申請先学部　基礎工学部　採択番号　No.2

令和５年度大阪大学未来基金「学部学生による自主研究奨励事業」研究成果報告書

ふりがな 氏名	ギ チホ WEI QIFENG	学部 学科	基礎工学部 システム科学科	学年	2 年
アドバイザー教員 氏名	内藤智之　准教授	所属	医学系研究科		
研究課題名	深層心理の鏡：機械学習を用いた主観的美のデジタル化 ～AI 絵師が描く感性の風景～				
研究成果の概要	研究目的、研究計画、研究方法、研究経過、研究成果等について記述すること。必要に応じて用紙を追加してもよい。（先行する研究を引用する場合は、「阪大生のためのアカデミックライティング入門」に従い、盗作剽窃にならないように引用部分を明示し文末に参考文献リストをつけること。）				

Abstract

本研究は従来の心理学的逆相関法（ICM）が社会属性の精神的表現を視覚化する任務において直面していた制約を克服するためのアプローチとして新たな手法 Diffusion Mental Model（Diffusion-MM）を提案した。以前の研究では、スパイクトリガー平均（STA）に基づきコサインノイズを用いて、人間の顔を使ったステレオタイプや社会的知覚を表現する精神画像を生成するのに効果的であった。近年の新しい研究では、この STA 法を StyleGAN モデルに拡張し、高品質な精神表現画像の生成を可能にした。しかし、Noise Based および StyleGAN 法はどちらでも、基礎画像とノイズへの依存、生成画像の品質、実世界画像を潜在ベクトルに逆マッピングが難しいなどの課題に直面している。当モデルは、知覚的損失で訓練された Autoencoders と Diffusion Model Encoder を組み合わせた 2 段階 Decoder を組み込んでおり、潜在空間の歪みや非対称性の問題を対処した。この進歩により、より高品質な精神表現画像の生成が可能となった。実験では DMM が生成した画像が被験者の好み・感性と高度一致している事を示した。さらに、当アプローチは、様々なカテゴリーにまたがる感情的判断のための Semantic 潜在空間内の共通メカニズムを示唆し、社会心理学研究におけるより広範な応用の可能性を浮き彫りにする。社会的知覚とステレオタイプに基づく精神スキーマを調査するための、より堅牢で多用途なツールを提供することにより、この分野における重要な進歩を遂げた。

1. Introduction

社会心理学の分野では、心理学的逆相関法[8]が一般的に社会的属性の知覚を調査するために用いられている。この方法は、個々人や集団が持つさまざまな現象に関するステレオタイプを視覚化する手段として機能し、生成された画像は「精神的表徴（心的テンプレート）」として参照される。これらの精神的表現は、参加者に与えられた印象に基づいて社会的属性を評価する基準を表し、これらの精神的表現間の類似性は、社会的属性を評価する上で重要であるとされている[1,2,3]。逆相関法[6]は、さまざまな職業の異なる社会集団が持つ温かさや能力のある顔の精神的表現を作成し、これらの社会集団内での知覚を定量的に分析するために使用された。

さらに、Ethier-Majcher et al.は、若者と高齢者の集団に関する信頼の精神的表現を視覚化し、これらの表現と幸福や怒りの表現との関係を調査し、若者と高齢者の間の信頼の判断における違いを

明らかにした[11]。一方で、この心理学的逆相関法には、基本画像やノイズに強く依存するなどいくつかの限界がある。この課題に対応するため、StyleGAN[7]を利用した新しい逆相関法が導入された[5]。Imai et al.は、StyleGAN のマッピングネットワークと、初期のランダム空間からマッピングされた抽象的な表現を含む空間と、人間の心理学的表現を含む空間との類似性を実証した。

しかし、StyleGAN の潜在空間の多次元正規性を保証することは難しい。その結果、人間の感情的判断の基準として機能する心理的表現画像または「心的テンプレート」は、歪みや不完全さを示す可能性がある。さらに、StyleGAN による画像生成は潜在空間から一方向であり、実際の画像を対応する潜在空間ベクトルに戻すことは不可能である。これらの問題に対処するため、本研究では Diffusion Models[4]に基づく Diffusion Mental Model を提案する。このモデルは、StyleGAN よりも正規化された潜在空間を持ち、高品質な画像を生成することができる。

Diffusion Mental Model は、2 段階のデコーダで構成されている。最初に、perceptual loss で訓練された autoencoder[10]を使用して画像を知覚的潜在空間に変換する。次に、拡散モデルにエンコーダを使用して、知覚的潜在空間をさらに抽象的な意味的潜在空間にマッピングする。2 段階のデコーダにより、実際の画像を潜在空間（意味的空間）にマッピングすることが可能になる。実験では、生成された画像だけでなく、実際の画像に対しても、意味的空間にマッピングされた潜在ベクトルに対する逆相関法が有効であることが示された。さらに、生成された精神的表現画像における歪みの問題に対処するため、逆相関法を通じて計算されたベクトルを微調整するための diffusion based latent Denoiser を提案する。StyleGAN との対照実験では、生成された画像が参加者の基準に基づいてより高品質であることが示された。

さらに、異なる Diffusion Mental Model モデルから生じる精神的テンプレートから別のカテゴリーに対する意味的潜在空間内の精神的表現を予測する実験が行われた。これは、異なるカテゴリーにわたる感情的判断の基準として機能する意味的潜在空間内の共通メカニズムの存在を示唆している。本研究で提案される Diffusion Mental Model の 2 段階デコーダは、従来の Latent Diffusion Model[10]や Diffusion Autoencoders[9]と比較して、より抽象的な表現を持つ diffusion Model の学習（モデル構築）をより迅速に行うことを可能にする。

2. Related works

2.1. Reverse Correlation Method

社会心理学における逆相関法は、心理物理実験で用いられ、刺激画像に対する評価を行う手法である。具体的には、ベース画像にノイズフィルターを重ね、その値を反転させた刺激画像を提示し、被験者はこれらの刺激に対して二択または四択で評価を行う。評価されたスコアに基づいてノイズフィルターを重み付けし、加算平均を計算することで心的表象が生成されるが、この方法はベース画像に強く依存する。Gosselin と Schyns et al.は、ベース画像なしで文字の「S」を逆相関法で可視化する実験を行ったが、刺激サイズが小さく、試行回数を大幅に増やす必要があった。また、Nester と Tarr et al.は、ジェンダーフェイスの可視化タスクにおいて、同様に小さい刺激サイズにもかかわらず、多数の試行が必要であった。今回提案した手法はこれらの問題を解決し、高い解像度（256×256）の自然に近い画像を生成する。

2.2. Diffusion Model

Diffusion Model[4]は、画像やテキストなどのデータを生成するための新しいアプローチである。Diffusion Model は、実際のデータに徐々にノイズを加えていく順過程（Forward Process）と、このノイズを段階的に取り除いていく逆過程（Reverse Process）の 2 つのプロセスに基づいている。

Forward Process では、実際の画像に対して徐々にノイズを加え、最終的には完全なノイズデータに変換する。この過程はマルコフ連鎖で定義され、各ステップでのノイズの加え方は確率的に決定される。このプロセスを数学的には以下のように表現することができる：

$$q(X_t \mid X_0) = N\left(X_t; \sqrt{\bar{a}_t} X_0, (1 - \bar{a}_t) I\right)$$

ここでX_0は元のデータ、X_t は時刻 t でのデータ状態、a_t はノイズの強さを制御するパラメータである。

Reverse Process では、ノイズが加えられたデータから徐々にノイズを取り除き、元のデータを再構築する。このプロセスは、各時刻でのデータのノイズを予測し、それを取り除いていくことによって行われる。この過程もマルコフ連鎖で定義され、以下の式で表される：

$$\begin{cases} q_\theta(X_T) = N(X_t; 0; I) \\ q_\theta(X_{0-T} \mid X_0) := q(X_T) \Pi q_\theta(X_{t-1} \mid X_t) \\ q_\theta(X_{t-1} \mid X_t) := N(X_{t-1}; \mu_\theta(X_t, t); \Sigma(X_t, t)) \end{cases}$$

ここで、$\mu_\theta(X_t, t)$ は時刻 t での画像のノイズを予測するネットワークであり、$\Sigma(X_t, t)$ はノイズの共分散行列である。

拡散モデルのネットワークとして U-Net が使用される。U-Net は、ある時刻の画像を入力とし、その画像のノイズを出力する。損失関数は予測ノイズと実際ノイズの L2 正則である。また、ノイズの強さは時刻によって異なるため、時刻の埋め込みが必要となる。

3. Diffusion Mental Model

当モデルは、Latent-Diffusion-Model[10]及び Diff-ae[9]のアーキテクチャに基づいて、Diffusion の U-net と他に、画像をより低い次元に圧縮する Semantic エンコーダを追加している。Semantic Encoder で圧縮された情報は U-net 内に組み込める。

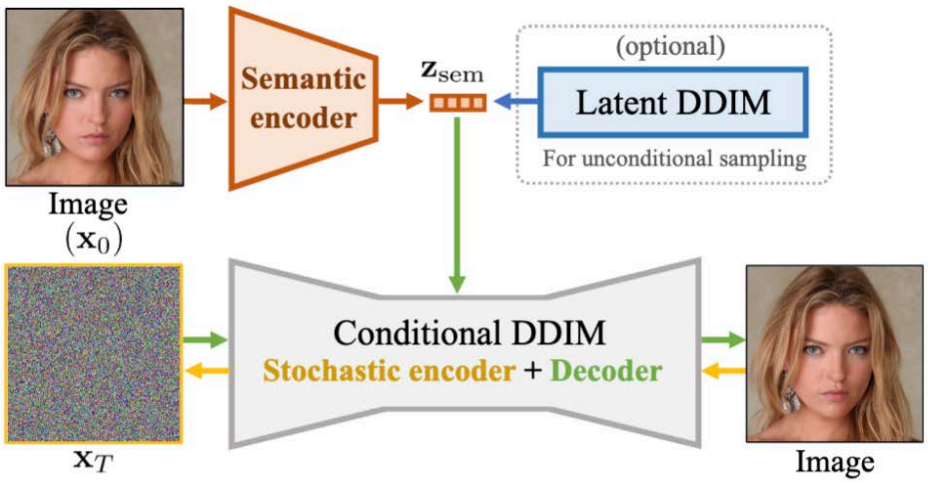

Figure 1

Diffusion-AE [9]の構造

また、画像を Diffusion プロセスに送る前に、知覚的損失で訓練されたオートエンコーダを使用して、画像を Perceptual Space(Z_{sem})に変換して学習するを行う。この手法を用いることにより、画像の冗長な情報を捨て、より本質的な情報だけを学習することで計算コストが大幅に軽減される。

Diffusion-Unet の学習が完了すると、データセット内の全データを一度 Semantic エンコーダを用いてエンコードし、得られた低次元ベクトルを使用して、Semantic Sampler を学習する。このプロセスにより、モデルの構成が完成する。この手法により、Reverse Correlation を行うための正規分布的な Semantic 空間が保たれ、さらに実際の画像を Semantic Space に効果的にマッピングし、より高度な画像生成が可能にした。

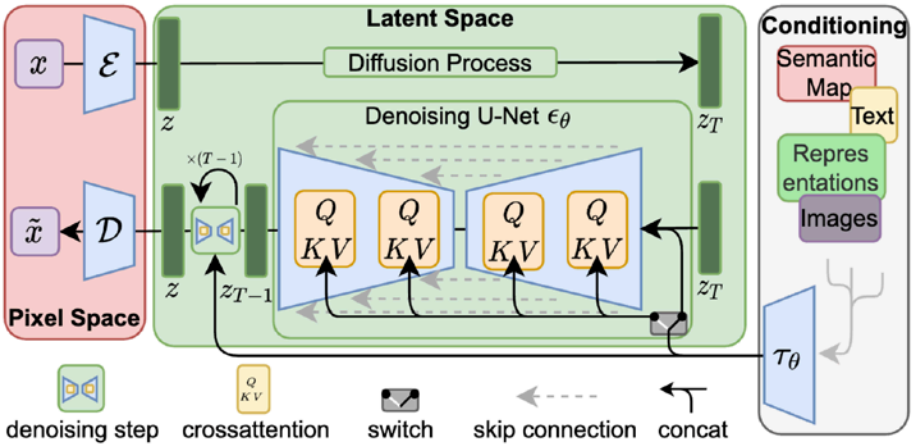

Figure 2

Latent Diffusion Model [10]のモデル構造

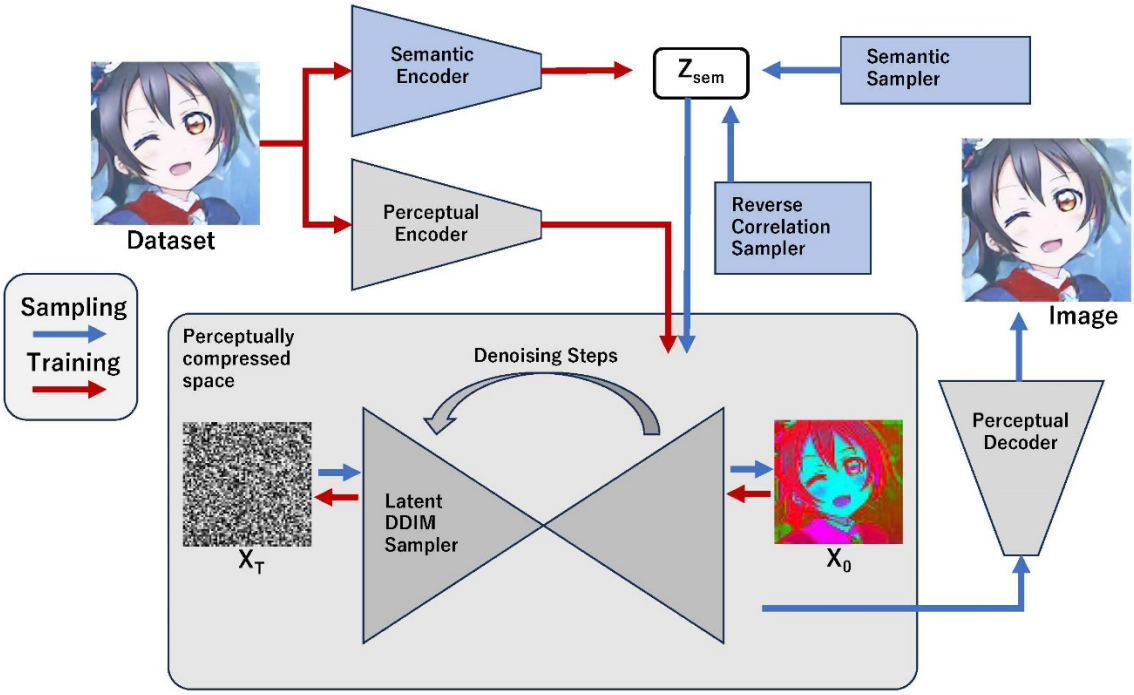

Figure 3

本研究が提案した Diffusion Mental Model(Diff-MM)の構造

4. Experiment

4.1. Visual Stimuli

本研究の実験で使われる画像は、データセット画像の特徴ベクトルを用いて学習 Semantic Sampler を用いて正規分布に従う初期ノイズからz_{sem}を Sampling する。その後、Latent DDIM Sampler に通じて画像を生成する。

Latent DDIM Sampler における生成時の特徴ベクトルと無関係なランダム性を避けるため、異なる初期ノイズマップを用いて同一の潜在ベクトルに基づいて画像を生成し、得られた画像の平均を計算する。この手法により、ランダム性に起因するバリエーション(高周波数成分)を抑制する。

本研究では、男性と女性、アニメと実写の 4 つのジャンルに分け、それぞれ 600 枚の画像を生成した、生成画像は Figure 2, Figure 3 で示す。

Figure 4

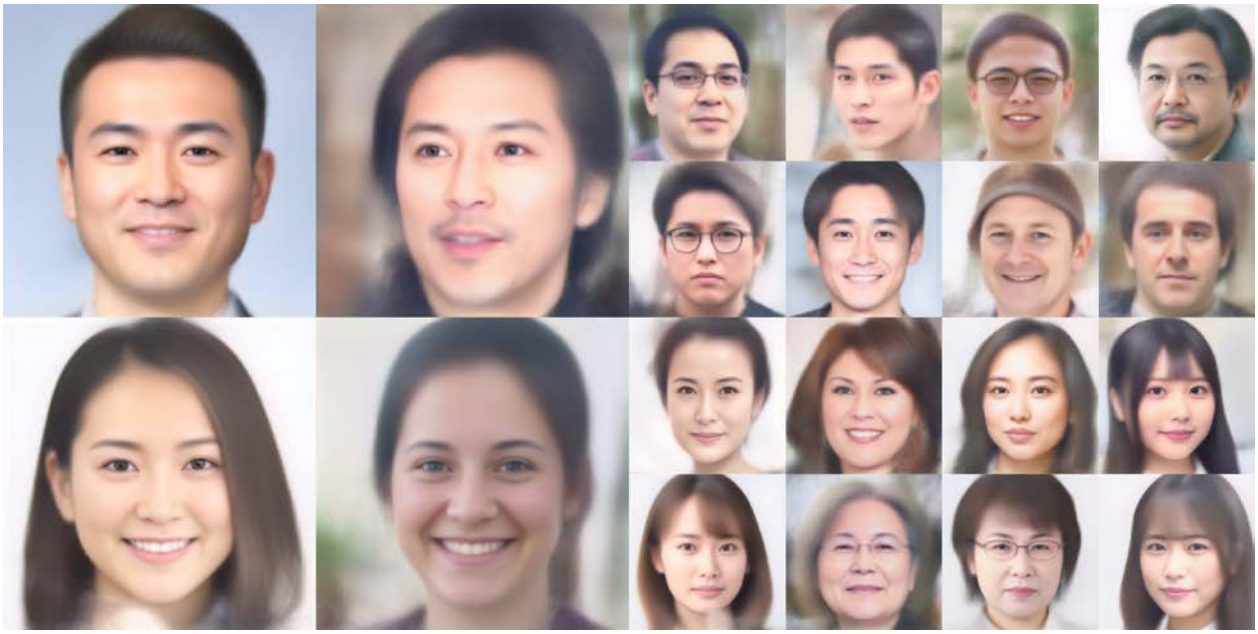

Figure 5

4.2. Experiment setup

4.2.1. Participants

本実験の参加者は、大阪大学/大阪大学大学院に在籍する学生で、キャンパス内のメーリングリストを通じて募集された。参加者の総数は 50 名（男性 28 名、女性 22 名）で、年齢分布の詳細は **Figure 6** に示されており、実験は一部対面、一部オンラインで行った。実験手順について口頭で説明を受けた後、参加者は **Figure 4** と **Figure 5** に示される美しさの評価タスクを実施した。

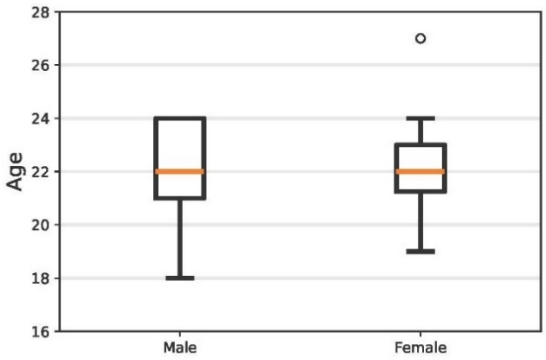

Figure 6

実験参加者の性別毎年齢分布

4.2.2. Tasks

実験参加者は当研究室の実験システムで行った。被験者は画面上に映し出した画像（600 枚・組）に対して美しさを基準に 0 点（醜）～100 点（美）の 101 段階で評定を行った **Figure 7**。600 枚の画像評定が終わったあと、評価された 600 枚の画像の潜在ベクトルを 4.3.3 で説明した逆相関法を適用し、心的表徴ベクトルを計算する。そして心的表徴ベクトルに対応する画像を生成し、601 枚目の画像として表示され、実験参加者に評定を行う。

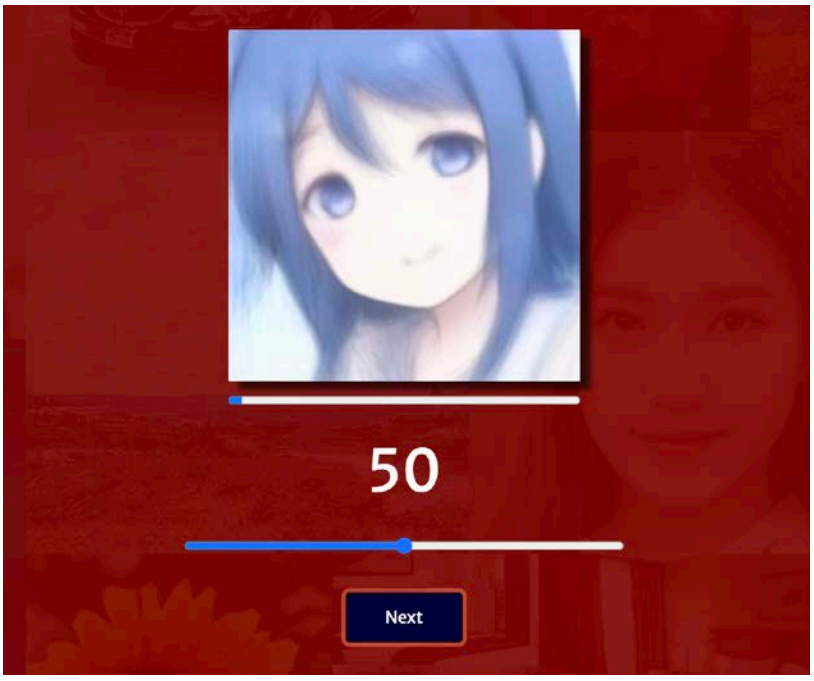

Figure 7

実験用刺激画像評定システム

4.2.3. Reverse Correlation Method

作成に使用する評定枚数を N、実験参加者が付けた魅力度評定値を x_i ($x_i \in R^N$) とし、Diffusion における画像の潜在ベクトルを z_{sem} ($z_{sem} \in R^N * R^{512}$) とする。

まずは実験参加者が評価した 600 の魅力度評定値 x_i に対して正規化を行い、偏差値に換算する。

$$x_i' = 10 * \frac{x_i - mean(x_i)}{std(x_i)} + 50$$

次に Diffusion における画像の特徴ベクトルに対して以下のように重み付け平均を計算し、心的テンプレートベクトルを作る。

$$z_{avg} = \frac{1}{N}\sum z_i$$

$$z_{beauty} = \frac{1}{N}\sum x_i'(z_i - z_{avg})$$

$$z_{CI} = z_{avg} + \alpha * z_{beauty}$$

α は z_{beauty} の L2 ノルムの α 倍と z_{avg} の L2 ノルムの和が画像に対応するベクトル z_i の L2 ノルムに等しくなるように調整された。

Reverse Correlation 法が正しく適用される条件として、神経科学の分野において Paninski et al.が試行回数を増やすと STA の計算結果が真の値に収束する性質を持つための必要十分条件は刺激分布が 1,2 の条件を満たすことであるとした[12]。

1. 平均が 0
2. 球対称性である

Diffusion-MM での逆相関法においてもこれらの特徴を得るためには潜在変数を定義する分布の平均が 1,2 の条件に従い、そこから生成される画像がランダムであることが条件となる。この二つの条件を検証するために、Diffusion Mental Model における z_{sem} が以上の条件を満たすかの検証を行う。

そのために z_{sem} に encode した潜在ベクトルから無作為に 2 つを取り出し、その間の関連性を確認する。その関係性は Figure 8 で示され、z_{sem}が高次元球対称性を持っていることを確認できる。

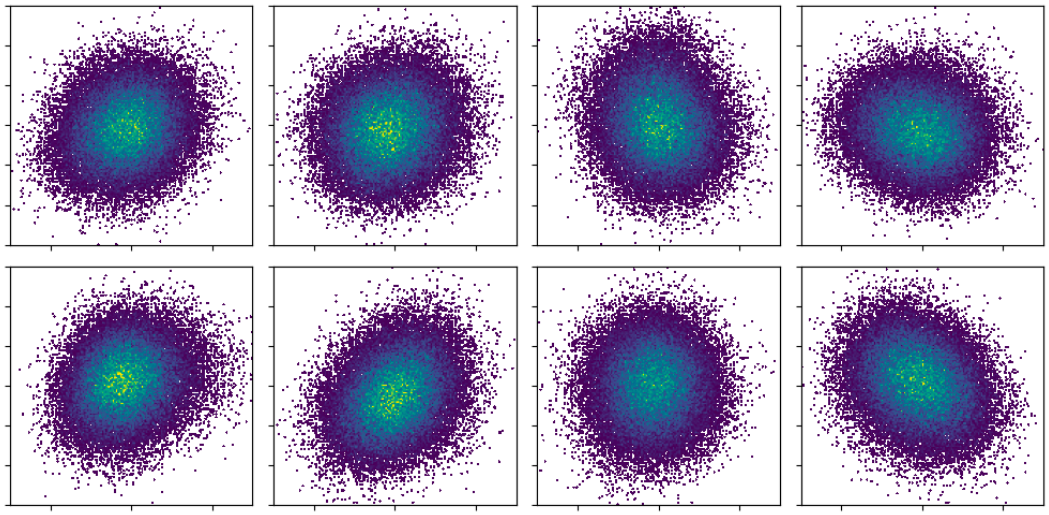

Figure 8

4.3. Result & Discussion

4.3.1. Evaluation of Mental Images

50 名の実験参加者に対し実験を行った。評価する画像のうち 600 枚は z_{sem} から生成された画像であり残りの一枚は 600 枚の画像の評定結果を基に Reverse Correlation Method を用いて生成された心理的表象（心的テンプレート）である。心的テンプレート画像は 600 枚の画像に対する評価が終わった時点で計算され被験者に通達されることなく提示される。実験の結果生成された心的イメージを **Figure 9, Figure 10** で示す。

Figure 9

Figure 10

次に最終的に心的テンプレートに対する評価スコアの分布と 600 枚の画像に対するスコアの分布を調べる。50 名の被験者において、全てのジャンル（男性と女性、アニメと実写）に対して心的テンプレートのスコアが遥かに上位になっていることが分かる（平均上位 0.93%・2.35SD）

また、対照実験として同時期に同被験者群に対して StyleGAN からの生成画像を用いて同様な評定実験を行った（アニメ女性・実写女性の 2 ジャンル）。StyleGAN と比べ、本研究が提案した Diffusion-MM がより被験者の好みの画像を生成したことが確認された。

	StyleGAN		DiffusionMM	
Type	Avg-SD	Avg-Score	Avg-SD	Avg-Score
Anif	0.78	64.48	**2.34**	**82.54**
Anim	-	-	2.60	84.96
Facef	0.54	55.44	**1.88**	**76.56**
Facem	-	-	2.60	70.04

Table 11

被験者が最終的に生成された心的テンプレート画像に対する評価。Type は（男性と女性、アニメと実写）4 ジャンルの評定画像。Avg-SD は被験者それぞれに対してスコアを正規化し、心的テンプレート画像の点数が平均から離れた距離（SD）の平均。Avg-Score は被験者が心的テンプレートに対しての評価スコアの平均（正規化なし）。

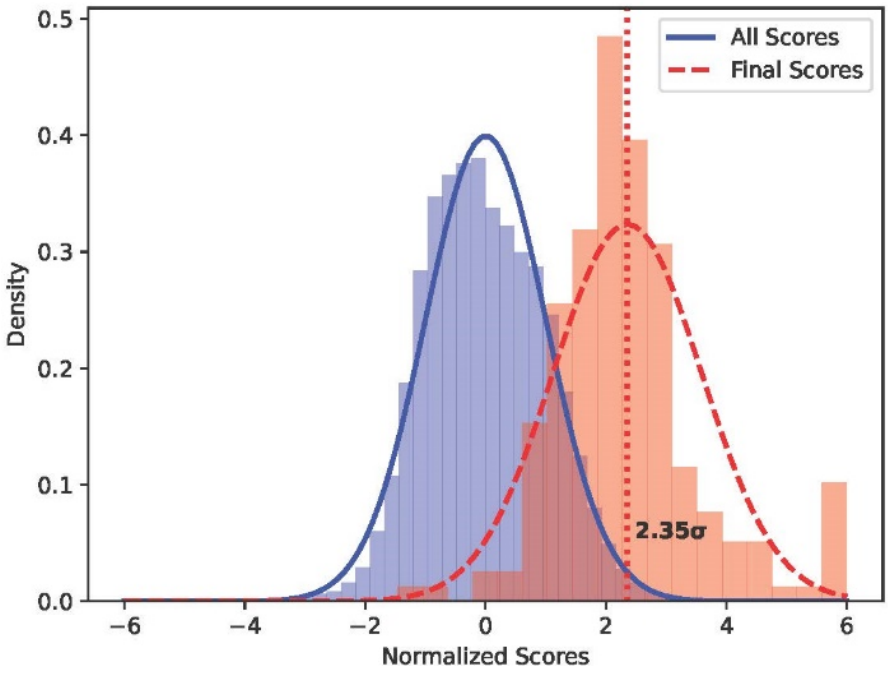

Figure 12.1

被験者が画像に対してつけたスコアの分布。Final Scores のヒストグラムは心的イメージに対するスコアの分布を示し、All Scores はそれ以外の画像に対する分布を示す。横軸は平均からの距離を標準偏差で示している。またヒストグラムは面積の合計が 1 になるように正規化した。それぞれの平均値は 2.35,0.00 であった。

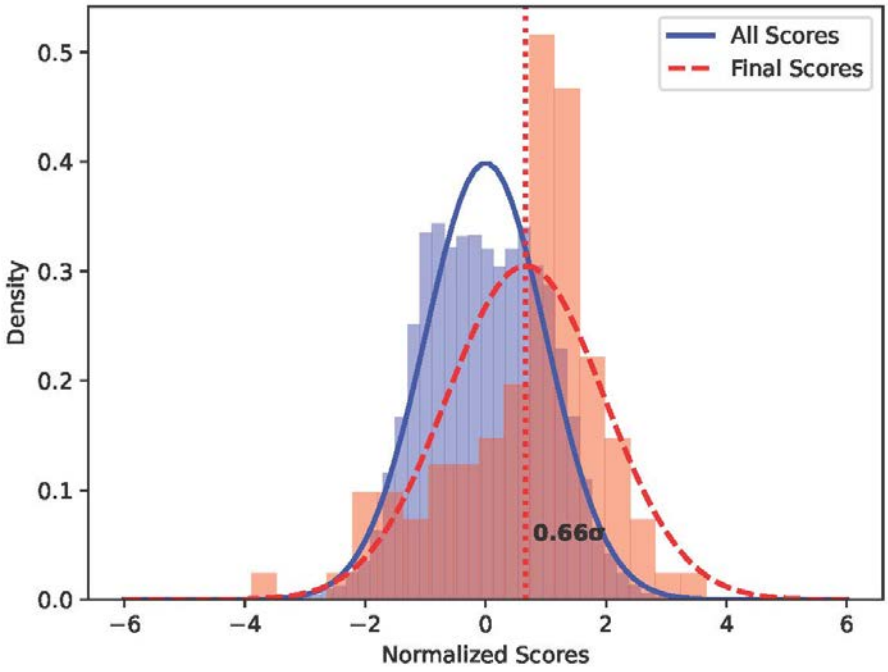

Figure 12.2

同時期同被験者群で StyleGAN を用いた対照群。Final Scores のヒストグラムは StyleGAN で生成した心的イメージに対するスコアの分布を示し、All Scores はそれ以外の画像に対する分布を示す。横軸は平均からの距離を標準偏差で示している。またヒストグラムは面積の合計が 1 になるように正規化した。平均値は 0.66,0.00 であった。

次に実験参加者の心的テンプレートを 600 枚の画像評定により効果的に近似していることを心的テンプレート仮説を用いて検証する。もし実験参加者の心的テンプレートが美しさの判断基準として機能しているならば、心的テンプレートと画像の類似度と評定点の間に正の相関が見られることが予想される。また、美しさの判断基準に一貫性があり、ブレが少ないほど、この相関係数は高くなると考えられる。Diffusion-MM で生成された画像には、対応する特徴ベクトル z_{sem} が存在し、これを用いて画像間の類似度を定量的に評価することが可能である。ベクトル間の類似度の測定には、コサイン類似度が使用された。

Figure 13 が実験参加者の結果が示されている。横軸は心的テンプレートと評定画像のコサイン類似度を、縦軸は正規化された評定画像スコアを示している。ジャンルに対して全参加者の相関係数の平均値及び有意な相関をもつ人数を **Table 14** で示す。

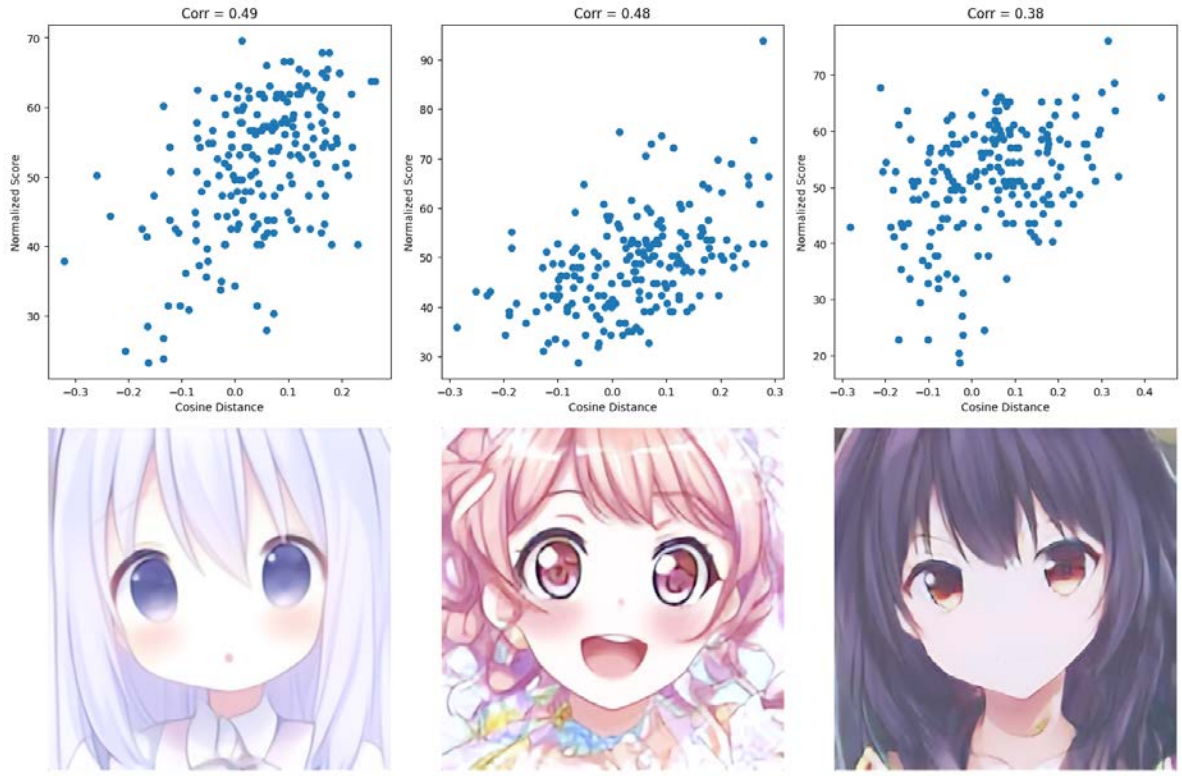

Figure 13

実験参加者が評定した 600 枚の画像のうち 400 枚を用いて心的テンプレートを作成し、残りの 200 枚のベクトルで心的テンプレートベクトルとのコサイン距離（横軸）、残りの 200 枚の正規化評定スコア（縦軸）

	Correlation	
Type	Avg-Corr	Significance
Anif	0.332	48/50
Anim	**0.394**	48/49
Facef	0.373	**48/48**
Facem	0.307	43/44

Table 14

（男性と女性、アニメと実写）の 4 ジャンルに対して、被験者ごとに心的テンプレートを計算した時の平均相関係数。及び相関検定をかけ有意な相関が確認された人数。

4.3.2. Convergence of Mental Images

同一ジャンルの異なる画像群に対して心的テンプレートの同一性について検証を行う。被験者が評価した 600 枚の画像から無作為に N 個取り出し、この N 枚に対して Reverse Correlation を行い、計算された$(z_{CI})_N$と 600 枚の評価画像全て使って作成した$(z_{CI})_{600}$の間の二乗誤差（MS を計算する。**Figure 15** ではベクトル空間における収束を表した。

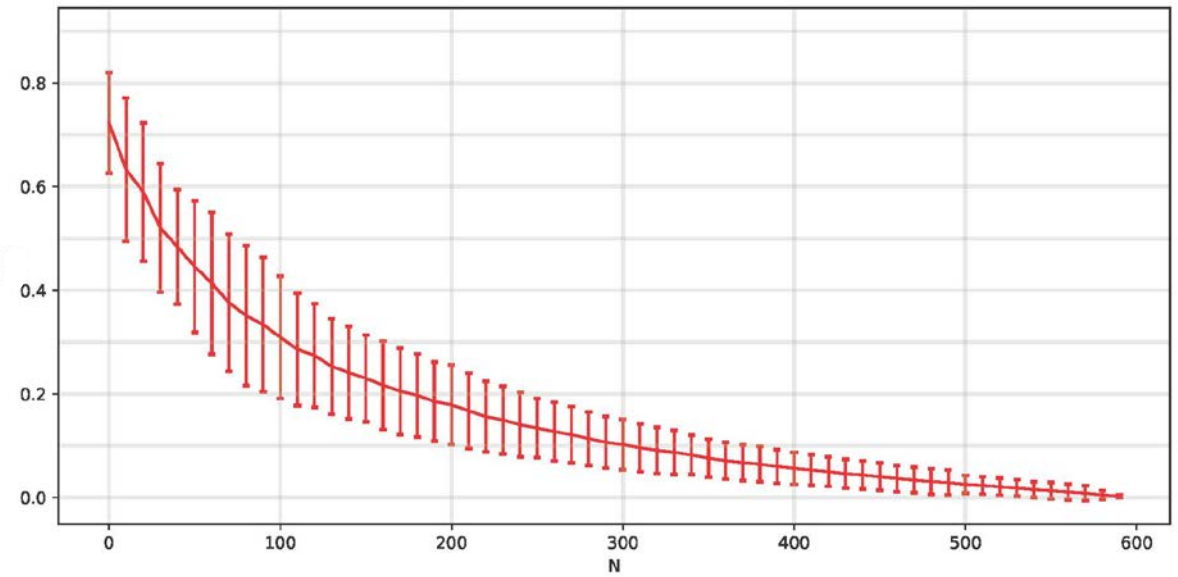

Figure 15

N 枚の画像に対して計算した$(z_{CI})_N$と 600 枚の評価画像全て使って作成した$(z_{CI})_{600}$の間の二乗誤差。横軸は N の数で、縦軸は二乗誤差（MSE）である。

次は生成画像で同一ジャンル・異なる画像群での心的テンプレートの収束を確認する。一人の被験者が評価した 600 枚の画像から無作為に N 個取り出し、この N 枚に対して Reverse Correlation を行い、計算された$(z_{CI})_N$を用いて$(心的テンプレート)_N$を生成する。**Figure 16** では心的テンプレートが画像空間における収束を表した。

Figure 16

N 枚の画像に対して計算した$(z_{CI})_N$を用いて生成した心的テンプレート画像。縦はそれぞれの被験者。

横は左（N=50）～右（N=600）を示し $\Delta N = 50/$コマ でN が左から右へ増加する。

4.3.3. Mental Images for Group

4.3.1 では、まず評定者ごとの心的テンプレートが作成及び可視化可能であることを確認した。本節では実験参加者 50 人の集団的好み及びその多様性を分析する。

集団的心的テンプレートを作成するためには、4.2.3 で計算した偏差値スケールのスコア x_i' を用いて、600 枚の評価画像に対してそれぞれの平均スコアを計算する。

$$x_{mean_i} = \frac{1}{50}\sum x_i'$$

x_{mean_i} を用いて Reverse Correlation を行い、集団心的テンプレートを生成する。

Figure 17

集団心的テンプレート

実験に参加した 50 人の集団心的テンプレートは **Figure 17** に示す。ジャンルごとで被験者に対して集団心的テンプレートとの類似度と正規化評定スコアの平均相関は **Table 18** に示す。集団心的テンプレートにおける平均相関が低いことは、集団においてより個体の多様性が存在していることを示唆する。

	Correlation	
Type	Avg-Corr	Significance
Anif	0.279	48/50
Anim	0.338	45/49
Facef	**0.351**	45/48
Facem	0.328	**43/44**

Table 18

（男性と女性、アニメと実写）の４ジャンルに対して、被験者が集団心的テンプレートにおける平均相関係数。及び相関検定をかけ有意な相関が確認された人数。

4.3.4.　Mental Hyperplane using PCA.

49 名の被験者によって作成された心的テンプレートに対し、主成分分析（PCA）を実施した。PCA は、多次元データの変動を最もよく捉える成分を抽出し、データの構造を理解するための強力な統計的手法である。この分析により、心的テンプレートのデータセットに含まれるパターンやトレンド、および被験者間の違いを効率的に可視化し理解することが可能となる。

PCA の結果は、7x7 の画像グリッド上にプロットされる。このグリッドにおいて、横軸は第一主成分を、縦軸は第二主成分を表す。第一主成分と第二主成分は、心的テンプレートデータの分散を最も大きく説明する二つの方向であり、これらの成分により心的テンプレートの主要な特徴が捉えられる。画像グリッド上のプロットにより、被験者ごとの心的テンプレートがどのように分布しているか、またそれらが全体的なデータセットの中でどのような位置にあるかが視覚的に理解できる Figure 19。

Figure 19

5. Future Expectation

本研究において提案された Diffusion Mental Model（Diff-MM）は、社会心理学における心的テンプレートの研究において重要な一歩を踏み出した。今後の研究においては、以下のような展望を持って進められることが期待される。

- モデルの改善と最適化：

現在の Diff-MM は、心的テンプレートの生成において顕著な結果をもたらしたが、さらなる改善の余地が存在する。Scratch から Diffusion-MM を学習するより、事前に大規模データで学習された Stable Diffusion モデルなどに Semantic Encoder を追加することで、より安価かつ多ジャンルを作成することが期待される。これは、異なるジャンルやカテゴリーに対する心的テンプレートの統合を向うために重要である。

- 心的テンプレートの応用：

Diffusion-MM は、社会心理学だけでなく、広範な分野に応用することが可能である。たとえば、マーケティング、広告、芸術など、人々の好みや嗜好を理解し、それに基づくコンテンツの生成に利用することが考えられる。

- 意味的潜在空間のさらなる研究：

Diffusion-MM において重要な役割を果たす意味的潜在空間に関する理解を深めることは、心的テンプレートのより正確な表現と生成につながる。この空間内のダイナミクスや構造に関する研究を通じて、心理学の理論的基盤を強化することが望まれる。例えば、異なるジャンルの画像における心的テンプレートの間の写像を試すことで心理学において対象に依存しない普遍的なメカニズムの解明につながる。

- 多様な文化や背景を持つ被験者における実験：

現在の研究は特定の集団に限定されているが、異なる文化や社会的背景を持つ被験者によるデータを用いることで、モデルの汎用性と適用範囲を広げることができる。このような多様性を考慮に入れた研究は、より普遍的な心的テンプレートの理解に寄与する。

本研究は、心的テンプレートのデジタル化という新しい領域において、多くの可能性を提示している。今後の研究においても、分野のさらなる発展が期待される。

6. Reference

[1]. Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualizing mental representations: A primer on noise-based reverse correlation in social psychology. European Review of Social Psychology, 28(1), 333-361. doi:10.1080/10463283.2017.1381469.

[2]. Dotsch, R., Wigboldus, D. H. J., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. Journal of Personality and Social Psychology, 100(6), 999-1014. doi:10.1037/a0023026.

[3]. Freeman, J. B., & Ambady, N. (2014). A dynamic interactive theory of person construal. Psychological Review, 118(2), 235-248. doi:10.1037/a0022327.
[4]. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. arXiv preprint arXiv:2006.11239.
[5]. Imai, R. (2020). Visualization of mental representation using reverse correlation and GAN and application. Master's thesis, Osaka University Graduate School of Life Function Studies.
[6]. Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. Frontiers in Psychology, 4, 386.
[7]. Karras, T., Laine, S., & Aila, T. (2018). A style-based generator architecture for generative adversarial networks. arXiv preprint arXiv:1812.04948.
[8]. Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. Cognitive Science, 28(2), 209-226. doi:10.1016/j.cogsci.2003.11.004.
[9]. Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[10]. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). High-resolution image synthesis with latent diffusion models. arXiv preprint arXiv:2112.10752.
[11]. Éthier Majcher, C., Joubert, S., & Gosselin, F. (2013). Reverse correlating trustworthy faces in young and older adults. Frontiers in Psychology, 4, 592. doi:10.3389/fpsyg.2013.00592.
[12]. Paninski, L. (2003). Convergence properties of some spike-triggered analysis techniques. Advances in Neural Information Processing Systems, (1). doi:10.1088/0954-898x/14/3/304
[13]. Suzuki, A. (2022). 異なる敵対的生成ネットワーク潜在空間間における心的テンプレート写像の試み. Master's thesis, Osaka University Graduate School of Life Function Studies.
[14]. Gosselin, F., & Schyns, P. G. (2003). Superstitious Perceptions Reveal Properties of Internal Representations. Psychological Science, 14(5), 505–509. https://doi.org/10.1111/1467-9280.03452
[15]. Nestor, A., & Tarr, M. J. (2008). Gender recognition of human faces using color. Psychological Science, 19(12), 1242–1246. Doi:10.1111/j.1467-9280.2008.02232.x

7. Acknowledgement

本研究を進めるにあたり、多大な支援とご指導を賜りました内藤智之准教授に深く感謝申し上げます。先生の専門的な知見と経験は、研究の指針を決定づけるものであり、またモデル学習における計算資源の提供により、本研究の実験と分析を円滑に進めることができました。内藤先生ご支援がなければ、本研究は成し遂げることができなかったでしょう。

また、同研究室の今西渉氏、石山敦史氏にも心からの感謝を表します。今西渉氏は、研究の進行において貴重な助言をくださり、また実験用システムの開発にも尽力してくださいました。石山敦史氏は、実験被験者の募集とその実施において大きな助力をしてくださいました。

最後に、本研究に関わるすべての方々に心から感謝申し上げます。皆様のご支援とご協力があったからこそ、本研究を無事に完成させることができました。心より御礼申し上げます。