

【IS3-47】 クラスラベルを仮定しない画像集合の多様性評価

岡本夏旺¹, 品川政太郎¹, 中村 哲¹

1.奈良先端科学技術大学院大学

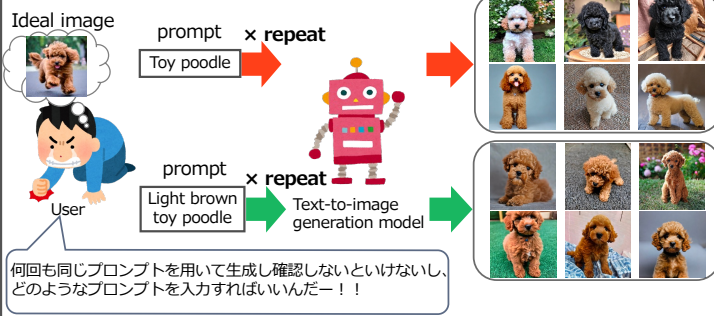


■ 目標

- プロンプトがどの程度多様性がある画像を生成することができるかを予測できる予測器の構築

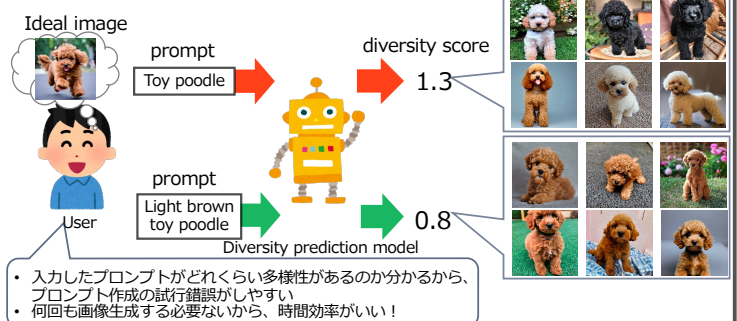
現在のプロンプトエンジニアリング

■ 時間がかかる



多様性予測器

- 画像生成を何度も行わなくて良い



■ 課題

- 既存の多様性評価指標IS(Inception Score)[1]では、**クラス分類モデル**を用いており、クラスラベルに属していない画像集合の多様性を正しく評価できない

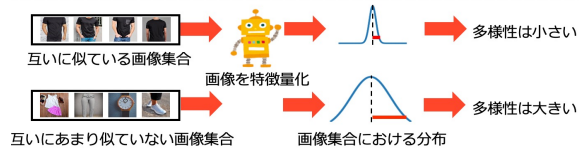
■ 本研究の貢献

- クラス分類モデルを用いない多様性評価指標の提案し、既存手法より正しく多様性評価が行えた

提案手法

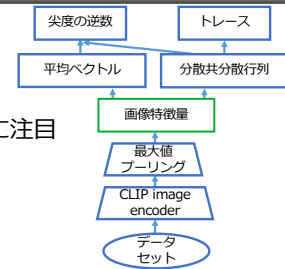
■ 提案手法の仮定

- 似ている画像で構成されている画像集合では画像特徴量がそれぞれ類似したベクトルになる



■ 提案手法の詳細

- CLIP[2](contrastive language-image pre-training)を用いて、画像集合における分布の裾の長さに注目
- 2つの提案手法
- 尖度の逆数を計算
- トレースを計算



実験 & 結果

■ 実験設定

- 画像集合cを上位概念, 中位概念, 下位概念クラスと定義
- 上位概念 > 中位概念 > 下位概念の順に多様性があると仮定

■ 評価方法

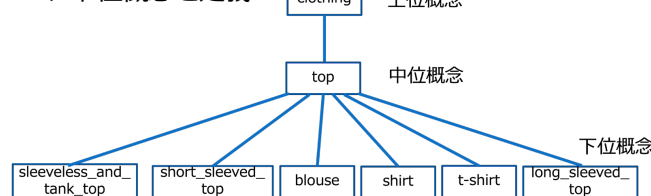
- それぞれのクラスにおいて多様性の値を計算し、多様性の値が大きい順に並び替えランク(順位)を計測する
- 上位概念, 中位概念の理想のランクは1,2であり、上位概念, 中位概念のランクが1,2に近いかを比較し、評価を行う



- **実験1**:提案手法は既存手法よりも多様性を正確に評価をすることができるのか?

■ データセット

- 実験1 :Fashion-200kデータセット[3]を用いてそれぞれのクラスに対して、上位概念、中位概念、下位概念を定義



■ 結果

手法	上位概念ランク	中位概念ランク
Inception v3entropy	1.8	4.0
CLIP entropy	2.6	4.2
トレース	1.0	3.2
尖度の逆数	1.2	2.4

- 提案手法がより多様性を正しく評価できている

- **実験2**:提案手法は生成画像集合においても多様性を測る有効な手法だといえるのか?

■ データセット

- Fashion-200kデータセットで用いられているクラス名を入力テキストとし、画像生成モデルを用いてデータセットを構築 入力例: "a **クラス名** item,real"

■ 結果

手法	上位概念ランク	中位概念ランク
Inception v3entropy	1.6	3.4
CLIP entropy	3.4	3.8
トレース	1.6	4.6
尖度の逆数	3.4	4.0

- 理想のランクから提案指標はかけ離れている
- 検討の余地あり
- 背景の影響..?
- 生成画像のクオリティー不足..?

[1] Salimans et al., Improved techniques for training gans.

[2] Radford et al., Learning Transferable Visual Models From Natural Language Supervision

[3] Han et al., Automatic Spatially- aware Fashion Concept Discovery