

NVIDIA Introduces Open-Source Project to Accelerate Computer Vision Cloud Applications

CV-CUDA combines accelerated image pre- and post-processing algorithms and tools to process 10x the number of images at the same cost.

Author: Michael Boone

Promising to help process images faster and more efficiently at a vast scale, NVIDIA introduced CV-CUDA , an open-source library for building accelerated end-to-end computer vision and image processing pipelines.

The majority of internet traffic is video. Increasingly, this video will be augmented by AI special effects and computer graphics.

To add to this complexity, fast-growing social media and video-sharing services are experiencing growing cloud computing costs and bottlenecks in their AI-based imaging processing and computer vision pipelines.

CV-CUDA accelerates AI special effects such as relighting, reposing, blurring backgrounds and super resolution.

NVIDIA GPUs already accelerate the inference portion of AI computer vision pipelines. But pre- and post-processing using traditional computer vision tools gobble up time and computing power.

CV-CUDA gives developers more than 50 high-performance computer vision algorithms, a development framework that makes it easy to implement custom kernels and zero-copy interfaces to remove bottlenecks in the AI pipeline.

The result is higher throughput and lower cloud-computing costs. CV-CUDA can process 10x as many streams on a single GPU.

All this helps developers move much faster when tackling video content creation, 3D worlds, image-based recommender systems, image recognition and video conferencing.

Video content creation platforms must process, enhance and moderate millions of video streams daily and ensure mobile-based users have the best experience running their apps on any phone.

For those building 3D worlds or metaverse applications, CV-CUDA is anticipated to enable tasks to help build or extend 3D worlds and their components.

In image understanding and recognition, CV-CUDA can significantly speed up the pipelines running at hyperscale, allowing mobile users to enjoy sophisticated and responsive image recognition applications.

And in video conferencing, CV-CUDA can support sophisticated augmented reality-based features. These features could involve complex AI pipelines requiring numerous pre- and post-processing steps.

CV-CUDA accelerates pre- and post-processing pipelines through hand-optimized CUDA kernels and natively integrates into C/C++, Python and common deep learning frameworks, such as PyTorch.

CV-CUDA will be one of the core technologies that can accelerate AI workflows in NVIDIA Omniverse , a virtual world simulation and collaboration platform for 3D workflows.

Developers can get early access to code in December, with a beta release set for March.

For more, visit the early access interest page .

Original URL: <https://blogs.nvidia.com/blog/2022/09/20/computer-vision-cloud/>