June 21, 2023

Intel Labs Introduces AI Diffusion Model, Generates 360-Degree Images from Text Prompts

LDM3D is industry's first generative AI model to deliver depth mapping. It has the potential to revolutionize content creation, metaverse and digital experiences.

SANTA CLARA, Calif.--(BUSINESS WIRE)-- Intel (Nasdaq: INTC):

This press release features multimedia. View the full release here: https://www.businesswire.com/news/home/20230621842353/en/

Intel Labs, in collaboration with Blockade Labs, has introduced Latent Diffusion Model for 3D (LDM3D), the industry's first diffusion model to deliver depth mapping to create 3D images with 360-degree views that are vivid and immersive. LDM3D has the potential to revolutionize content creation, metaverse applications and digital experiences, transforming a wide range of industries, from entertainment and gaming to architecture and design. (Credit: Intel Corporation)

What's New: Intel Labs, in collaboration with Blockade Labs, has introduced Latent Diffusion Model for 3D (LDM3D), a novel diffusion model that uses generative AI to create realistic 3D visual content. LDM3D is the industry's first model to generate a depth map using the diffusion process to create 3D images with 360-degree views that are vivid and immersive. LDM3D has the potential to revolutionize content creation, metaverse applications and digital experiences, transforming a wide range of industries, from entertainment and gaming to architecture and design.

> *"Generative AI technology aims to further augment and enhance human creativity and save time. However, most of today's generative AI models are limited to generating 2D images and only very few can generate 3D images from text prompts. Unlike existing latent stable diffusion models, LDM3D allows users to generate an image and a depth map from a given text prompt using almost the same number of parameters. It provides more accurate relative depth for each pixel in an image compared to standard post-processing methods for depth estimation and saves developers significant time to develop scenes."*
> —Vasudev Lal, AI/ML research scientist, Intel Labs

Why It Matters: Closed ecosystems limit scale. And Intel's commitment to true democratization of AI will enable broader access to the benefits of AI through an open ecosystem. One area that's seen significant advancements in recent years is in the field of computer vision, particularly in generative AI. However, many of today's advanced generative AI models are limited to generating only 2D images. Unlike existing diffusion models, which generally only generate 2D RGB images from text prompts, LDM3D allows users to generate both an image and a depth map from a given text prompt. Using almost

the same number of parameters as latent stable diffusion, LDM3D provides more accurate relative depth for each pixel in an image compared to standard post-processing methods for depth estimation.

This research could revolutionize how we interact with digital content by enabling users to experience their text prompts in previously inconceivable ways. The images and depth maps generated by LDM3D enable users to turn the text description of a serene tropical beach, a modern skyscraper or a sci-fi universe into a 360-degree detailed panorama. This ability to capture depth information can instantly enhance overall realism and immersion, enabling innovative applications for industries ranging from entertainment and gaming to interior design and real estate listings, as well as virtual museums and immersive virtual reality (VR) experiences.

On June 20th, LDM3D won the Best Poster Award at the 3DMV workshop at CVPR.

How It Works: LDM3D was trained on a dataset constructed from a subset of 10,000 samples of the LAION-400M database, which contains over 400 million image-caption pairs. The team used the Dense Prediction Transformer (DPT) large-depth estimation model (previously developed at Intel Labs) to annotate the training corpus. The DPT-large model provides highly accurate relative depth for each pixel in an image. The LAION-400M dataset has been built for research purposes to enable testing model training on larger scale for broad researcher and other interested communities.

The LDM3D model is trained on an Intel AI supercomputer powered by Intel® Xeon® processors and Intel® Habana Gaudi® AI accelerators. The resulting model and pipeline combine generated RGB image and depth map to generate 360-degree views for immersive experiences.

To demonstrate the potential of LDM3D, Intel and Blockade researchers developed DepthFusion, an application that leverages standard 2D RGB photos and depth maps to create immersive and interactive 360-degree view experiences. DepthFusion utilizes TouchDesigner, a node-based visual programming language for real-time interactive multimedia content, to turn text prompts into interactive and immersive digital experiences. The LDM3D model is a single model to create both an RGB image and its depth map, leading to savings on memory footprint and latency improvements.

What's Next: The introduction of LDM3D and DepthFusion paves the way for further advancements in multi-view generative AI and computer vision. Intel will continue exploring the use of generative AI to augment human capabilities and build a strong ecosystem of open source AI research and development that democratizes access to this technology. Continuing Intel's strong support for open ecosystem in AI, LDM3D is being open sourced through HuggingFace. This will allow AI researchers and practitioners to improve this system further and fine-tune it for custom applications.

More Context: Intel's research will be presented at the IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR) June 18-22. For more information, refer to "**LDM3D: Latent Diffusion Model for 3D**" or view the **LDM3D Demo**.

About Intel

Intel (Nasdaq: INTC) is an industry leader, creating world-changing technology that enables global progress and enriches lives. Inspired by Moore's Law, we continuously work to advance the design and manufacturing of semiconductors to help address our customers' greatest challenges. By embedding intelligence in the cloud, network, edge and every kind of computing device, we unleash the potential of data to transform business and society for the better. To learn more about Intel's innovations, go to newsroom.intel.com and intel.com.

View source version on businesswire.com:
https://www.businesswire.com/news/home/20230621842353/en/

Laura Stadler
laura.stadler@intel.com

Source: Intel