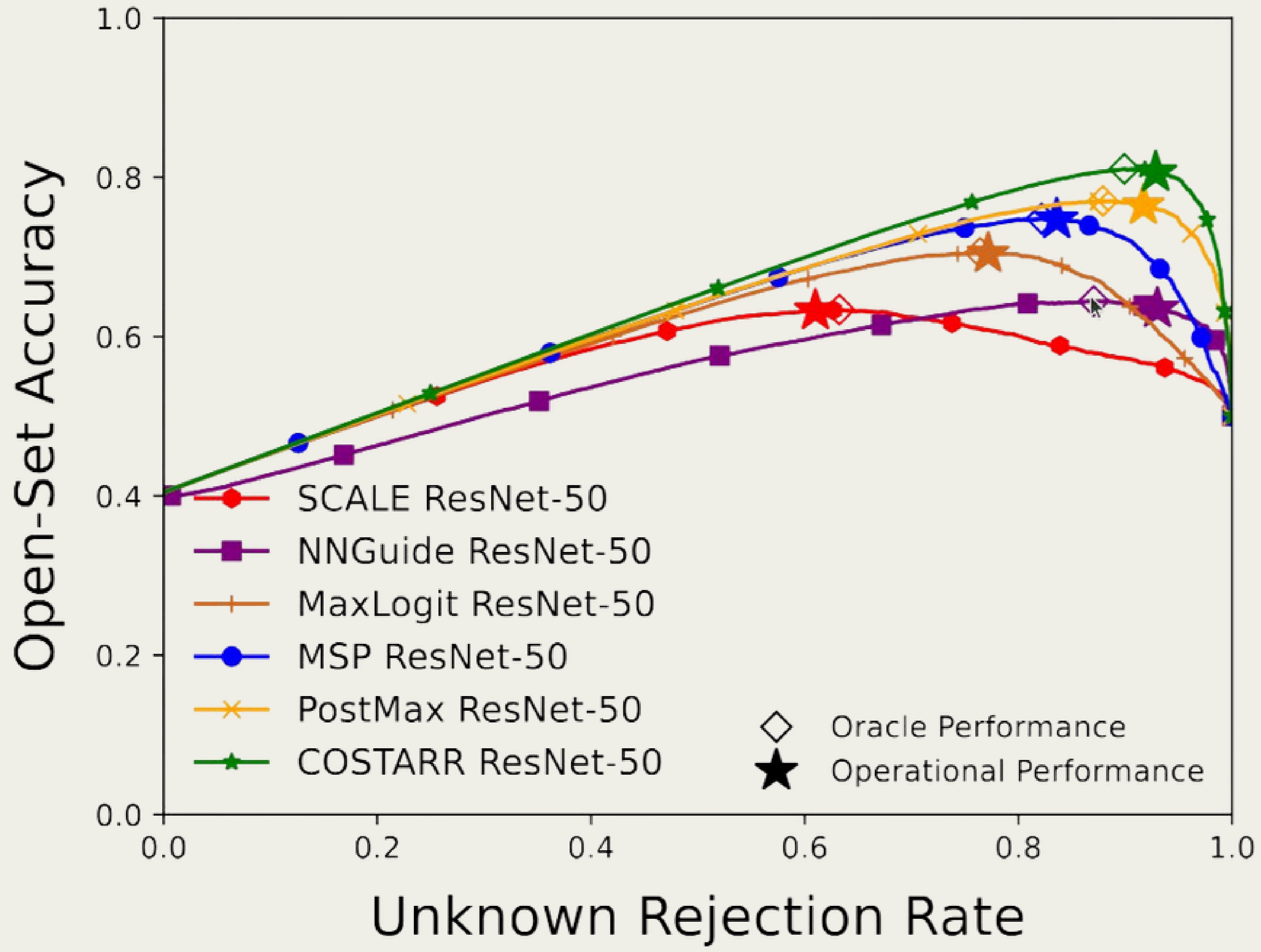
OPERATIONAL OPEN SET ACCURACY (OOSA)

$$\mathrm{CCR}(\theta) = \frac{\left|\{x \mid x \in D_c \wedge \arg\max_c P_c(x) = \hat{c} \wedge P_c(x) \geq \theta\}\right|}{|D_c|}$$

$$\mathrm{URR}(\theta) = \frac{\left|\{x \mid x \in D_a \wedge \max_c P_c(x) < \theta\}\right|}{|D_a|}$$

$$\mathrm{OSA}(\theta) = \alpha \cdot \mathrm{CCR}(\theta) + (1-\alpha) \cdot \mathrm{URR}(\theta)$$

Characterizes Operational Open Set Accuracy by selecting an optimal threshold on a validation dataset and reporting the resulting Open Set Accuracy on the test set.



Cruz, S., Rabinowitz, R., Günther, M., & Boult, T. E. (2024, September). Operational Open-Set Recognition and PostMax Refinement. In European Conference on Computer Vision (pp. 475-492). Cham: Springer Nature Switzerland.

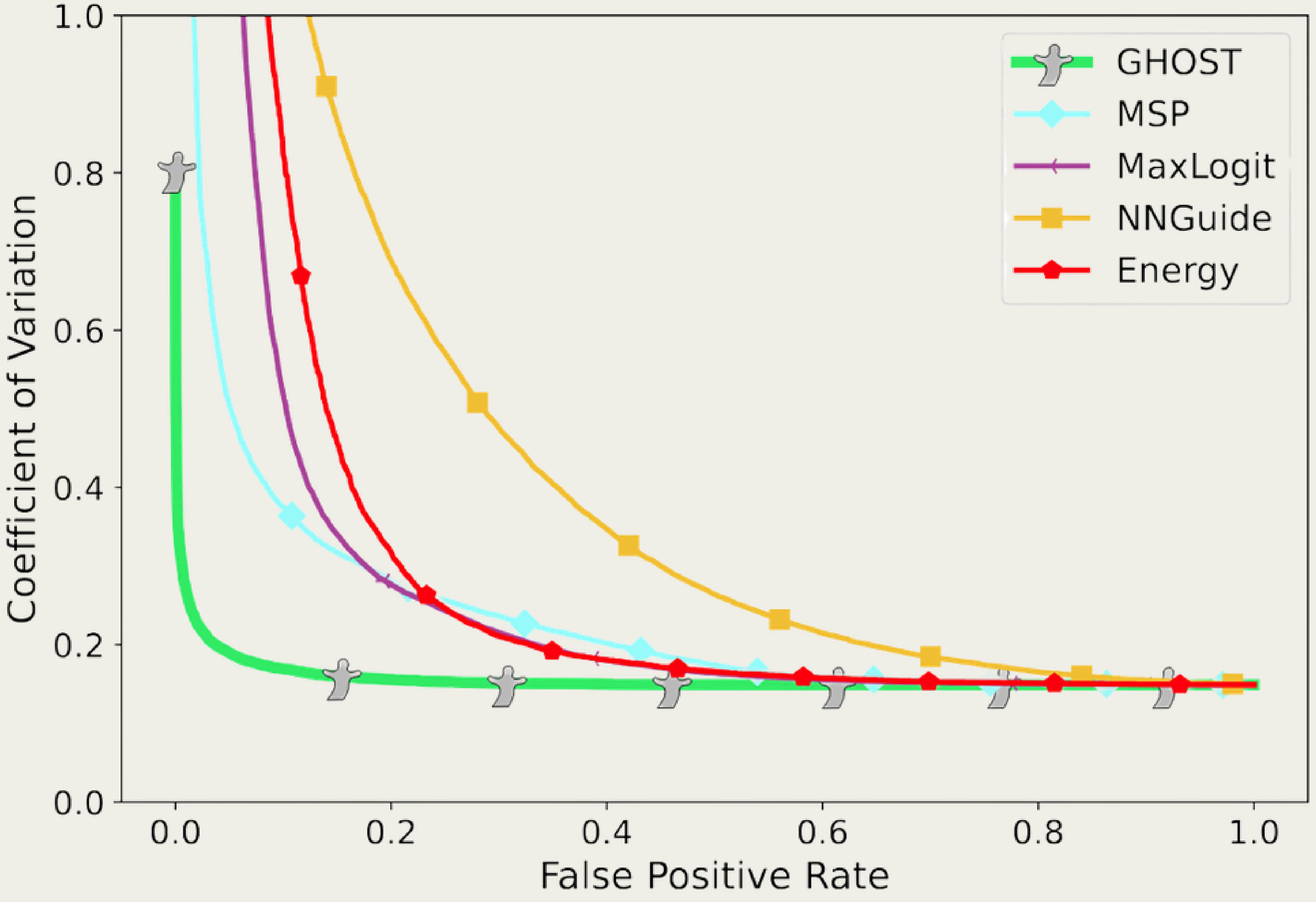
COEFFICIENT OF VARIATION OF CORRECT CLASSIFICATION RATE

$$\mathrm{CCR}_k(\theta_\tau) = \frac{\left|\{x_n \in \mathcal{K}_k \wedge \hat{k} = k \wedge \gamma \geq \theta_\tau\}\right|}{|\mathcal{K}_k|}$$

$$\mu_{\mathrm{CCR}}(\theta_\tau) = \frac{1}{K}\sum_k \mathrm{CCR}_k(\theta_\tau)$$

$$\sigma^2_{\mathrm{CCR}}(\theta_\tau) = \frac{1}{K-1}\sum_k \left(\mathrm{CCR}_k(\theta_\tau) - \mu_{\mathrm{CCR}}(\theta_\tau)\right)^2$$

$$\mathcal{V}_{\mathrm{CCR}}(\theta_\tau) = \frac{\sigma_{\mathrm{CCR}}(\theta_\tau)}{\mu_{\mathrm{CCR}}(\theta_\tau)}$$



As we select an FPR threshold, we would expect to sacrifice correct classification rate uniformly. Coefficient of variation shows that many post-processors eliminate known classes unequally. In some cases, entire classes are rejected just to achieve a moderate FPR.

Rabinowitz, R., Cruz, S., Günther, M., & Boult, T. E. (2025, April). Ghost: Gaussian hypothesis open-set technique. In Proceedings of the AAAI Conference on Artificial Intelligence(Vol. 39, No. 6, pp. 6666-6674).

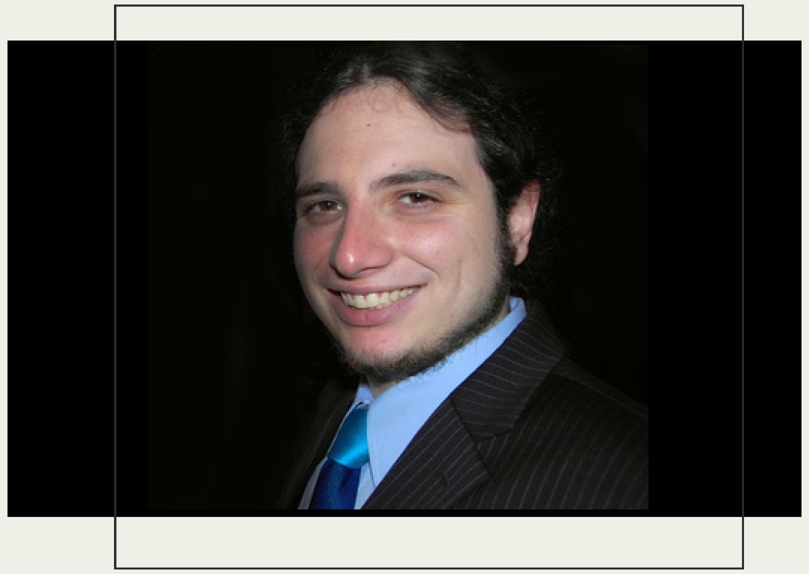
TAKEAWAYS

- Shift away from expensive computation, small-scale data, & traditional metrics
- Improve over good closed set classifiers
- Operationalize open set recognition
- System-wide thresholding has unequal effects on known classes

ACKNOWLEDGEMENTS

Notre Dame

Steve Cruz



Notre Dame

Ryan Rabinowitz

UCCS

Terry Boult

Questions?