

X:AI 5th Seminar 2024

DiffStyler : Diffusion-based Localized Image Style Transfer

Shaoxu Li

arxiv 2024.03 출판

2024.07.16 2주차 ADV Session 발표

국민대학교 AI빅데이터융합경영학과
유광열 kt01054197938@gmail.com



국민대학교
인공지능학회 X:AI

CONTENTS

01

Introduction

02

Method

03

Experiment

04

Conclusion

05

Reference

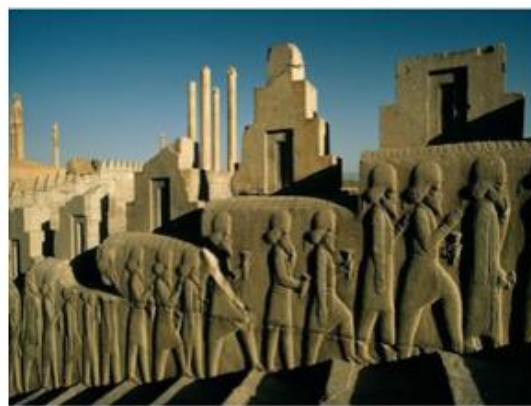


Introduction

Image Style transfer

- 디지털 이미지에 스타일 타겟의 독특한 속성(색상, 붓 자국, 형상 등)을 적용하는 과정
- 콘텐츠의 의미적 완전성을 동시에 유지하는 것이 목표
- 한계점 : 임의 스타일 전송 방법의 발전에도 불구하고, 콘텐츠 의미와 스타일 속성 간의 균형 유지가 어려움

content image



Ancient city of Persepolis

+

style image



The Starry Night (Van Gogh)

=

generated image



Persepolis
in Van Gogh style

Introduction

DiffStyler

- 효율적이고 정밀한 이미지 스타일 전송을 지원하는 새 방법론
- Stable Diffusion 모델 기반의 LoRA(Low-Rank Adaptation) 사용
 - 스타일 타겟의 본질을 요약하고 캡슐화
- LoRA가 UNet의 공간적 특성 일관성 유지를 기반으로 한 마스크 방식의 스타일 전송 기술 개발
- 총 3가지 방식 제안

Introduction

DiffStyler

- 전반적인 스타일 변환:
 - content 이미지와 style 이미지를 사용하여 content 이미지의 속성을 지정된 style 대상으로 변환
- 프롬프트 기반 로컬 스타일 변환:
 - 특정 영역만 스타일 변환을 원할 경우, 프롬프트를 사용하여 해당 영역에 스타일을 적용



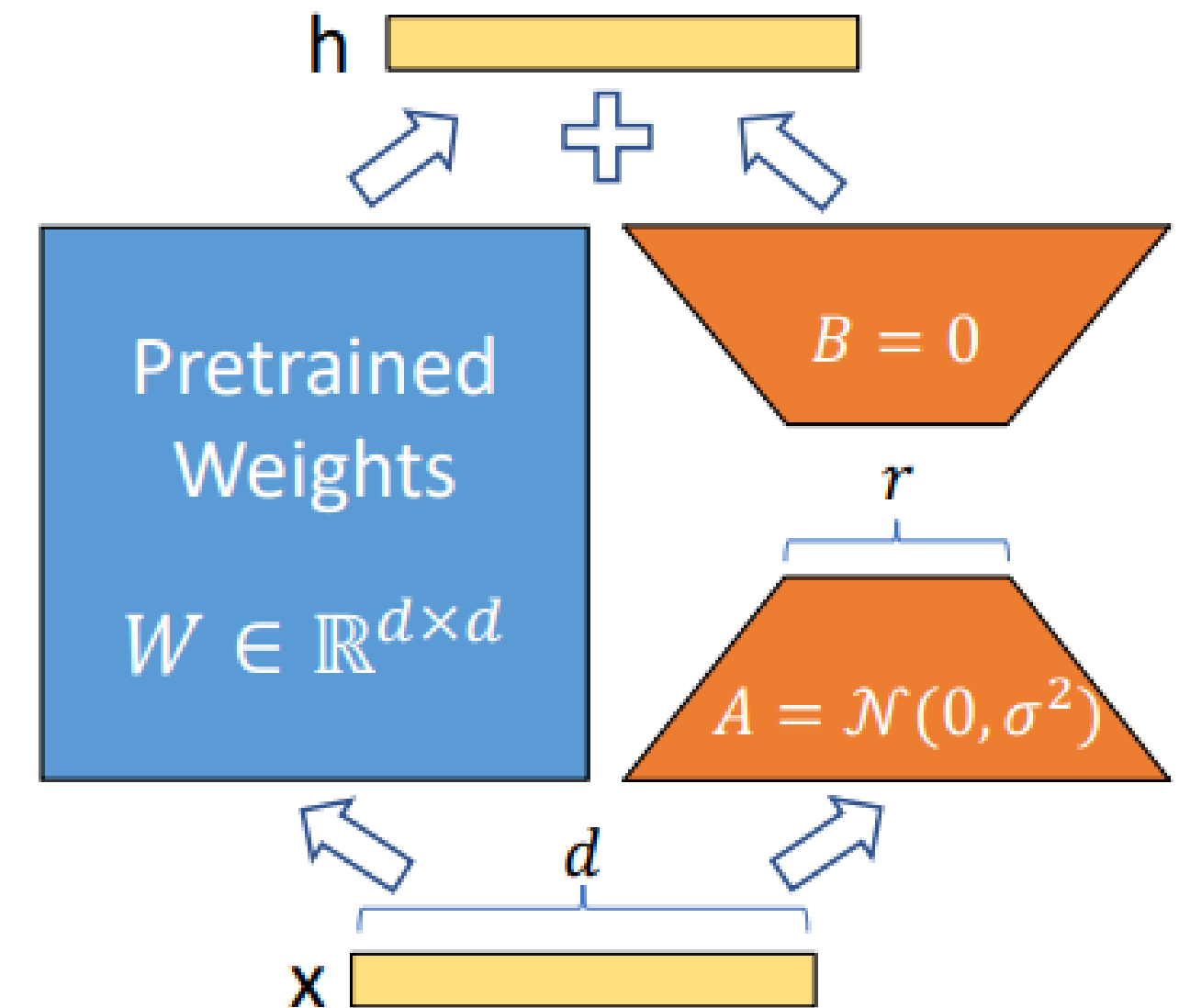
DiffStyler

- 마스크 프롬프트와 스타일 대상 통합:
 - 여러 개의 마스크 프롬프트와 스타일 대상을 통합할 수 있는 기능을 제공
 - 예를 들어, 한 이미지의 특정 부분에 각각 다른 스타일을 적용하는 가능성을 제시함.



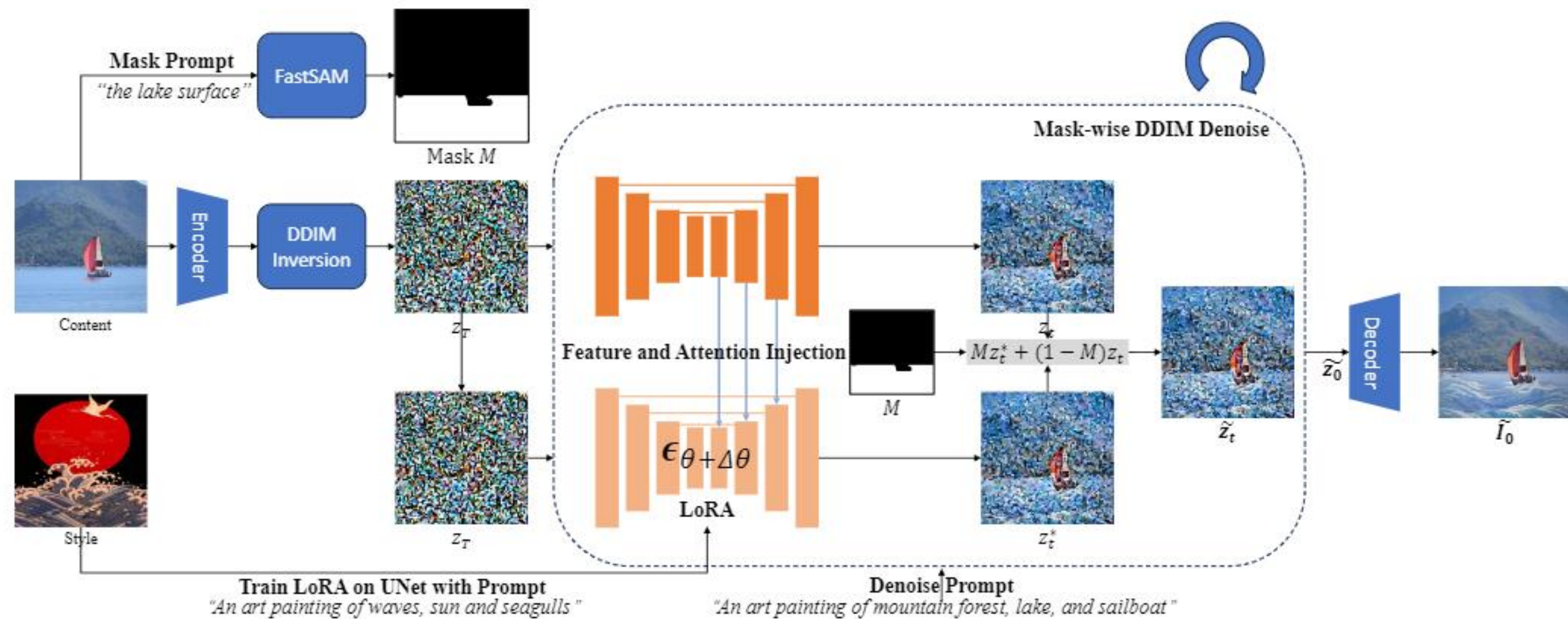
LoRA (Low-Rank Adaptation)

- PEPT(Parameter-Efficient Fine-Tuning) 기술 중 하나
- 대형 사전 학습 모델을 특정 작업에 맞게 효율적으로 튜닝하는 방법
- 왼쪽 파란색은 Pretrained Weights로 학습에 참여X
- 오른쪽 주황색은 LoRA Adaptor → Fine tuning 학습 되는 파라미터
- A와 B의 파라미터 수는 W대비 매우 적음 (보통은 1% 미만)



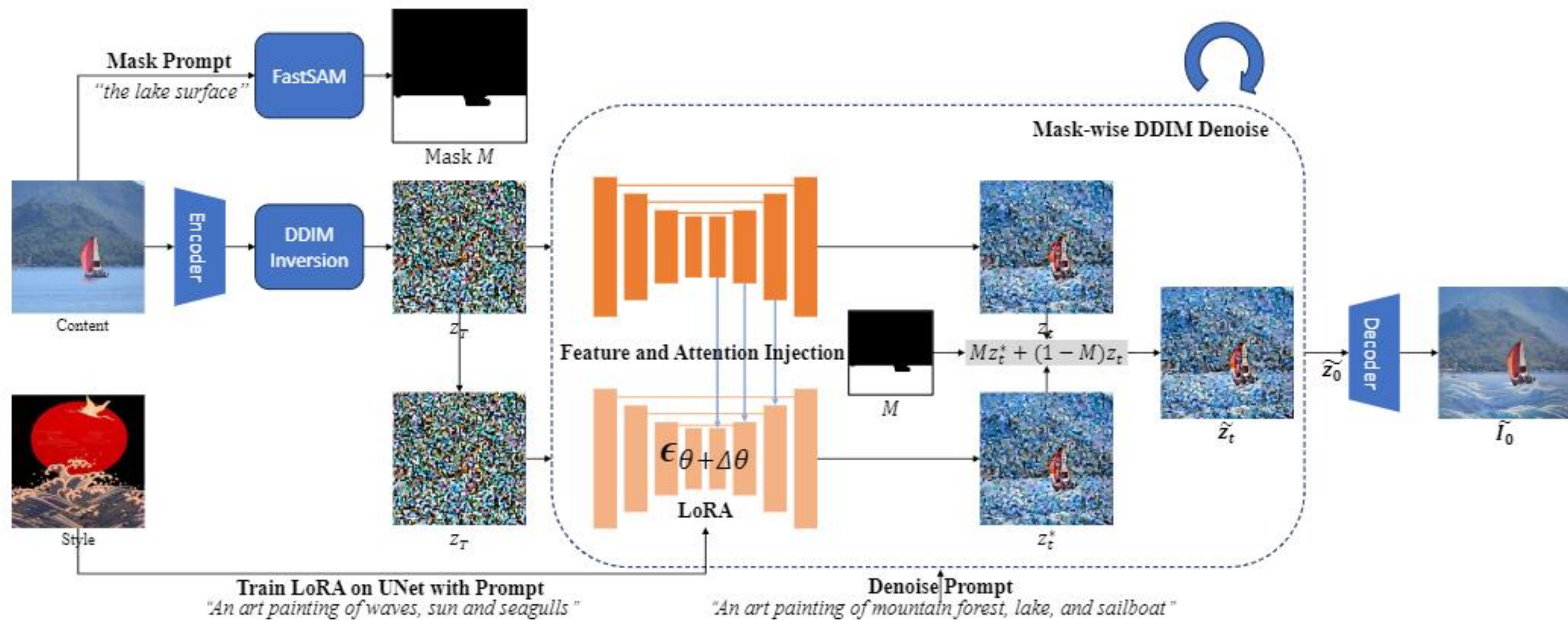
Content와 Style 이미지 입력

- 왼쪽 상단의 Content 이미지와 하단의 Style 이미지를 사용



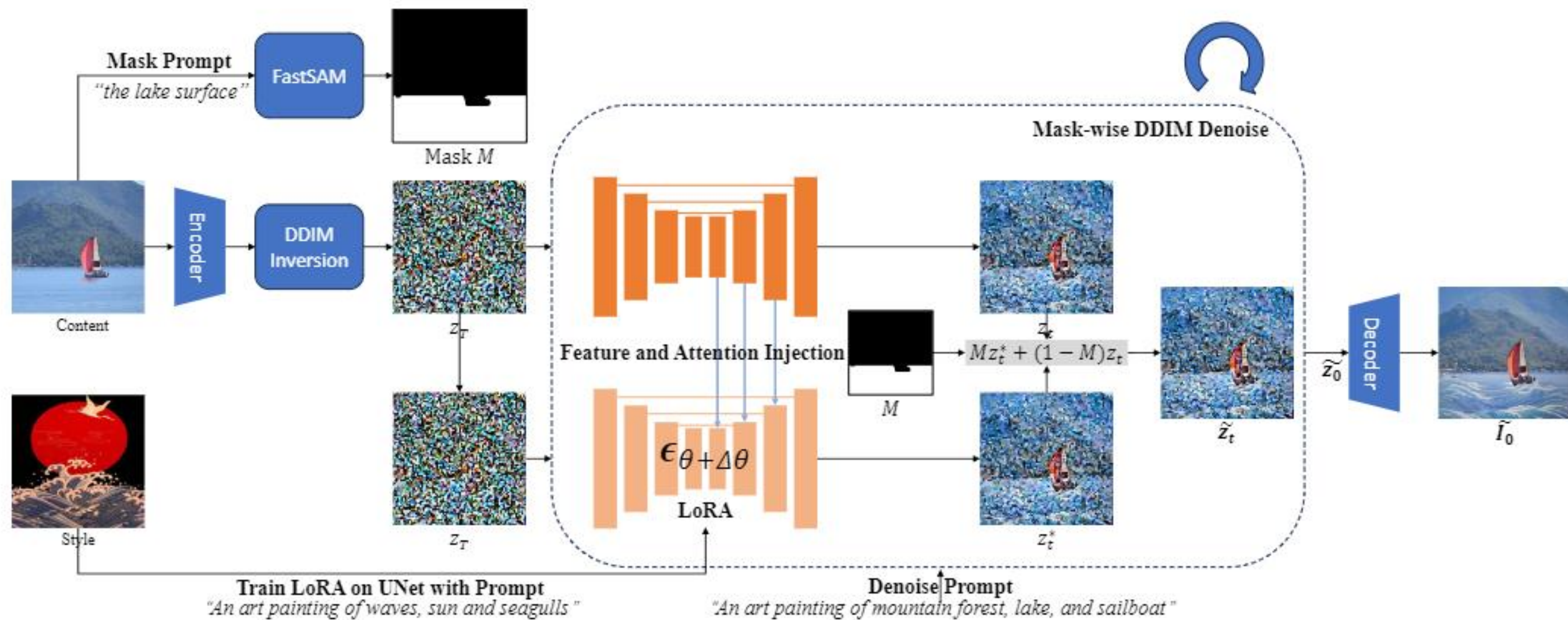
Mask Prompt와 Mask 생성

- Mask Prompt를 통해 특정 영역을 선택
- FastSAM을 사용하여 Mask M을 생성



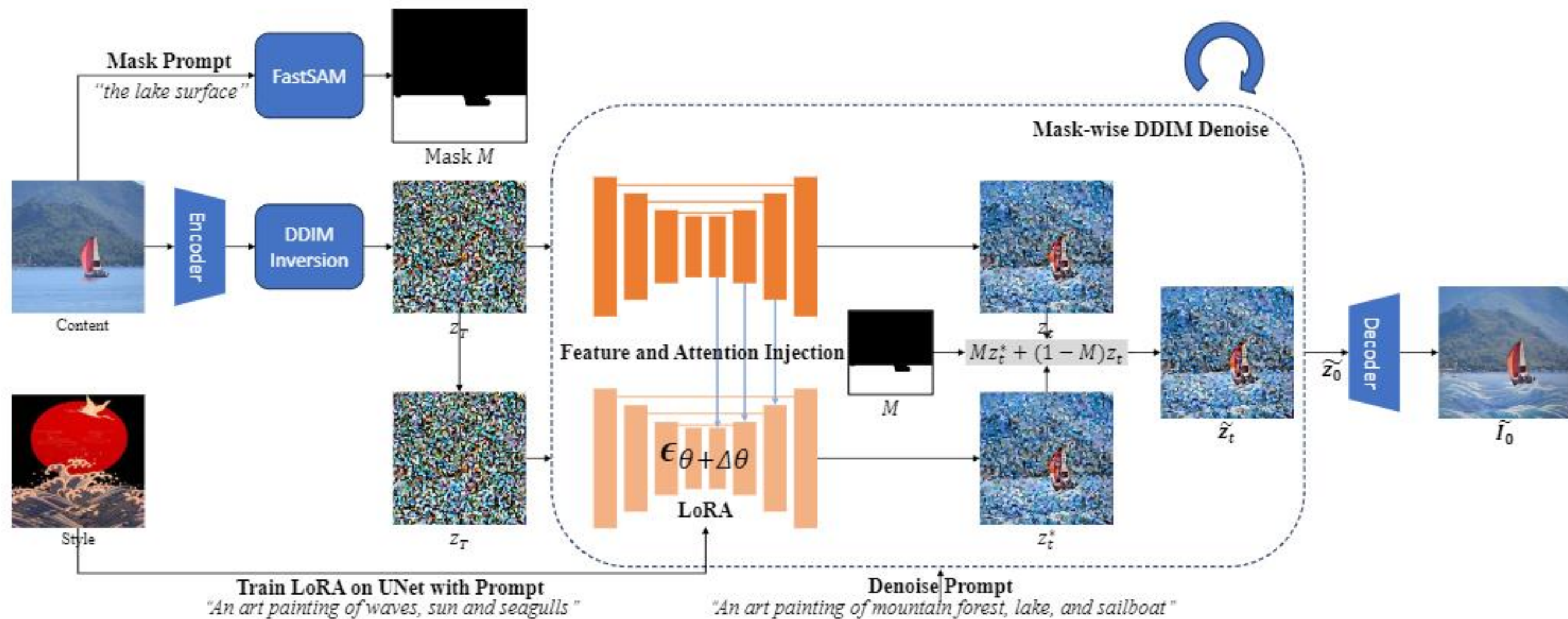
DDIM Inversion

- DDIM Inversion을 통해 Content 이미지에서 z_T 와 Style 이미지에서 z_T 를 추출



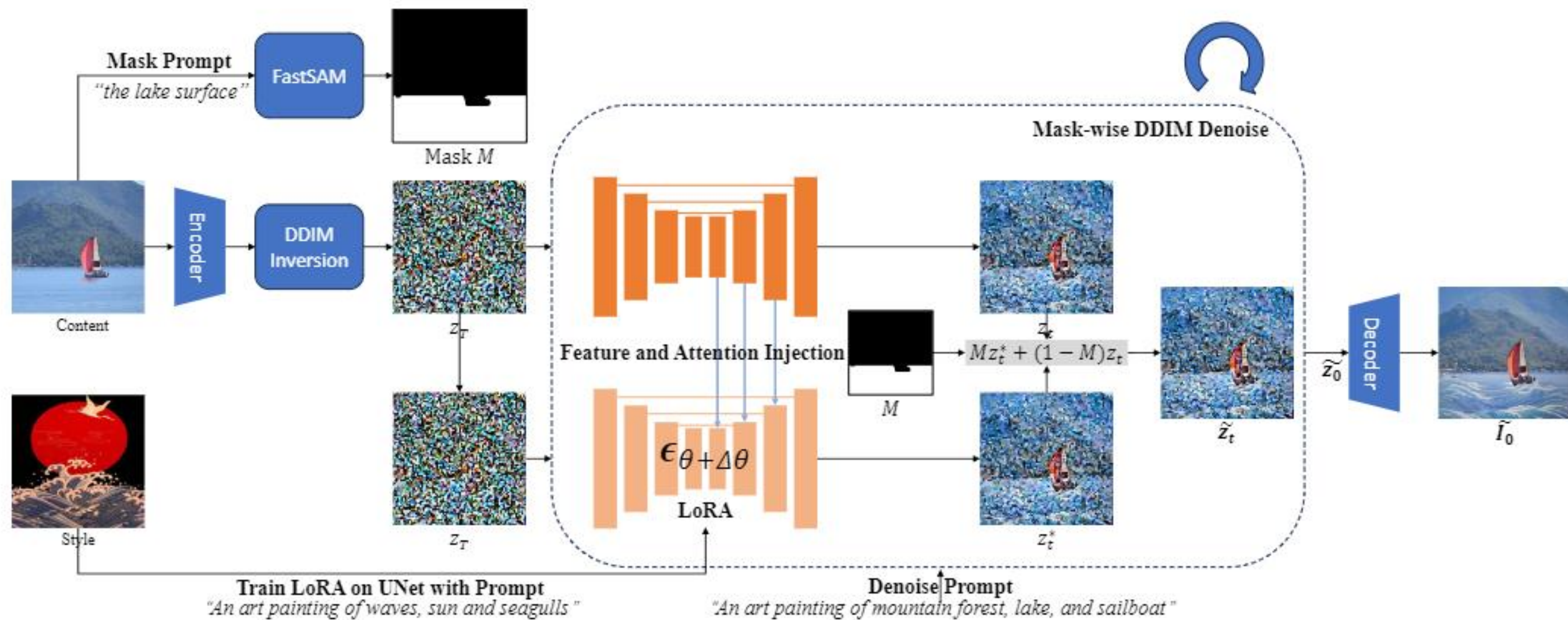
Feature와 Attention Injection

- LoRA 유닛은 U-Net 아키텍처 내에서 사용되어, Style 이미지의 특징을 Content 이미지의 선택된 영역에 주입
- 이 단계에서 Mask M 을 통해 선택된 영역에만 스타일 변환이 적용



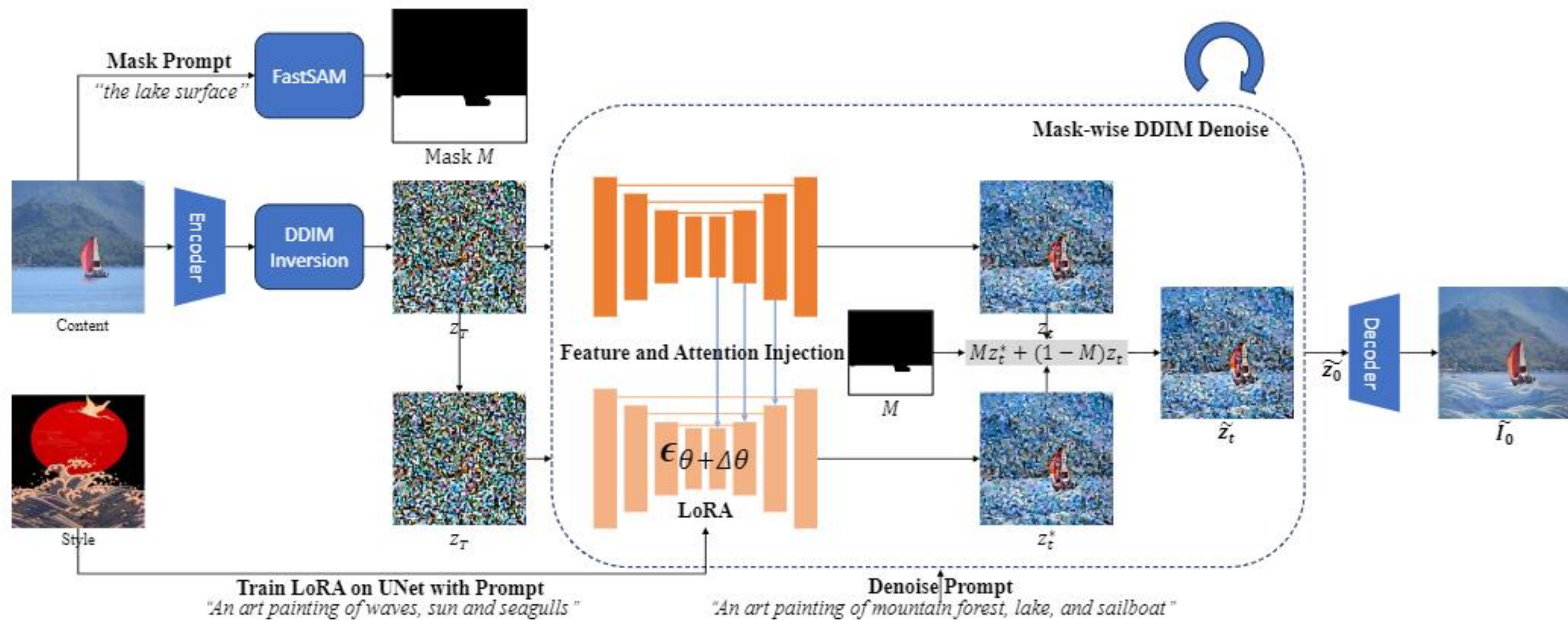
Mask-wise DDIM Denoise

- 선택된 영역과 선택되지 않은 영역을 구분하여 각각에 맞게 DDIM Denoising을 수행
- 해당 식을 사용하여 Mask 영역과 결합

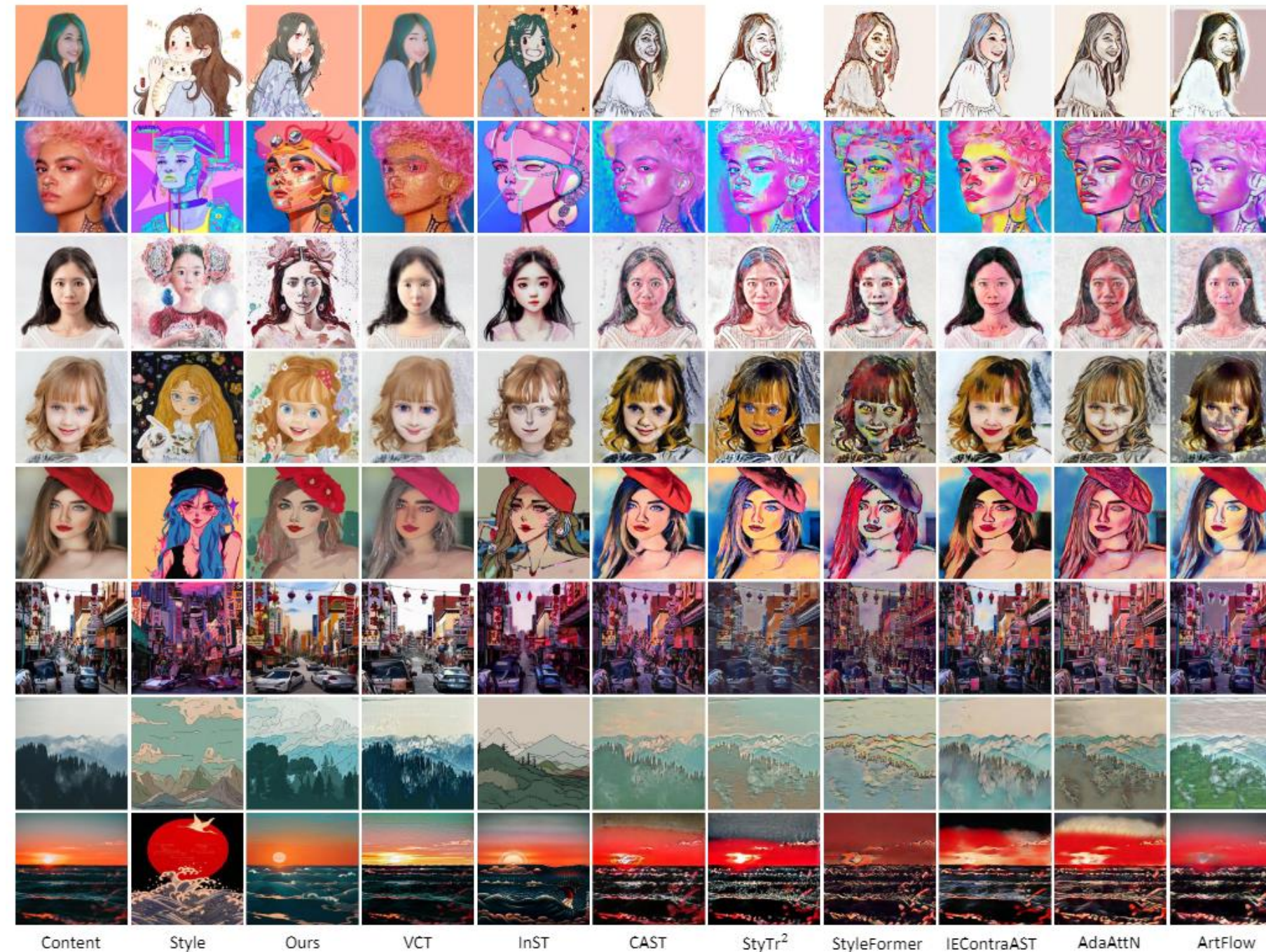


Decoder

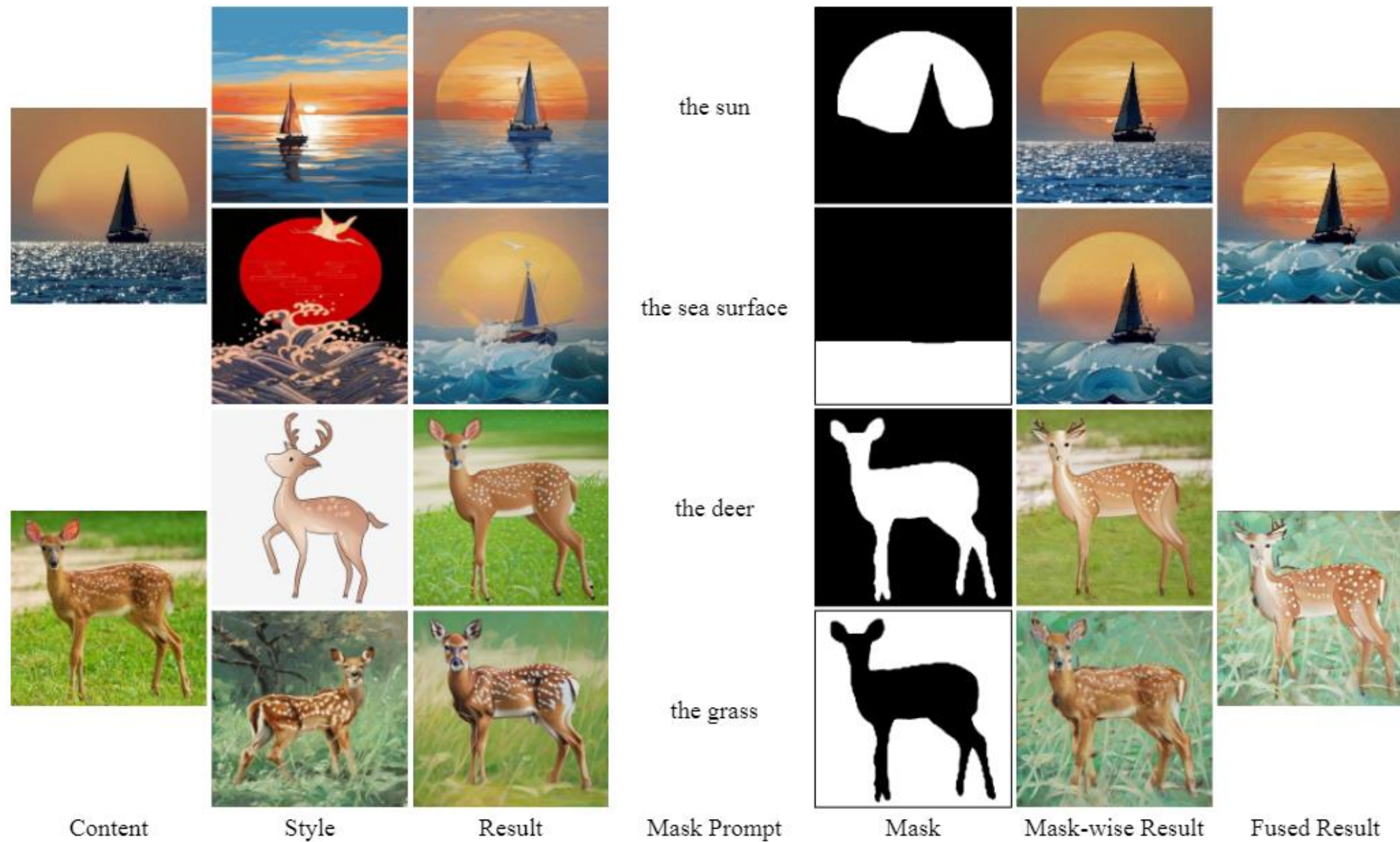
- 최종적으로 Decoder를 사용하여 z_T 에서 변환된 이미지를 생성



Comparison with Style Transfer Methods



Mask-wise Style Transfer Results



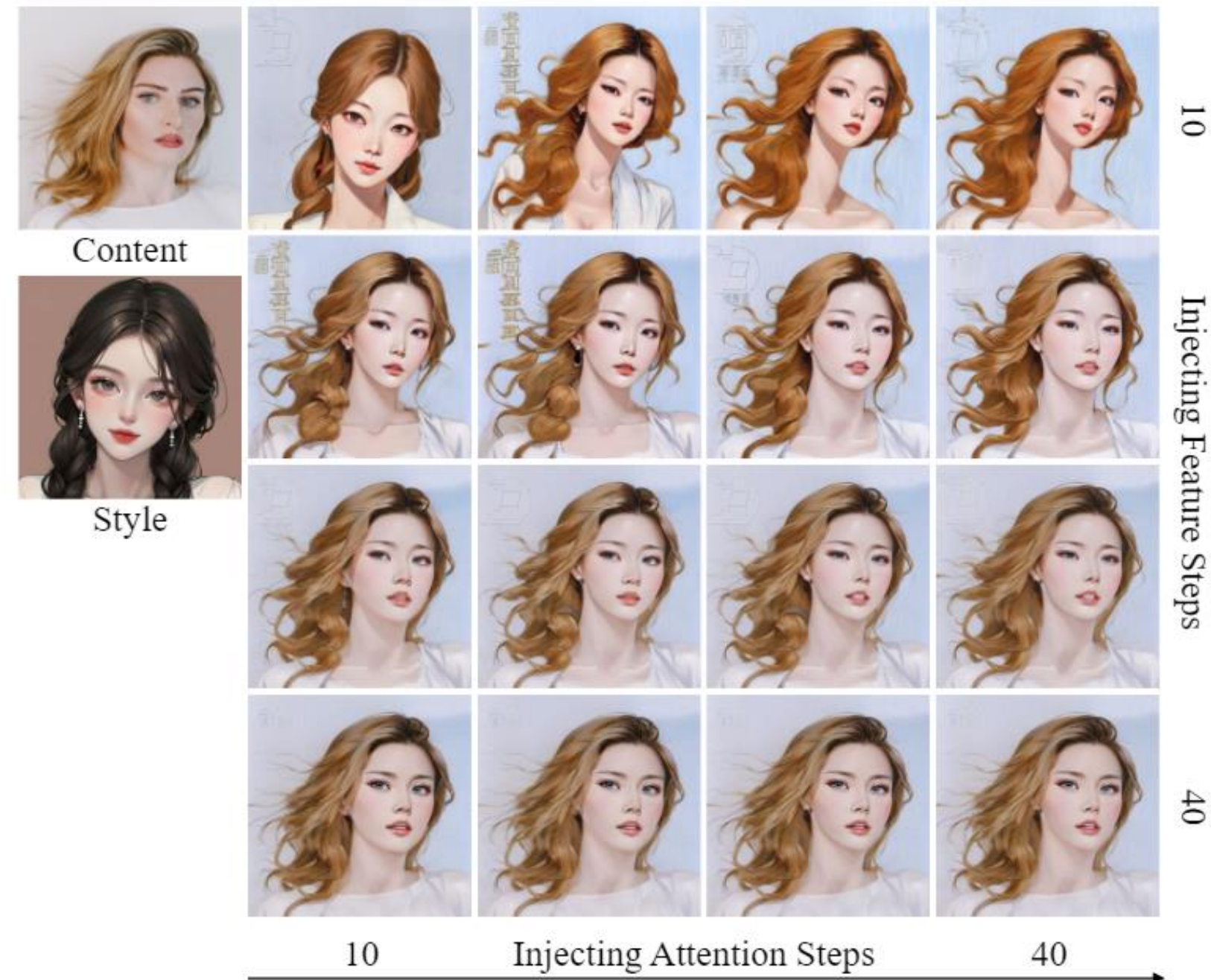
Experiment

Cross-LoRA feature and attention injection



Experiment

Feature and attention injection steps



- ✓ LoRA를 활용한 새로운 Image Style Transfer 모델 제안
- ✓ 학습한 속성을 기반으로 Feature와 Attention 주입을 통해 이미지 합성 과정을 안내

Reference

<https://arxiv.org/abs/2403.18461>

<https://www.youtube.com/watch?v=0lf3CUIUQtA>

X:AI 5th Seminar 2024

T h a n k Y o u

국민대학교 AI빅데이터융합경영학과
유광열 kt01054197938@gmail.com



국민대학교
인공지능학회 X:AI