

T2Iモデルにおける評価指標の開発

岡山理科大学 理工学研究科 房ハルノ

目的 Attention mapsの情報を活用し、汎用的で精度の高い評価指標の研究開発

内容 Stable diffusion (SD) など画像生成AIモデルの普及に伴い、テキストと生成された画像の適合性の評価は重要な課題である。Text-Image Alignment Metric (TIAM) は、テンプレートベースの手法を用いて、指定されたオブジェクトの数、属性を多角的に評価できるが、属性の評価は色に限定されている。SDsは色だけでなく、形状やサイズなど、多様な属性を表現できるため、より包括的な評価が必要である。本研究では、TIAMを拡張し、画像生成過程で得られるAttention Mapsと画像言語モデルを組み合わせることで、色以外のサイズ、年齢、形状、素材の属性をテンプレートベースで評価できる新たな評価手法を提案する。

結果 提案手法の有効性を検証するため、アンケート調査による評価と比較を行った。その結果、提案手法はTIAMより評価対象が広く、ベースライン（BLIP2）よりも人の評価に近いことが示され、より正確に生成画像の属性を評価できることが明らかになった。さらに、複数のSDsに対する性能評価を行い、その特性をまとめた。

利用した計算機 GPUノード群
ノード時間 40 時間
使用メモリ 30 GB

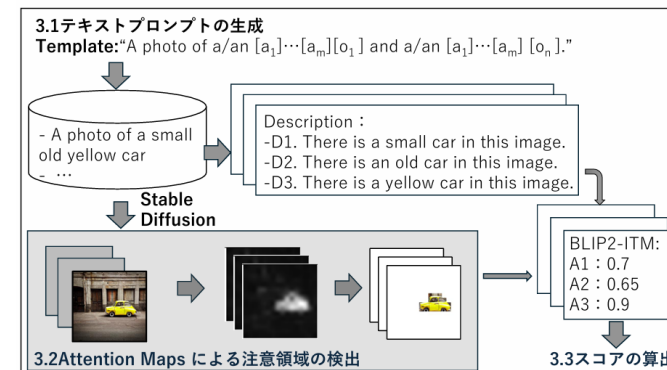


図 1: 提案手法のフロー