

Abstract

Diffusion models have become one of the most powerful frameworks in modern generative AI, achieving state-of-the-art performance in image, audio, and text synthesis. Inspired by physical diffusion processes, these models transform random noise into meaningful data through a step-by-step denoising process, offering superior quality and stability compared to earlier methods like GANs and VAEs. This seminar explores the theoretical foundations of diffusion models, their mechanism of operation, and their rapid advancements in image generation tasks such as style transfer, restoration, super-resolution, and text-to-image synthesis. It also highlights LayoutDiffusion, a cutting-edge controllable diffusion model for layout-to-image generation that uses structural image patches, Layout Fusion Modules, and object-aware cross-attention to deliver precise control over object placement and scene composition. In addition to technical insights, the seminar examines ethical and societal considerations—such as data privacy, authenticity of generated content, and responsible AI practices—underscoring the transformative impact and challenges of diffusion models across creative industries, entertainment, and healthcare. Participants will gain a comprehensive understanding of this evolving technology and its potential to shape the future of generative AI.