

Paper summary

AI VISION Lab

1. 공부한 논문의 제목, 게재된 학회 혹은 저널 등 논문 기본 정보를 적으세요.
 - A. 이름: **Diffusion Enhancement for Cloud Removal in Ultra-Resolution Remote Sensing Imagery**
 - B. 저널: **TGRS**
 - C. 도메인: **Cloud Removal**
 - D. 출판연도: **2024**
 - E. 저자:
2. 논문에서 제안한 알고리즘 및 프레임워크에 대해 본인이 이해한대로 다이어그램을 그려보세요. **논문 Figure를 그대로 따라 그리면 안됩니다.**
 - A. Diffusion Network의 Output (Noise Prediction 혹은 x_0 Prediction)은 여전히 Refinement를 진행할 여지가 존재함을 이야기하고 있음.
 - B. 따라서, 해당 Output을 Reference Model(학습 X – ADM 모델 load)의 Output에 의거하여 Refinement를 진행함. 이때, Diffusion Model의 Loss(Pixel Optimal) 과 Refinement의 Target은 다른 형태다 보니, 이를 계수 형태의 Linear한 구조로 최적화하고 있음 (WA-Weight Average, Trainable)
 - C. 결과적으로, Refinement를 통해 구한 값을 통해 – Spatial한 정보를 보완해 줄 수 있음을 이야기하고 있음. (L2 기반의 Diffusion의 경우 fine detail 측면에서 부족한 성능을 보이는 한계가 있다보니, 이를 보완해주고자 함)
3. 본인이 생각하는 이 논문의 장점이 무엇이라고 생각하나요? **논문 Contribution bullet을 그대로 따라 적으면 안됩니다.**
 - A. 고해상도 데이터셋 (CUHK-Thin Cloud, CUHK-Thick Cloud 0.5m) 데이터셋 제공
 - B. Diffusion model Output 과 Reference Model Output을 통한 Refinement Network 설계
4. 이 논문을 읽으면서 느낀 점, 혹은 배운 점이 있으면 적어보세요.

- A. 다양한 Loss가 주어졌을 때, 최적화하는 과정에 대해서 추가적인 고민을 하게 만드는 부분이 있었음.
 - B. 또한, ADM의 아키텍처를 토대로 만들었다는 점에서 모델을 설계하는데 참고할 수 있을 것이라 판단됨.
5. 이 논문의 한계점이 있다면 무엇이라고 생각하나요?
- A. Evaluation 때 사용된 Dataset (Rice, CUHK)을 실제 적용한 모델 혹은 연구가 극히 적음. 뿐만 아니라 이를 Experience에서 적용한 모델들은 상당히 오래된 것들(GAN, CNN 기반)이 대다수다 보니, 모델의 성능이 어느정도인지 파악하기 어려움.
 - B. 뿐만 아니라, 정성적 평가를 진행하는 과정에서 Thin Cloud (겉보기로는 구름이 거의 없는)에 대한 비교군만 존재하다 보니, 주요 Task라 할 수 있는 Thick에 대한 Recover성능을 확인하기 어려움
6. 본인의 연구에 접목시켜볼 점이 있을지 생각하고 적어보세요.
- A. 최종 Loss에 대해서, 두가지 Loss를 결합한 형태를 사용하고 있는데, 일반적으로는 두 Loss의 목적이 다르다 보니 ($L2 - \text{Pixel}$, $L_{wa} - \text{Reference}$ 와 predicted 차이 줄이기. $L2 + L_{wa} = L_{\text{total}}$) 최적화가 어려움. 따라서, 두 Loss를 모두 최적화시킬 수 있는 Weight를 찾는 Weight Average를 적용했고, 이는 다양한 Loss를 적용하는 경우 고려해볼 만한 개념이라 판단됨.
7. 본 Summary를 작성하는 과정에서 생성형AI를 사용했나요?
- A. 아니요

날짜: 2025-08-18

이름: 신준원

Diffusion Enhancement for Cloud Removal in Ultra-Resolution Remote Sensing Imagery

Introduction

- 1. L_1, L_2 is the loss of θ (in cloud removal) \rightarrow Semantic Loss.
 - ⊕ bad result.
- 2. Data \Rightarrow CV4K (0.5m) - VR / Detailed.
- 3. Architecture \Rightarrow DE (Diffusion enhancement) - WA (Weight Allocation)
 - ⊕ LDM Based. (i.e. - Image process)
- 4. Coarse-to-fine training strategy.

Architecture

- DDPM is:
 - 1. forward process.
 - 2. reverse process.

Reference \hat{x}_t (SR) \Rightarrow Refinement \hat{x}_t !

문제 \Rightarrow Global Image structure \hat{x}_t

\rightarrow \hat{x}_t hallucination \downarrow

이제 \Rightarrow reverse step

x_{t-1} : Generated with (x_0, t)

\rightarrow timestep

$x_0, x_{0,t}$ is
noisy image

$\rightarrow x_{0,t}$ 생성 시에는, \hat{x}_t 필요.

$\hat{x}_t: \hat{\epsilon}_\theta(x_t, t, y)$

정확한 step. t 일때 예측한 x_0 를 계산함.

(가정) $\epsilon \approx 0$ 일때.

$$x_{0,\hat{\epsilon},t} = (x_t - \sqrt{1 - \alpha_t} \hat{\epsilon}_t) / \sqrt{\alpha_t}$$

$$x_{0,t} = x_{0,\hat{\epsilon},t}$$

이제, $x_{0,\hat{\epsilon},t} \rightarrow x_{0,t}$ refine: $\hat{\epsilon} / \hat{\epsilon}(y) = x_{0,\hat{\epsilon}}(\text{cloud removal output})$

$x_{0,\hat{\epsilon},t} \Rightarrow$ Diffusion output.

$$x_{0,t} = T(x_{0,\hat{\epsilon}}, x_{0,\hat{\epsilon},t})$$

result = (cloud removal + refinement)

이제, T function을 정의.

$$X_{0,t} = (1-W) \odot X_{0,\epsilon,t} + W \odot X_{0,\epsilon}$$

3 항이 됨.

- \odot : element-wise mult
- W : pixelwise fusion ratio
- ϵ 는 diffusion steps $\epsilon \rightarrow \mu, \sigma$ trainable.

WA

dynamic size \Rightarrow fusion (constant linear time, $(1-W)$, $W \rightarrow W$) 적용

WA block $\rightarrow X_{t,y}, X_{0,\epsilon}, t$
 \rightarrow U-net based (CNP \rightarrow) \rightarrow temporal, spatial 정보 제공

training speed

step + block $\leftarrow \uparrow \Rightarrow$ Coarse-to-fine training strategy.

image size: $1/4$ 정도 + 1차 (diffusion model)

WA는 full size image. (initial \rightarrow trained diffusion)

$$L_{WA} = | \tilde{x}_0 - x_{0,t} |$$

$$= | \tilde{x}_0 - (1-W) \odot (x_{0,\epsilon,t})_{\epsilon} + W \odot x_{0,\epsilon} |$$

결과: $L_{joint} = \lambda \cdot L_{DDPM} + L_{WA}$.

Metrics

PSNR
SSIM
LPIPS

evaluation \rightarrow (C4H4-CR
RICE (12x512 RA/3))

thick \rightarrow thin + (증가 양자 \rightarrow thin cloud on the plane)

