

**CÁ NHÂN HÓA MÔ HÌNH TẠO SINH ẢNH  
TỪ VĂN BẢN VỚI ÍT MẪU DỮ LIỆU**

**FEW-SHOT PERSONALIZATION OF  
TEXT-TO-IMAGE GENERATION MODELS**

**Trần Gia Nghĩa - 240101019**

# Tóm tắt

- Họ và Tên: Trần Gia Nghĩa
- Lớp: CS2205.CH183
- Link Github của nhóm:  
<https://github.com/SoulOfWindTGN/CS2205.CH183>
- Link YouTube video: <https://youtu.be/fKLgz0VTSFM>



# Giới thiệu

**Cá nhân hóa mô hình tạo sinh ảnh** hướng đến việc tinh chỉnh hoặc điều chỉnh có điều kiện các mô hình khuếch tán để chúng có thể tạo ra hình ảnh theo sở thích cá nhân.



# Giới thiệu

## Mục tiêu bài toán:

- Học và duy trì được đặc trưng, phong cách (style) hoặc chủ thể (subject) từ dữ liệu của người dùng/nhóm người dùng.
- Vẫn giữ được tính khái quát của mô hình gốc (không đánh mất quá nhiều khả năng sinh ảnh đa dạng).

## Hạn chế của các phương pháp hiện tại:

**Ảnh tham chiếu**



**+** *"beside a swimming pool."*

**Chủ thể - Nền**



**Chủ thể - Chủ thể**



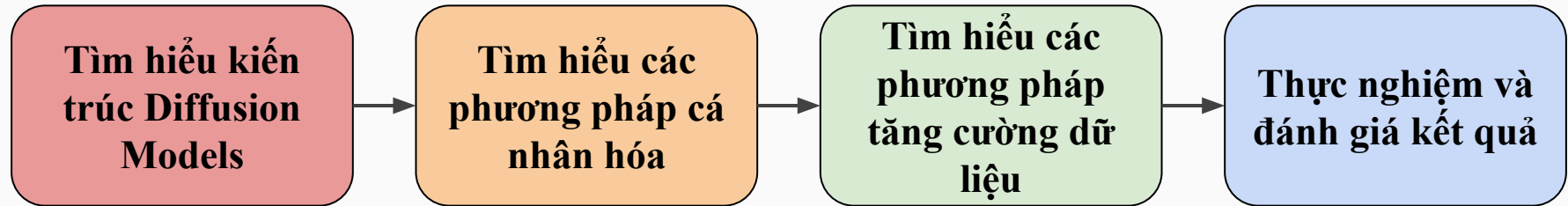
**Quá khớp**

**Trộn lẫn thuộc tính**

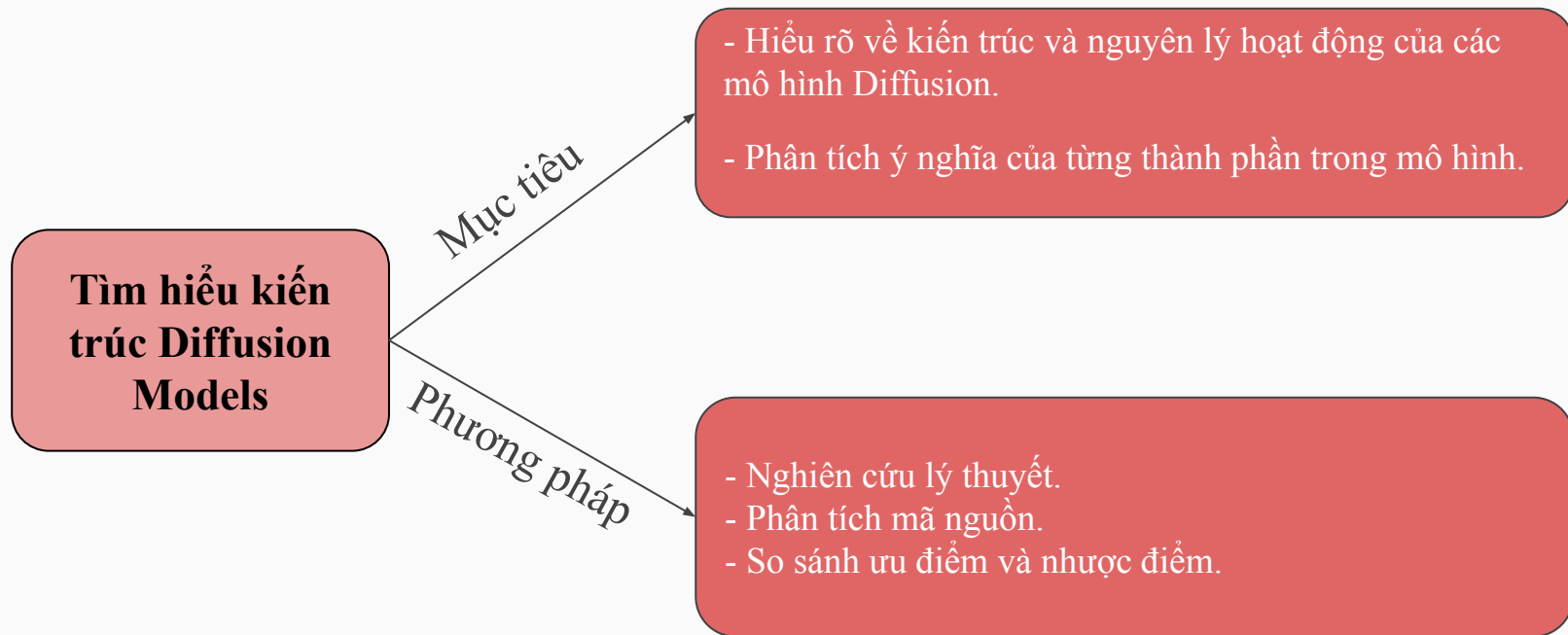
# Mục tiêu

- Nghiên cứu quy trình tinh chỉnh tham số cho mô hình khuếch tán (ví dụ Stable Diffusion) nhằm nhúng các khái niệm (concept) cá nhân hóa một cách hiệu quả từ một tập ảnh tham chiếu nhỏ vào quá trình sinh ảnh.
- Xây dựng, tích hợp và đánh giá hiệu quả của kỹ thuật tăng cường dữ liệu để hạn chế quá khớp (overfitting) trong điều kiện ít dữ liệu.
- Triển khai và thử nghiệm và đánh giá phương pháp đề xuất cho cả tác vụ tạo sinh đơn hoặc đa khái niệm.

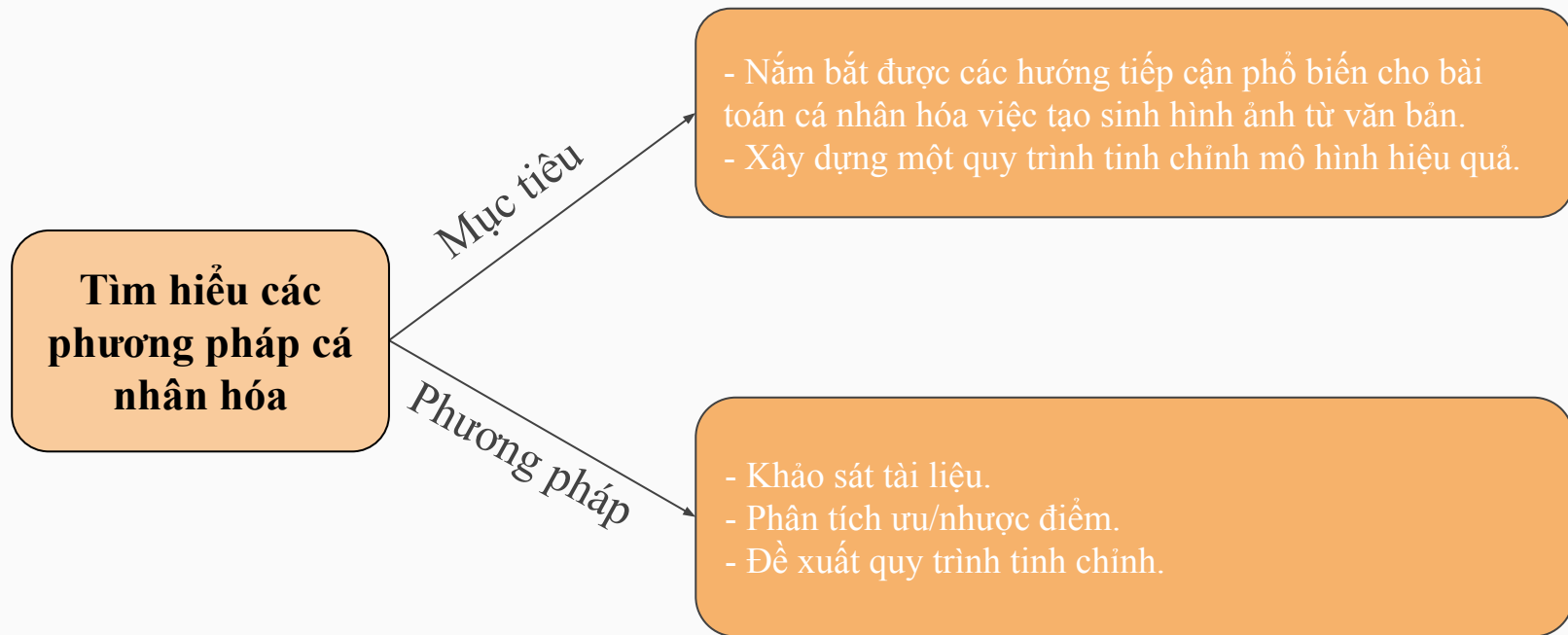
# Nội dung và Phương pháp



# Nội dung và Phương pháp

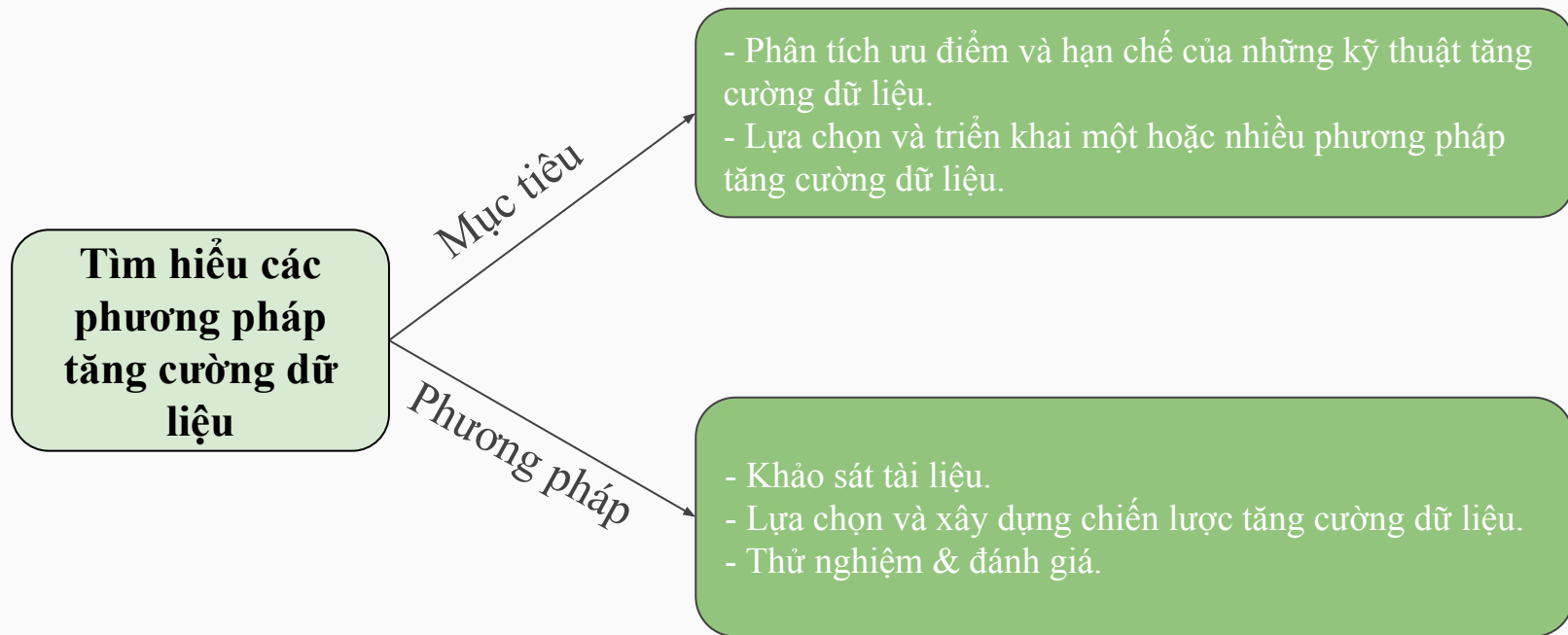


# Nội dung và Phương pháp

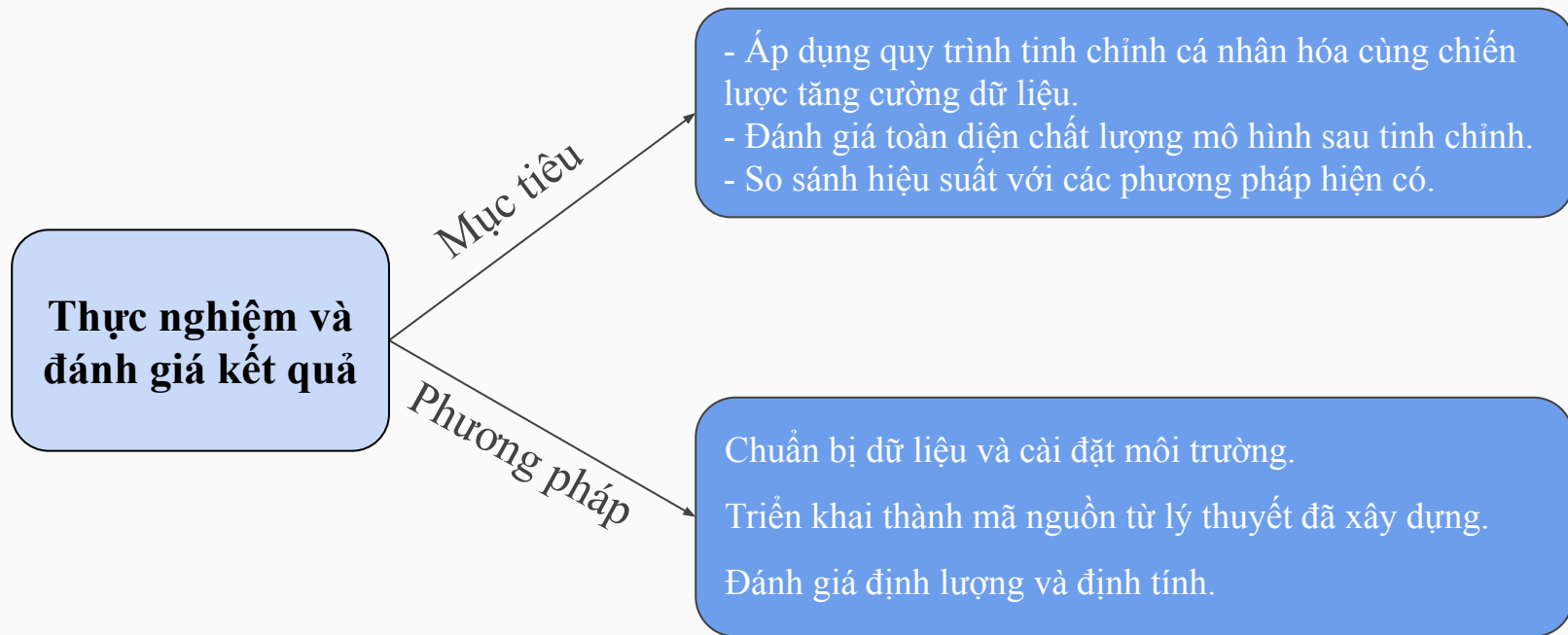




# Nội dung và Phương pháp



# Nội dung và Phương pháp



# Kết quả dự kiến

- Tài liệu báo cáo chi tiết cơ sở lý thuyết về phương pháp đề xuất.
- Chi tiết kết quả thực nghiệm, đưa ra các đánh giá về hiệu năng, số liệu định lượng và định tính và tính hiệu quả của phương pháp.
- Kết quả so sánh với các phương pháp hiện nay trên cùng một tập dữ liệu.

# Tài liệu tham khảo

- [1]. Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer: High-Resolution Image Synthesis with Latent Diffusion Models. CVPR 2022: 10674-10685
- [2]. Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman: DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. CVPR 2023: 22500-22510
- [3]. Rinon Gal, Yuval Alaluf, Yuval Atzmon, Or Patashnik, Amit Haim Bermano, Gal Chechik, Daniel Cohen-Or: An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion. ICLR 2023
- [4]. Lvmin Zhang, Anyi Rao, Maneesh Agrawala: Adding Conditional Control to Text-to-Image Diffusion Models. ICCV 2023: 3813-3824
- [5]. Yuchao Gu, Xintao Wang, Jay Zhangjie Wu, Yujun Shi, Yunpeng Chen, Zihan Fan, Wuyou Xiao, Rui Zhao, Shuning Chang, Weijia Wu, Yixiao Ge, Ying Shan, Mike Zheng Shou: Mix-of-Show: Decentralized Low-Rank Adaptation for Multi-Concept Customization of Diffusion Models. NeurIPS 2023