

創造性から実用性まで

—さまざまなカテゴリを通じた AI に基づく画像生成ツールの比較—

呉 起東

本研究は、AI に基づく画像生成ツール（DALL-E、Artbreeder、Adobe Firefly）の性能と実用性を、創造性から実用性までのさまざまなカテゴリにわたって比較・評価することを目的とする。これを達成するために、まず、各ツールの理論的背景と技術的構造を分析し、次に、特定の画像生成タスクにおけるパフォーマンスを評価した。評価の方法としては、5つのカテゴリ（ロゴ作成、リアリティイメージ、デザイン応用、イラストレーション、3D グラフィック）において一貫したプロンプトを用いて画像生成を行い、生成された画像の創造性、品質、および用途への適合性を専門家による評価とユーザー調査を通じて比較した。これらの結果を基に、各ツールの特性、強み、限界、そしてデジタルクリエイティブ分野における今後の改善方向について考察した。最終的に、本研究は、AI 画像生成ツールの選択と応用における実践的なガイドラインを提供することを目指している。

キーワード：AI Image Generation Tools CG Design Digital Media
Digital Creative Fields

はじめに

Artificial Intelligence（以下 AI とする）に基づく画像生成技術の進歩は、デジタル時代の創作とコミュニケーション方法に革命的な変化をもたらした。この技術は、コンピュータビジョン、デジタルアート、グラフィックデザイン、広告、教育など、多岐にわたる分野で創造的かつ実用的な用途に利用されている。

Generative Adversarial Networks (GANs)、Variational Autoencoders (VAEs)、Transformer ベースのモデルなどの進展は、リアルで詳細な画像を生成する能力を向上させ、従来の創作方法を変革すると同時に新たな可能性を開いている。

この技術の重要性は、新たな画像を生成する能

力を超えて、創造性とイノベーションを促進し、コミュニケーションを改善し、データ増強やシミュレーション、パーソナライズされたコンテンツの生成など、多様な応用分野での利用可能性を提示している点である。

例えば、AI によって生成された画像は、教育的目的で複雑な概念を視覚化し、パーソナライズされたマーケティング資料を生成するために使用されることがある。また、データが不足している状況、例えば、新規市場や製品に関するデータがない場合や、希少イベントや特殊条件のデータが集められない場合、さらにはプライバシー保護やコストの制約によりデータ収集が困難な場合などに、AI モデルを訓練するための仮想データを生成する上で重要な役割を果たしている。

この背景を踏まえ、本研究は、AI に基づく画像生成技術、現在の適用例、そして将来の発展可能性を探索することにより、この分野の学術的理

解を深め、実用的な応用方策を模索することを目的とした。この技術の多様な潜在力を評価し、創造的な応用例を通じて人間の創作活動に与える影響を分析することで、技術と芸術、コミュニケーションの交点において新たな知識と洞察を提供する。

1. AI に基づく画像生成技術

AI の革新性の中でも、特に、Generative Adversarial Networks (GANs)、Variational Autoencoders (VAEs)、および Transformer モデルなどのアルゴリズムは、複雑な画像生成タスクを実行する上で重要な役割を果たしている。これらの技術は、画像のリアリティや詳細さ、生成の効率を大幅に向上させ、従来の方法では困難だった創造的かつ実用的な画像生成を可能にしている。

1-1 Generative Adversarial Networks (GANs)

Goodfellow et al. (2014) によって紹介された GANs は、相互に競合する 2 つのニューラルネットワーク（生成者と識別者）を通じて、実物と区別がつかない画像を生成するフレームワークである。この方法では、識別者が本物の画像と偽の画像を区別する過程を通じて、生成者が徐々により現実的な画像を生成するように学習を行う^{1) 2)}。

1-2 Variational Autoencoders (VAEs)

Kingma と Welling (2013) によって提案された VAEs は、入力データを効率的に圧縮し、復元する過程で新しい画像を生成する方法論である。これは、データの潜在的な特性をモデリングし、さまざまな変形された画像を生成する³⁾。

1-3 Transformer モデル

最近、自然言語処理 (NLP) 分野で広く使用される Transformer モデルは、画像生成にも適用されている。例えば、OpenAI の DALL-E は、テキスト説明から高品質の画像を生成するために Transformer モデルを使用している。これは、複雑な関係やパターンを学習し、さまざまな視覚的表現を生成する能力を示している^{4) 5)}。

2. 研究方法

本研究の目的は、DALL-E、Artbreeder、Adobe Firefly という 3 種類の AI 画像生成ツールを用いて、異なるカテゴリにおける画像生成の特性を比較することにある。本節では、研究における各ツールの使用方法、共通のプロンプトの設計、画像生成手順、および比較方法について詳細に説明する。

2-1 AI ツールの選定と概要

本研究で使用する AI 画像生成ツールは以下の通りである。

2-1-1 DALL-E (OpenAI) : テキストから画像を生成するための Transformer ベースのモデル。特に複雑な関係や詳細な視覚表現を学習する能力に優れている^{2) 6)}。

2-1-2 Artbreeder: GANs を活用してユーザーが画像を混合し、変形させることができるプラットフォーム。画像の「遺伝子」を「交配」することにより、新しい画像を生成する³⁾。

2-1-3 Adobe Firefly: Adobe が開発した、カスタム画像、テキストチャ、パターンを生成するツール。ユーザーインターフェースがシンプルで、さまざまなデザイン応用に適している⁷⁾。

各ツールは、その技術的特性と実用性に基づいて選定された。

2-2 プロンプトの設計

画像生成の評価に一貫性を持つ行うため、各ツールに対して同一のプロンプトを使用した。プロンプトは以下の 5 つのカテゴリにおいて設定された。

2-2-1 ロゴ作成: 「持続可能なファッションブランドのためのミニマリストロゴデザイン。水色と白を主色として使用し、地球と葉のモチーフが含まれたデザイン。」

2-2-2 リアリティイメージ: 「朝日が差し込む山中の木造キャビンのリアルなイメージ。キャビンの周りには密林と山岳地形が広がり、かすかに霧がかかっている風景。」

2-2-3 デザイン応用:「現代的なホームオフィス空間のためのインテリアデザインの模擬イメージ。木材素材と緑の植物を主要な要素として使用し、自然光がよく入るデザイン。」

2-2-4 イラストレーション:「魔法の森を背景にしたファンタジーイラストレーション。様々な色の魔法の花、きらめく光の要素、そして森を探検する少女が含まれたシーン。」

2-2-5 3D グラフィック:「未来都市の3Dモデル。高層ビル、浮遊する車、そしてLEDライトで照らされた道路が特徴的なサイバーパンクスタイルの都市scape。」

これらのプロンプトは、各カテゴリにおける一般的な要求と、ツールの能力を試すための多様な視覚的要素を含むように設計された。

2-3 画像生成手順

各ツールに対して以下の手順で画像生成を行った。

2-3-1 プロンプトの入力: 各ツールの指定された入力形式に従い、共通のプロンプトを入力する。

2-3-2 画像生成: プロンプトに基づき、ツールが画像を生成する。ツールが提供するデフォルトの設定を使用し、特別な調整は行わない。

2-3-3 画像の保存: 生成された画像はそのまま保存し、後の評価に使用するためのデータとして記録する。

3. 画像比較と評価方法

本節では、生成された画像の比較と評価の方法について説明する。

3-1 評価基準

生成された画像は、以下の基準で比較・評価を行った。

3-1-1 創造性: 画像の独創性や新規性。プロンプトの意図に対する独自の視覚的解釈。

3-1-2 品質: 画像の解像度、ディテールの再現度、色彩の豊かさ。

3-1-3 プロンプトへの忠実度: 生成された

画像がプロンプトの内容にどれだけ忠実か。プロンプトの意図をどれほど再現しているか。

これらの評価基準は、画像生成ツールの性能を総合的に判断するために設定された。

3-2 評価プロセス

評価プロセスは以下の通りである。

3-2-1 視覚的評価: カテゴリごとに生成された画像を視覚的に比較し、創造性、品質、プロンプトへの忠実度について評価する。

3-2-2 創造性: 生成された画像の独創性や、新規性を評価。プロンプトの意図に対する独自の解釈がどれほど創造的であるかを見る。

3-2-3 品質: 解像度やディテールの再現度、色彩の豊かさを評価。画像が高品質であるか、詳細な部分まで表現されているかを確認。

プロンプトへの忠実度: 画像がプロンプトの内容にどれだけ忠実であるかを評価。プロンプトの要件にどれほど正確に従っているかを判断した。

ツールごとの比較: 各カテゴリで生成された画像をツールごとにまとめ、視覚的な違いや特性を比較。

各ツールで生成された画像の違いを明確にし、それぞれの特性や強み、弱みを分析。

同一プロンプトによる画像生成結果から、ツールの表現特性や限界を抽出。

総合的な評価: すべてのカテゴリにおける評価結果を基に、各ツールの総合的なパフォーマンスを評価した。

生成された画像の全体的な質やプロンプトへの忠実度を総合的に判断した。

各ツールが特定のカテゴリにおいて強みを持つか、全体的なバランスを評価した。

3-3 ツールの特性の比較

得られた評価結果を基に、以下の点で各ツールの特性を比較した:

3-3-1 DALL-E: 複雑なプロンプトの解釈力と現実的な画像生成能力が高い。特にイラストレーションとリアルな画像生成において、高い創造性と現実感を示した。

3-3-2 Artbreeder: 画像の混合と変形に特化しており、ユーザーが既存の画像を基に新しい作品を探索する際に便利。特に、画像の融合において独特の表現を提供する。

3-3-3 Adobe Firefly: ユーザーフレンドリーなインターフェースを通じて、迅速にさまざまなデザイン要素を適用できる機能を提供。デザイン応用において使いやすさと効率性を発揮する。

4. イメージ生成と比較

4-1 DALL-E (OpenAI) を用いて作成したイメージ

DALL-E (OpenAI) は、テキストの説明に基づいて画像を生成する強力な AI ツールであり、創造的で多様な画像を生成できる能力が特徴である⁸⁾。

4-1-1 DALL-E を用いたロゴ作成



図1. DALL-E (OpenAI) を用いて作成したロゴ

さまざまなスタイルや創造的なアイデアを基にロゴを生成する。ただし、ロゴデザインに必要な特定の詳細や明確なブランドメッセージの伝達の面では、制限がある場合がある。カスタマイズにおいても、テキストプロンプトの正確性に大きく依存することがある (図1)。

4-1-2 DALL-E を用いたリアリティイメージ

テキストの説明を基に非常に現実的でディテールの優れた画像を生成する能力を有する。特に自然風景、都市の風景、物体のリアルな画像生成に

おいて、高いリアル感と細やかな質感をうまく再現する (図2)。

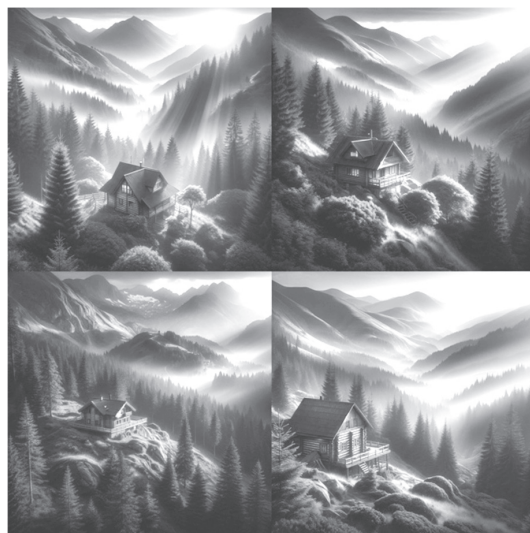


図2. DALL-E (OpenAI) を用いて作成したリアリティイメージ

4-1-3. DALL-E を用いたデザイン応用



図3. DALL-E (OpenAI) を用いて作成したデザイン応用

デザインプロジェクトのさまざまな要件に対応する柔軟性を提供する。創造的なアイデアの実現とさまざまなデザインスタイルの適用に優れ、若干制限されることがあるカスタマイズをプロンプ

トの調整によって克服できる（図3）。

4-1-4 DALL-Eを用いたイラストレーション



図4. DALL-E (OpenAI) を用いて作成したイラストレーション

複雑なテキストの説明を基に、非常に創造的で芸術的なイラストを生成する能力に優れている。ユーザーの想像力を基にした視覚的表現で非常に高い創造性を発揮する（図4）。

4-1-5 DALL-Eを用いた3Dグラフィック



図5. DALL-E (OpenAI) を用いて作成した3Dグラフィック

基本的な3D効果と立体感を持つ画像を生成することができるが、専門的な3Dモデリングソフトウェアの複雑さや精度には及ばない。3Dモデリングに関する具体的な要求がある場合、限定的な場合がある（図5）。

DALL-Eはさまざまなカテゴリにおいて優れたパフォーマンスを発揮するが、特定の要件や専門的な用途に対応する能力には制限がある場合がある。

4-2 Artbreederを用いて作成したイメージ
Artbreederは、ユーザーが創造的なイメージを生成できるAIベースのプラットフォームで、異なるイメージを混合して変形することにより、ユーザーが簡単に「遺伝子」を「交配」して新しいイメージを作り出す機能を提供している。

4-2-1 Artbreederを用いたロゴ作成

様々なイメージの混合機能を提供するが、ロゴデザインに必要な明確さとカスタマイズの面で制限があると思われる。特定のブランドアイデンティティに合致する高度にカスタマイズされたロゴを生成するには適していない可能性がある。（図6）。

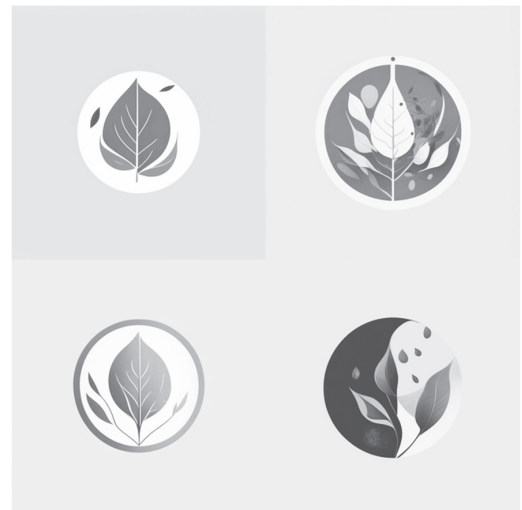


図6. Artbreederを用いて作成したロゴ

4-2-2 Artbreederを用いたリアリティイメージ

混合機能を通じて様々なリアリティイメージを生成することができるが、特定のシナリオに適した高度にリアルなイメージを生成するには限界があるが自然風景や一部のシナリオに対しては満足のいく結果を得られると思われる（図7）。



図 7. Artbreeder を用いて作成したリアリティイメージ

4-2-3 Artbreeder を用いたデザイン応用



図 8. Artbreeder を用いて作成したデザイン応用

ユーザーがイメージを簡単に混合し、調整することで、様々なデザイン応用分野に利用できる。しかし、非常に具体的なデザイン要求を満たすには多少制限がある。(図 8)。

4-2-4 Artbreeder を用いたイラストレーション

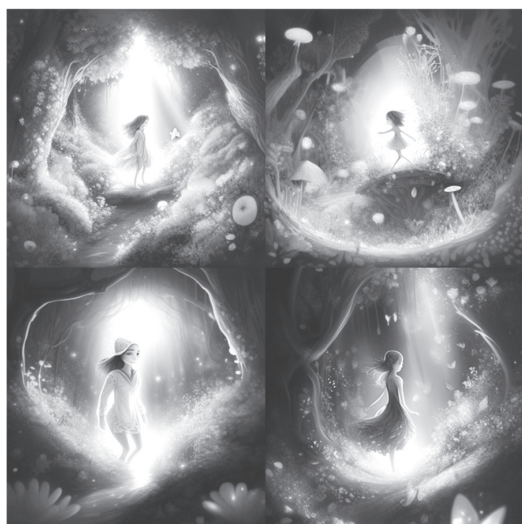


図 9. Artbreeder を用いて作成したイラストレーション

Artbreeder の強みの一つは、ユーザーが容易に創造的でアーティスティックなイラストレーションを生成できることである。異なるイメージの混合を通じて独特の視覚表現を探索でき、個性的なイラストレーション作業に有用である(図 9)。

4-2-5 Artbreeder を用いた 3D グラフィック

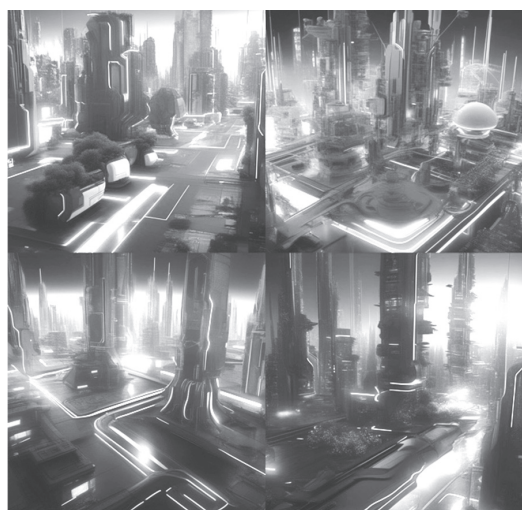


図 10. Artbreeder を用いて作成した 3D グラフィック

Artbreeder は主に 2D イメージの生成に焦点を当てており、専門的な 3D グラフィック作業に必

要なモデリングやレンダリング機能を提供していない。そのため、複雑な 3D プロジェクトには適していない可能性がある。

4-3 Adobe Firefly を用いて作成したイメージ

Adobe Firefly は、Adobe Systems によって開発されたデザインおよびグラフィックソフトウェアである。このソフトウェアはさまざまなデザイン作業をサポートし、ロゴ作成、画像編集、イラストレーション、3D モデリングなどの機能を提供している。Adobe Firefly は、ユーザーがさまざまなデザインプロジェクトを進行する際に、創造的で効果的なツールとして活用できと思われる。

4-3-1. Adobe Firefly を用いたロゴ作成



図 11. Adobe Firefly を用いて作成したロゴ

さまざまなテンプレートや機能を提供して創造的でカスタマイズ可能なロゴを作成するのに役立つ。デザインの明確さと伝達力を強調する要素もあり、ブランドメッセージを効果的に伝えることができる (図 11)。

4-3-2 Adobe Firefly を用いたリアリティイメージ



図 12. Adobe Firefly を用いて作成したリアリティイメージ

リアルな画像を生成するのにいくつかの制限がある。一部のディテールや質感が不足していることがあり、リアル感を表現するのに若干の難しさがあると思われる (図 12)。

4-3-3 Adobe Firefly を用いたデザイン応用



図 13. Adobe Firefly を用いて作成したデザイン応用

多数のパラメーターによりさまざまなデザイン応用に柔軟に対応できる機能を提供している。さまざまなテンプレートやツールを通じてデザイナー

が自分の創造性を発揮できるだけでなく、プロジェクトの統合性を高めるのにも役立つと思われる（図13）。

4-3-4 Adobe Firefly を用いたイラストレーション



図14. Adobe Firefly を用いて作成したイラストレーション

さまざまな芸術的な表現に対応する機能を提供し、創造的なイラストを作成するのに役立つと思われる。ただし、ユーザーが完全にカスタマイズしたり、細かく調整したりするにはやや制限がある（図14）。

4-5-5 Adobe Firefly を用いた3Dグラフィック

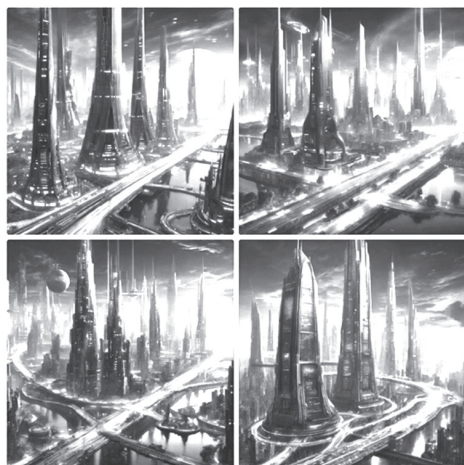


図15. Adobe Firefly を用いて作成した3Dグラフィック

Adobe Firefly は、3D グラフィックを生成するための便利なツールや機能を提供するが、専門的な3Dモデリングソフトウェアに比べてやや制限的である。立体感や奥行きを表現するのに苦労する場合がある。

5. 結果

各ツールによって生成された画像は、予想通りさまざまな特性を示した。DALL-E は特にイラストレーションとリアルな画像生成において高い創造性と現実感を示した。Artbreeder は画像の混合と変形を通じて独自のデザイン応用とイラストレーションを生成するのに強みを示した。Adobe Firefly は、ユーザーが簡単にアクセスして操作できるインターフェースを通じて、さまざまなデザイン要件を満たす画像を作成するのに役立つと思われる。

この研究を通じて、各AI画像生成ツールの利点と特性を明確に把握することができた。

DALL-E は、複雑なプロンプトの解釈力と現実的な画像生成能力が優れていることが確認された。これは特にリアルな画像とイラストレーションのカテゴリで顕著である。

一方、Artbreeder は、画像の混合と変形に特化しており、ユーザーが既存の画像を基に新しい作品を探索したいときに非常に便利である。

Adobe Firefly は、ユーザーフレンドリーなインターフェースとともに、迅速にさまざまなデザイン要素を適用できる機能を提供することで、デザイン応用と3Dグラフィック生成で強みを示した。

各ツールのパフォーマンス評価結果は、AI技術を活用したデジタルクリエイティブ分野の可能性を拡大すると同時に、特定の目的に最も適したツールの選択が重要である。たとえば、ロゴデザインや非常に具体的なデザイン要件を持つプロジェクトの場合、カスタマイズと細かい調整が可能なツールの選択が必須である。一方、創造的なアイデア発想やコンセプトの視覚化が主な目的の場合、DALL-E など、高い創造力を発揮できるツールが有利になることがある。

今回の研究は、同一のプロンプトによる画像生

成の結果を比較するものであり、ツールの性能評価には一定の限界があることを認識する必要がある。各ツールに適したプロンプトの設計方法によって、生成される画像の表現が変化する可能性があるため、ツールの実際の能力を正確に評価するには、プロンプト設計の多様性を考慮したさらなる研究が必要である。

6. 結論

AI ベースの画像生成ツールの発展は、デジタルクリエイティブ分野に革新をもたらし、さまざまな用途でこれらのツールが提供する可能性は無限である。この研究で評価されたツール DALL-E、Artbreeder、Adobe Firefly は、ロゴ制作、リアリティ画像、デザインアプリケーション、イラストレーション、3D グラフィック生成など、多岐にわたるカテゴリで有用なリソースであることを示した。

各ツールは、特定の機能とユーザーの要求に応じて独自の利点を持ち、このような評価を通じてデジタルアーティストとデザイナーが自分の創造的活動に最も適したツールを選択するための重要なガイドラインを提供する。

DALL-E の限界と改善方向

限界：DALL-E はテキストベースの詳細なプロンプトを必要とし、時にはユーザーの意図と完全に一致しない結果を生成することがある。

改善方向：ユーザーインターフェースを改善してプロンプトに対するフィードバックを提供し、結果の正確さを高めるための繰り返し修正プロセスを簡素化する必要がある。

Artbreeder の限界と改善方向：

限界：Artbreeder は、ユーザーが特定の画像に対して細かい調整を求める際に限定的である可能性がある。

改善方向：より細かい画像調整オプションを追加し、生成プロセス中にユーザーがより多くの制御を行えるようにインターフェースをアップデートする必要がある。

Adobe Firefly の限界と改善方向：

限界：Adobe Firefly は、3D モデリングなどの高度な機能において、他の専門ソフトウェアに比べ

て限定的である可能性がある。

改善方向：より進化した 3D 機能と、ユーザーが自分のプロジェクトにより多くのクリエイティブな要素を追加できるようにするさまざまなツールとテンプレートを含めることが重要である。

これらのツールの限界と改善方向を考慮すると、将来の研究では、新しい技術の発展とともに登場するツールの性能と適用可能性を継続的に追求することが重要である。これは、デジタルクリエイティブ分野を拡大し、創造的プロセスの効率性と創造性を向上させるために貢献できると思われる。この研究は、デジタルアーティストとデザイナーが現在および将来の創造的な課題に最も適したツールを識別し、選択するのを助ける貴重なリソースになると思う。

7. 参考文献

- 1) Goodfellow Ian et al. "Generative Adversarial Nets." *Advances in Neural Information Processing Systems*. 2014.
- 2) Kumar, A., Salimans, T., Ghosh, P., & Song, Y. (2021). "On the Design of Text-to-Image Diffusion Models."
- 3) Kingma Diederik P. and Max Welling. "Auto-Encoding Variational Bayes." *arXiv preprint arXiv:1312.6114* (2013).
- 4) Radford Alec et al. "Learning Transferable Visual Models from Natural Language Supervision." *arXiv preprint arXiv:2103.00020* (2021).
- 5) Hosseini-Asl, E., He, H., & Liu, Y. (2022). "Decoupling Knowledge Graph Reasoning from Text Generation via Pre-trained Language Models." In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*.
- 6) Ramesh Aditya et al. "Zero-Shot Text-to-Image Generation." *arXiv preprint arXiv:2102.12092* (2021).
- 7) Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021).

- 8) Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C.,
Radford, A., Chen, M., & Sutskever, I. (2021). "Zero-
Shot Text-to-Image Generation."
(受付 2024.3.22 受理 2024.7.29)
-