

令和 6 年度
静岡大学大学院 総合科学研究科
工学専攻数理システム工学コース
修士論文

文章から漫画を作成するための
プロンプト生成手法

指導教員

岡部 誠 准教授

提出者

51335021 南 貴明

提出日

2025 年 2 月 3 日

目次

第 1 章 はじめに.....	4
第 2 章 関連研究	6
2.1 ChatGPT を用いたアニメーション生成.....	6
2.2 画像生成 AI.....	6
2.3 漫画生成の試み	7
第 3 章 提案手法	10
3.1 プロンプト作成	12
3.2 画像生成.....	12
第 4 章 結果.....	15
4.1 作成結果.....	15
4.2 比較評価.....	21
4.2.1 提案手法と ChatGPT のみの比較.....	21
4.2.2 Midjourney を使った場合との比較.....	22
第 5 章 まとめと今後の展望	28

謝辞	29
参考文献	30
付録	32

第1章 はじめに

日本には、「漫画」という優れた文学表現がある。漫画とは、絵を用いて物語を表現する文学の一形態であり、情報を伝える手法として非常に適している。その用途は架空の物語の描写にとどまらず、実話をもとにしたエピソードの伝達にも活用されている。その例として、伝記や小説、舞台作品などが数多くコミカライズされている。また、漫画は日本国内だけでなく、海外でも広く認知され、世界を代表する娯楽の一つとなっている。漫画を用いることで、文学的な情報を視覚的に表現できるが、漫画の制作には高度なスキルが求められる。まず、物語を漫画化するためには、その文章を理解し、適切に処理する必要がある。物語のあらすじや詳細を把握し、それを基に漫画の構成を決定することが不可欠である。これは、物語の内容を十分に伝えるために必要な作業である。さらに、漫画を描くには高い画力が求められ、その制作は非常に困難な作業となる。これらの要因から、漫画の制作やコミカライズには、多くの時間と高度な技術が必要とされる。

そこで我々は、誰でも文章から漫画を作成できる手法の研究を行っている。その手法として、画像生成 AI を活用して、漫画を作成する手法を提案する。画像生成 AI とは、人工知能を活用して新しい画像を生成する技術であり、この技術を用いることで、絵を描くことが難しい人でも漫画を作成することが可能となる。本研究では、使用する画像生成 AI として Midjourney を選定した。Midjourney は、Midjourney Inc.によって開発された AI を活用した画像生成ツールであり、プロンプトを入力することで、それに基づいた画像を生成する(図 1)。しかし、Midjourney の利用は容易ではない。画像生成 AI を活用して適切な画像を得るためには、プロンプトを作成する必要がある。プロンプトとは、コンピュータや AI に特定の動作や回答を指示するための入力文であり[8][9]、一般的に膨大な文章となることが多いため、作成が困難な作業である。特に、漫画は視覚的に情報を伝えるメディアであるため、適切なプロンプトの作成はさらに難易度が高くなる。そこで、本研究ではプロンプトの作成を大規模マルチモーダルモデルで行う。大規模マルチモーダルモデルとは、テキスト、画像、音声などの異なるデータ形式を統合的に処理・理解し、生成できる AI モデルである。この技術を活用することで、文章の理解、プロンプトの作成だけでなく、漫画に適した画像の選定にが可能になると考えた。大規模マルチモーダルモデルを導入することで、汎用性の高い漫画作成手法の実現を目指す。また、近年では、大規模マルチモーダルモデルを使用して文章の下処理を行う研究が行われている[1][2]。本研究では、プロンプトを作成する大規模マルチモーダルモデルとして、ChatGPT を使用する。ChatGPT とは、OpenAI によって開発された AI サービスであり、情報提供、文章要約・作成、多言語翻訳、画像生成・認識などの多様な機能を備えている。また、これらの機能をチャット形式で利用でき、ユーザーのニーズに応じた柔軟な対応が可能である。

本研究では、大規模マルチモーダルモデルを用いて文章の理解・処理を行い、それに基づ

いてプロンプトを作成し、画像生成 AI に入力することで画像を生成する。このプロセスを統合し、漫画を作成する手法を提案する。本手法を用いることで、誰でも文章を用意するだけで漫画の作成が可能となる。



図1 Midjourney で生成した画像
左のプロンプトを入力した結果、右の画像が生成される。

第 2 章 関連研究

2.1 ChatGPT を用いたアニメーション生成

近年、ChatGPT を用いた文章や画像の処理に関する研究が進められている。Y. Li らは、ChatGPT などの AI を活用するアニメーションの自動生成手法の Anim-Director[1] を提案した。この手法では、アニメーション化したい文章を用意するだけでアニメーションの生成が可能となる。

ChatGPT を用いて文章を解析し、不足部分を補完した後、キャラクターや設定の詳細を抽出し、それに基づいて各場面の説明を生成する。場面の説明の妥当性を ChatGPT が確認し、正しいと判断された場合にアニメーションの生成を開始する。まず、画像生成 AI Midjourney を用いてキャラクターと背景の画像を生成する。その際、プロンプトは ChatGPT に作成させる。次に、生成したキャラクターと背景を基に、各場面の画像を Midjourney で作成する。生成された場面の画像は ChatGPT が評価し、必要に応じて修正を行う。画像が完成したら、動画生成ツール Pika を用いて画像からアニメーションを生成する。Pika には画像と説明テキストを入力し、複数のアニメーションを作成する。その中から ChatGPT が最適なものを選び、最終的なアニメーションとして採用する。

この研究は、文章を用意し、複数の生成 AI を用いてアニメーションを生成している。また、大規模マルチモーダルモデルを使用してプロンプトを生成することが可能であることを示しており、複数の生成 AI を用いることの有効性を確認できる。

2.2 画像生成 AI

画像生成 AI とは、特定の条件やプロンプトに基づいて画像を生成する手法である。プロンプトを画像に入力することで、画像が出力される。画像生成技術は複数あり、GAN や拡散モデルなどが挙げられる。

GAN とは I. Goodfellow [4] らが提案した機械学習モデルであり、画像や音声を生成するための深層学習技術である。生成モデルと識別モデルの 2 つのニューラルネットワークで構成されている。生成モデルは画像生成をランダムなノイズで画像の生成を行い、識別モデルは生成モデルが生成したデータがデータセットの画像と比較して本物か偽物かを判定するモデルである。また、生成モデルは識別モデルを騙すような画像を生成するように学習し、識別モデルは生成モデルが生成した画像とデータセットの画像を判別することができるように学習する。このように、互いのモデルが敵対して学習を行う手法である。また、様々な

研究[3][10][16][18][21]で応用されている手法である。

拡散モデルとは、標準正規分布に従うランダムなノイズを画像データに加え、徐々にノイズを取り除きながら画像を再構築することで画像を生成する手法である。代表的な拡散モデルとして、Stable Diffusion など[5][6][11]が挙げられる。

本研究では画像生成技術を活用して漫画の作成を行う。また、提案手法に至るまで、GAN を用いて漫画の自動生成を試みた。

2.3 漫画生成の試み

漫画の自動生成の試みをするにあたり、様々な手法を検討した。昨今では様々な画像生成 AI が存在するが、その中で、漫画生成の手法として最初に検討したのが、DF-GAN[3]である。DF-GAN は M. Tao らが提案したテキストから画像を生成する画像生成 AI の GAN の一種である。DF-GAN は、GAN の生成過程で生じる損失を抑え、より高品質な画像を生成することが可能な画像生成 AI である。

DF-GAN と、その識別器を利用することで、漫画を作る手法を検討した(図 2)。図 2 は文章 y から一部の文章を取り出したストック x より画像を生成する手法である。まず、文章 y の最初の文を文章 x に追加する。次に、文章 x を画像生成機に入力し、画像 g_1 を生成する。その後、識別機を用いて、画像 g_1 が文章 x から正しく生成されているかを判別する。正しく生成されていた場合、文章 y の最初の文を削除し、画像 g_1 を画像 g_2 に保存する。続いて、再び文章 y の次の 1 文を x に追加し、同様に画像 g_1 を生成する。このプロセスを繰り返し、画像 g_1 が正しく生成された場合は、文章 y の追加した文を削除し、画像 g_1 を g_2 に保存しながら進めていく。もし画像生成が失敗した場合、文章 x をリセットし、これまで保存した画像 g_2 を漫画 C に追加する。そして、文章 y の最初の文に戻り、再び画像の生成を試みる。この手順を繰り返しながら、最終的に一連の画像を生成する。しかし、DF-GAN では望んだ画像は生成されなかった(図 3)。これらの画像は、日本昔話「桃太郎」の文章内の一文ごとに生成した画像の一例である。例えば、図 3 の 3 枚目の画像は、桃太郎と鬼が戦うシーンの文章を DF-GAN に入力して得られたものである。しかし、この画像には人のようなものは写っているものの、桃太郎のような少年や鬼の姿は確認できない。他の画像も同様に、入力した文章の特徴を適切に捉えられておらず、文章の内容を正しく表現できていない。このため、これらの画像から文章の内容を推測することは困難であり、漫画としての精度には達していないと言える。

この手法の結果から、文章をそのまま入力するだけでは画像の生成が難しいということが明らかとなった。また、文章全体を漫画で表すためには、文章全体を理解したうえで文章の処理が必要だとわかった。このことから、精度の高い画像生成 AI と文章を理解、処理することのできる大規模マルチモーダルモデルを利用することが必要だと確認できた。これ

を踏まえて、本研究では、高精度の画像を生成することできる Midjourney と、大規模マルチモーダルモデルであり、ユーザとチャット形式で処理を行うことのできる ChatGPT を用いて漫画の生成を試みた。

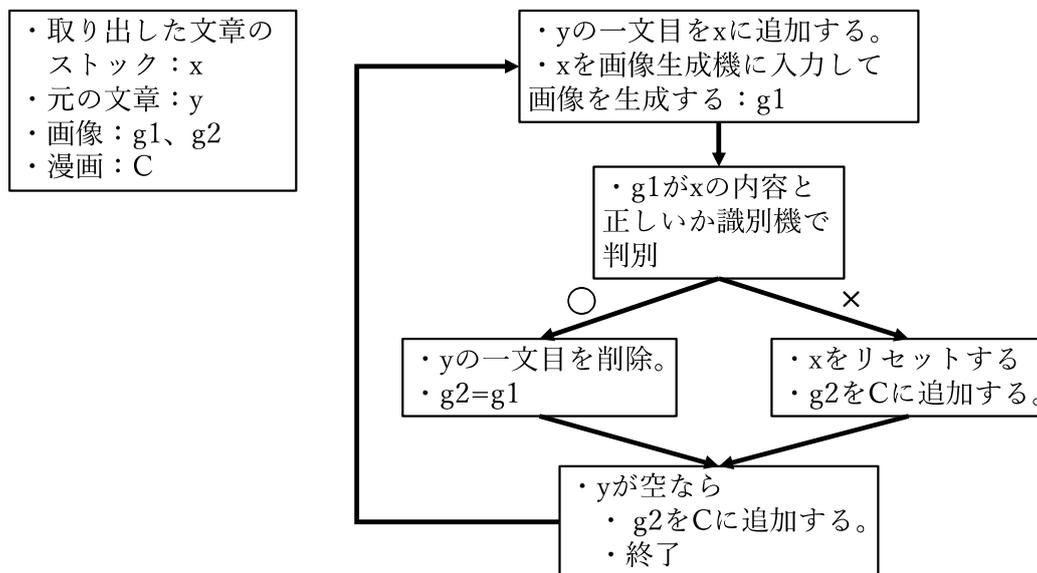


図2 DF-GAN を使用した漫画を作成する手法



図3 DF-GAN で生成した画像：

以下のプロンプトを入力して得られた画像。

以下のプロンプトは「桃太郎」を英訳し、文章の一部をそのまま入力したもの。

1. Long, long ago, in a small village, there lived an elderly man and woman.
2. She brought the peach home, and when she and the old man tried to cut it open, a healthy and cheerful boy emerged from inside.
3. They fought bravely and defeated the demons, restoring peace to the village.

第 3 章 提案手法

本研究では、大規模マルチモーダルモデルと画像生成 AI を用いて漫画の生成する手法を提案する(図 4)。大規模マルチモーダルモデルは ChatGPT、画像生成 AI には Midjourney を選定した。

まず、ChatGPT を用いて漫画となる画像を作成するためのプロンプトの作成を行う。漫画にする文章そのままを入力し、漫画を作る際に必要となる場面を選定させ、その場面を画像にするための各プロンプトを生成させる。次に生成されたプロンプトを Midjourney に入力し、漫画の画像を生成する。生成された複数の画像を ChatGPT に入力し、漫画に適している画像を選定する。また、生成がうまくいかない場合は、プロンプトの修正を ChatGPT にさせ、再度 Midjourney で画像の生成を行う。各場面のプロンプトを入力し、画像を生成することで、それぞれの画像を生成し、漫画を作成する。

また、上記の手法において、ChatGPT に与える指示は一定の形式に統一している(表 1)。これは誰でも漫画の場面選定やプロンプト生成を行える手法にするためである。さらに、同じ意味の指示であっても出力結果が異なる場合があるため、結果の一貫性を確保する目的も含まれている。

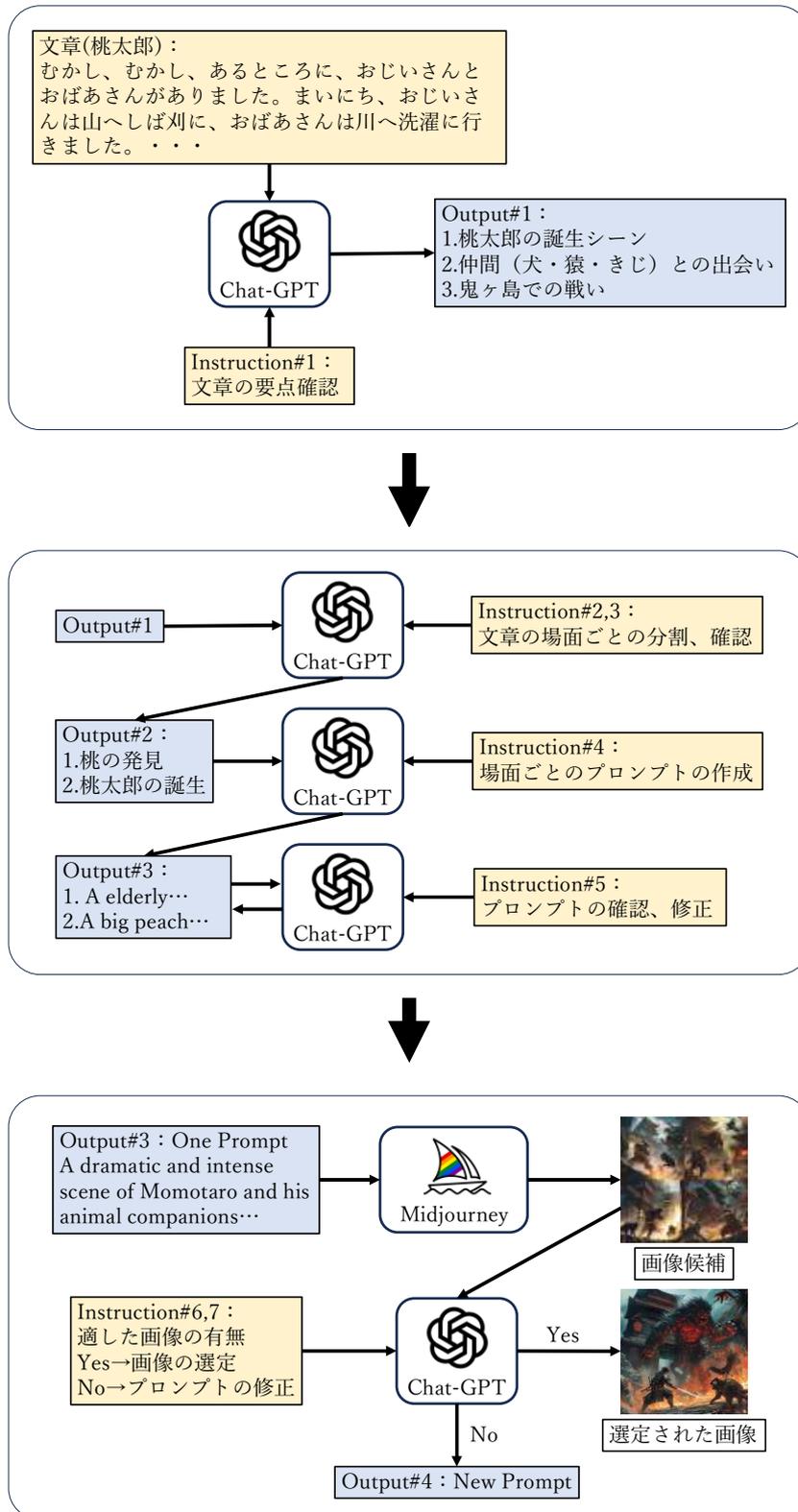


図4 提案手法のフロー

3.1 プロンプト作成

画像生成 AI を用いて画像を生成するにあたって、入力するプロンプトが必要となる。プロンプトの作成を、ChatGPT を用いてチャット形式で行う。まず、Chat GPT に漫画にする文章の見どころ、盛り上がる場面を出力させる (Instruction#1(昔話の「桃太郎」のあらすじです。この話の盛り上がる場面、見所はどこですか?))。この指示を与えることで、Output#1 のように文章の見どころが出力され、漫画にする際に重要な場面が抜けていることを防ぐことができる。また、文章を漫画にするメリットとして、より内容が明確となり、迫力や臨場感が増すことがあり、絵にすべき場面を把握することができる。それを踏まえた上で、ChatGPT に文章の場面を分けさせる (Instruction#2(では、それを参考に漫画を作成したい。漫画にすべき場面ごとにを分けてください。))。Instruction#1 の指示を出さずに場面分けをさせると、漫画として必要な場面が不足する場合があるため必要となる。Instruction#2 の指示を与えると、Output#2 のように分割した場面が出力される。次に、ChatGPT で場面分けを完了した後、それらの場面が漫画作成に適しているかを確認させる (Instruction#3(今場面分けしたもので十分ですか? 入力した文章全文を参考に確認して。))。場面分けが完了したら、その場面を表した画像を作成するために、ChatGPT に Midjourney に入力するプロンプトを生成させる (Instruction#4(場面を表す画像を Midjourney で作成したい。画像を生成するためのプロンプトを生成して。))。プロンプトは、背景、登場人物、表情、スタイルなど画像の詳細を記している。Instruction#4 の指示を与えると図 4 の Output#3 のように場面ごとのプロンプトが生成される。また、表 2 の Output#3 : Prompt は生成されたプロンプトの一例で、「桃太郎」の鬼ヶ島で桃太郎、犬、猿、雉が鬼と戦う場面のプロンプトである。このプロンプトは激しい戦いのシーンが生成されるように、一文目で記されている。また、鬼ヶ島の特徴や、桃太郎の格好などの細かい情報が指定されている。次に、ChatGPT にプロンプトが文章の内容と誤りがないかを確認させる (Instruction#5(プロンプトが文章の内容と誤りがないか確認してください。誤りがあれば、プロンプトを修正してください。))。また、誤りがあった場合は、プロンプトを修正させる。Midjourney にプロンプトを入力する前にプロンプトの確認、修正をすることで画像を文章の内容に適したものにすることができる。

3.2 画像生成

3.1 で生成されたプロンプトから、Midjourney を使用して漫画となる画像を生成する。Midjourney にプロンプトを入力すると、4 つの候補の画像が表示される (図 5)。この 4 つの画像は、表 2 の Output#3 : Prompt を Midjourney に入力して得られたものであり、プロン

プロンプトの詳細を満たすように生成されており、構図や細かい点が異なっている。この生成された画像を ChatGPT に入力し、生成された 4 つの画像の中から文章に最も適した画像を選定させる (Instruction#6 (これはプロンプトから作られた 4 つの画像です。この中で、最も場面の内容と近いものはどれですか？どれも適していない場合はないと教えてください。))。また、どの画像も適していない場合はそうであると出力させる。出力として、4 つの中で最も適した画像か、どれも適していないという返答が出る。4 つの画像の中に、適した画像がなければ、同じプロンプトで生成を繰り返す。Midjourney には画像を再生成する機能があり、同じプロンプトで異なった画像をもう一度生成することができる。また、他の機能として、生成された 4 つの画像のうち、1 つがユーザの求めている画像に近い場合、その画像とプロンプトから再度画像を生成することができる。例えば、画像の構図などは適しているが、納得のいかない画像の場合、このように画像を再生成することができる (図 6)。図 6 は図 5 の右下の画像を参考に、同じプロンプトで生成された画像の候補であり、図 7 の右下の画像に似ていることが見て取れる。そして、再生成された 4 つの画像と Instruction#6 を再度入力する。この方法を、適した画像が生成されるまで繰り返す。しかし、何度も同じプロンプトを入力して画像を生成しても、適しているものとして選定されない場合、ChatGPT にプロンプトの修正をさせる (Instruction#7 (このプロンプトでは、適した画像が生成されません。プロンプトを修正してください。))。プロンプトの修正ができたなら、そのプロンプトを Midjourney に入力して画像を生成させ、ChatGPT に画像の選定をさせる。

以上のように、ChatGPT と Midjourney を用いて各場面を表す画像を生成、選定する。全ての場面の画像が決定した時、それを順番ごとに並べることで、漫画を作成する。



図 5

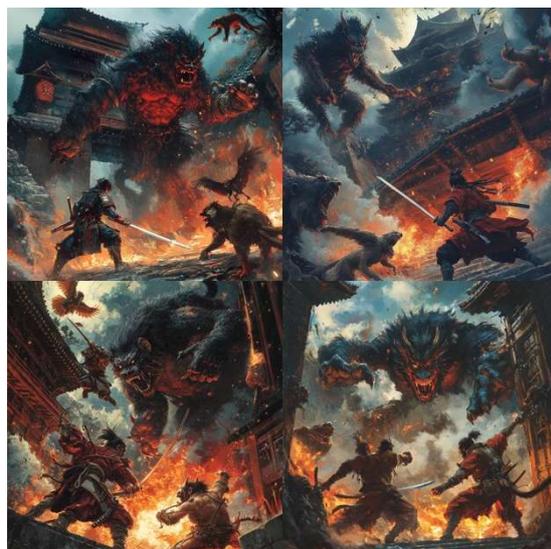


図 6

図 5 : Output#3 : Prompt を入力した画像候補

図 6 : 図 5 右下を元に再生成した画像候補

Instruction	ChatGPT に与える指示
Instruction#1 (文章の要点確認)	昔話の「桃太郎」のあらすじです。この話の盛り上がる場面、見所はどこですか？
Instruction#2 (文章の分割)	では、それを参考に漫画を作成したい。漫画にすべき場面ごとにを分けてください。
Instruction#3 (分割した場面の確認)	今場面分けしたもので十分ですか？入力した文章全文を参考に確認して。
Instruction#4 (プロンプトの作成)	場面を表す画像を Midjourney で作成したい。画像を生成するためのプロンプトを生成して。
Instruction#5 (プロンプトの確認)	プロンプトが文章の内容と誤りがないか確認してください。誤りがあれば、プロンプトを修正してください。
Instruction#6 (画像の確認)	これはプロンプトから作られた4つの画像です。この中で、最も場面の内容と近いものはどれですか？どれも適していない場合はないと教えてください。
Instruction#7 (プロンプトの修正)	このプロンプトでは、適した画像が生成されません。プロンプトを修正してください。

表1 プロンプトを作成するために ChatGPT に出す指示

Output#3 : Prompt(桃太郎と鬼が戦うシーン)
A dramatic and intense scene of Momotaro and his animal companions battling oni (demons) at the gates of a dark, iron fortress on Onigashima. Momotaro, wielding a gleaming sword, leads the charge. The dog bites a demon's leg, the monkey climbs a wall, and the pheasant dives, attacking demons from above. The oni are massive and fearsome, with muscular bodies, glowing red and blue skin, sharp fangs, and wild, tangled hair. They have menacing horns on their heads, and their eyes burn with fiery orange light. Each demon wields a massive spiked club and wears torn, animal-skin clothing. The background is fiery and chaotic, with smoke, sparks, and a stormy sky. The scene is highly detailed, action-packed, and inspired by Japanese folklore, with vivid colors and dynamic lighting.

表2 ChatGPT で作成したプロンプトの例

第4章 結果

4.1 作成結果

提案手法によって漫画を作成できることを確認するため、「桃太郎」の文章から漫画を作成した。「桃太郎」の文章は青空文庫[22]から引用した。まずは提案手法の ChatGPT の出力結果を評価した。ChatGPT に文章を入力し、Instruction#1 の指示を与え、その出力結果を得た(表3)。表3のように「桃太郎」の見どころを4つ箇条書きで出力されており、作品の見どころを抑え、文章の理解をしていることがわかる。次に、Instruction#2,3 の指示を与え、文章の場面を分けた出力結果を得た(表4)。この出力結果では、「桃太郎」を10場面に分けられた。これらの場面は、過不足なく文章を全体的に網羅できていた。次に、Instruction#3,4 を入力して得られたプロンプトを Midjourney に入力し、画像の出力結果を得た。図7左図は生成された画像の候補と ChatGPT が選定した画像である。この画像と Instruction#6 を ChatGPT に入力すると、図7右図が選定された。ChatGPT は画像の内容を分析し、場面との一致度を評価し、画像の選定を行った。表5は図7左図の評価結果である。入力は1つの画像であるが、4つの画像を識別し、評価している。画像と ChatGPT が評価した内容は一致しており、評価も正しいものとも言える。また、図7右図は場面と一致している。

次に作成した漫画図8は提案手法を使用して作成した「桃太郎」の漫画である。表4はページごとの場面である。作成した漫画の場面と画像の一貫性を図るため、それぞれの場面と画像を比較する。全体の総評としては、場面と画像は大きな相違は見られなかった。場面と画像を比較した際、少しだけ相違が見られた。例えば、6~9p で犬、猿、キジの3匹の一部がないことが挙げられる。しかし、分けられた場面の文章と画像を比較した際、文章が理解できる程度のため、漫画として十分な品質である。

<p>1. 桃太郎の誕生シーン</p> <p>おばあさんが川で拾った桃から赤ん坊（桃太郎）が生まれる場面は、物語の導入であり、驚きと喜びに満ちた象徴的なシーンです。</p> <p>見所：「桃が割れて子供が出てくる」という不思議な現象が、昔話特有のファンタジーを感じさせます。</p>
<p>2. 仲間（犬・猿・きじ）との出会い</p> <p>桃太郎が道中で動物たち（犬・猿・きじ）と出会い、きびだんごを与えて仲間にする場面は、物語の展開において重要です。</p> <p>見所：それぞれの動物が桃太郎の仲間になることで、チームワークや絆の要素が強調されます。</p>
<p>3. 鬼ヶ島での戦い</p> <p>クライマックスはやはり、桃太郎と家来たちが鬼ヶ島に乗り込み、鬼たちと戦うシーンです。鬼たちが強そうに見えながらも、犬・猿・きじがそれぞれの力を発揮し、桃太郎が鬼の大将を倒すところは見所です。</p> <p>見所：勇敢に鬼を征伐し、宝物を取り戻すことで「正義が勝つ」という爽快感が得られます。</p>
<p>4. 帰還とハッピーエンド</p> <p>宝物を持ち帰り、桃太郎と家来たちが凱旋し、おじいさんとおばあさんが喜ぶシーンは物語の締めくくりとして心温まる部分です。</p> <p>見所：「平和の回復」や「恩返し」の要素が描かれており、家族の幸福が表現されています。</p>

表3 Instruction#1 の出力結果(Outout#1)

1p	桃の発見と家への持ち帰り： おばあさんが川で大きな桃を見つけ、家に持ち帰る場面。
2p	桃太郎の誕生： おじいさんとおばあさんが桃を割ると、中から赤ちゃん(桃太郎)が現れる場面。
3p	桃太郎の成長： 桃太郎が成長し、村一番の力持ちになる場面。
4p	出発の準備： おばあさんがきびだんごを作り、桃太郎が出発の挨拶をする場面。
5p	仲間(犬、猿、キジ)との出会い： 犬、猿、キジが桃太郎に忠誠を誓う場面。
6p	鬼ヶ島への上陸： 鬼ヶ島に到着し、仲間と進軍する場面。
7p	鬼との戦い： 桃太郎と仲間たちが鬼たちと戦う場面。
8p	鬼の降参と宝物の回収： 鬼の大将が降参し、宝物を差し出す場面。
9p	凱旋帰郷： 桃太郎と仲間が宝物を持って帰る場面。
10p	幸せなエンディング： おじいさん、おばあさんが桃太郎の帰還を喜ぶ場面。

表4 「桃太郎」の場面の分割結果(Output#2”)

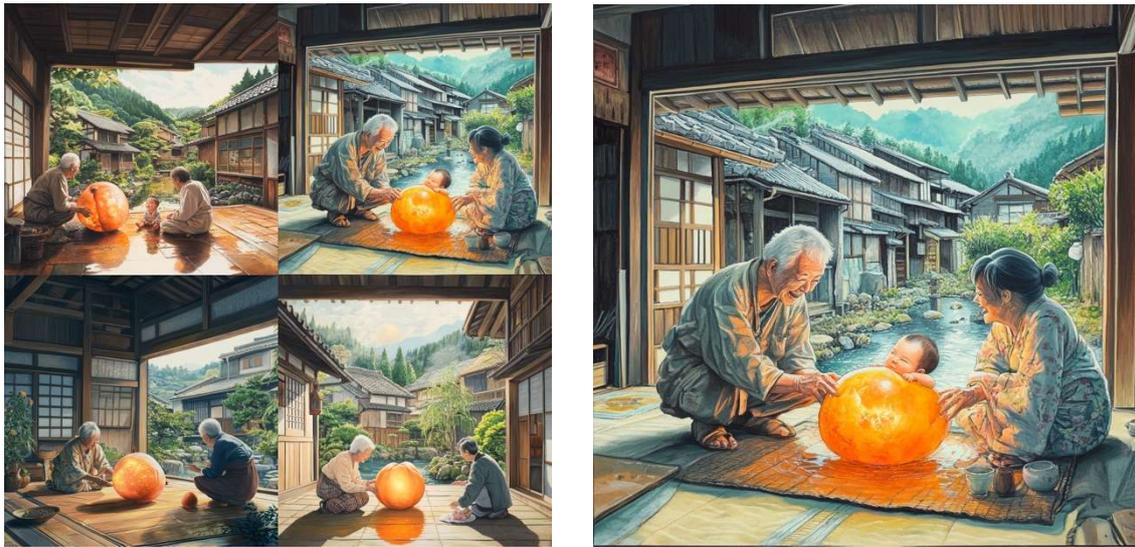


図7 生成された画像(左図)と ChatGPT によって選定された画像(右図)。
桃太郎が桃から生まれる場面 2p の画像

画像	評価内容
左上の画像	<p>内容: おじいさんとおばあさんが家の中で桃を見つめている様子。桃から赤ちゃんが現れている描写も含まれています。</p> <p>一致度: 原作の「桃を割ると赤ちゃんが出てくる」場面に非常に近いです。</p> <p>評価: 適している。</p>
右上の画像	<p>内容: おじいさんとおばあさんが家の中で桃を割り、赤ちゃんを取り出しているシーン。</p> <p>一致度: 桃太郎の誕生の瞬間がよく表現されています。</p> <p>評価: 非常に適している。</p>
左下の画像	<p>内容: おじいさんとおばあさんが桃を見つめているだけで、赤ちゃんが出てくる描写はありません。</p> <p>一致度: 重要な要素(赤ちゃんの誕生)が欠けています。</p> <p>評価: 適していない。</p>
右下の画像	<p>内容: おじいさんとおばあさんが庭先で桃を眺めている様子。赤ちゃんの誕生シーンは含まれていません。</p> <p>一致度: 赤ちゃんが出てこないため、物語との一致度が低いです。</p> <p>評価: 適していない。</p>

表5 ChatGPT による図7左図の評価の出力結果



図8 提案手法を使用して作成した漫画「桃太郎」

比較的簡単な文章である昔話の漫画化が可能であることを確認したため、より難易度の高い小説の文章を対象に漫画を作成した(図9)。図9は、短編小説「走れメロス」を提案手法を用いて作成した漫画である。「走れメロス」は、「桃太郎」などの昔話と比較して、人間の心理描写が多く含まれるため、より難易度の高い文章といえる。図9の漫画では、分割した場面と生成された画像の間に大きな相違はなく、文章との一貫性も保たれている。また、登場人物の表情なども適切に描写されている。これらの結果から、難易度の高い文章であっても、提案手法を用いることで漫画化が可能であることを確認した。



図9 提案手法で作成した「走れメロス」
 以下は図9の漫画の各ページの場面である。

1p	メロスの村からの出発： 背景説明として、妹の結婚準備とシラクス市への出発、そして市の異常さに気づく場面。
2p	メロスの決意と王との対峙： 暴君ディオニスの言葉とメロスの対決。信実と正義のテーマが描かれる場面。
3p	セリヌンティウスとの友情： セリヌンティウスが身代わりを引き受け、二人が抱き合う場面。
4p	妹の結婚式と未練： 村に戻り、妹の結婚式を強行するメロスの人間らしい迷いや未練が描かれる場面。 プロンプト：
5p	メロスの苦難の旅路（豪雨の川を渡る）：

	川の濁流との戦い。勇気と決意が際立つ場面。
6p	メロスの苦難の旅路（山賊との戦い）： 山賊との対峙。アクションと緊張感のある場面。
7p	メロスの挫折と再起： 体力の限界に達したメロスが泉の清水で再起する場面。内面的な葛藤が重要。
8p	メロスの刑場への到着： 刑場に間一髪で到着し、友を救う感動のクライマックスの場面。
9p	セリヌンティウスとの友情の確認 殴り合いを経て抱き合う二人の場面。
10p	暴君ディオニスの改心： 暴君ディオニスが信じる心を取り戻し、群衆の歓声の中で和解が描かれる。

4.2 比較評価

4.2.1 提案手法と ChatGPT のみの比較

提案手法を評価するため、既存手法を使用して作成した漫画と比較する。漫画を生成する先行研究がないため、ChatGPT を比較対象とする図 10 は ChatGPT に「桃太郎」の文章を入力し、漫画を作成するように命令した出力結果である。「桃太郎」の文章を ChatGPT に入力し、漫画を作成するように命令すると、漫画の構成が出力された(表 5)。次に、ChatGPT が出力した構成を元に漫画を生成するように命令した。この漫画は 5 ページで構成されており、それぞれのページで画像が生成された。これらの画像は、漫画のようにコマ割りが施されている。例えば、図 10 の 1 ページ目の画像は、「桃太郎」の冒頭からお婆さんが桃を拾うまでの場面を表しており、「お爺さんが芝刈りをしているコマ」、「お婆さんが川で洗濯をしているコマ」、「桃が流れてくるコマ」などが含まれている。しかし、「自転車が写っているコマ」もあり、これは「桃太郎」の文章には存在しない要素である。他の図 5 の画像も同様に、複数のコマで 1 つの場面を表しているが、原文の「桃太郎」とは異なる内容が含まれている。また、一部のコマでは何を表しているのかが不鮮明であり、情報を正しく伝える漫画として十分に機能していない。

提案手法によって作成された漫画の品質を評価するため、被験者を対象に評価実験を実施した(表 6)。被験者には漫画を読んでもらい、「一貫性・整合性」と「完全性・包括性」の観点から 10 点満点で評価を行った。「一貫性・整合性」とは、漫画の内容が一貫しているか、また漫画と文章の内容に違いがないかを指す。漫画の内容の一貫性とは、登場するキャラクターの見た目、背景、画風などが統一されているかどうかを意味する。「完全性・包括性」は、漫画が文章の内容を漏れなく、広く全体的に反映しているかを評価するものである。

評価は 20 人の被験者によって行われ、提案手法で作成した漫画と、ChatGPT を使用して作成した漫画を対象とした。表 6 からは、提案手法の漫画は高い評価を受けたことがわかる。また、既存手法で生成した漫画は、低い評価となり、漫画としては不十分な品質であることがわかる。



図 10 ChatGPT で作成した漫画「桃太郎」

	提案手法	ChatGPT のみ
一貫性・整合性	7.5	4.55
完全性・包括性	7.3	4.7

表 6 生成した漫画の評価実験

4.2.2 Midjourney を使った場合との比較

提案手法の有効性を検証するため、提案手法を使用しない場合との比較を行った。漫画化の対象作品として、短編小説「赤い蠟燭と人魚」を選定した。比較の手法として、被験者に文章を一読させた後、文章を基に漫画を作成してもらう方法を採用した。漫画の作成には Midjourney のみを使用し、プロンプトは英語で入力する必要があるため、翻訳アプリの使用を認めた。被験者は、短編小説「赤い蠟燭と人魚」を読んだ経験がなく、かつ Midjourney を使用したことがない人を対象とした。実験前に、被験者に対して Midjourney の使用方法や基本的な操作に関する説明を行った。Midjourney での作成は、作成時間のばらつきを考

慮するため、3人に漫画の作成をさせた。作成にあたって、漫画の枚数や制限時間は設けず、作成に要した時間および漫画の品質を提案手法と比較した。評価は5人の被験者によって行われ、提案手法で作成した漫画と、Midjourneyのみ使用して作成した漫画を対象とした。

表7は提案手法と使わなかった手法で漫画を作成した際の比較結果であり、作成に要した時間と漫画の品質を示している。また、Midjourneyのみ使用した際の作成所要時間は最短で作成した結果のものである。図11は提案手法によって作成した漫画、図12はMidjourneyだけを使用して最短時間で作成した被験者(以下、被験者A、他の被験者B、C、D、E)の漫画である。作成した漫画は提案手法が8枚、Midjourneyで作成した漫画は13枚と枚数が異なった。他の被験者の作成した漫画も8枚以上の画像であった。作成品質を評価する項目は、4.2.1と同じである。作成に要した時間に関しては、提案手法の方が約2倍以上速く作成できた。この要因として、Midjourneyのみで漫画を作成する場合、文章の理解や処理を最初から行う必要があることが挙げられる。また、被験者がMidjourneyの使用経験がなく、プロンプトの作成が困難であったことも影響していると考えられる。これらの点から、提案手法はプロンプトを高速で作成することができ、プロンプトの作成は難易度の高い作業であるといえる。一方で、漫画の品質に関しては、Midjourneyのみで作成した漫画の方が提案手法よりも優れた結果となった。両手法の比較では、「一貫性・整合性」の観点では大きな差は見られず、どちらも概ね良好な評価を得た。しかし、「完全性・包括性」の観点では、提案手法の評価がMidjourneyのみ使用した場合よりも低い結果となった。人が文章を理解し処理した上で場面を分割した方が、より適切な漫画化ができていていることが示された。

以上の結果から、提案手法のように大規模マルチモーダルモデルを活用することで、人間と同等の優れたプロンプトを作成できることが判明した。一方で、文章の場面分割に関しては、人による分割の方が優れていることが分かった。

	提案手法	Midjourneyのみ
作成所要時間	26:59	1:00:52
一貫性・整合性	6.4	6.8
完全性・包括性	6.2	8.2

表7 「赤い蠟燭と人魚」の漫画の作成時間と品質の比較



図 11 提案手法で作成した漫画「赤い蠟燭と人魚」

以下は図 11 の漫画の各ページの場面である。

1p	北の海での人魚の孤独： 北の海で人魚が岩の上に座り、孤独を感じつつも子供を人間に託す決意をする場面。
2p	老夫婦が赤ん坊を見つける： 老婆が神社の石段で赤ん坊を見つけ、家に連れ帰る場面。
3p	娘が育ち、才能を見せる： 人魚の娘が成長し、美しい絵を蠟燭に描き始め、それが町で評判になる場面。
4p	老夫婦の裏切り： 香具師が老夫婦に娘を売るよう持ち掛け、老夫婦が迷いながらも金に目がくらみ、娘を売ることを決める場面。
5p	娘の悲しみと赤い蠟燭： 娘が自らの運命を嘆きつつ、赤い蠟燭に自分の感情を込めるように塗り残す場面。
6p	香具師が娘を連れ去る 内容: 香具師が娘を檻に入れて連れ去る場面。娘が無力感を抱えながらも、最後に赤い蠟燭を残す。
7p	赤い蠟燭が災いの象徴となる： 赤い蠟燭が神社に灯されると同時に嵐が起り、船や町に災難をもたらす場面。

8p	廃墟となった町と神社： 廃墟となった町と神社に灯る赤い蠟燭が、娘の復讐や祈りの象徴として燃え続ける場面。
----	---

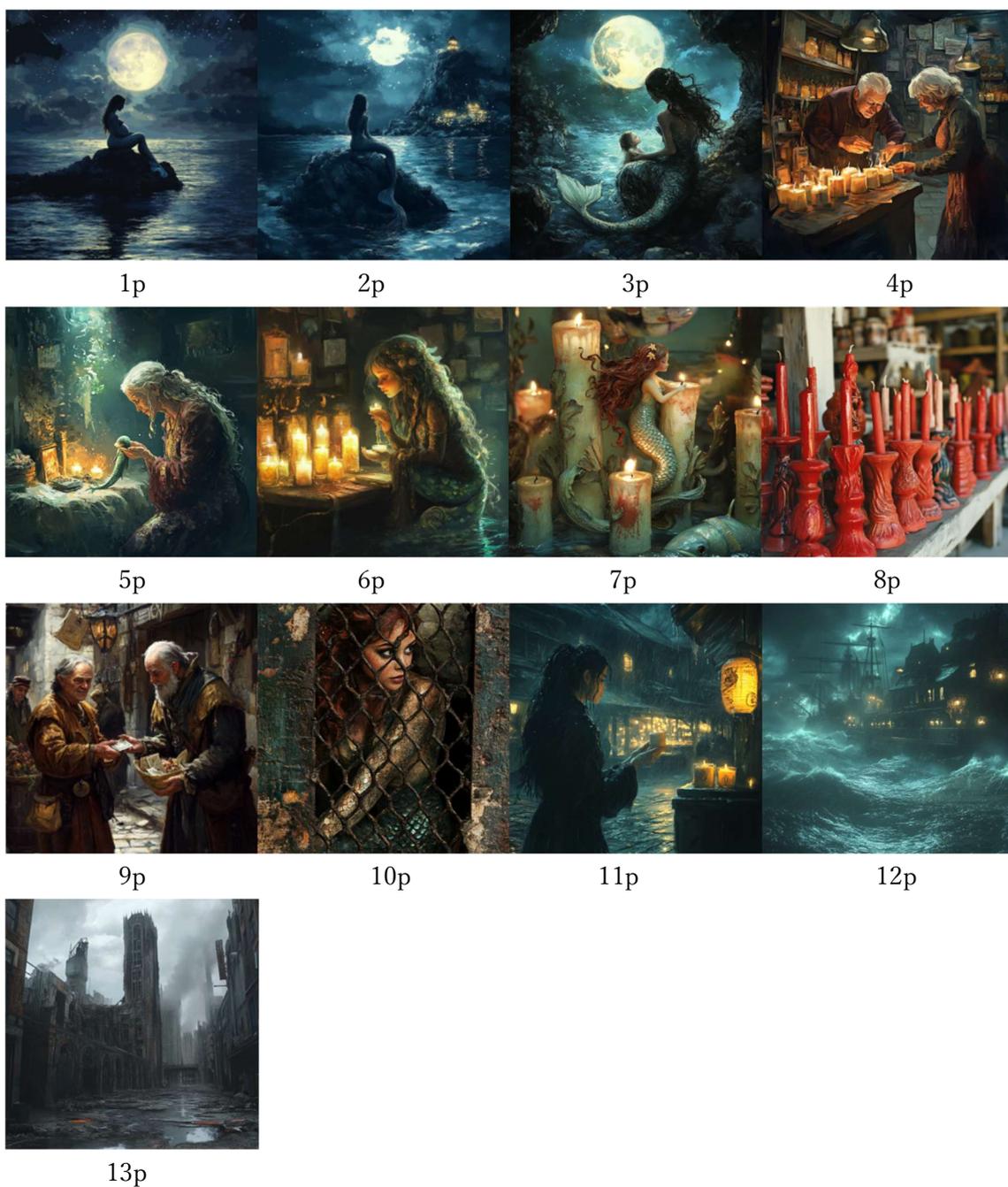


図 12 被験者 A が Midjourney のみで作成した漫画「赤い蝋燭と人魚」
 以下は被験者 A が考案した図 12 の漫画の各ページの場面である。

1p	孕った人魚の母親が登場する場面
2p	人魚の母親が子供をどうするか考える場面
3p	人魚の母親が子供を置いていく場面
4p	お爺さんとお婆さんが蝋燭を売る場面

5p	お婆さんが人魚の赤ん坊を拾う場面
6p	人魚の赤ん坊が育ち、美しい人魚になる場面
7p	人魚が蠟燭に絵を描く場面
8p	蠟燭が大繁盛する場面
9p	お爺さんとお婆さんが香具師に人魚の娘を売る話をされる場面
10p	人魚の娘が檻に閉じ込められ、香具師に売られる場面
11p	人魚の母親が蠟燭屋に行く場面
12p	嵐で船が転覆する場面
13p	町が廃れ、滅ぶ場面

第5章 まとめと今後の展望

本研究では、生成 AI を活用することで、誰でも文章から漫画を作成できる手法を提案した。また、提案手法の一環として、画像生成に必要なプロンプトを大規模マルチモーダルモデルを用いて作成する手法を導入した。その結果、文脈的に一貫性のある漫画を安定して生成できることが確認された。文章の場面に即したプロンプトの生成は可能である一方で、場面の分割などの処理に関しては人の判断に比べて精度が劣ることが明らかとなった。

一方で、作成された漫画は、大規模マルチモーダルモデルと画像生成 AI の出力結果に依存するため、すべての試行において完全に同一の作品が生成されるとは限らない。今後の課題として、同じ文章の入力に対して一貫した漫画を安定的に生成できる手法の開発が挙げられる。

また、本研究で生成された漫画は、一般的な漫画とは異なり、コマ割りが施されず、1 ページに 1 つの絵が配置される形式であった。今後の展望として、一般的な漫画のようにコマ割りを取り入れた漫画を自動生成する手法の開発が求められる。さらに、セリフの吹き出しの追加は手作業で行われており、負担の大きい作業となっている。そのため、コマ割りの自動構成やセリフの吹き出しの自動配置を実現することを今後の目標とする。最終的には、小説や伝記などのコミカライズを自動で作成できる技術の確立を目指す。

謝辞

本研究及び論文の作成にあたり、研究の着想や論文執筆等、多くのご指導、ご助言を頂きました静岡大学工学部の岡部誠准教授に心から感謝申し上げます。また、ご助力頂いた修士課程学生及び学部生の皆様に深く感謝致します。

参考文献

- [1] Y. Li, H. Shi, B. hu, L. Wang, J. Zhu, J. Xu, Z. Zhao, M. Zhang, “Anim-Director: A Large Multimodal Model Powered Agent for Controllable Animation Video Generation”, In SIGGRAPH Asia 2024.
- [2] Z. Zhang, D. Chen, J. Liao, “SGEdit: Bridging LLM with Text2Image Generative Model for Scene Graph-based Image Editing”, In SIGGRAPH Asia 2024.
- [3] M. Tao, H. Tang, F. Wu, X. Jing, B. Bao, C. Xu, “DF-GAN: A Simple and Effective Baseline for Text-to-Image Synthesis”, In CVPR 2022.
- [4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, “Generative Adversarial Nets”, In CVPR 2014.
- [5] J. Ho, A. Jain, P. Abbeel, “Denoising Diffusion Probabilistic Models”, In arXiv:2006.11239.
- [6] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models” In CVPR 2022.
- [7] S. Zheng, J. Huo, Y. Wang, Y. Fu, “Intelligent Director: An Automatic Framework for Dynamic Visual Composition using ChatGPT”, In arXiv:2402.15746.
- [8] J. Yao, Y. Liu, Z. Dong, M. Guo, H. Hu, K. Keutzer, L. Du, D. Zhou, S. Zhang, “PromptCoT: Align Prompt Distribution via Adapted Chain-of-Thought” In CVPR 2024.
- [9] X. Li, Z. Zhang, X. Tan, C. Chen, Y. Qu, Y. Xie, Li. Ma, “PromptAD: Learning Prompts with only Normal Samples for Few-Shot Anomaly Detection” In CVPR 2024.
- [10] H. Liu, Z. Wan, W. Huang, Y. Song, X. Han, J. Liao, “PD-GAN: Probabilistic Diverse GAN for Image Inpainting”, In CVPR 2021.
- [11] P. Dhariwal, A. Nichol, “Diffusion Models Beat GANs on Image Synthesis”, In arXiv:2105.05233.

- [12] C. Wang, W. Luo, Q. Chen, H. Mai, J. Guo, S. Dong, X. Xuan, Z. Li, L. Ma, S. Gao, “MLLM-Tool: A Multimodal Large Language Model For Tool Agent Learning” In arXiv:2401.10727.
- [13] M. Zhang, Z. Wang, L. Chen, K. Liu, J. Lin, “Dialogue Director: Bridging the Gap in Dialogue Visualization for Multimodal Storytelling” In arXiv:2412.20725.
- [14] J. Bai, S. Bai, S. Yang, S. Wang, S. Tan, P. Wang, J. Lin, C. Zhou, J. Zhou, “Qwen-VL: A Versatile Vision-Language Model for Understanding, Localization, Text Reading, and Beyond” In arXiv:2308.12966.
- [15] N. Kumari, B. Zhang, R. Zhang, E. Shechtman, J. Zhu, “Multi-Concept Customization of Text-to-Image Diffusion” In CVPR 2023.
- [16] Y. Li, Z. Gan, Y. Shen, J. Liu, Y. Cheng, Y. Wu, L. Carin, D. Carlson, J. Gao, “StoryGAN: A Sequential Conditional GAN for Story Visualization” In CVPR 2019.
- [17] D. Zhu, J. Chen, X. Shen, X. Li, M. Elhoseiny, “MiniGPT-4: Enhancing Vision-Language Understanding with Advanced Large Language Models” In ICLP 2024.
- [18] T. Xu, P. Zhang, Q. Huang, H. Zhang, Z. Gan, X. Huang, X. He, “AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks” In CVPR 2019.
- [19] Z. Yang, J. Wang, Z. Gan, L. Li, K. Lin, C. Wu, N. Duan, Z. Liu, C. Liu, M. Zeng, L. Wang, “ReCo: Region-Controlled Text-to-Image Generation” In CVPR 2023.
- [20] S. Ge, T. Park, J. Zhu, J. Huang, “Expressive Text-to-Image Generation with Rich Text” In ICCV 2023.
- [21] J. Zhu, T. Park, P. Isola, A. Efros, “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks” In CVPR 2017.
- [22] 青空文庫, <https://www.aozora.gr.jp/>.

付録

1p	<p>桃の発見と家への持ち帰り：</p> <p>"A traditional Japanese countryside river scene with an elderly woman washing clothes by the water. She is clapping her hands and singing to a giant peach floating gently downstream toward her, surrounded by lush green trees and sparkling water under a sunny sky."</p>
2p	<p>桃太郎の誕生：</p> <p>"Inside a traditional Japanese house, an elderly couple joyfully reacts as they open a giant peach, revealing a baby boy glowing with a mystical aura. The room is warmly lit with wooden beams and a tatami floor."</p>
3p	<p>桃太郎の成長：</p> <p>"A strong young boy standing confidently in a Japanese village, smiling kindly as he helps an elderly couple carry a heavy load. Villagers watch in amazement, highlighting his strength and kind-hearted nature."</p>
4p	<p>出発の準備：</p> <p>"Inside a traditional Japanese kitchen, an elderly couple work together to prepare millet dumplings using a large mortar and pestle. The young boy dressed in samurai-style travel attire watches excitedly, holding a small travel bag."</p>
5p	<p>仲間(犬、猿、キジ)との出会い：</p> <p>"A young Japanese warrior with a traditional samurai look standing on a forest path, surrounded by three animals: a loyal dog, a mischievous monkey, and a vibrant pheasant, each bowing or showing allegiance to him. The forest is lush and alive with natural light."</p>
6p	<p>鬼ヶ島への上陸：</p> <p>"A dramatic view of a rocky island fortress with a dark iron gate. A young samurai, accompanied by a dog, monkey, and pheasant, confidently steps off a boat onto the shore under a cloudy sky, preparing for battle."</p>
7p	<p>鬼との戦い：</p> <p>"A chaotic battle at a dark fortress. A brave young samurai fights a massive red oni with a spiked club, while a dog bites another oni's leg, a monkey scratches its face, and a pheasant swoops in to peck at their eyes. Sparks and debris fill the dramatic scene."</p>
8p	<p>鬼の降参と宝物の回収：</p> <p>"Inside a dimly lit castle with iron walls, a massive red oni bows humbly, tears</p>

	streaming down its face, as a victorious young samurai pins it down. Nearby, treasure chests filled with gold, jewels, and mystical artifacts await collection by the dog, monkey, and pheasant."
9p	凱旋帰郷： "A cheerful procession through a Japanese village. A young samurai leads the way, followed by a dog pulling a cart filled with treasures, a monkey pushing from behind, and a pheasant flying above. Villagers cheer and wave, celebrating their hero."
10p	幸せなエンディング： "A warm and joyful reunion in a traditional Japanese countryside house. An elderly couple hugs a young samurai, while the dog, monkey, and pheasant play happily nearby. Cherry blossoms bloom in the garden, with petals gently falling under a clear blue sky."

表8 「桃太郎」の提案手法のプロンプトの生成結果

これは、4.1 作成結果の提案手法を用いて「桃太郎」の漫画を作成する過程で生成されたプロンプトである。提案手法によって分割された各場面に対し、その場면을画像化するためのプロンプトを生成し、これらを Midjourney に入力して画像を生成した。その結果、図8に示す漫画が作成された。

1p	<p>メロスの村からの出発： 背景説明として、妹の結婚準備とシラクス市への出発、そして市の異常さに気づく場面。 プロンプト： "A determined man wearing a simple tunic departs alone from a peaceful rural village at dawn, carrying a small bundle for his sister's wedding. The serene countryside is dotted with grazing sheep, and the man strides toward the distant city with a serious expression. Vibrant and hopeful atmosphere, medieval Greek style."</p>
2p	<p>メロスの決意と王との対峙： 暴君ディオニスの言葉とメロスの対決。信実と正義のテーマが描かれる場面。 プロンプト： "A brave man with rugged features, wearing a simple tunic, stands defiantly in front of a pale-faced king with a crown and deep frown lines, inside a grand but ominously shadowed throne room. The atmosphere is tense, the king's face exudes mistrust and disdain, while the man's eyes burn with determination, holding a dagger tightly in his hand. Medieval Greek style, highly detailed, cinematic lighting."</p>
3p	<p>セリヌンティウスとの友情： セリヌンティウスが身代わりを引き受け、二人が抱き合う場面。 プロンプト： "Two men embracing tightly under the moonlit sky, one wearing a tunic with a determined expression, and the other with ropes tied around his hands, showing quiet acceptance. The background shows the walls of a Greek city. The mood is solemn yet deeply emotional, conveying trust and unspoken words of loyalty. Renaissance painting style, soft lighting, and detailed expressions."</p>
4p	<p>妹の結婚式と未練： 村に戻り、妹の結婚式を強行するメロスの間らしい迷いや未練が描かれる場面。 プロンプト： "A wedding celebration inside a modest village home, with joyful guests and a young bride in a simple yet elegant dress. The scene transitions to an outdoor view, where heavy rain begins to pour, soaking the decorations as guests huddle indoors, singing and clapping. A rugged man stands apart, his expression torn between pride for his sister and sorrow for his impending departure. Emotional and atmospheric."</p>
5p	<p>メロスの苦難の旅路（豪雨の川を渡る）： 川の濁流との戦い。勇気と決意が際立つ場面。 プロンプト： "A lone man kneels at the edge of a fierce, turbulent river under a stormy sky, raising</p>

	<p>his hands in desperate prayer before diving into the raging current. The water is dark and wild, with crashing waves and foam. Moments later, he is shown swimming with straining muscles and a determined expression, his soaked tunic clinging to his body. Dynamic and dramatic lighting, highly detailed water and storm effects."</p>
6p	<p>メロスの苦難の旅路（山賊との戦い）： 山賊との対峙。アクションと緊張感のある場面。 プロンプト： "A fierce battle in a rocky mountain pass, with a lone man in a tattered tunic fighting off several rugged bandits wielding clubs. The man wields a stolen club, shouting 'For justice!' as he strikes down his attackers. Dust and rocks fly in the air, the scene charged with action and tension. Cinematic composition, dynamic poses, and intense lighting."</p>
7p	<p>メロスの挫折と再起： 体力の限界に達したメロスが泉の清水で再起する場面。内面的な葛藤が重要。 プロンプト： "A weary man collapsed near a rocky spring, scooping fresh water with his hands and drinking it with relief. The background shows a glowing sunset, with light filtering through trees. The man's face reflects exhaustion transforming into hope, his wet hair clinging to his forehead. The scene is serene yet inspirational, with golden lighting and warm tones."</p>
8p	<p>メロスの刑場への到着： 刑場に間一髪で到着し、友を救う感動のクライマックスの場面。 プロンプト： "A determined man running through a large gathering at sunset, his body dusty and exhausted, arms raised in urgency as he pleads with the crowd. The onlookers, dressed in ancient Greek-style clothing, react with surprise and curiosity. In the background, a wooden structure stands silhouetted against a blazing orange sky, emphasizing the tension of the moment. Highly detailed, with dramatic lighting and expressive emotions, focusing on the man's desperate attempt to convey his message."</p>
9p	<p>セリヌンティウスとの友情の確認 殴り合いを経て抱き合う二人の場面。 プロンプト； "Two men in a dramatic embrace, their faces streaked with tears, under the watchful eyes of a silent crowd in an ancient Greek execution square. One man's cheek is red from a recent slap, while the other's hand is raised, ready to return the gesture. The</p>

	scene is deeply emotional, with warm lighting highlighting their bond and the surrounding crowd's awe. Renaissance painting style, detailed expressions."
10p	<p>暴君ディオニスの改心：</p> <p>暴君ディオニスが信じる心を取り戻し、群衆の歓声の中で和解が描かれる。</p> <p>プロンプト：</p> <p>"A humbled king in a royal robe, his crown slightly tilted, approaches two men standing side by side, both looking worn but victorious. The background shows a cheering crowd under the golden light of dawn. The king's expression is sincere, his hand outstretched as he seeks acceptance. The atmosphere is triumphant and emotional, with rich colors and intricate details."</p>

表9 「走れメロス」の場面の分割結果とプロンプトの生成結果

これは、4.1 作成結果における、提案手法を用いた「走れメロス」の場面分割結果とプロンプトの生成結果である。「走れメロス」の文章を入力すると、表9 のように場面が分割され、それぞれの場면을画像化するためのプロンプトが生成された。生成されたプロンプトをMidjourneyに入力し、漫画の画像が作成された。表9は図9の画像の各ページと対応しており、画像の場面と、その場면을生成する際にMidjourneyに入力したプロンプトを示している。そして、最終的に図9に示す漫画が完成した。

1p	<p>北の海での人魚の孤独：</p> <p>北の海で人魚が岩の上に座り、孤独を感じつつも子供を人間に託す決意をする場面。</p> <p>プロンプト：</p> <p>"A serene and melancholic scene of a mermaid sitting on a rocky shore under the moonlight, gazing at the vast blue ocean. The sky is cloudy, and the moonlight reflects gently on the water. The mermaid has long flowing hair and a shimmering fish tail, a sorrowful and determined expression on her face."</p>
2p	<p>老夫婦が赤ん坊を見つける：</p> <p>老婆が神社の石段で赤ん坊を見つけ、家に連れ帰る場面。</p> <p>プロンプト：</p> <p>"An elderly woman with kind, weathered features finds a baby wrapped in soft cloth at the bottom of stone shrine steps under a bright moonlit night. The shrine in the background is surrounded by pine trees, and the woman looks astonished and compassionate as she picks up the baby."</p>
3p	<p>娘が育ち、才能を見せる：</p> <p>人魚の娘が成長し、美しい絵を蠟燭に描き始め、それが町で評判になる場面。</p> <p>プロンプト：</p> <p>"A beautiful young girl with a human upper body and a shimmering fish tail paints intricate designs of fish, shells, and seaweed on white candles using red paint. The room is dimly lit, and candles with her artwork glow softly, revealing her delicate features and long black hair."</p>
4p	<p>老夫婦の裏切り：</p> <p>香具師が老夫婦に娘を売るよう持ち掛け、老夫婦が迷いながらも金に目がくらみ、娘を売ることを決める場面。</p> <p>プロンプト：</p> <p>"An intense indoor scene with an elderly couple and a sly merchant. The merchant, dressed in dark exotic clothing, offers a bag of gold coins on a wooden table, while the couple looks conflicted and hesitant. The room is dimly lit by candlelight, creating a tense and shadowy atmosphere."</p>
5p	<p>娘の悲しみと赤い蠟燭：</p> <p>娘が自らの運命を嘆きつつ、赤い蠟燭に自分の感情を込めるように塗り残す場面。</p> <p>プロンプト：</p> <p>"A sorrowful mermaid girl with long black hair paints a red candle in a dimly lit room. Tears roll down her cheeks as she paints, her fish tail curled on the wooden</p>

	floor. Nearby, candles with red and white designs glow faintly, emphasizing the melancholic mood."
6p	香具師が娘を連れ去る 内容: 香具師が娘を檻に入れて連れ去る場面。娘が無力感を抱えながらも、最後に赤い蠟燭を残す。 プロンプト : "A heart-wrenching scene of a mermaid girl being placed into a cage on a wooden cart. Her long hair flows as she looks back at the house tearfully, leaving behind a few red candles. The cart is surrounded by shadowy figures, and the atmosphere is tense and sorrowful."
7p	赤い蠟燭が災いの象徴となる : 赤い蠟燭が神社に灯されると同時に嵐が起こり、船や町に災難をもたらす場面。 プロンプト : "A haunting scene of a red candle glowing brightly in a shrine atop a hill, surrounded by stormy skies and violent winds. Below, the village appears dark and abandoned, with waves crashing against the cliffs. The red candlelight seems ominous, casting eerie shadows over the shrine and the landscape."
8p	廃墟となった町と神社 : 廃墟となった町と神社に灯る赤い蠟燭が、娘の復讐や祈りの象徴として燃え続ける場面。 プロンプト : "A desolate and eerie landscape with the ruins of a shrine on a hill. The remains of the village lie in ruins below, and the ocean in the background is dark and tumultuous. On the shrine's altar, a single red candle burns brightly, its light piercing through the surrounding darkness."

表 10 「赤い蠟燭と人魚」の提案手法の場面の分割結果

これは、4.2.2Midjourney を使った場合との比較における、提案手法を用いた「赤い蠟燭と人魚」の場面分割結果とプロンプトの生成結果である。「赤い蠟燭と人魚」の文章を入力すると、表 10 のように場面が分割され、それぞれの場면을画像化するためのプロンプトが生成された。生成されたプロンプトを Midjourney に入力し、漫画の画像が作成された。表 10 は図 11 の画像の各ページと対応しており、画像の場面と、その場面を生成する際に Midjourney に入力したプロンプトを示している。そして、最終的に図 11 に示す漫画が完成した。



図 13 被験者 B が Midjourney のみで作成した漫画「赤い蠟燭と人魚」

作成所要時間：1:22:40

以下は被験者 B が考案した図 13 の漫画の各ページの場面である。

1p	人魚の母親が登場する場面
2p	人魚の母親が子供を置いていく場面
3p	お爺さんとお婆さんが蠟燭を売る場面
4p	お婆さんが人魚の赤ん坊を拾う場面
5p	人魚の赤ん坊が育ち、美しい人魚になる場面
6p	人魚が蠟燭に絵を描く場面
7p	蠟燭が大繁盛する場面
8p	お爺さんとお婆さんが香具師に人魚の娘を売る話をされる場面
9p	人魚の娘が檻に閉じ込められ、香具師に売られる場面

10p	嵐で船が転覆する場面
11p	町が廃れ、滅ぶ場面



図 14 被験者 C が Midjourney のみで作成した漫画「赤い蠟燭と人魚」

作成所要時間：1:25:15

以下は被験者 C が考案した図 13 の漫画の各ページの場面である。

1p	人魚の母親が子供を置いていく場面
3p	お爺さんとお婆さんが蠟燭を売る場面
4p	お婆さんが人魚の赤ん坊を拾う場面
5p	人魚の赤ん坊が育ち、美しい人魚になる場面
6p	人魚が蠟燭に絵を描く場面
7p	お爺さんとお婆さんが香具師に人魚の娘を売る話をされる場面
8p	人魚の娘が檻に閉じ込められ、香具師に売られる場面
9p	嵐で船が転覆する場面
10p	町が廃れ、滅ぶ場面

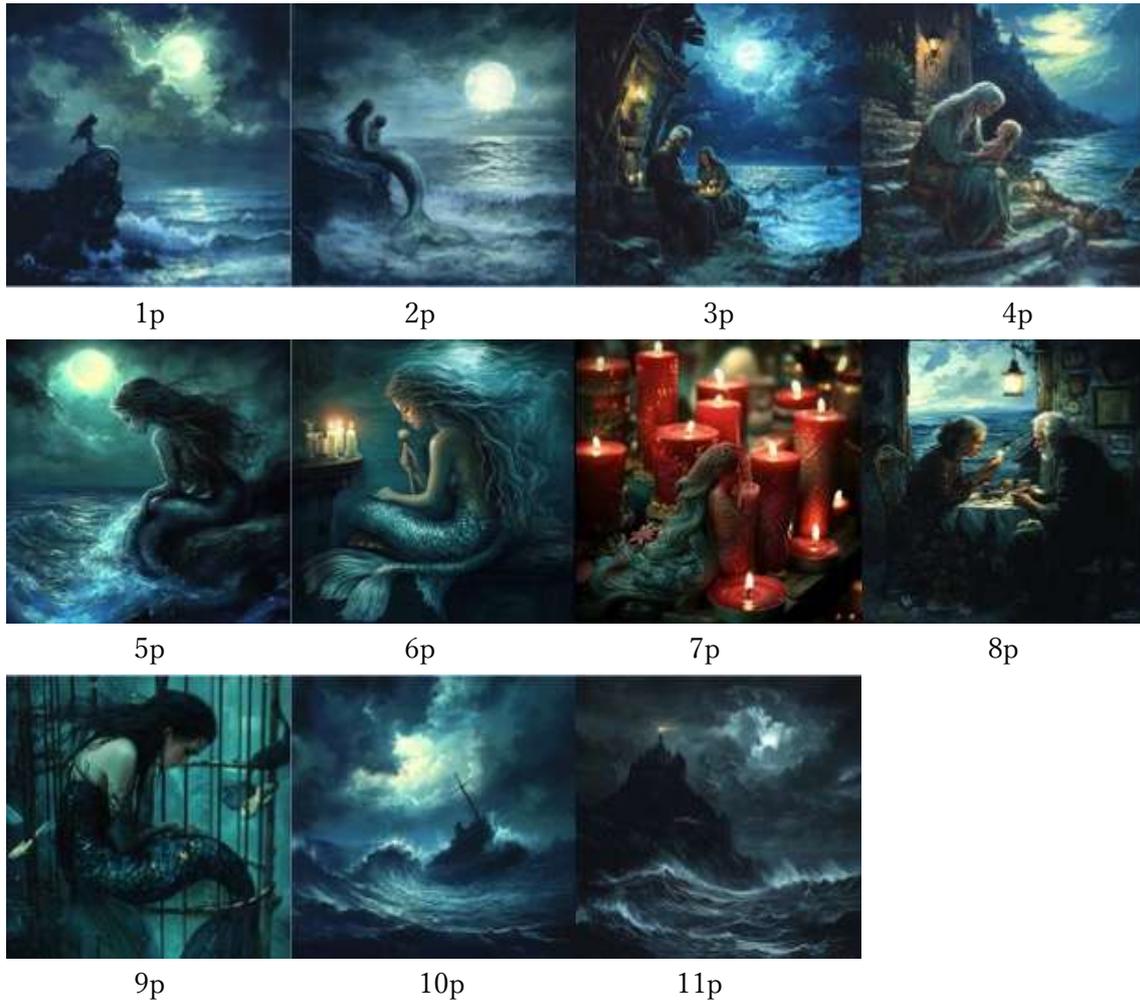


図 15 被験者 D が Midjourney のみで作成した漫画「赤い蝋燭と人魚」

作成所要時間：1:22:40

以下は被験者 D が考案した図 15 の漫画の各ページの場面である。

1p	人魚の母親が登場する場面
2p	人魚の母親が子供を置いていく場面
3p	お爺さんとお婆さんが蝋燭を売る場面
4p	お婆さんが人魚の赤ん坊を拾う場面
5p	人魚の赤ん坊が育ち、美しい人魚になる場面
6p	人魚が蝋燭に絵を描く場面
7p	蝋燭が大繁盛する場面
8p	お爺さんとお婆さんが香具師に人魚の娘を売る話をされる場面
9p	人魚の娘が檻に閉じ込められ、香具師に売られる場面

10p	嵐で船が転覆する場面
11p	町が廃れ、滅ぶ場面



図 16 被験者 E が Midjourney のみで作成した漫画「赤い蠟燭と人魚」

作成所要時間：1:30:02

以下は被験者 E が考案した図 16 の漫画の各ページの場面である。

1p	人魚の母親が登場する場面
2p	人魚の母親が子供を置いていく場面
3p	お爺さんとお婆さんが蠟燭を売る場面
4p	お婆さんが人魚の赤ん坊を拾う場面
5p	お爺さんとお婆さんが人魚の赤ちゃんを育てると決心する場面
6p	人魚の赤ん坊が育ち、美しい人魚になる場面
7p	人魚が蠟燭に絵を描く場面
8p	お爺さんとお婆さんが香具師に人魚の娘を売る話をされる場面
9p	人魚の娘が檻に閉じ込められ、香具師に売られる場面

10p	人魚の母親が蝋燭屋に行く場面
11p	嵐で船が転覆する場面
12p	町が廃れ、滅ぶ場面