Can I see an Example? Active Learning the Long Tail of Attributes and Relations

Tyler L. Hayes, Maximilian Nickel, Christopher Kanan, Ludovic Denoyer, & Arthur Szlam
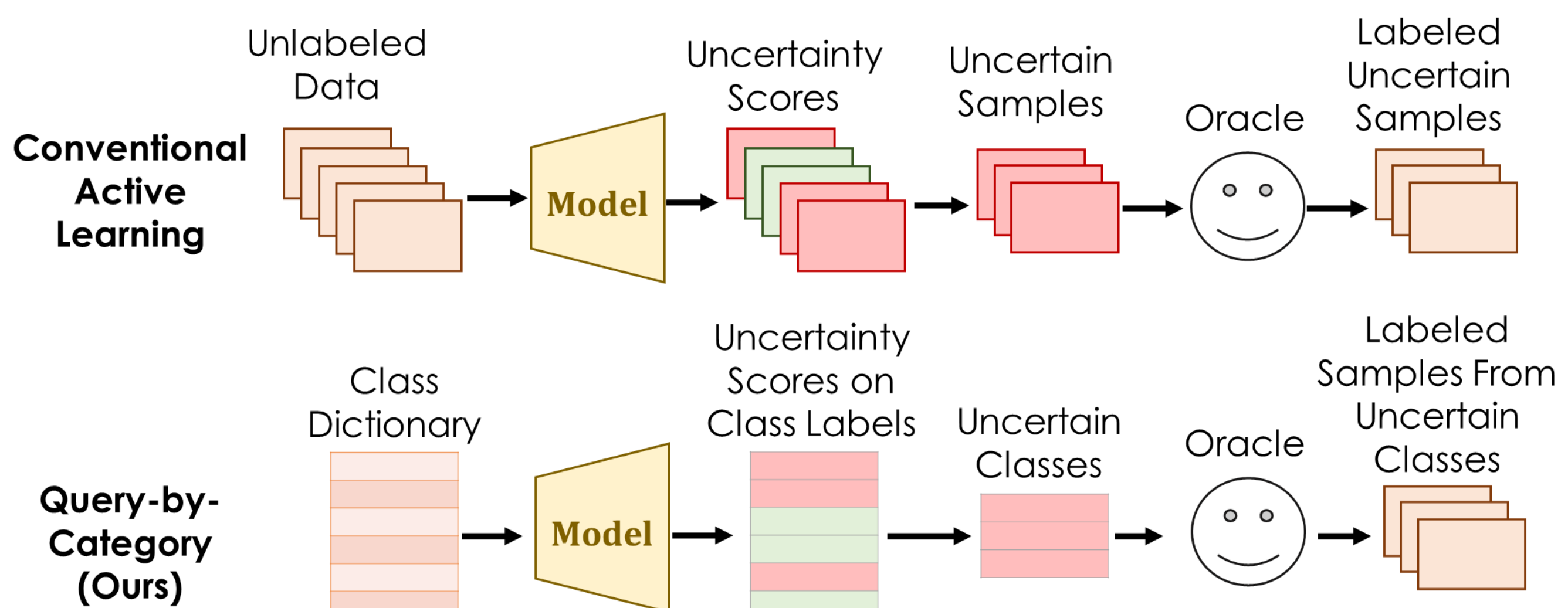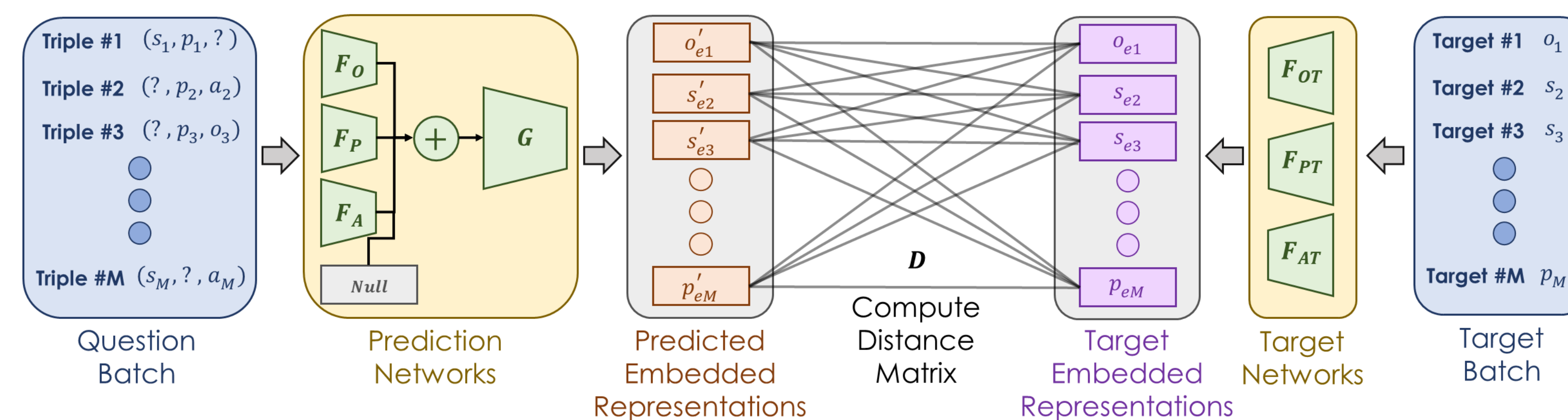
BMVC 2022

Overview

❖ An agent could better understand a scene by asking **questions** about particular **objects**, **attributes**, or **relationships** in a scene
❖ We introduce an active learning framework allowing agents to ask questions at the **category level** instead of the example level
❖ **Challenge:** The distribution of attributes and relations in the natural world is **long-tailed**, causing overfitting
    ❖ Problem is exacerbated in active learning settings
❖ **Goal:** Train an agent to better understand visual scenes while minimizing the number of questions it asks an annotator

Query-by-Category (QBCat) Framework
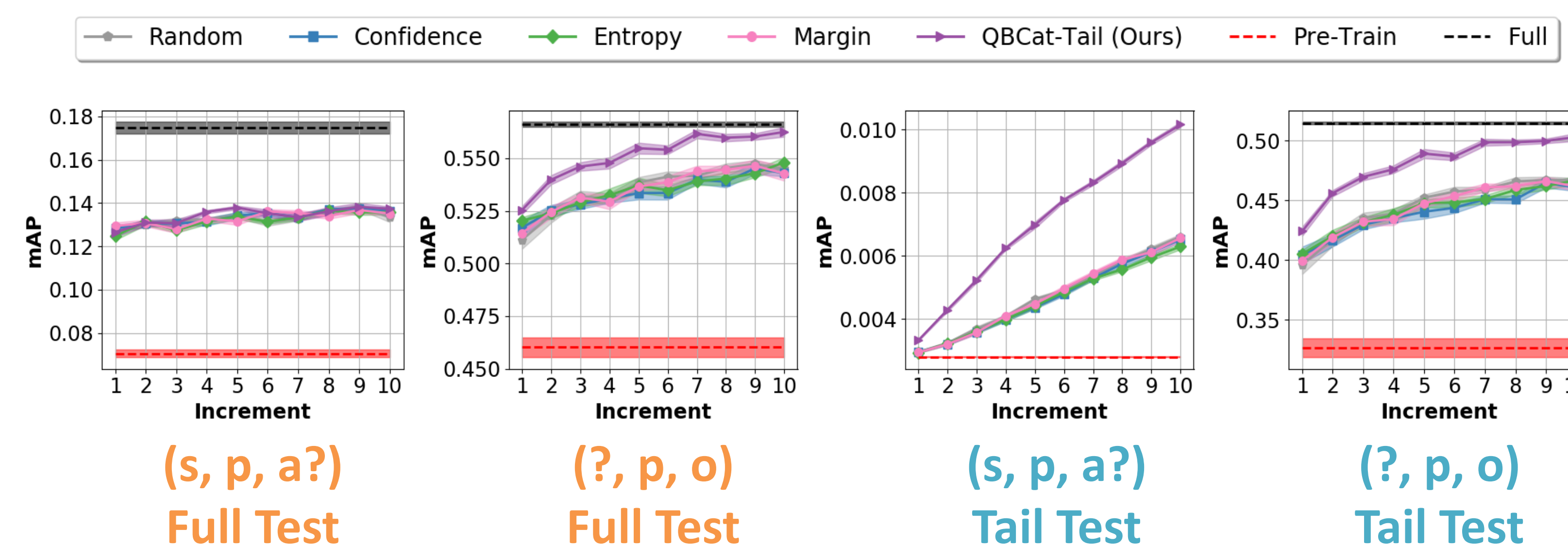

Model Architecture

❖ We perform **Visual Triple Completion** where an agent is provided with two elements in a triple, and it must predict the missing element
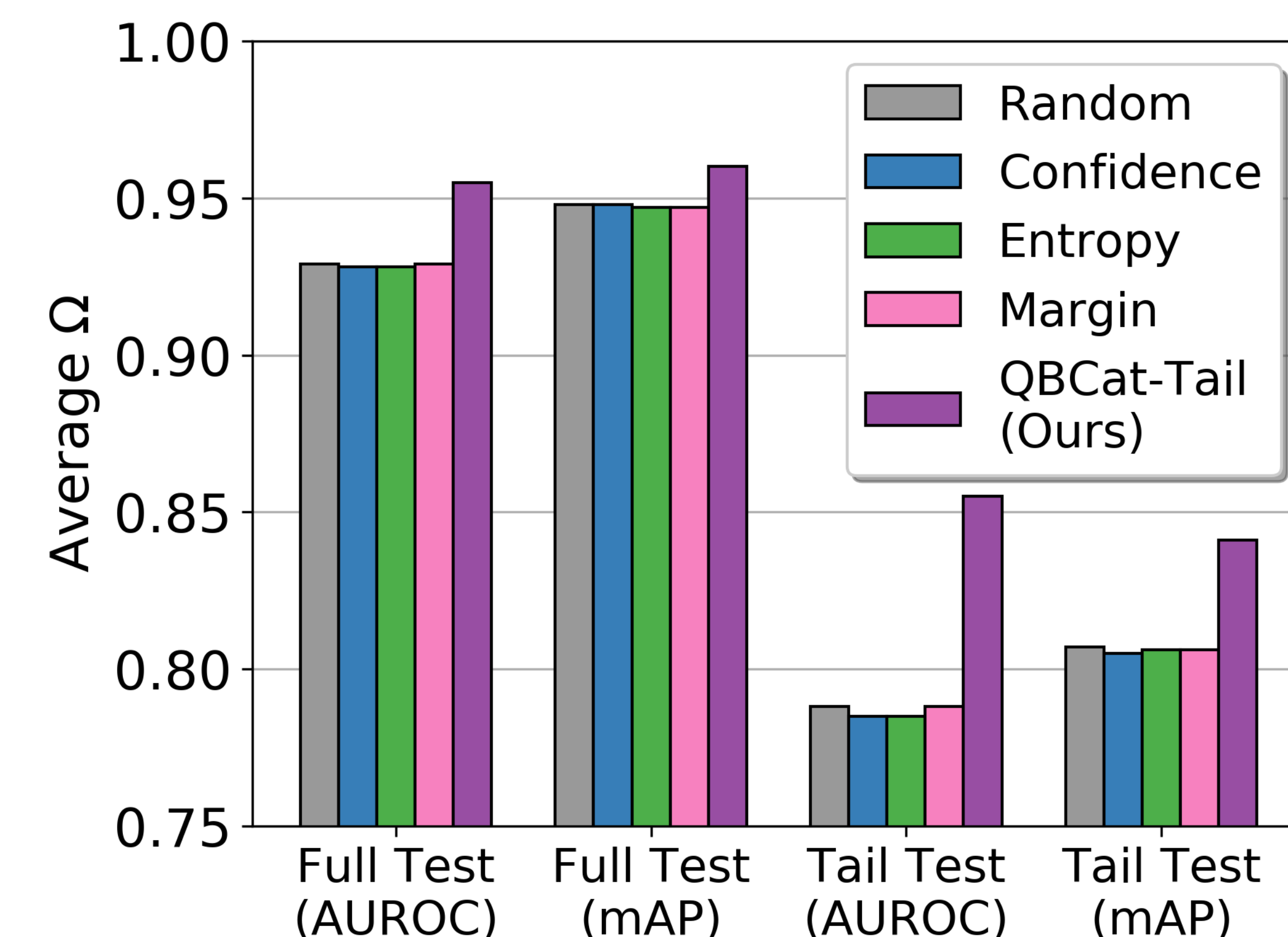❖ We use a Neighborhood Component Analysis Loss (Metric Learning) for training

Experimental Setup

❖ We propose the **QBCat-Tail** active sampling method that selects data from **tail classes uniformly at random**
❖ We compare QBCat-Tail to four conventional active learning baselines: **Random, Confidence, Entropy,** and **Margin**
❖ We perform incremental active learning over 10 increments
❖ We compare performance on the **Visual Genome** dataset using a **full test set** and a **tail test set**

Incremental Results

(s, p, a?)
Full Test

(?, p, o)
Full Test

(s, p, a?)
Tail Test

(?, p, o)
Tail Test

Overall Results

✓ QBCat-Tail has *strong performance on tail data* without sacrificing performance on the natural data distribution
✓ QBCat-Tail *outperforms all baselines*
✓ Performance differences among conventional active learners are *minimal*

Summary

❖ We introduced the **Query-by-Category** framework to train agents to predict objects, predicates, and attributes in **visual scenes**
❖ We introduced a simple yet effective active sampling approach (**QBCat-Tail**) that asks for examples from tail classes
❖ We demonstrated the effectiveness of QBCat-Tail on the Visual Genome dataset compared to active learning baselines