

特定知に適応するマルチモーダル基盤モデルの構築

オムロンサイニックエックス・シニアリサーチャー・齋藤邦章

目的 多様な環境でロバストに画像認識可能な基盤モデルの構築

内容 1. 多様なVLMのベンチマーキングおよびベンチマーキングに必要なデータの生成

2. VLMのハルシネーションを検出可能なモデルの構築

結果 1. VLMのベンチマーキング結果についてarxivに論文を投稿済

2. 現在、モデルを学習中であり、2026年度に論文投稿予定

利用した計算機 SQUID GPUノード群
ノード時間 700~1000 時間
使用メモリ 40 GB x 8GPUメモリ
並列化 1ノード

