

プロンプト・プログラミングによる小説挿絵の自動生成

久永 龍星[†] 牛尼 剛聡^{††}

[†]九州大学芸術工学部 〒815-8540 福岡県福岡市南区塩原 4-9-1

^{††}九州大学大学院芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: [†]hisanaga.ryusei.455@s.kyushu-u.ac.jp, ^{††}ushiana@design.kyushu-u.ac.jp

あらまし 近年, DALL-E 2, Midjourney, Stable Diffusion をはじめとした Text2Image モデルが多数出現してきたことで, 一般の人でも簡単に自分がイメージする画像を生成できるようになった. そこで重要になるのは, 自分のイメージした場面や人物を的確にプロンプトとして表現し, それらを Text2Image モデルに入力することである. 本研究では小説の内容に対応した画像を Text2Image モデルを用いて生成することに着目し, 文章の理解支援や物語への没入感の向上などの体験の向上につながるように小説の内容に応じた画像を提示することを研究の目的とする. 本稿では, 小説の挿絵画像を生成するため, 質問応答モデルを用いた場所や人物などの場面特徴抽出機構, それらを組み合わせたプロンプト, 利用した Text2Image モデルについて説明し, テキストと同時に生成した画像を提示する実験の結果について報告する.

キーワード 挿絵, 文書理解, Text2Image, プロンプト・プログラミング

1 はじめに

1.1 研究の背景

近年, 人々が読書に費やす時間が減ってきている傾向がある. 平成 30 年度の文化庁が全国の 16 歳以上の男女 3590 人に行った調査では, 1 ヶ月に本を 1 冊も読まないと回答した者が 47.3 パーセント存在する. また, 「読書量が以前に比べて減っているか, それとも増えているか」という質問に対して, 67.3 パーセントが「読書量が減っている」と回答している [1].

しかし, 2022 年にインプレス総合研究所が行った電子書籍ビジネス調査報告書によると, 電子書籍の利用者は増加しており [2], 「kindle」や「eBook Japan」を例として多くの電子書籍サービスが展開されている. 近年の電子書籍の売上は増加しているが, その多くがコミックによるものであり, 小説などの文字を中心とした書籍の売り上げの増加割合は低い (図 1). つまり, コミックという絵とテキストで表現される電子書籍より, 文字だけで表現される文字ものの魅力が低いと考えられる. しかし, 世の中に存在するすべての小説のテキストに対して, それらを説明するイラスト・画像を付けることはコストが大きく現実的ではない. この問題を解決するために, AI を利用して自動的に小説の挿絵を生成することが考えられる.

また, 最近 DALL-E 2 [3] や Midjourney [4], Stable Diffusion [5] をはじめとした Text2Image モデルが多数出現している. Text2Image モデルは, テキストを与えると, その内容を描写した画像や, テキストの内容から連想される画像を生成可能な機械学習モデルである. Text2Image モデルに入力するテキストは AI への要求指示 (プロンプト) と呼ばれる. Text2Image モデルを利用することにより専門的な知識のない一般の人々でも, 簡単に高品質な画像を生成可能である. そこで重要になるのは, ユーザが欲求する場面や人物を的確に Text2Image に入

力するプロンプトとして表現することである. また, テキストのみから画像を生成できるため, 必要なテキストを自動的に決定することができれば, Image2Text モデルを利用することで, 様々なアプリケーションが実現できる可能性がある. その一つとして, 小説における挿絵の自動生成が考えられる.

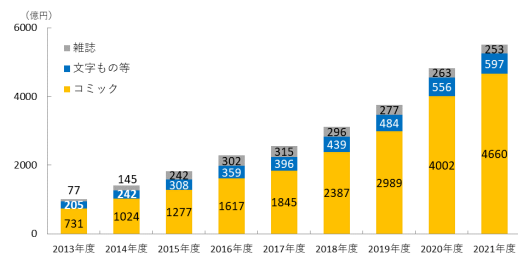


図 1 電子書籍市場規模のジャンル別内訳 [2]

1.2 研究の目的

本研究では, Text2Image モデルを用いて小説の内容を表現する画像を生成する手法を提案する. 本手法では, 書籍の場面ごとにその場面の特徴を記述するプロンプトを作成し, それを Text2Image モデルに入力することで, テキストの内容を表した画像を生成することを実現する (図 2).

本研究の貢献は, 以下の通りである.

- 小説の描写内容から質問応答モデルを用いて場面を構成する特徴を抽出し, Text2Image モデルに入れるプロンプトを作成, 挿絵画像を生成する手法を開発した.
- 評価実験により, 提案手法の有効性を評価した.



図2 システムの出力例

1.3 本論文の構成

本論文の構成は以下のようになっている。第2章では、関連研究の紹介を行う。第3章では、使用するモデルやデータセットについて述べる。第4章では、提案手法について述べる。第5章では、提案手法の有効性を評価するための実験内容について述べる。第6章では、結果と考察について述べる。第7章では、本論文のまとめについて述べる。

2 関連研究

2.1 Text2Image モデル

近年の Text2Image モデルの普及に伴って、どのようなテキストプロンプト、キーワードやモデルハイパーパラメーターが一貫した出力、良い出力を得るのに役立つのかについての研究が行われている。Vivian Liu ら [6] は Text2Image モデルに入力するプロンプトの順列やスタイル、生成モデルにおけるシードの値を多様に変えて調査し、Text2Image モデルにおいての良い結果を生み出すために役立つガイドラインを示している。この研究より、Text2Image モデルに入れるプロンプトの順列やスタイルの重要度がわかる。

2.2 小説の情報抽出

本研究では、Text2Image モデルに入力するテキストプロンプトを作成するために、小説から情報を抽出するが、これまでも自動的に小説から情報を抽出する研究が行われている。

米田ら [7] は小説中の単語の局所出現頻度や述語情報を利用して小説から登場人物を抽出する手法を提案している。局所出現頻度による判別では、小説の一定文数における特定の主語の出現頻度に基づいて算出、述語情報による判別では着目する述語の主語が人物になる確率を求めたデータを作成し評価実験を行なった。実験の結果、局所出現頻度と述語情報データを組み合わせた判別方法が最も良い評価を得られることを示している。米田らの研究では登場人物の抽出しかできておらず、小説の挿絵を構成するための場所や人物の動作などを抽出することはできていない。

2.3 小説に対しての情報付加

読書という体験に対して音を付加情報として付与することで

短い文や会話文において、ユーザ体験が向上するという事例が報告されている。西川ら [8] は、読者が文字を視認するタイミングで効果音を提示することで、短い文や会話文に限り「読みやすい」「気持ちがいい」といった意見が得られたことを示している。

また、上野ら [9] は VR 画面で小説本文に関連する映像を提示する環境の提案をしている。また、小説の内容に合わせた文字の背景色を提示することで感情移入をうながす提案も行われている。上浦ら [10] は、小説の本文からキーワードを抽出しそれらをキーとした yahoo 画像検索を行いその検索結果 Top20 の画像を取得し画像の画素値を平均色として抽出し取得した単語毎の画像の平均をとり単語色として取得し文字の背景として提示している。評価実験として青空文庫の芥川龍之介「沼地」、新美南吉「赤とんぼ」島崎藤原「二人の兄弟」小川未明「春がくる」を対象に提案手法を適用し大学生 16 人に対し通常の白背景と比べ、読みやすさ・感情移入できるかを調べた。結果として、感情移入では一定の効果が見られたが、読みやすさと文章と背景色との合致はタイトルによって変化するという結果を得ている。

このように、読書体験においてテキストの内容に関する付加情報を付与し提示することで一定の効果があることの検証はなされているが、そのどれもが画像を表示するための具体的な手法を提案できておらず、色や画像などの既存のものを提示しており、新たな画像を生成することはできていない。本研究では、テキストの内容に応じた画像を生成するために、まず小説を入力として場面特徴を記述する場面特徴生成機構により特徴を抽出し、それで出力した場面の特徴を Text2Image モデルに入れることで画像を出力する。

3 使用するモデルとデータセット

3.1 BERT

本研究で使用する自然言語処理モデルは BERT (Bidirectional Encoder Representations from Transformers) と呼ばれる Google の Jacob Devlin ら [11] が 2018 年 10 月に論文で発表したモデルである。BERT の学習は事前学習とファインチューニングに分かれている。事前学習とは、ある言語の構造や単語の意味といった言語の基礎を学習させることである。BERT の事前学習で用いるデータは生の文章データのみでデータとそれに対応する出力の関係を人間が付与したラベル付きデータを用いていない。BERT ではそれらラベルのついていない単語の分散表現を BERT の持つ Transformer と呼ばれる深層学習モデルが以下で述べる Mask Language Model と Next Sentence Prediction で文章から、文脈を双方向に学習する。

3.1.1 Mask Language Model

Mask Language Model では入力文の 15% の単語を確率的に別の単語で置き換えし、文脈から置き換える前の単語を予測させる。選択された 15% の単語のうち 80% は [MASK] という新しいトークンに置き換え、10% は別の単語に、残りの 10% はそのまま単語を置き換えずにタスクを解く。置換された単語

を周りの文脈から当てるタスクを解くことで単語に対応する文脈情報を学習する。

3.1.2 Next Sentence Prediction

上記の Mask Language Model では、単語に関する学習はできるが、文単位の学習はできない。そのため Next Sentence Prediction において「2つの文は隣り合っているか」を正解できるように学習することで、2つの文の関係性を学習させる。具体的な処理としては、2文のうち片方を50%の確率で他の文に置き換える。ここで、入力された2文が連続したものである(IsNext)か、連続していない(NotNext)かを判定するタスクを繰り返し学習を行う。

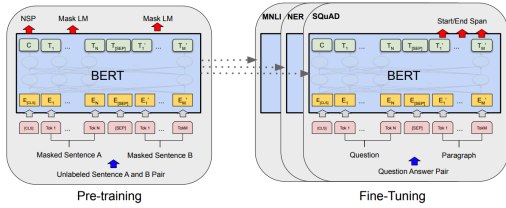


図3 bertの事前学習とファインチューニング[11]

3.2 質問応答モデル

BERTの事前学習済みモデルは、多くの場合ファインチューニングを行い、解きたいタスクに応じてラベル付きのデータを学習し特化させて使われる。

本研究では、東北大学乾研究所による日本語 Wikipedia 事前学習済み BERT モデル [12] を黒橋らの研究室が公開している運転ドメイン QA データセット [13] を用いてファインチューニングを行い質問応答モデルを作成する。質問応答モデルは、図4のようにテキストとそれに対しての質問文を入力することで、質問文に対する解答を出力することができる。

ここで、テキストを t 、質問テキストを q とするとき解答 a を生成する。モデルは以下の QA として表現される。

$$QA(t, Q) = a \quad (1)$$

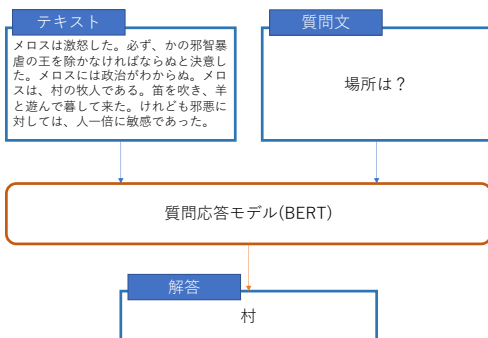


図4 質問応答モデル

3.3 Japanese-Stable-Diffusion

本研究では Text2Image モデルとして Japanese-Stable-Diffusion [14] を利用する。Japanese-Stable-Diffusion とは Stability AI 社がオープンソースで公開している Text2Image モデル Stable Diffusion [5] に対して日本語のキャプション付き画像を用いて追加学習することで日本語に特化した画像生成モデルとして開発されたものである。この Japanese-Stable-Diffusion によって既存の Stable-Diffusion では入力できなかった日本語のプロンプトの入力や日本語固有の表現についても画像に反映可能である。

4 提案手法

4.1 提案手法の概要

本節では、目的とするシステムと提案手法について説明する。本研究で目的とするシステムは電子書籍における読書体験向上のために、場面を表した挿絵画像を自動生成することを目的とする。具体的には小説のテキストをシステムに入力するとそれらを自動的に適切な場面に分割し、それらの場面に登場する人物や場所、人物の動作なども特徴を抽出し、それらを Text2Image モデルに入力することで挿絵を出力する。図5では提案手法の概要図を示しており、本システムは主に場面特徴生成機構、プロンプト作成機構、画像生成機構に分かれている。

4.2 場面特徴生成機構

小説のテキストから場面特徴を抽出するための場面特徴生成機構について説明する。小説の本文テキストをそのまま Text2Image モデルに入力するだけでは、描写内容についての多くの情報が含まれているために適切に場面を表した画像を生成することができないという。本研究では、質問応答モデル (BERT) を用いて場面の特徴を抽出する。

小説 nv を文の s_i 系列として以下の様に定義する。

$$nv = \langle s_1, s_2, \dots, s_{|nv|} \rangle \quad (2)$$

ここで $|nv|$ は系列 nv を構成する要素数を表す。

本機構では小説のテキストを n 文ごとに区切って質問応答モデルに入力する。これは BERT の入力系列長の問題でありに長すぎるテキストは入力できないためである。

いま、 nv の部分系列を抽出する窓関数 $w(nv, i, n)$ を以下の様に定義する。

$$w(nv, i, n) = \langle s_i, s_{i+1}, \dots, s_{i+n-1} \rangle \quad (3)$$

次に、それぞれの n 文ごとのまとまりとした部分系列を質問応答モデルに入力する。本論文では $n = 4$ とする。ここで、窓関数で抽出された小説中の部分系列をセグメントと呼ぶ。セグメント seg_i を以下の式で定義する。

$$seg_i = w(n, i, n) \quad (4)$$

それぞれのセグメントに対して、入力する際の4つの質問文

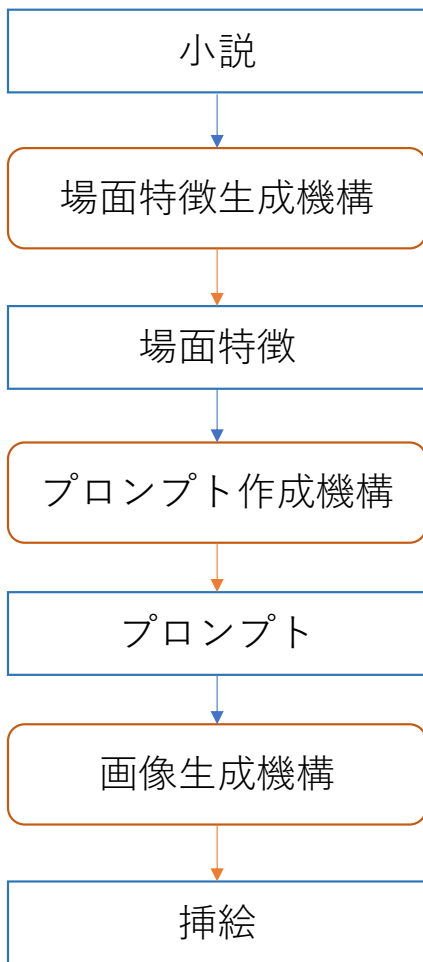


図 5 提案手法概要

をプロンプトとして入力し、それに対応した回答を得る。

q_1 : 時間はいつですか?

q_2 : 場所はどこですか?

q_3 : 誰ですか?

q_4 : 何をしていますか?

そして上記の質問に対して全ての解答が得られるまで次の文章を質問応答モデルに入力し、全ての解答が得られた時点でそれらの文章をひとまとまりにして1つの場面とし、得られた場面特徴4つを出力する。図6は、上記の機構の仕組みを表している。

なお、場面特徴生成機構では、 $w(nv, i, n)$ を1つのセグメントとして扱うが、次に入力されるセグメントは $w(nv, i+n/2, n)$ である。これは、文章の途中で場面の区切りがあった場合にその区切りを検出できるようにするためである。

質問文集合を $Q = \{q_1, q_2, q_3, q_4\}$ として表したとき、セグメントに関する質問 q_j に対する特徴 f_i^j は以下の式で表す。

$$f_i^j = QA(seg_i, q_j) \quad (5)$$

ここで QA モデルは、質問文に解答できない場合は null を返すものとする。

提案手法では、対象とする小説の先頭のセグメントから特徴を抽出する。そして、すべての特徴が null でない最短の部分系列を場面とする。そして、小説を場面の系列に変換してそれぞれの場面に対して画像を生成する。

表 1 場面特徴例

q_1	q_2	q_3	q_4
結婚式も間近 いつか	村 カムパネルラ	メロス 先生	村の牧人 星図

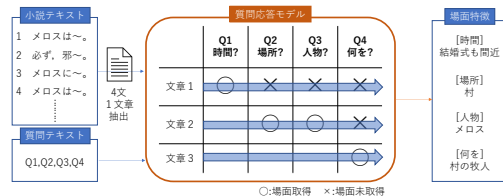


図 6 場面特徴生成機構図

4.3 プロンプト生成機構

次に、それぞれの問いで得られた場面特徴をその小説の時代そして欲しい画像の画風という小説固有の固定入力と組み合わせる。固定入力はユーザが手動で与えるものとする。

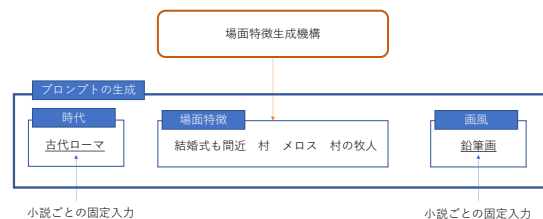


図 7 プロンプト生成機構図

4.4 画像生成機構

先ほど得られたプロンプトを Text2Image モデルに入力して画像を得る (図 8)。

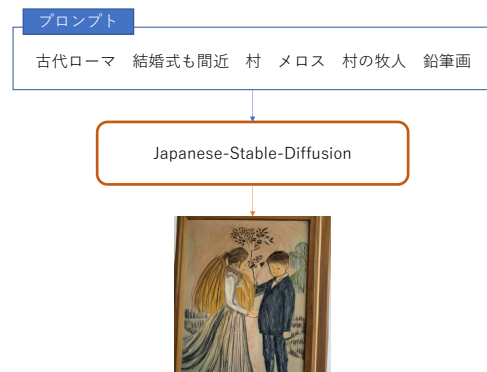


図 8 画像生成機構図

5 評価実験

提案手法の有効性を評価するために、クラウドソーシングサービス¹を用いて被験者によるアンケートを行い実験を行った。対象とした小説は「青空文庫²」、「小説家になろう³」のサイトより、公開されている以下4つのタイトルの小説である。

- 青空文庫「走れメロス」 著者名：太宰 治
- 青空文庫「銀河鉄道の夜」 著者名：宮澤 賢治
- 小説家になろう「魔女と傭兵」 著者名：超法規的かえる
- 小説家になろう「無能才女は悪女になりたい～義妹の身代わりで嫁いだ令嬢、公爵様の溺愛に気づかない～」 著者名：一分咲

本研究の提案手法を適用して得られた挿絵と、小説のテキストをそのまま入力として Text2Image に入力し挿絵を得る (比較手法) 2つの手法により得る挿絵を比較して、提案手法の有意性を検証した。このとき、各小説3場面に対し提案手法と比較手法を適用して得られた画像を対応するテキストと組み合わせ表示した。

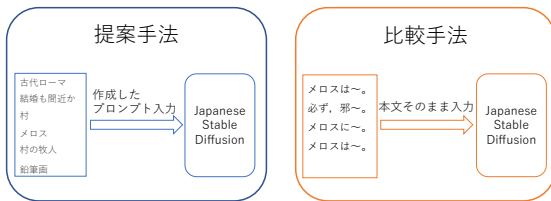


図9 提案手法と比較手法

場面は4.2節で説明している4つの場面特徴が得られた時点を区切りとしている。生成した画像を実際のテキストの挿絵として利用し、被験者の主観を求める。質問した項目は以下の通りである。

質問1 この挿絵が対応する小説の内容を適切に表せていると思いますか？

質問2 挿絵が小説の内容理解に役立つと思いますか？

質問1, 質問2では5段階評価(5が最も良いとする)で回答を求めた。図5.1に、実際にアンケートに用いたフォームの一部を示す。

図10 アンケート

6 結果と考察

6.1 実験結果

評価実験のアンケートを行い、被験者30人の有効な結果を得た。図11, 図12では5章で説明しているフォーム内での質問2つに対する提案手法と比較手法の回答結果を示している。質問1では、評価値の平均が小数点第三位を四捨五入して提案手法では3.10、比較手法では2.40であった。また、このデータに対してt検定を行なった結果2つの手法の有意な差が得られた。すなわち提案手法は、比較手法と比べて小説の内容を適切に表していることが示された。

質問2では、評価値の平均が小数点第三位を四捨五入して提案手法では3.06、比較手法では2.37であった。また、このデータに対してt検定を行なった結果2つの手法の有意な差が得られた。すなわち提案手法は、比較手法と比べて小説の内容理解に役立つことが示された。

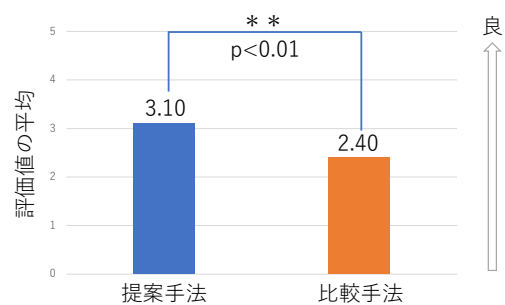


図11 質問1に対する回答結果

1: クラウドワークス: <https://crowdworks.jp/>

2: 青空文庫: <https://www.aozora.gr.jp/>

3: 小説家になろう <https://syosetu.com/>

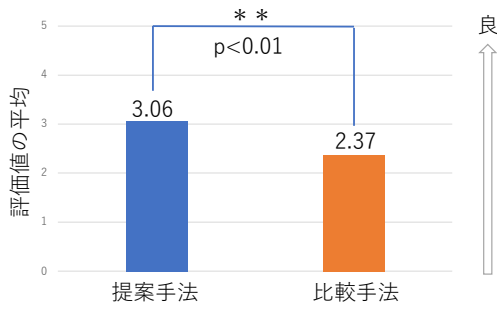


図 12 質問 2 に対する回答結果

図 13, 図 14 では質問 1 と質問 2 に対しての 15 までの評価値の総数を提案手法と比較手法に分けて表している。

質問 1 の回答では、評価値 1 は、比較手法の総数が提案手法と比べて多く、4 と 5 の比較的良い評価値に対しては、提案手法の総数が比較手法と比べ多いという結果を得た。

質問 2 の回答では、質問 1 への回答と同様に評価値 1 は、比較手法の総数が提案手法と比べて多く、4 と 5 の比較的良い評価値に対しては、提案手法の総数が比較手法と比べ多いという結果を得た。

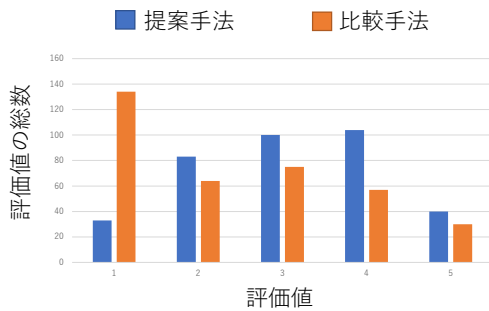


図 13 質問 1 に対する各評価値の総数

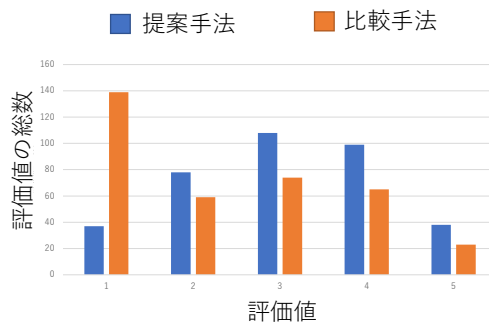


図 14 質問 2 に対する各評価値の総数

図 15, 図 16 では各テキストに対しての質問 1,2 の評価値の平均を表している。

それぞれのテキストへの評価を見た時には、テキスト番号 2 と 8 では比較手法の方が高い評価値を出しており、その他のテキストに対しては、提案手法の評価値が高いという結果を得た。この結果は質問 1 と質問 2 の両方の評価値の結果に当てはま

る。また、テキスト番号 2 と 8 について t 検定を行ったところ質問 1 ではテキスト 2,8 ともに $p > 0.01$ となり有意差は得られなかった。質問 2 についても同様に $p > 0.01$ となり有意差は得られなかった。

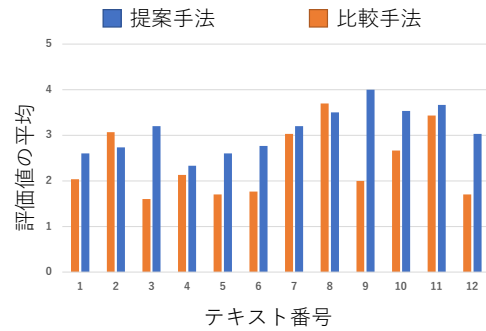


図 15 各テキストに対する質問 1 の評価値の平均

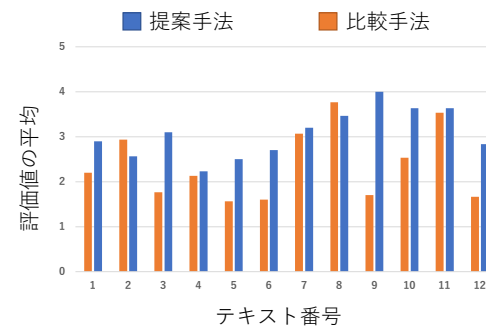


図 16 各テキストに対する質問 2 の評価値の平均

6.2 考 察

本研究での評価実験を通して、小説の内容を適切に表せているか、小説の内容理解に役立つか、の 2 点について提案手法の有効性が示された。提案手法が比較手法と比べ優れている理由としては、特徴を表したプロンプトにあると考える。比較手法では、対応する場面のテキストが全てプロンプトとして入力される。そのため、場面を表す特徴を含まない人物同士の会話や感嘆詞などをテキストに含まれている。また、提案手法と比べあまりにも多くの場面特徴を含んでいるためとも考えられる。提案手法では、4 つの質問テキストに対して 4 つの解答が抽出されるため、正しく質問に答えることができるか、できないかに関わらず 4 つの場面特徴と 2 つの固定特徴をプロンプトとして入力される。しかし、比較手法では入力される場面特徴に制限はなくテキストに含まれている限り多くの特徴がプロンプトとして入力される。そのため、1 つの場面として画像を出力することが難しかったのだと考えられる。

また、評価実験において被験者の任意自由回答として

- 場面を表している単語/文節/文章などあれば教えてください

文 献

という質問をしたところ表 2 のような回答結果を得た。回答例として、最も得られた回数が多い回答を挙げている。この結果から、本研究の場面特徴生成機構で用いた質問応答モデルによってある程度適切に場面特徴を得られたのではないかと考えられる。理由としてはテキスト 1 の例で挙げられている「村の牧人」やテキストの 2 での「結婚式」、他のタイトルでは、「魔女」「理科の時間」など実際に場面特徴として抽出されプロンプトに入っているからである。しかし、「邪智暴虐」や「いちいちの光るつぶ」など被験者が考える場面を表す単語を抽出できていない場合もあった。

また、質問テキストは少しの変化を加えるだけでも得られる場面特徴が変わるためより適切な質問テキストを検討する必要がある。

7 おわりに

本論文では、小説の内容に対応した画像を Text2Image モデルを利用して生成する手法を提案した。提案手法では、小説テキストと質問テキストを質問応答モデルに入力して場面特徴を抽出しそれらと 2 つの固定入力を組み合わせてプロンプトを作成、Text2Image モデルに入力する手法を提案した。提案手法の有効性を示すために、1 場面全てのテキストをそのままプロンプトとする手法 (比較手法) と提案手法の 2 つを 4 つの小説に適用して被験者による評価実験を行った。その結果、提案手法は以下の二つの点で有効であることがわかった。

- 小説の場面を適切に表しているか
- 小説の理解支援に役立つか

これらの結果は小説における挿絵の自動生成において本手法が有効であることを示している。

今後の課題として、場面特徴抽出時の精度改善が挙げられる。本論文では、時間や場所、人物、動作を尋ねるような質問テキストを入力し特徴を得ているが、実際にはそれぞれの質問に対応しない応答が得られることがある。他にも、人物と動作を尋ねている部分で抽出した人物と動作の主が同一人物でない場合なども挙げられる。今後それらの改善についても検討したいと考える。

謝 辞

本研究は JSPS 科研費 19H04219 の助成を受けたものです。

- [1] 文化庁：平成 30 年度「国語に関する世論調査」の結果について (2018).
- [2] インプレス総合研究所：電子書籍ビジネス調査報告書 2022 (2022).
- [3] Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. and Chen, M.: Hierarchical text-conditional image generation with clip latents, *arXiv preprint arXiv:2204.06125* (2022).
- [4] Midjourney, : <https://www.midjourney.com/home/>.
- [5] Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B.: High-resolution image synthesis with latent diffusion models, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695 (2022).
- [6] Liu, V. and Chilton, L. B.: Design guidelines for prompt engineering text-to-image generative models, in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–23 (2022).
- [7] 米田崇明, 篠崎隆宏, 堀内靖雄, 黒岩真吾: 述語情報を利用した小説の登場人物の抽出, 言語処理学会第 18 回年次大会発表論文集, Vol. 18, pp. 855–858 (2012).
- [8] 西川尚志, 橋本直他: 文字表示に同期した音の提示が読書体験に与える影響, 研究報告エンタテインメントコンピューティング (EC), Vol. 2020, No. 27, pp. 1–6 (2020).
- [9] 上野浩平, 内田知巳, 羽岡浩二, 中島智晴, 吉田典弘: 人工知能の自然言語処理を利用した VR による本読解アプリの提案, 情報システム学会.
- [10] 上浦拓人, 山田昇平他: 電子書籍アプリケーション用背景色生成システムの提案, 映像情報メディア学会誌, Vol. 72, No. 10, pp. J183–J188 (2018).
- [11] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding, *arXiv preprint arXiv:1810.04805* (2018).
- [12] 東北大学乾研究室: Pretrained Japanese BERT models : <https://github.com/cl-tohoku/bert-japanese>.
- [13] 黒橋・村協研究室: 運転ドメイン QA データセット (2019), <https://nlp.ist.i.kyoto-u.ac.jp/?Driving+domain+QA+datssets>.
- [14] HuggingFace, : *japanese-stable-diffusion* (2022), <https://huggingface.co/ri/stable-diffusion>.

表 2 場面を表している単語・文節. 文章についての回答例

	得られた回答例	
走れメロス	テキスト 1	村の牧人, 邪智暴虐
	テキスト 2	結婚式, 羊と遊んで暮らしてきた
	テキスト 3	竹馬の友, セリヌンティウス
銀河鉄道の夜	テキスト 4	川
	テキスト 5	先生, 雑誌
	テキスト 6	理科の時間, いちいちの光るつぶ
魔女と傭兵	テキスト 7	剣, 傭兵
	テキスト 8	魔女
	テキスト 9	魔女, 戦闘
無能才女は悪女になりたい	テキスト 10	結婚式
	テキスト 11	ドレス, 地下に続く階段室
	テキスト 12	ドレス, 扉

小説タイトル	走れメロス	銀河鉄道の夜	魔女と傭兵	無能才女は悪女になりたい
小説テキスト	メロスには政治がわからぬ。メロスは、村の牧人である…	カムパネルラが手をあげました。それから四五人手をあげました…	彼は未知の敵との戦闘を想像しながらその対処に頭を悩ませてい…	一瞬だけ静まり返った教会に響いた、鈍い…
プロンプト	古代ローマ 結婚も間近か 村 メロス 村の牧人 鉛筆画	昭和時代 いつかカムパネルラ先生 星図 鉛筆画	現代 彼 魔女 ファンタジー	現代 自分の結婚式 教会 扉に体当たり ファンタジー
提案手法				
比較手法				

図 17 出力結果例

表 3 本文テキスト

小説タイトル	場面に含まれるテキスト
走れメロス	<p>メロスは激怒した。必ず、かの邪智暴虐の王を除かなければならぬと決意した。メロスには政治がわからぬ。メロスは、村の牧人である。</p> <p>笛を吹き、羊と遊んで暮して来た。けれども邪悪に対しては、人一倍に敏感であった。きょう未明メロスは村を出発し、野を越え山越え、十里はなれた此のシラクスの市にやって来た。メロスには父も、母も無い。女房も無い。十六の、内気な妹と二人暮しだ。この妹は、村の或る律気な一牧人を、近々、花婿として迎える事になっていた。結婚式も間近かなのである。</p> <p>メロスには竹馬の友があった。セリヌンティウスである。今は此のシラクスの市で、石工をしている。その友を、これから訪ねてみるつもりなのだ。久しく逢わなかったのだから、訪ねて行くのが楽しみである。歩いているうちにメロスは、まちの様子を怪しく思った。ひっそりしている。もう既に日も落ちて、まちの暗いのは当りまえだが、けれども、なんだか、夜のせいばかりでは無く、市全体が、やけに寂しい。のんきなメロスも、だんだん不安になって来た。路で逢った若い衆をつかまえて、何かあったのか、二年まえに此の市に来たときは、夜でも皆が歌をうたって、まちは賑やかであった筈だが、と質問した。</p>
銀河鉄道の夜	<p>「ではみなさんは、そういうふう川だと云われたり、乳の流れたあとだと云われたりしていたこのぼんやりと白いものがほんとは何かご承知ですか。」先生は、黒板に吊した大きな黒い星座の図の、上から下へ白くけぶった銀河帯のようなところを指しながら、みんなに問をかけました。カムパネルラが手をあげました。それから四五人手をあげました。</p> <p>カムパネルラが手をあげました。それから四五人手をあげました。ジョバンニも手をあげようとして、急いでそのままやめました。たしかにあれがみんな星だと、いつか雑誌で読んだのですが、このごろはジョバンニはまるで毎日教室でもねむく、本を読むひまも読む本もないので、なんだかどんなこともよくわからないという気持ちがするのです。ところが先生は早くもそれを見附けたのです。「ジョバンニさん。あなたはわかっているのでしょうか。」ジョバンニは勢よく立ちあがりましたが、立って見るともうはっきりとそれを答えることができないのです。ザネリが前の席からふりかえって、ジョバンニを見てくすくとわらいました。ジョバンニはもうどぎまぎしてまっ赤になってしまいました。先生がまた云いました。「大きな望遠鏡で銀河をよく調べると銀河は大体何でしょう。」</p> <p>「このいちいちの光るつぶがみんな私どもの太陽と同じようにじぶんで光っている星だと考えます。私どもの太陽がこのぼぼ中ごろにあって地球がそのすぐ近くにあるとします。みなさんは夜にこのまん中に立ってこのレンズの中を見まわすとしてごらん下さい。こっちの方はレンズが薄いのでわずかの光る粒即ち星しか見えないのでしょ。こっちやこっちの方はガラスが厚いので、光る粒即ち星がたくさん見えその遠いのはぼうっと白く見えるというこれつまり今日の銀河の説なのです。そんならこのレンズの大きさがどれ位あるかまたその中のさまざまの星についてはもう時間ですからこの次の理科の時間にお話します。」</p>
魔女と傭兵	<p>ずいぶん集めたもんだな野営地で準備をする傭兵、領主の私兵、果ては民兵まで勢揃いしていた。雑多な寄せ集めともいえる彼らを見て呆れ半分につぶやいた。何とはなしに漏れ出た彼のつぶやきに近くで剣の手入れをしていた男が笑う。「そりゃあそうだろ。」</p> <p>何とはなしに漏れ出た彼のつぶやきに近くで剣の手入れをしていた男が笑う。「そりゃあそうだろ。準備しすぎるに越したことはねえ。なんせ—魔女を狩ろうってんだからな」魔女 魔術と呼ばれる未知の業を操りその力は天候すら動かすといわれている。かつて魔女の怒りを買った国が一夜にして滅んだ、戯れに洪水で村を押し流した等その危険性を言い伝える噂は数々。</p> <p>彼は未知の敵との戦闘を想像しながらその対処に頭を悩ませていた。酒場で聞いた与太話では何もないところから道具も使わずに火の玉を出したり風を巻き起こすようだが。ジグにはそのような芸当が本当にできるのか甚だ疑問であった。「とはいえ魔女が実在するのは事実。魔術とやらが本当なのかどうかは置いておくとして、噂相応の芸当ができると考えておいたほうがいいか。」彼は魔術に懐疑的であったが、魔女の脅威自体は現実のものとして受け止めていた。おとぎ話で片づけるには被害にあった事例が多い上に大きな国は魔女の討伐に躍起になっている。これから討伐に向かう魔女も過去幾度も討伐隊を差し向けられたことがあるが、すべて失敗に終わっているという。とてもではないが個人でどうにかできる範囲を超えている。彼は魔女というのは何かの集団を表すもので、個人名ではないと考えていた。「何かしらの国家絡みの集団か、あるいは犯罪組織かなにかか……。」そんなところだろうと考えていた。</p>

表 4 本文テキスト続き

小説タイトル	場面に含まれるテキスト
無能才女は悪女になりたい	<p>今日は公爵様とエイヴリルの結婚式だ。エイヴリルの格好をした義理の妹コリンナは、エイヴリルの夫となる公爵閣下、ディラン・ランチェスターのもとに向かって、バージンロードをゆっくりゆっくと歩いていく。なぜか、それを当のエイヴリルは地下に続く階段室の隙間から見守っていた。どうしてエイヴリルが自分の結婚式をこんな場所から見守っているのかというと、今まさにバージンロードを歩いているコリンナによってドレスを奪われ閉じ込められたからである</p> <p>なぜか、それを当のエイヴリルは地下に続く階段室の隙間から見守っていた。どうしてエイヴリルが自分の結婚式をこんな場所から見守っているのかというと、今まさにバージンロードを歩いているコリンナによってドレスを奪われ閉じ込められたからである。こんな残念なことになるまでには紆余曲折あったし今すぐに説明できないのが惜しいくらいなのだが、少なくとも本意ではないことだけは声を大にして言いたい。(あのドレス、やっぱり素敵……。さすが、ディラン様が私のために仕立ててくださったドレスだわ。真っ白なシルクの生地に散りばめられた刺繍とラインストーンがきれい……。!) こんな状況なのに、自分がディランからの贈り物であるドレスに夢中になっていたことに気がついたエイヴリルは、ハッと蒼くなる。(いけないわ、また私ったら。ここはドレスの美しさに感激するところではなかった。ディラン様にごめんなさい、だわ。ディラン様からの贈り物のドレスが……。!) 隙間から見守っていたエイヴリルは、口にあてられた布紐を噛みしめ、扉に体当たりをする。「ふい……。ふいはひは、はらはは」(い……。いたいわ、からだが) 一瞬だけ静まり返った教会に響いた、鈍い音</p>