text-to-image 拡散モデルにおける 誘導 attention map を用いた画像生成手法の提案

水口 徳人[1], 北田 俊輔[1], 守田 竜梧[1], 彌冨 仁[1] [1]法政大学大学院 理工学研究科
{yasuto.mizuguchi.4e@stu., iyatomi@}hosei.ac.jp

❐ Summary

拡散モデルにおけるcross attention機構に着目した新たな画像生成手法の提案

- 各tokenに対応するattention mapのグループ化、attention誘導処理により画像の生成性能が向上
- 提案手法をStable Diffusion(SD)へ適用することで指定した複数要素に対する生成性能の向上を確認

❐ Background

特定の構成をしたプロンプトの画像生成が困難

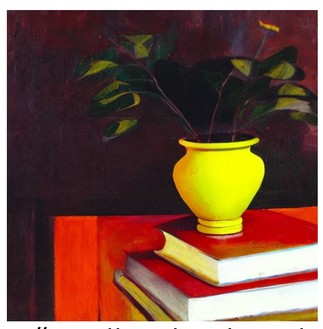
- 複数の対象物を異なる色や位置に指定した時に指示どおりの画像が生成されない問題が存在[Chefer+ SIGGRAPH'23]

画像生成においてcross attentionは重要！

- テキストの埋め込みを潜在表現に組み込むことで物体の位置や要素の決定に関与

"a red car and a white sheep"

"a blue bird and a brown bear"



"a yellow book and a red vase"

SD v1.5 による画像生成の失敗例

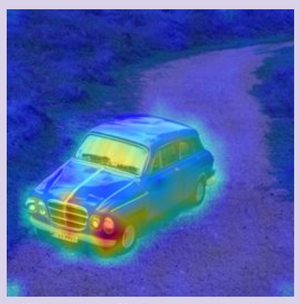
❐ Method

事前分析：attentionの観察

- 生成されなかった物体のattentionは他の物体のattentionに混合していることを確認

"a red car and a white sheep"の生成画像

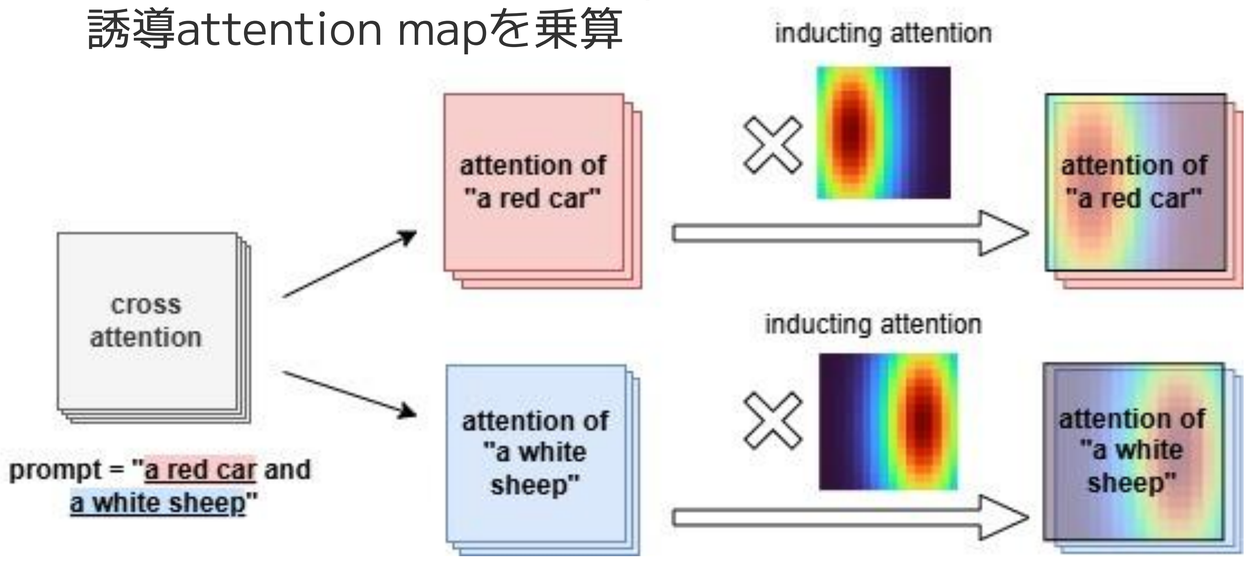
"car"のattention



"sheep"のattention

物体の特徴が欠落または混合するのを防ぐための attention map誘導

1) 同じ物体の特徴を示すtokenをグループ化
2) 中心や範囲を適切な位置に指定した、ガウス分布による誘導attention mapを作成
3) 各グループのattention mapにそれぞれ異なる誘導attention mapを乗算



提案手法の構成図

❐ Experiment & Result

実験方法

- ベースラインモデル：Stable Diffusion(SD)v1.5
- 評価用プロンプト: DrawBench [Saharia+ NeurIPS'22] に含まれる異なる色の物体を2つを有する以下を使用
  - *"a red car and a white sheep"*
  - *"a blue bird and a brown bear"*
  - *"a green apple and a black back pack"*
  - *"a green cup and a blue cell phone"*
  - *"a yellow book and a red vase"*

:group A :group B

- 各プロンプト100枚の画像を生成（計500枚）

評価方法

- YOLOv8 [Varghese+ ADICS'24] を用いて物体の有無を確認
- 評価基準を以下のように設定
  - 成功：両方の物体を検出
  - 不十分：一方の物体のみを検出
  - 失敗：該当物体の検出なし

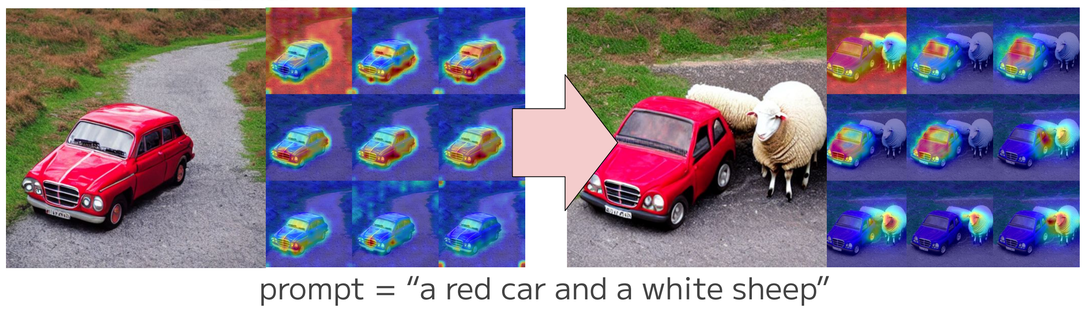
平均スコアの結果

	SD [枚]	SD+提案手法 [枚]
成　功	38.6	**42.4**
不十分	37.0	39.4
失　敗	24.4	**18.2**

➡ ***スコアが改善され、提案手法が有効に作用***

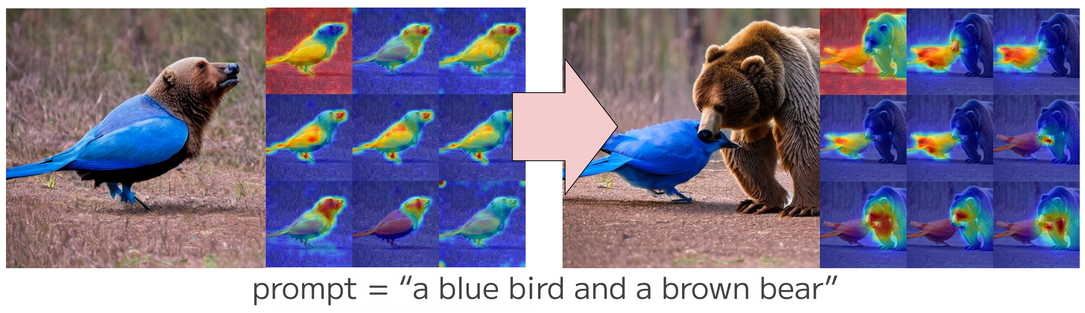
適用前 適用後

提案手法の成功例



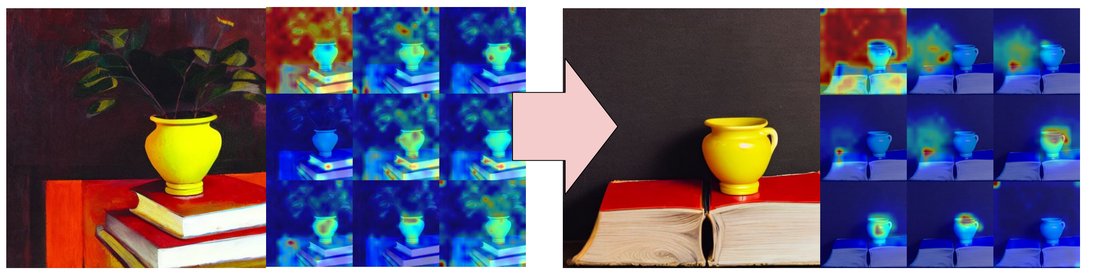
prompt = "a red car and a white sheep"

提案手法の成功例



prompt = "a blue bird and a brown bear"

提案手法の失敗例



prompt = "a yellow book and a red vase"

生成画像とattentionの比較結果

❐ Discussion & Future Work

- プロンプトによって改善の幅にばらつきがあり、物体のattention mapの面積や形に差があることが起因
- SDの学習データによるバイアスが存在し、バイアスに影響されない attention誘導と画像生成が困難
- それぞれの画像に臨機応変に対応した適切な位置への誘導法の考案
- tokenのグループ化の自動化
- 色が正しく付けられているかの評価手法の考案