

# 画像生成 AI の現状とドラえもん

ニヤースのパーティ

## 1. はじめに

数々の生成 AI が誕生し、社会の目が一気に生成 AI に集まっている[1]。LLM（大規模言語モデル）だけでなく画像生成 AI も続々とリリースされ、世の中は生成 AI ブームになっている。Meta 社の LLaMA や Google の PaLM2 など、高性能の LLM モデルが次々に登場し、また、画像生成 AI も開発が盛んになっており、StabilityAI が開発した「Stable Diffusion」、Midjourney による「Midjourney」など様々な生成 AI が開発、リリースされている。その中でも「ChatGPT」の影響は大きい。ChatGPT は Open AI が 2022 年 11 月 30 日に公開した生成 AI である。現在は無料公開中で、公開から 1 週間で 100 万ユーザ、2 ヶ月で 1 億ユーザに到達した。その性能の高さから各所で話題になり、従来の大規模言語モデルよりも高度な意味理解と会話(チャット)が可能である。GPT-3（とその後継の GPT-3.5）をベースに、現在は GPT-4 が有料会員向けに公開されている。

さらに 2023 年 11 月現在、OpenAI は新たに DALL·E3 が公開された。こちらは画像生成 AI の一種で高品質な画像が生成できる。今回は、DALL·E3 を元にした Microsoft が運用している Bing image creator[2]を用いて、生成 AI の現状を紹介しようと思う。

## 2. Bing image creator

Bing image creator は Microsoft Bing の image creator であり、現在無料で公開されている。筆者は Microsoft edge 環境下で使用した。例えば、図 1 にこのエンジンで生成されたクロネコの画像を示す。このようにプロンプトと呼ばれる、AI との対話においてユーザが入力する指示や質問を投げかけることで生成された画像が出力される。このとき、画像生成 AI は乱数を使って異なるパターンの画像を生成しているが、このシード値は内部的に隠されているため、同じプロンプトを入力したとしても同じ結果が得られないことに留意する。この画像を定量的に評価することはしないが、非常に質の高い画像が生成されているように思う。また、Bing image creator の特徴として、単純なプロンプトでも精度の高いプロンプトを内部的に補って生成されるというものがある。詳しい説明は割愛する。



図 1 プロンプト: "A spooky and cute illustration of a black cat with yellow eyes and a purple collar sitting on a wooden windowsill with a jack-o-lantern and a candle beside it, looking at a full moon and a swarm of bats in the dark blue sky, Halloween themed drawing" の生成画像(color online)

### 3. 画像生成 AI とドラえもん

#### 3.1 画像生成 AI の変化

ここでドラえもんをプロンプトに入れた結果を実験させることを考える。また、生成 AI の発展というところにも着目したい。2023 年 4 月に実施した生成 AI の画像を図 2 に示す。これは、Stable Diffusion Online という Stable Diffusion をブラウザで動かすことができるサービスで実行した。



図 2 プロンプト:"Doraemon"の Stable Diffusion Online による画像生成結果 (2023 年 4 月) (color online)

次に Bing image creator を用いて生成した画像を図 3 に示す。



図 3 ドラえもんの絵と入力したときの Bing image creator による出力結果(color online)  
比較すると図 3 の方がドラえもんの特徴を抑えられていることが一目瞭然ではないだろうか。図 2 は手が 5 本指になっているなど、ドラえもんの特徴を抑えていないことが見受けられる。生成 AI エンジンが変わっているため単純な比較は難しいが、半年間での生成 AI の変化が見受けられると思う。一方、図 3 でもドラえもんの鈴の部分は完全再現されていないなどまだ改善の余地が見られる。

### 3.2 ドラえもんの画像生成出力結果

プロンプトにドラえもんという単純な入力ではなく、他の入力についても生成した。その画像を図 4 に示す。

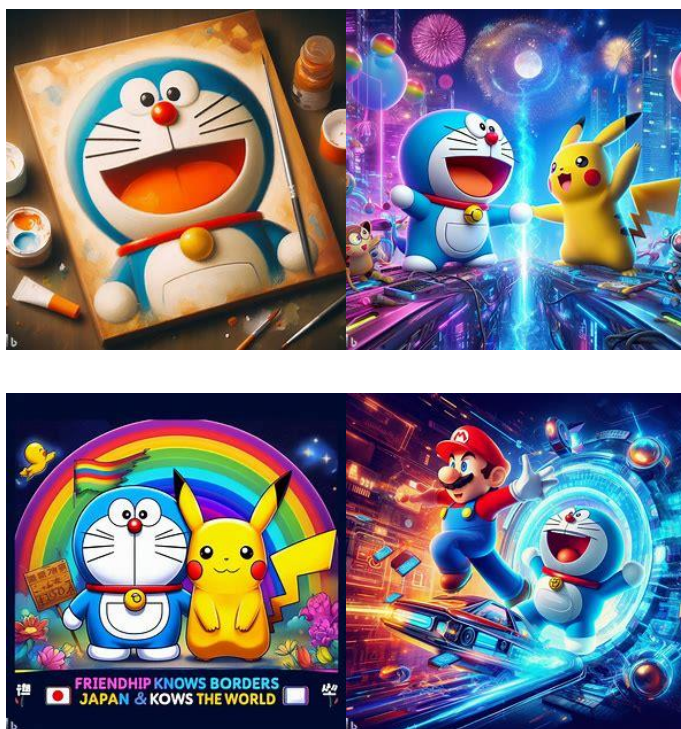


図 4 プロンプト(左上)"ドラえもんを油絵風に描く" (右上)"ドラえもんとピカチュウをコラボした絵"  
(左下)"ドラえもんとピカチュウのコラボ絵" (右下)"ドラえもんとマリオのコラボ画像(color online)

様々なプロンプトでもドラえもんの特徴を良く抑えていることと考える。また、右上と下の画像はドラえもん以外のキャラクターが出力結果に含まれているものである。比較すると、

ドラえもんよりピカチュウやマリオの出力結果の質が高いと感じた。これは学習した画像の量が原因だと考える。例えば、Doraemon を Google 検索した場合、約 205,000,000 件の検索結果であるが、Pokemon は約 1,080,000,000 件、Mario は約 2,420,000,000 件と 5 倍から 10 倍以上の差がある。GPT が発展した経緯を見ても、学習量がある量を超えたときに出力結果が良くなったという結果が出ており [3]、学習量の差は生成画像の質を左右する重要なパラメータだと言える。

#### 4. まとめ

今回は画像生成 AI、特に Bing image creator に絞って、出力結果を見てきた。画像生成 AI の質の高さ、つまりドラえもんの特徴を捉えた画像を生成できていることがわかった。今回はプロンプトをより細かく指示することはしなかったが、背景の指定など改善の余地は大きい。読者の方にも一度生成 AI に触れてみてほしい。感想も歓迎する。

まだ DALL-E3 が実装されて 1、2 ヶ月である。これから生成 AI が発展していくことにつれて、画像生成のさらなる改善が期待される。

最後に 2024 年公開予定の『映画ドラえもん のび太の地球交響楽』をプロンプトに入力した画像生成結果を図 5 に示してこのまとめを終える。



図 5 プロンプト"『映画ドラえもん のび太の地球交響楽』というタイトルの絵"(color online)

#### 参考文献

- [1] 松尾 豊、「生成 AI の技術動向と影響」、  
<https://www.mhlw.go.jp/content/11601000/001125241.pdf>
- [2] Bing image creator, <https://www.bing.com/images/create?FORM=GENILP>
- [3] Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

カラー版は京都大学藤子不二雄同好会公式 HP (<https://kyodaifujiko.web.fc2.com/>)  
に後日アップロードされるものをご覧ください。

