SILO AI

Stable Diffusion for Spatio-Temporal Consistent Generations

Join us at Silo AI and SiloGen to be at the forefront of AI innovation and European digital sovereignty.

About Silo AI
Silo AI is Europe's largest private AI lab. We partner with industry leaders to build smart devices, autonomous vehicles, Industry 4.0, and smart cities. Silo AI recently launched SiloGen, which is now part of a groundbreaking consortium aimed at building the world's largest open-source Large Language Model. With SiloGen, our specialized arm in generative AI, we are redefining the future of AI and ensuring European digital sovereignty.
We offer an unparalleled opportunity to work on cutting-edge generative AI research projects. You'll be part of a world-class team, working on applied research that pushes the boundaries of Generative and Multimodal AI.

The Project: Stable Diffusion for Spatio-Temporal Consistent Generations
Diffusion models have achieved remarkable success in generating highly realistic images. However, they often fall short when it comes to 3D awareness and spatio-temporal consistency. This project aims to extend the capabilities of stable diffusion models to generate not just images but also spatio-temporally consistent videos or 3D shapes.

Goals
- Conduct a rigorous literature review focusing on Stable Diffusion, Generative and Multimodal AI techniques.
- Identify innovative gaps, particularly in extending generative models to 3D and video generation.
- Design, develop, and benchmark your methodologies using our state-of-the-art computational resources.
- Engage in collaborative brainstorming sessions with our tech teams to refine and integrate your findings.
- Strive to publish your groundbreaking research in esteemed AI conferences.

Relevant Material
- [DreamFusion: Text-to-3D using 2D Diffusion] (https://dreamfusion3d.github.io/)
- [SceneScape: Text-Driven Consistent Scene Generation] (https://scenescape.github.io/)
- [MVDream: Multi-view Diffusion for 3D Generation] (https://mv-dream.github.io/)
- [SyncDreamer: Generating Multiview-consistent Images from a Single-view Image] (https://arxiv.org/abs/2309.03453v1)

Next Steps
This project demands a high level of expertise in Deep Learning, along with strong mathematical and coding skills, akin to what you'd acquire in the ML or SCR programs at KTH. Previous experience in Stable Diffusion, Generative, and Multimodal AI will be a strong advantage. Given the ambitious nature of this project, be prepared for a commitment that may go beyond the typical scope of MSc thesis projects.

To apply, submit the following in English to Pier Luigi Dovesi (pierluigi.dovesi@silo.ai):
- CV
- Cover letter
- Transcripts (with grades)