

複雑な表情を生成可能なAIの開発

産業能率大学 情報マネジメント学部 川野邊研究室 3年 野川 祐心 指導教員:川野邊 誠

研究概要

研究背景

画像生成AIの発展により、プロンプトから笑いや怒りといった
単独の感情を含む表情を生成することが可能



感情と表情が一致しない「複雑な表情」の生成は困難

【表情編集】

- ChatGPT-4o, Gemini2.5 proを用いて画像生成
- プロンプト「表面上は笑っているにもかかわらず、内面では怒りを抱いている」
- 結果として、ChatGPT-4oでは、眉間にしわを寄せ、口元は笑っている足し算的な表情が生成(図1)
Gemini2.5 proでは、影の濃さで怒りを表現した表情が生成(図2)された



図1. ChatGPT-4oで出力した複雑な表情

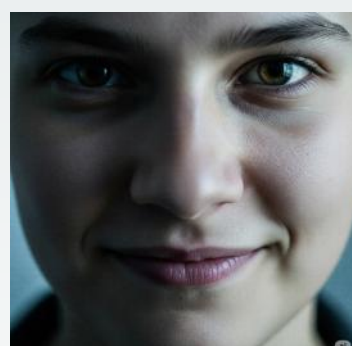


図2. Gemini2.5 proで出力した複雑な表情

先行研究

- Saarらの「Image Generation from Contextually-Contradictory Prompts」^[1]では、通常同時に現れない矛盾した概念がプロンプトに入力された場合、AIがプロンプトの意図を汲み取れず、片方の概念を無視してしまう現象や、両方の特徴を不自然に合成してしまう現象が発生すると述べている

- Heらの「Towards Boosting Facial Expression Recognition with Synthetic Data」^[2]では、顔表情データセットが、プライバシーの問題、アノテーションの主観性、データ収集の労力といった理由で数的規模が限られていると述べている

- Tuomasらの「Towards Localized Fine-Grained Control for Facial Expression Generation」^[3]ではAU(Action Unit)の活性状態や強度を操作可能なモデルを作成し、従来の感情カテゴリに収まらないより細やかな表情の生成を実現

研究目的

- 従来の感情カテゴリに収まらない「複雑な表情」を生成可能なAIの開発

※複雑な表情：感情と表情が異なる状態を意味する

研究構想

感情と表情が異なる複雑な表情を選定し
MagicFaceモデルを用いてAUの強弱、組み合わせを設定して複雑な表情を生成



生成した複雑な表情を被験者に提示し、人間が複雑な表情を認識可能か検証



複雑な表情のデータセットを作成し
Contrastive Language-Image Pre-trainingを用いて
複雑な表情の画像とテキストの対応関係を学習



複雑な表情の学習データセットをStable Diffusionモデルに導入し
テキストプロンプトから複雑な表情を生成



生成された複雑な表情を人間が適切に認識可能か検証

期待効果

- AIが生成できる表情の幅が広がり、画像生成AI全体の表現力向上に貢献
- イラストレーターやコンテンツクリエイターが、表現が難しかった複雑な表情を効率的に描写することが可能になり、制作コストの削減が可能

研究成果

MagicFaceモデルを用いた表情編集

＜概要＞

- 元の顔画像から変更するAUと強度を指定し、元画像の特徴を残しつつ表情のみを変更し生成
- AUの強度は-10～10の間で指定し、強度が負の値の場合はAUを抑制、正の値の場合はAUを活性化
- 強度が±5を超えると表情が崩れる場合があるため、本研究では-5～5の範囲で指定

＜開発環境＞

- PC: Windows 11 Enterprise
- Google Colab Pro
- Python3.12

＜使用ライブラリ＞

- diffusers ■opencv_python ■scipy ■numpy
- einops ■packaging ■tqdm
- imageio ■transformers ■Pillow
- insightface ■torchvision ■torch

表1. 本研究で表情の編集に使用するAU

AUコード	動作
AU4	眉をひそめる
AU6	頬を上げて目を細める
AU12	口角を上げる

＜AUの選定＞

- FACS(顔面動作符号化システム)に基づいて選定
- 喜びの表情のAUは、喜びの表出動作であるAU6+AU12を選定し、怒りの表情はAU4を選定

【表情編集】

- 顔画像ファイル、顔画像から抽出した背景、輪郭保持の画像ファイルのパスを指定し、AUと強度を入力して実行。生成結果はMagicFace-outputディレクトリへ保存

```
1 output_image_path = "/content/drive/MyDrive/Magicface-output/AU4_v5.png" # 例: /content/drive/MyDrive/
2
3 #生成画像の保存先指定, 編集したい画像を変更したい場合パス指定を間違えないように
4 !python /content/drive/MyDrive/kenkyuu0926/MagicFace/inference.py %
5 --img_path "/content/drive/MyDrive/kenkyuu0926/MagicFace/test_images/00381.png" %
6 --bg_path "/content/drive/MyDrive/kenkyuu0926/MagicFace/test_images/00381_bg.png" %
7 --au_test 'AU4' --AU_variation '5' --saved_path "{output_image_path}"
8
9 print(f"生成された画像は '{output_image_path}' に保存されました!") #保存したファイル先と生成結果の確認
```

図3. 指定した顔画像のAU編集コード

【テスト画像を用いた表情編集】

- テスト画像として提供されていた顔画像を用いて、AU4、AU6+AU12を強度5で生成



図4. 提供されている顔画像^[3]

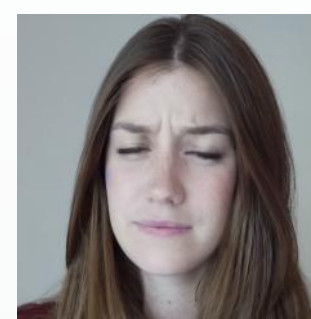


図5. AU4/強度5

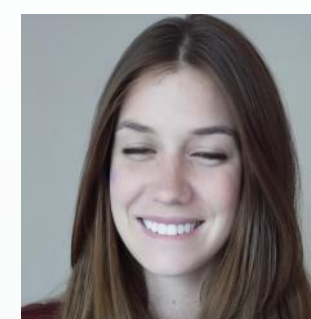


図6. AU6/強度5 + AU12/強度5

【東洋人の顔を用いた表情編集の検証】

＜目的＞

- 東洋人(日本人)で表情変更を行い、MagicFaceモデルが西洋人に偏らず表情を正しく反映できるかを検証

＜手法＞

- 東洋人男性と西洋人男性の顔画像を入力し、AU4の強度を4に指定して表情生成を実行
- 出力画像における眉、目、口、鼻の変化を比較、考察

＜結果＞

- 東洋人男性の場合、口を開けて目を瞑り、眉をひそめた画像が生成されたが、鼻が不自然に長く変形した(図7)
- 西洋人男性の場合、同様に口を開けて目を瞑り、眉をひそめた画像が生成されたが破綻は見られなかった(図8)
- 東洋人男性のAU強度を5に上げた場合は表情が崩れ(図9) AU強度を2に下げた場合は眉の操作はわずかで一方目を閉じて口を開ける動作が現れた(図10)

＜考察＞

- 表情のデータセットが西洋人に偏っており、高い鼻、目のくぼみといった西洋人特有の顔形状を基準に表情を生成したのではない
- AU4の強度を下げてでも閉眼や開口が残ったことから、AU4と同時に現れやすい他のAU動作をモデルが同時に再現している可能性がある

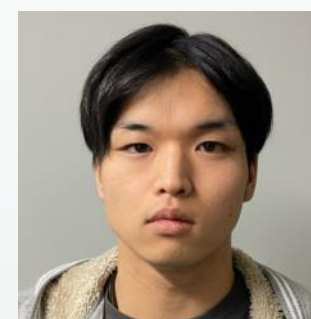


図7. 東洋人 AU4/強度4

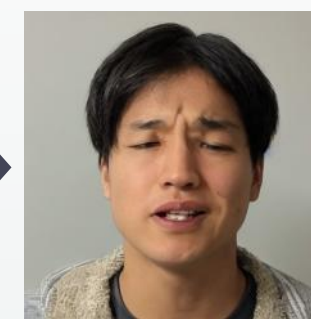


図8. 西洋人 AU4/強度4

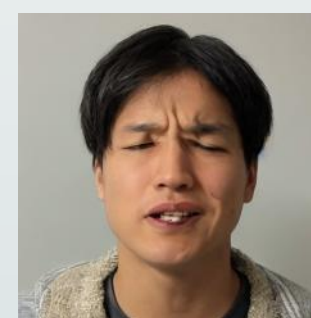


図9. 東洋人 AU4/強度5

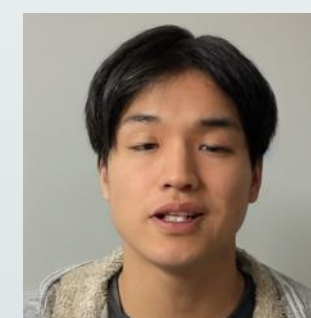


図10. 東洋人 AU4/強度2

参考文献

- [1] S. Huberman, O. Patashnik, O. Dahary, R. Mokady, D. Cohen-Or. Image Generation from Contextually-Contradictory Prompts. 2025.
- [2] X. He, C. Luo, X. Xian, B. Li, S. Song, M.H. Khan, W. Xie, L. Shen, Z. Ge. SynFER: Towards Boosting Facial Expression Recognition with Synthetic Data. 2024.
- [3] M. Wei, T. Varanka, X. Jiang, H.Q. Khor, G. Zhao. MagicFace: High-Fidelity Facial Expression Editing with Action-Unit Control. 2025.

今後の予定

- 各AUの表情変化傾向の検証
- MagicFaceモデルを用いた東洋人の顔における表情編集の再検証
- 東洋人顔への適応に向けた学習データ拡張の必要性の検討