SAGIVTECH IMVC 2026

15th Israel Machine Vision Conference (IMVC) 2026

Monday | April 27, 2026 | Pavilion 10, EXPO Tel Aviv

CARES: Context-Aware Resolution Selector for Efficient VLM Deployment

Moshe Kimhi, Technion - Israel Institute of Technology

Large vision–language models (VLMs) commonly process images at native or high resolution to remain effective across tasks. This inflates visual tokens up to 97-99% of total tokens, resulting in high compute and latency, even when low-resolution images would suffice. We introduce **CARES** - a **C**ontext-**A**ware **R**esolution **S**elector, a lightweight preprocessing module that, given an image–query pair, predicts the *minimal* sufficient input resolution. CARES uses a compact VLM (350M) to extract features and predict when a target pretrained VLM's response converges to its peak ability to answer correctly. Though trained as a discrete classifier over a set of optional resolutions, CARES interpolates continuous resolutions at inference for fine-grained control. Across five multimodal benchmarks spanning documents and natural images, as well as diverse target VLMs, CARES preserves task performance while reducing compute by up to 80%.