

コンテンツ生成 AI が見せる拡張的なデザイン —認知拡張実験・異なる AI の比較・活用意向から— Expansive Design of Contents Generation AI - Expansion of Cognition, Comparison of AI, Intension for Utilization -

○原田 真衣^{*1}, 山田 悟史^{*2}
Mai HARADA^{*1}, Satoshi YAMADA^{*2}

*1 立命館大学 大学院 理工学研究科 博士課程前期課程

Graduate Student, Graduate School of Science and Engineering, Ritsumeikan Univ.

*2 立命館大学 理工学部 建築都市デザイン学科 准教授・博士 (工学)

Associate Professor, Dept. of Architecture and Urban Design, Ritsumeikan Univ., Dr.Eng.

キーワード: Deep Learning, AI, コンテンツ生成, 認知

Keywords: Deep Learning, AI, Contents Generation, Cognition

1. はじめに

人間が新しいアイデアを大量に考え出す場合、時には発想に限界を感じたり時間がかかったりするが、一方で AI は学習済みモデルがあればパラメータを変更するだけで大量の結果を出力することができる。人間は他者のデザインを複数参照し参考にしたり、あるいは組み合わせたりすることでデザインを生み出している。このとき参照物として AI が生成するアイデアやデザインを活用できれば、AI によって人間の発想を拡張できる可能性がある。

これまでの筆者の研究¹⁾では、AI の生成するデザインが人間の考えるものと異なり、AI によるデザインを見ることで人間の認知や発想が拡張される可能性があることを被験者実験で検証した。本研究ではこれに追加実験をおこなった上、最新のコンテンツ生成 AI 3つの生成物を比較した。また、最新の AI 技術をふまえた建築でのコンテンツ生成 AI の活用について、建築実務者を対象にアンケート調査をおこなった。

2. DCGAN を用いた認知拡張に関する実験

2.1. 実験の概要

本項は既発表報告の実験¹⁾の被験者数を増やし、さらに新しい実験として 2.5 章に述べる分類実験 C を追加したものである。AI の生成するデザインは人間の考えるものと異なっているのか、AI が生成したデザインを見ることで人間の認知や発想は拡張されるのかを、円と三角形という単純な図形を用いて検証した。(Figure1)

本研究ではコンテンツ生成 AI として DCGAN²⁾を採用し、Figure2 に示すような円または三角形のみが描かれた画像合計 400 枚をデータセットとし、画像を生成した。

2.2. DCGAN 生成画像を用いた描画実験

描画実験の対象は建築学科の学生 50 名である。「円と三角形を組み合わせる新しい図形をつくる」という題材のもと、時間内に思いつく限りの図案を描きだしてもらっ

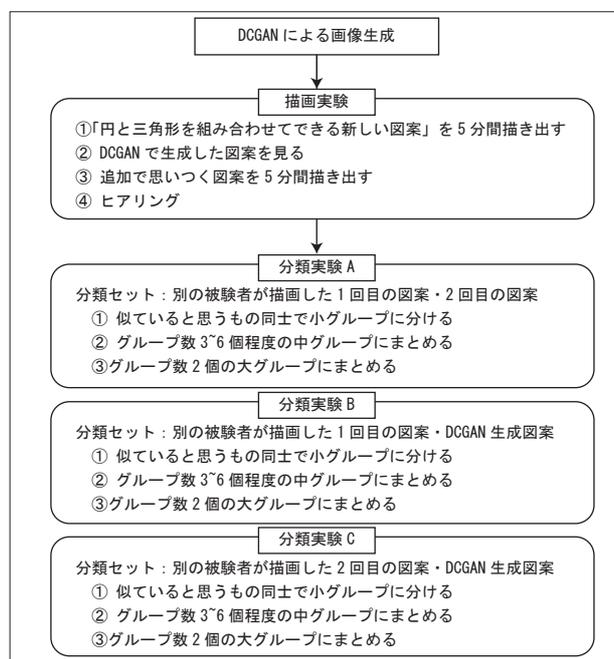


Figure 1. Research flow

た。描画 1 回目 (Figure2) では完全な円や三角形を複数個用いて作った図案が大多数であった。DCGAN 生成画像を見てもらった後の描画 2 回目では、円や三角形の一部を切り出して組み合わせた図案や、線が途切れた部分を持つ図案が見られ、実験後のヒアリングでは AI が描く図案は意外性があり自分の思考を広げる参考になったという意見を得た。一方で、描画 1 回目と同様に完全な円と三角形を組み合わせた図案を描き続ける被験者も見られ、一部の被験者は自分の考える図案と AI の描く図案がかけ離れすぎていて参考にできなかったと述べている。

2.3. 分類実験 A: 描画 1 回目と 2 回目

各分類実験の結果を Figure3 に示す。分類 A の被験者数は 20 名で、描画 1 回目の図案と 2 回目の図案を混ぜ合わせたセットを用いて、形が似ているものをグループ分けしてもらった。最終的に 2 つのグループに分け、

それぞれに1回目と2回目の図案がどのような割合で含まれるかを調べる。分類Aでは合計の項目を見ても、1回目の図案と2回目の図案が概ね分類された結果となった。

2.4. 分類実験B：描画1回目とGAN生成画像

分類Bでは、人間が描いた1回目の図案とDCGAN生成画像を混ぜ合わせたセットを用いて20名に分類してもらった。なかには人間が描く図案とDCGAN生成画像がはっきりと分けられる例もあり、合計したものを見ても同様であるといえる。これより、人間が考える図案とDCGAN生成画像には見た目からわかる違いがある可能性がある。

2.5. 分類実験C：描画2回目とGAN生成画像

分類Cの被験者数は12名である。DCGAN生成画像が参考になったと述べた被験者の描画2回目の図案と、DCGAN生成画像を混ぜたセットを用いた。人間とDCGANの図案でおよそ半分ずつに分類された例(セット⑪, ⑫)もあり、AIを参考にした被験者の図案はAIの図案と見分けがつかない場合があるといえる。

3. DeepSVGを用いた図形の生成

3.1. DeepSVGの概要とネットワーク構造

ラスター画像を用いて学習しラスター画像を生成するDCGANは、はっきりとした直線を描くことが困難であった。そこでベクター画像を扱うDeepSVG³⁾と呼ばれる生成モデルを用いて、DCGANの実験と同じ題材である円と三角形の中間のような図形を生成した。

DeepSVGはVAEの一種で、エンコーダーとデコーダーで構成される。エンコーダー、デコーダーともにすべてのパスで処理をおこなうパートと、ひとつひとつのパスを別々に処理するパートの2段階に分かれている。

3.2. DeepSVGによる生成結果

生成結果をFigure4に示す。円から三角形へ緩やかに形態が推移する様子を見ることが出来た。DeepSVGはsvg画像をパスごとに分解して処理がおこなわれるため、Figure7下段のように3つの円と3つの三角形の画像を補間すると、それぞれの円がそれぞれの三角形に混じることなく個別に変形していることが確認できる。

4. 高精度なコンテンツ生成AI

4.1. DALL-E 2

テキストから画像を描画する高精度コンテンツ生成AIとして、DALL-E 2⁴⁾、Midjourney⁵⁾、Stable Diffusion⁶⁾の3つが挙げられる。

DALL-E 2はOpenAIによって開発され、同じくOpenAIが開発したCLIPという技術を用いている。テキストが与えられると、CLIPが画像とテキストの埋め込みに変換し、それを用いて自己回帰モデルまたは拡散モデルによって画像を生成する。最後に拡散デコーダーを用いて最終画像を出力する。

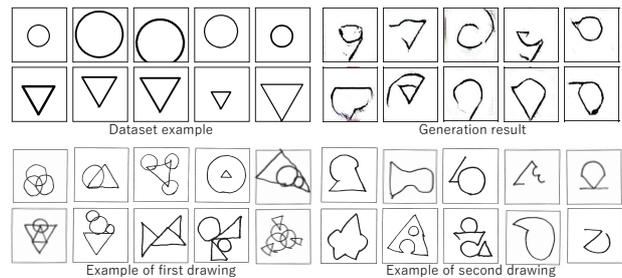


Figure 2. Example of dataset, generation result, drawing

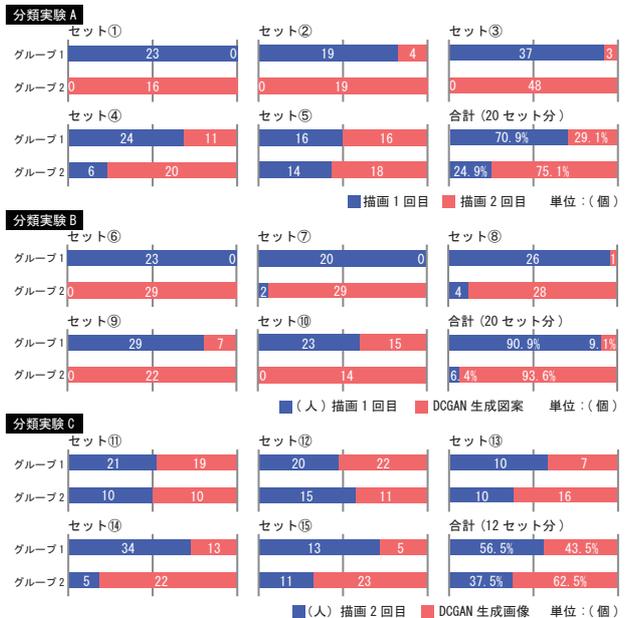


Figure 3. Result for classification experiment A, B, C

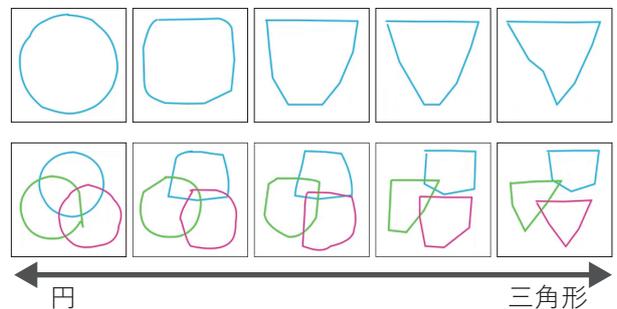


Figure 4. Generation result for DeepSVG

4.2. Midjourney

MidjourneyはDavid Holz氏を始めとする研究チームが開発したモデルで、ソースコードなどは公開されていないが、登録制で画像生成を試すことができる。

4.3. Stable Diffusion

Stable DiffusionはCompVis, LAION, Stability AIによって作成されたモデルである。学習データの作成にはLAION Aestheticsというデータセットを使用しており、ソースコードや学習モデルは公開されている。ネットワーク構造としては、ランダムノイズから画像を生成するLatent Diffusion Modelで生成した画像と、入力テキストとの類似度をCLIPを用いて検討する。

Prompt : circle and triangle

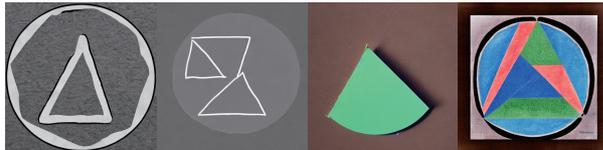
DALL-E 2



Midjourney



StableDiffusion

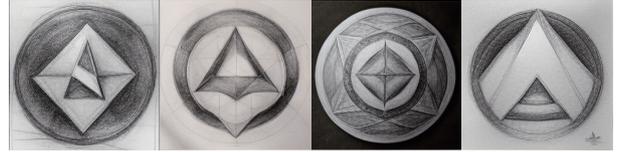


Prompt : circle and triangle sketch drawn by pencil

DALL-E 2



Midjourney



StableDiffusion

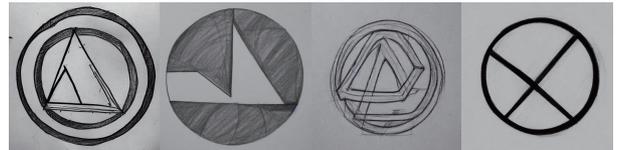


Figure 5. Generation result for 3 models

4. 4. 3つのコンテンツ生成AIの生成物比較

Figure5は、3つのモデルでそれぞれ2種類の文字列(prompt)から生成した画像の例である。

“circle and triangle”という文字列では、DALL-E 2の出力画像は完全な円や三角形を組み合わせたような画像が多く、色使いがはっきりしたものが生成された。一方Midjourneyは円が強調された画像が多く見られ、色をぼかして混ぜ合わせたような表現があった。またStable Diffusionでは、写真のようリアルな作風の画像が比較的多く生成され、不完全な円と三角形が融合しているような画像や、線が途中で途切れた図も描画された。

“circle and triangle sketch drawn by pencil”という文字列を与えた場合は、DALL-E 2は完全な円や三角形を複数個組み合わせたスケッチが非常に多く見られた。3つのモデルの中では、2.2章の描画実験で人間が描いた図案に一番近い作風といえる。Midjourneyでは円の中に様々な図形を入れ子構造のように重ねた図が生成された。Stable Diffusionにおいては、円や三角形の一部を切り取って組み合わせたようなスケッチが見られた。

単純な文字列のみで、作風などの詳細は指定せずに生成をおこなったが、3つのモデルにおいて見た目で判別できる特徴の違いが現れた。

5. 実務者を対象としたアンケート

5. 1. アンケート概要

4章で挙げたような“他者が作成した高精度なコンテンツ生成AI”を、建築設計の分野で活用する場合どのような評価がなされるのかを、実務で建築に携わる方を対象としたアンケートを用いて調査した。回答者は41人で、うち男性37人(90.2%)、女性4人(9.8%)であった。職種は建築設計、構造設計、IT、管理職など多岐にわたる。

5. 2. 質問内容

回答者には回答の前に、他者が作成した高精度なコンテンツ生成AIの一例として、Figure6のStable Diffusionを用いた生成画像を回答者に示した。アンケートの質問内容はTable1に示す。回答者には設計者・意思決定者・施主の3つの立場を想定してもらい、「自分ではない他者が作成し学習をおこなったコンテンツ生成AI」を設計で用いた建築デザインへの評価をヒアリングした。

5. 3. アンケート結果

選択式設問の回答をFigure7に示す。2.1.1と2.3.1より、設計者の立場ではコンテンツ生成AIが生成したデザインをそのまま提案はしないという意見が多かった一方で、施主の立場では否定的な意見の割合が減り、デザインの作者がAIであるか人であるかは評価に影響しないとする回答が増えた。また2.2.1より、デザインプロセスに他者が開発したAIを使用していることを開示するとした回答者が多かったが、その開示の積極性や程度については意見が分かれた。2.3.1、2.3.4、2.3.7の3問からは、建築以外のものに比べて建築の方が、デザインの作者が誰であるかが比較的重視されていると読み取れる。なお自由記述回答の項目については本稿では言及しない。

8. おわりに

本研究では、コンテンツ生成AIによる人間の認知の拡張について簡単な図形を用いて実験した。さらに3つの最新のコンテンツ生成AIの生成画像を比較し、その一つであるStable Diffusionを例に挙げながら、建築実務者を対象とした実務での活用と評価に関するアンケート調査をおこなった。今回の結果等をもとに、急速に発展してきているコンテンツ生成AIの建築分野での活用について考察を続けていきたい。

Table 1. Generation result for 3 models

【属性】	
1.1	性別
1.2	年代
1.3	職種
【設計者の立場】 “他者が作成したコンテンツ生成AI” が「あなた自身より優れたデザインを生成（提案）した場合」を想定。	
2.1.1	AIが生成したデザインをそのまま提案するか（5段階、択一式）
2.1.2	2.1.1の回答の理由（記述式）
2.1.3	設計・提案の一部として、AIのデザインを取り入れるか（5段階、択一式）
2.1.4	2.1.3の回答の理由（記述式）
2.1.5	AIによるデザインについて設計者としての受け止め方や考え方（記述式）
【意思決定者の立場】 意思決定をする立場（設計会社の経営陣など）であり、会社やチームでコンテンツ生成AIを活用する場合を想定	
2.2.1	デザインプロセスでのAI使用をクライアントに開示するか（5段階、択一式）
2.2.2	2.2.1の回答の理由（記述式）
2.2.3	コンテンツ生成AIを活用することは業務プロセス・人事評価・組織構成などどのように影響するか、またその理由（記述式）
【施主の立場】 施主の立場である場合を想定	
2.3.1	提案されたデザインが「AIが生成したもの そのもの」である場合、提案の評価に影響するか（5段階、択一式）
2.3.2	提案されたデザインが「AIが生成したもの そのもの」である場合の支払額へに影響するか（5段階、択一式）
2.3.3	2.3.1, 2.3.2の2問での回答の理由（記述式）
2.3.4	提案されたデザインが「AIによるデザインを設計・提案の一部として取り入れたもの」である場合、提案の評価に影響するか（5段階、択一式）
2.3.5	提案されたデザインが「AIによるデザインを設計・提案の一部として取り入れたもの」である場合、支払額に影響するか（5段階、択一式）
2.3.6	2.3.4, 2.3.5の2問での回答の理由（記述式）
2.3.7	建築以外で何かを購入する際、気に入ったデザインがAIによるものか人間によるものかによって購買意欲は変化するか（5段階、択一式）
2.3.8	2.3.7での回答の理由（記述式）
【AIの作成者があなた自身か他者にかかわらず、コンテンツ生成AIについて】	
3.1	建築分野でのコンテンツ生成AI活用によって生じる変化や期待（記述式）
3.2	AI技術が進歩した場合、自身の設計業務などに用いてみたいか（記述式）
3.3	「AIに与える指示」のオリジナリティや創造性とは何か、人間の創造性や創造の楽しさについて思うこと（記述式）

【Stable Diffusionに入力した言葉】

- ①... design concept sketch, pencil, master, simple, shape, zoning
- ②, ③... design concept sketch, pencil, master, simple, zoning, round , linear
- ④... architecture by Zaha Hadid

⑤, ⑥...

Built-in seating in the house. Photo, Hyperrealistic, sharp focus, 35mm, Architectural magazine, internationally transcendent works of contemporary architecture, a masterpiece, Pritzker Prize winner

⑦... Modern Architecture

⑧, ⑨... Japanese modern house

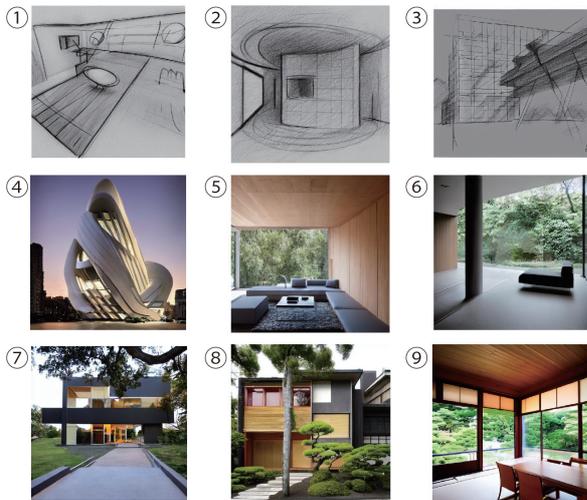


Figure 6. Generation result for 3 models

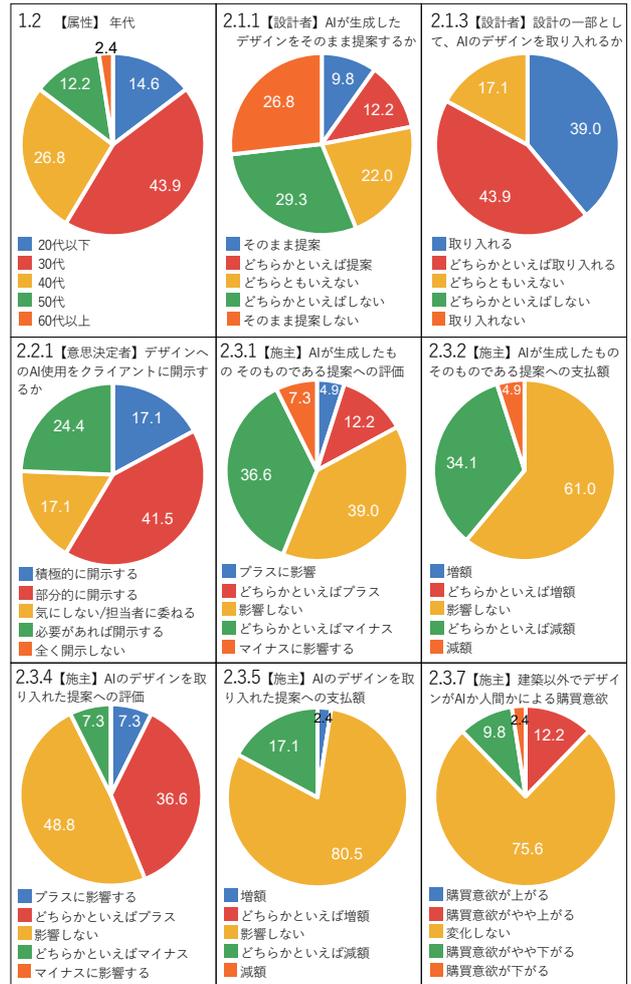


Figure 7. Generation result for 3 models

【参考文献】

- 1) 原田真衣, 山田悟史: Deep Learningを用いたデザイン生成AIに対する意識調査と認知拡張の検証, 第44回情報・システム・利用・技術シンポジウム論文集: 報告 pp. 467-470, 2021. 12, 日本建築学会
- 2) Alec Radford, Luke Metz, Soumith Chintala: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, arXiv, <https://arxiv.org/abs/1511.06434>, 2015
- 3) Alexandre Carlier, Martin Danelljan, Alexandre Alahi, Radu Timofte: DeepSVG: A Hierarchical Generative Network for Vector Graphics Animation, arXiv, <https://arxiv.org/abs/2007.11301>, 2020
- 4) Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, Mark Chen: Hierarchical Text-Conditional Image Generation with CLIP Latents, arXiv, <https://arxiv.org/abs/2204.06125>, 2022
- 5) David Holz: Midjourney, <https://www.midjourney.com/>, (2022年09月30日閲覧)
- 6) Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer: High-Resolution Image Synthesis With Latent Diffusion Models, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 10684-10695