# NOVA: A Benchmark for Rare Anomaly Localization and Clinical Reasoning in Brain MRI

Cosmin Bercea

NEURAL INFORMATION PROCESSING SYSTEMS

Munich Center for Machine Learning

HELMHOLTZ MUNICH

# Closed vs. Open World AI: The Medical Challenge

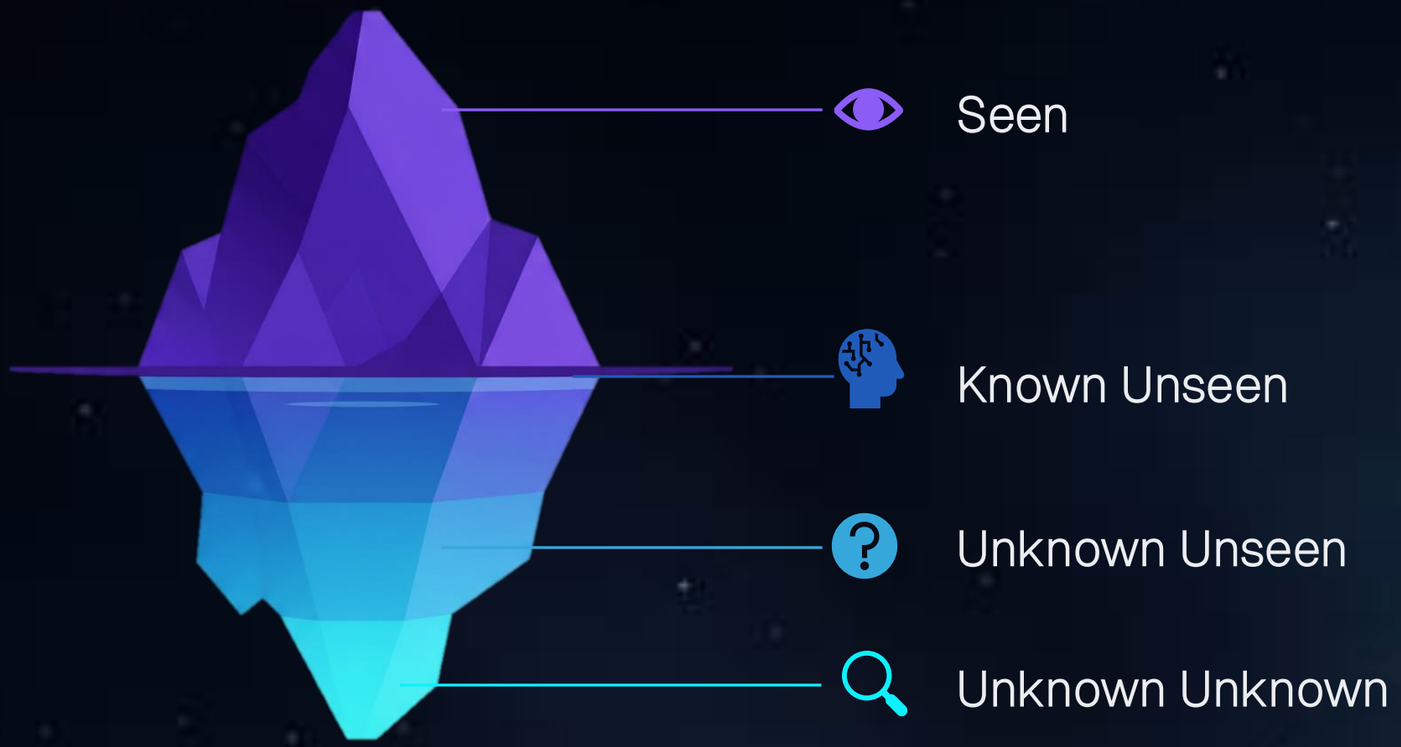
## "Closed World" of Labeled Data

Traditional AI excels with its training data; a "closed world" of meticulously labeled examples it already knows.

## "Open World" of Unseen Variations

The real world, especially in medicine, is an "open world" filled with novel and unlabelled variations unseen by AI.

# Closed vs. Open World AI: The Medical Challenge




**100M**

Americans with neurological disorder

Approximately 1 in 3 people

**15-20M**

Americans with rare diseases

~2,000 rare brain diseases models never encounter

# Closed vs. Open World AI: The Medical Challenge

Medical benchmarks mimic a closed-world setting

BraTS

ADNI

FastMRI+

Padchest

MIMIC-CXR

AI promises open-world recognition

VLM Grounding

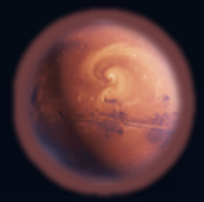
Open-Set Object Detection

OoD Detection

Novelty Detection

Anomaly Detection

# Closed-Set Evaluation

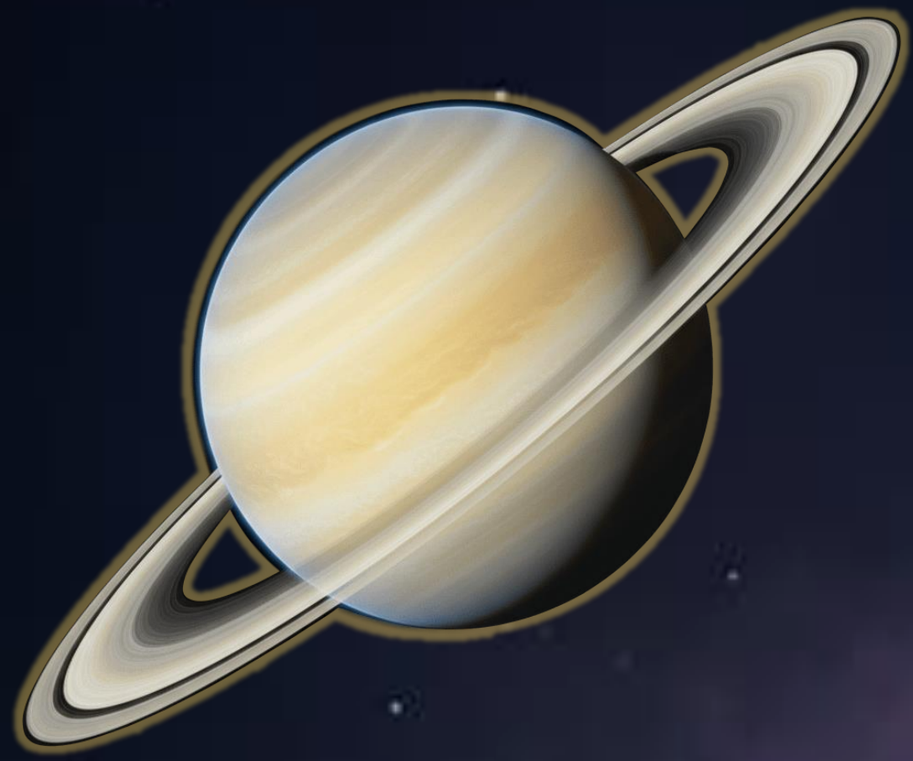


1 Disease

BraTS

14-25 Disease Classes

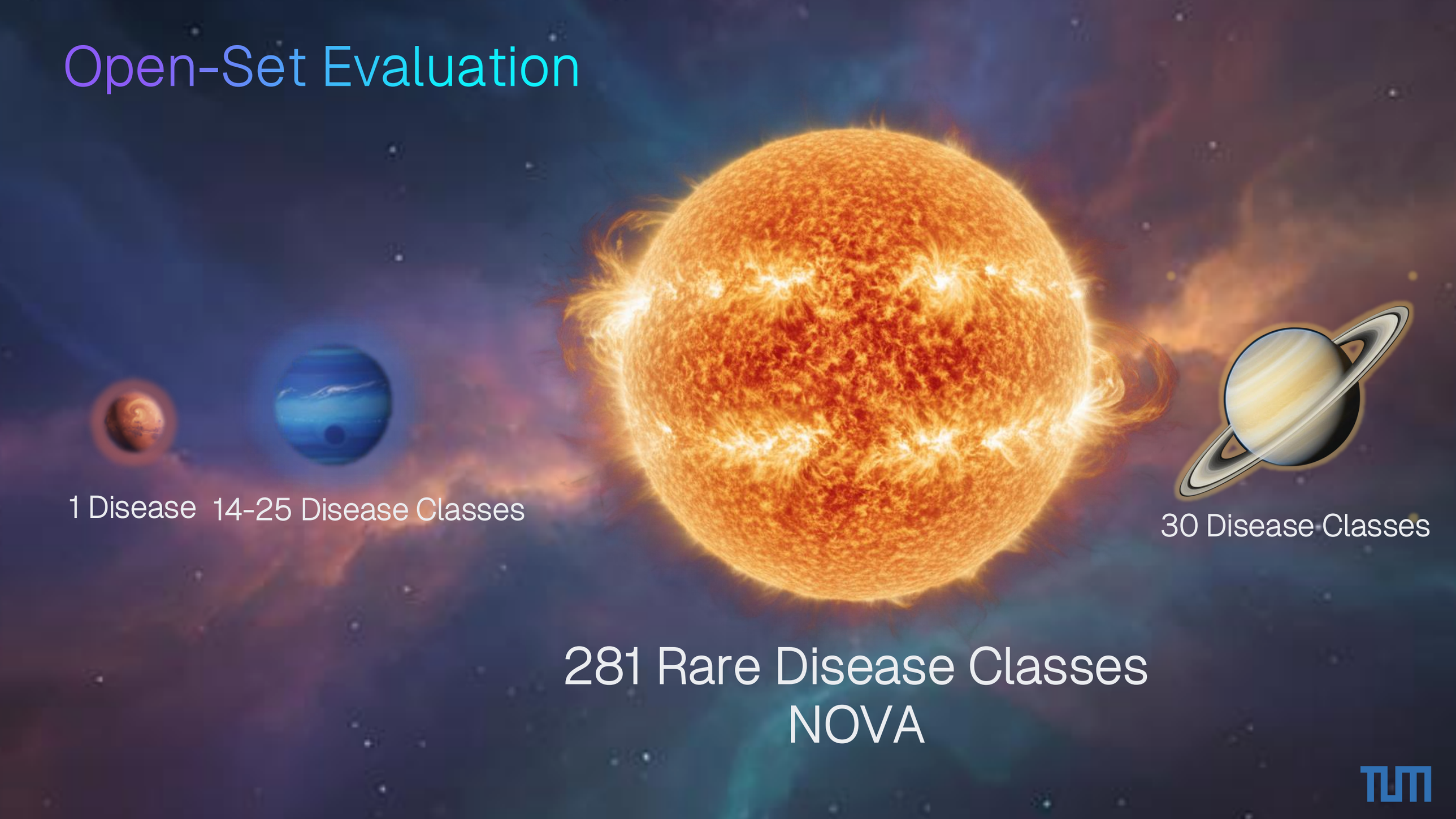
CXR-14, PAdchest



30 Disease Classes

FASTMRI+

# Open-Set Evaluation



NOVA

# A benchmark for *open-world generalization* and reasoning in vision–language models.

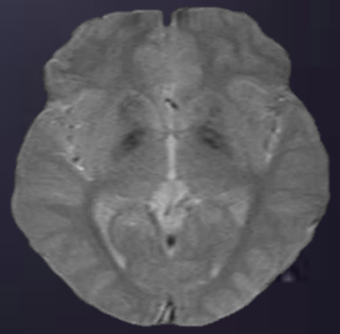
*Models must bridge a domain shift to brain MRI and a semantic shift to diseases never seen before.*

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*

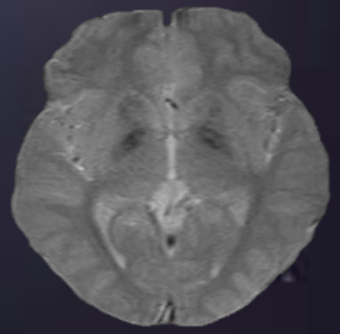
NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*



NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*

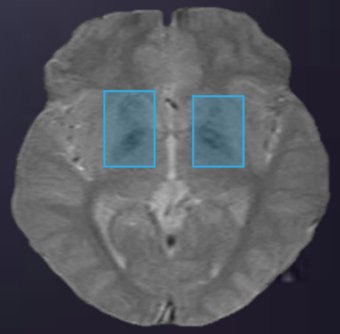


## Open-set detection

Detect unseen and rare anomalies under real distribution shift.

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*

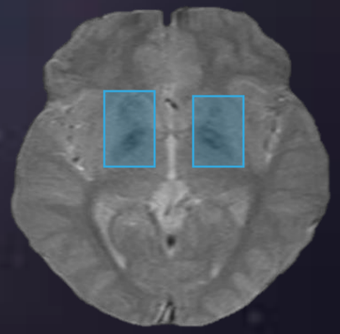


## Open-set detection

Detect unseen and rare anomalies under real distribution shift.

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*



*‘T2* image through brain showing drop out signals in **bilateral basal ganglia**, corresponding to **calcifications** seen on CT’*

## Open-set detection

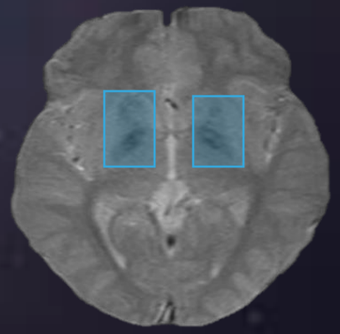
Benchmark for real distribution shift under uncertainty

## Captioning

Generate meaningful descriptions aligned with visual findings.

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*



*‘T2* image through brain showing drop out signals in **bilateral basal ganglia**, corresponding to **calcifications** seen on CT’*

*‘Birth anoxia’*

## Open-set detection

Benchmark for real distribution shift under uncertainty

## Captioning

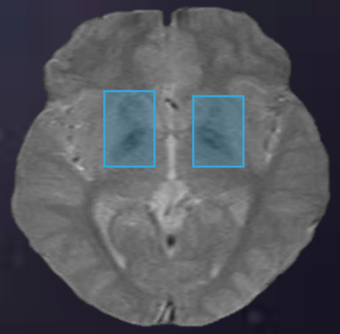
Generate meaningful descriptions aligned with visual findings.

## Reasoning

Integrate visual evidence and context to infer possible diagnoses.

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*



*‘T2* image through brain showing drop out signals in **bilateral basal ganglia**, corresponding to **calcifications** seen on CT’*

*‘Carbon monoxide intoxication’*

## Open-set detection

Benchmark for real distribution shift under uncertainty

## Captioning

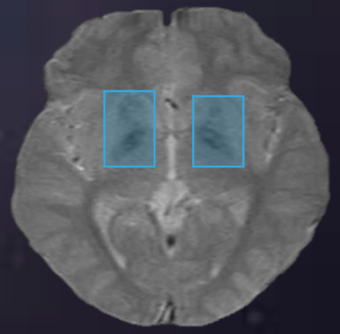
Generate meaningful descriptions aligned with visual findings.

## Reasoning

Integrate visual evidence and context to infer possible diagnoses.

NOVA

*‘A 15-year-old male patient with history of seizures and limb spasm presented at our hospital.’*



*‘T2* image through brain showing drop out signals in **bilateral basal ganglia**, corresponding to **calcifications** seen on CT’*

*‘Basal ganglia calcifications secondary to hypoparathyroidism‘*

## Open-set detection

Benchmark for real distribution shift under uncertainty

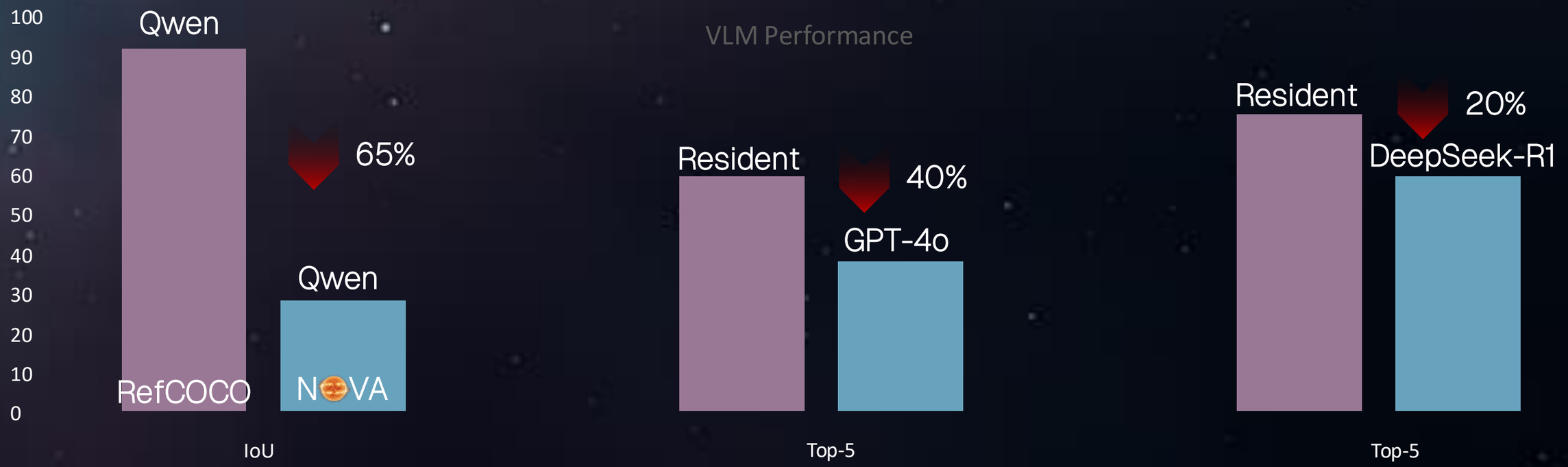
## Captioning

Generate meaningful descriptions aligned with visual findings.

## Reasoning

Integrate visual evidence and context to infer possible diagnoses.

NOVA




## Open-set detection

Benchmark for real distribution shift under uncertainty

## Captioning

Generate meaningful descriptions aligned with visual findings.

## Reasoning

Integrate visual evidence and context to infer possible diagnoses.

NOVA

# Join the Open-World Challenge

## Open-set detection

Benchmark for real distribution shift under uncertainty

## Captioning

Generate meaningful descriptions aligned with visual findings.

## Reasoning

Integrate visual evidence and context to infer possible diagnoses.

```
from datasets import load_dataset
ds = load_dataset("parquet", data_files=f"hf://datasets/c-i-ber/Nova/data/nova-v1.parquet", split="train")
```

NOVA is open. If your model can localise what it's never seen, describe it, and reason about it, you're pushing the boundaries of open-world AI.