

1. Introduction

The field of computer vision has witnessed significant advancements with the integration of generative AI and foundational models.

Generative models like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and transformers have revolutionized the field.

2. State of the Art

Recent advancements in generative AI and foundational models have introduced several state-of-the-art techniques.

Generative Adversarial Networks (GANs): GANs are capable of generating realistic images by learning the underlying data distribution.

Variational Autoencoders (VAEs): VAEs learn a probabilistic representation of the data, enabling the generation of diverse samples.

Transformers in Computer Vision: Models like Vision Transformers (ViT) and CLIP (Contrastive Language-Image Pre-training) have shown remarkable performance.

Multimodal Models: Combining vision and language models has led to the development of powerful tools for tasks like image captioning and visual question answering.

3. Research Project

This research project aims to explore innovative approaches in generative AI and foundational models to enhance computer vision tasks.

Development of Efficient Generative Models: Investigate techniques to improve the efficiency and scalability of generative models.

Domain Adaptation and Fine-Tuning: Explore methods for fine-tuning generative models on domain-specific data to improve performance.

Ethical and Responsible AI: Develop frameworks for bias detection and mitigation in generative models, ensuring transparency and accountability.

The didactic-scientific path of the proposal is divided into the following training modules:

First Year:

Module 1: Acquisition of in-depth knowledge on generative AI technologies and their applications in computer vision.

Module 2: Acquisition of skills in deep learning and foundational models, focusing on their implementation and optimization.

Second Year:

Module 3: Application of acquired skills to develop efficient generative models for specific computer vision tasks.

Third Year:

Module 4: Experimentation and enhancement of research results through collaborations with research centers or industry partners.

Module 5: Writing the Ph.D. thesis and publications, documenting the research findings and contributions to the field.

Conclusion

This research project aims to advance the field of computer vision by developing efficient, scalable, and ethically sound generative models.