

小説を対象とした物語の場面を表現する挿絵自動生成方式

大西 力登[†] 岡田 龍太郎[†] 峰松 彩子[†] 中西 崇文[†]

[†]武蔵野大学データサイエンス学部 〒135-8181 東京都江東区有明 3 丁目 3-3

E-mail: [†]s2022069@stu.musashino-u.ac.jp, {ryotaro.okada, takafumi.nakanishi, ayako.minematsu}@ds.musashino-u.ac.jp,

あらまし 本稿では、小説を対象とした物語の場面を表現する挿絵自動生成方式を示す。一般的に、小説において挿絵は物語の状況を読者が想像し理解する助けとするために挿入される。しかしながら、小説の各場面を表現した挿絵を作成するコストが大きいたことが問題となる。本方式は、テキストで表現された小説コンテンツ全体を入力として、その小説のそれぞれの場面を表現する挿絵を出力する。本方式の挿絵は、ノベルスタイルとシーンオブジェクトによって決定される。ノベルスタイルは、事前にいくつかの小説コンテンツのテキストデータと表紙絵の組を知識ベースとして準備し、入力小説のテキストデータとの類似度が高い知識ベース内の小説の表紙絵の画風を抽出することで選択される。シーンオブジェクトは、入力された小説コンテンツを場面ごとに分割し、それぞれの場面を表現する画像を検索することで選択される。シーンオブジェクトをノベルスタイルに応じた画風変換 (Style Transfer) することにより、小説の場面を表現する挿絵を自動生成することが可能となる。本方式が実現されることにより、小説の著者が簡単に挿絵を作成出来るだけでなく、読者にとっても場面把握のために挿絵を作成し、確認することが可能となる。

キーワード Style Transfer, 自然言語処理, メディアミックス, Doc2Vec, GAN

1. はじめに

小説における挿絵とは物語の場面を表現する補助的な絵のことである。また近年ではインターネットの普及により小説はだれでも執筆し共有できるようになった。それに伴い小説家になろう[1]などのオンライン小説を共有する小説投稿サイトが誕生し人気を博している。このようなサイトでは挿絵が挿入されていることが少ない。またカクヨム[2]のように著者が小説本文中に挿絵を挿入できないようなサイトも存在する。

しかし挿絵が読者に対して与える効果は確かに存在している。

田中の研究 [3]では 18 歳以上の男女を対象として、調査参加者をテキストに挿絵があるグループとないグループにわけ、理解度、創造性、関心度を問う調査を行った。本調査ではテキストの理解度及び関心度に関しては挿絵の有無によって結果は異ならなかった。しかし創造性に含まれる登場人物などの具体的なイメージ像形成に関しては挿絵の影響を受けていると示されていた。

また和田の研究[4]では心理学の授業を受講している生徒に対して挿絵の有無によって状況モデルの構築に寄与する状況的次元にいかなる差異が見られるかについて実験を行った。この研究における状況モデルとはテキストの意味や文章の流れに読み手の推定過程や既有知識に由来する情報が統合された高次の心的表現であり、豊かな状況モデルの構築が文章理解の到達目標とされている[5]。また読者が状況的モデルを作製する際に状況的次元を用いている。状況的次元とは

Event-Indexing モデル[6,7]によって定義されており、読者がもつ状況モデルとは、登場人物の同一性、空間性、時間性、因果性、意図性の 5 つの状況的次元によって構成されていると仮定された物である。たとえば、物語において同じ場所で起きた出来事は空間性の状況的次元が同じであると表現される。実験の結果より物語文に挿絵が含まれる場合はイベント同士の因果関係を示す因果性次元が状況モデルの構築に影響を与えていることが示された。この影響は挿絵が含まれない場合には存在しないものであった。また挿絵が文章と異なる位置に挿入されている場合はこれらの影響がなくなることにも実験されている。

以上の研究により挿絵の有無により読者に対して与える効果は存在しており、読書における挿絵は重要であると考えられる。

しかし Web 上の作品を含めたすべての小説作品に対して挿絵を挿入することは難しいと考える。小説投稿者が同時に絵を書くことが困難であるといった問題がある。市販されている小説においても著者とイラストレーターが同人物であることは非常に少ない。また他者に絵の作製依頼を出すことが金銭的な面や手間の面から容易でないことも起因している。しかし、現在ではライトノベルという小説のジャンルが流行している。このジャンルは主に若年層向けの小説を想定しており、アニメ調の挿絵を利用し会話をやや多めにおくことで気楽に読めることが多い。このジャンルにおいて挿絵を多用することが多いため本ジャンルにおける挿絵の重要性は高いと考えられる。

以上のことより小説における挿絵は、読者の小説に対する想像力を補う点や、状況の因果関係を理解する上で有用であり、挿絵を必要としている読者も多いが、一般的な著者においては挿絵を挿入することは容易ではないという状況があると言える。

本稿では小説の著者が挿絵を作製するのを容易にする目的とし、小説のテキストを利用して自動で挿絵を生成するシステムを示す。

本方式では、小説の挿絵を自動生成するにあたり、挿絵を構成するための要素としてノベルスタイルとシーンオブジェクトを我々が新たに提案する。ノベルスタイルとは小説全体における挿絵の画風のことであり、シーンオブジェクトとは小説の各場面におけるその場面情報を表す画像のことであり、これら二つを抽出することにより、小説の画風にあった場面ごとの挿絵を適切に作製することが出来る。

本方式では、小説文章からノベルスタイルを認識しシーンオブジェクトを取得する。それらを要素としてシーンオブジェクトが指定する小説内の場面を表す画像に対して、ノベルスタイルが指定する画風への変換を行うことにより、小説内の場面を表現する挿絵を生成する手法を提案する。

2. 関連研究

本節では関連研究について述べる。2.1 節ではキーワードから画像の自動生成に関する研究について述べる。2.2 節では小説内のシーンを特定するためのキーワードを抽出する研究について述べる。2.3 節では小説への挿絵の挿入に関する研究について述べる。

2.1. キーワードから画像の自動生成に関する研究

菅生ら[8]は感性を考慮した自然言語文からの風景画像生成システムを提案している。この提案は入力として文章を受け取り、合成画像を出力するシステムである。この提案において画像の作成には事前に複数種類の画像を風景オブジェクトとして作成し、その風景オブジェクト名と合致した単語がある際に風景オブジェクトを複数組み合わせることで画像を作成するという手法を用いている。また受け取った文章における単語に対して連想語を抽出することにより、画像作製の際に事前に作成した風景オブジェクト名と合致させられるようにしている。加えて色の感性的イメージを考慮しており、西山ら[9]の文献を参考に色の感性的な印象と入力名詞との共起度をいもとに入力された文章にたいして代表的な色を決定している。

また村山[10]らは Word2Vec[11]を用いた挿絵自動挿入手法の提案を行っている。この提案では初めに、ロイヤリティフリー素材を提供しているサイトよりスクレイピングを行うことで画像とその画像を説明するテキストの組である挿絵候補画像データベースを作成

する。最後にユーザがテキストデータを入力するとそのデータに Word2Vec を利用し、挿絵候補画像データベース内部のどの画像と類似しているかコサイン類似度を用いて計算し選択するといった手法である。

また海外では GAN を利用し、文章から画像を作成する研究も盛んである AttenGAN[12]では与えられた文章全体からの画像の生成ではなく、各単語の表現に関して考慮した画像生成を行うシステムを構築している。また MirrorGAN[13]では各単語の表現に加え文章全体の表現を活用することで画像の多様性と意味の一貫性を強化した。加えて、生成された画像に対してテキストを生成する過程を追加することで以前までの GAN と比べて優位性を示した。

本方式ではシーンオブジェクトを決定する際にはシーンを表す単語と一致した単語をシーンオブジェクト単語として取得し、その単語より画像検索を行うことでシーンオブジェクトを取得している。このシーンオブジェクトに対して小説全体のスタイルであるノベルスタイルを適用することで小説全体としてまとまりのある挿絵を生成している。

2.2. 小説内のシーンを特定するためのキーワードを抽出する研究

五十嵐[14]らは小説内の挿絵を構成するためのキーワードを抽出する際に、単語出現頻度を用いて重要な小説内の位置を特定し、その周辺の名詞 10 語を抽出するとともに、付加的な情報として 5W のうちの Where, What, Who, When の 4 つに着目して抽出を行っている。

本方式では、挿絵の元となる要素としてノベルスタイルとシーンオブジェクトシーンという要素を抽出し、これらの組み合わせから挿絵を自動生成している。シーンオブジェクトの抽出には、小説内の文章からシーンを表す語を抽出し、分類語彙表を用いて典型的な表現に変換している。

2.3. 小説への挿絵の挿入に関する研究

五十嵐 [14]らは挿絵の挿入箇所を主人公に着目し、主人公の登場する文章に対して単語出現頻度を利用し数値化し重要な文章の推定を行っている。結果挿小説内で重要と判断されたページの前後 2 ページ分の文章を用いて挿絵は生成されていると推定された。

本方式ではノベルスタイルおよびシーンオブジェクトを利用して画風変換を行うことで挿絵を自動生成する方式に主眼を置いている。挿絵を自動生成する際には後述する一定の規定により分割したシーンに対してシーンオブジェクトを取得し、各シーンの挿絵を自動生成しているため挿入箇所は該当のシーン付近になると考察される。

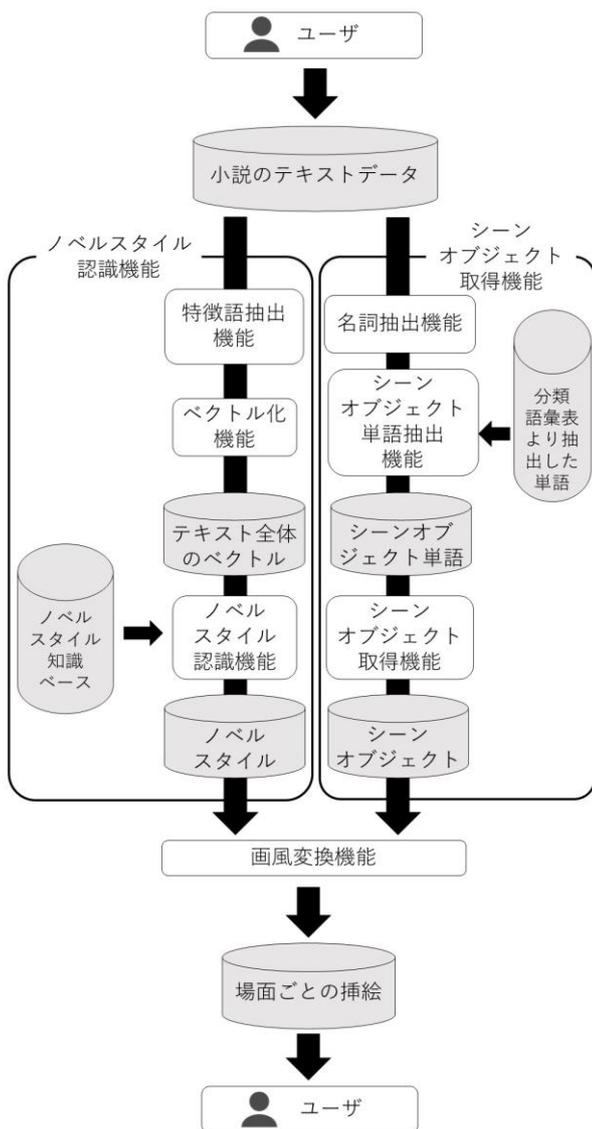


図1 提案方式の全体像

3. 小説を対象とした挿絵自動生成方式

本方式では、小説を対象とした物語の場面を表現する挿絵自動生成方式を示す。本方式の挿絵は、ノベルスタイルとシーンオブジェクトによって決定されそれらを組み合わせることによって生成する。

3.1. ノベルスタイルとシーンオブジェクト

ノベルスタイル及びシーンオブジェクトとは本稿において新しく定義した概念である。本稿においては挿絵を構成する要素としてノベルスタイルおよびシーンオブジェクトの2つを用いる。

ノベルスタイルとは小説全体における挿絵の画風のことである。各小説にはそれぞれホラーや恋愛といった小説のジャンルが存在している。また小説に挿入された挿絵に関しても小説の内容であるテキストデータにあった画風が用いられていることが多い。そのた

め挿絵を自動生成する際に内容にあった画風を定めるべきである。本方式では画風を定めるに当たり、事前にいくつかの小説コンテンツのテキストデータと表紙絵の組をノベルスタイル知識ベースとして準備したものを定める。そのノベルスタイル知識ベースを利用してユーザが入力したテキストデータと類似度の高い知識ベース内の小説を取得し、知識ベースにおける小説の表紙絵を本稿ではノベルスタイルと定め、ユーザが入力した各小説においてそれぞれノベルスタイルを作成し本方式における挿絵自動生成の際の画風として利用する。

シーンオブジェクトとは小説の各場面におけるその場面情報を表す画像のことである。挿絵とは挿絵が挿入された近辺に書かれているテキストデータと強い関係のある絵のことである。そのため挿絵の自動生成の際にそれぞれの場面であるシーンごとの内容と合った挿絵を生成する必要がある。これらは該当するシーンにおける場所や登場人物、登場人物の表情などが含まれている。本稿ではこれらシーンにおける場面情報を表す単語をシーンオブジェクト単語とし、その単語を利用して後述する手法を用いて生成したシーンオブジェクト単語と合致した画像に関してシーンオブジェクトと定義する。

これら手法を用いることで、各シーンと合致した挿絵を小説としてまとめたスタイルで生成することができる。

3.2. 提案方式の全体像

図1に提案方式の全体像を示す。本方式は入力としてユーザが挿絵を生成したいと思っている小説のテキストデータを受け取り、それぞれのシーンにあった挿絵となる画像を出力する。本方式は、形態素解析機能、ノベルスタイル認識機能、シーンオブジェクト取得機能、画風変換機能によって構成される。

3.3. ノベルスタイル認識機能

3.3.1. 特徴語抽出機能

ノベルスタイルを作成する際に入力されたテキストデータのジャンルを推定する必要がある。そのため本方式ではノベルスタイルに影響があると思われる名詞、形容詞、動詞、形容動詞の品詞を形態素解析エンジン MeCab[15]を利用し、形態素解析および分かち書きを行う事で抽出している。MeCab[15]に対応する単語区切り用の辞書として NEologd[16]を用いる。NEologd[16]は語彙の多さや更新頻度、新語の採録の早さを特徴した単語区切り用の辞書であり、新語を利用している確率の高い Web 小説に用いるのに適切であるためと考えられるため利用している。

3.3.2. ベクトル化機能

挿絵自動生成の際にユーザが入力したテキストとノベルスタイル知識ベース内のテキストデータとの類似度を測ることで、ユーザが入力したテキストデータに対するノベルスタイルを定めることが出来る。この類似度を測るに当たり、テキストデータを類似度の測れるデータに変換する必要がある。本方式ではテキスト同士の類似度を測るため、Python の機械学習ライブラリである Gensim[17]の Doc2Vec[18]を用いベクトル化する。Doc2Vec[18]では Word2Vec[11]とは違い単語の並び順を考慮してベクトル化を行うことができるため小説におけるベクトル化として適切であると考えられる。

3.3.3. ノベルスタイル知識ベース

ノベルスタイルの選定基準となるノベルスタイル知識ベースを定めるにあたり本方式においてはテキストデータの取得が容易であるという点や、新語等を利用した比較的新しい文章が多い点、サイトの主な趣向として若年層向けのライトノベルの取り扱いが多いため挿絵を必要としている可能性が高いと考えられる点から、小説投稿サイトである小説家になろう[1]におけるテキストデータを利用する。このサイトに投稿されている小説の内、表紙絵の取得が可能な書籍化作品の中から様々なスタイルやジャンルを網羅できるように11種類を選びノベルスタイル知識ベースとして定めている。本稿では11種類を選定する際に、画風が明らかに異なる作品を選択し、幅広いジャンルを網羅することを目的としている。そのため、様々な画風やより幅広いジャンルを網羅するために、ノベルスタイルを増強することも必要である。また一般的な小説1冊の分量として掲載されている文章のうち先頭約100,000文字を本方式におけるノベルスタイル知識ベースに利用するテキストデータと定める。

表1は本方式において定めたノベルスタイル知識ベースの一覧である。表1に記載されている正解ジャンルは小説家になろう[1]における著者によるタグ設定を引用しており、本方式における評価実験にのみ用いる。文字数に関しては2つの条件である100,000文字を超えること、小説家になろうにおける話の区切りであることを満たした箇所までをテキストデータとして取得しているためばらつきが存在している。

またノベルスタイル知識ベースに保存するテキストデータに関してベクトル化機能を用いる。

3.3.4. ノベルスタイル認識機能

ベクトル化機能を用いてユーザが入力したテキストデータに適用することで、ノベルスタイルデータベースにおけるテキストデータと類似度を測ることを可能にしている。類似度計算にはコサイン類似度を用い

る。求めた類似度を利用しユーザが入力したテキストデータと最も類似したノベルスタイル知識ベース内の小説の表紙絵をユーザが入力したテキストデータにおけるノベルスタイルと定める。

3.4. シーンオブジェクト取得機能

3.4.1. 名詞抽出機能

シーンオブジェクトを作成する際に入力されたテキストデータから場所等の情報を含む名詞を抽出する必要がある。そのため本方式では場所を示す単語が含まれている品詞である名詞を形態素解析エンジン MeCab[15]を利用し、形態素解析および分かち書きを行う事で抽出している。MeCab[15]に対応する単語区切り用の辞書として NEologd[16]を用いる。

3.4.2. 分類語彙表より抽出した単語

シーンオブジェクト生成の際に、そのシーンにおける情報を表す単語を抽出する必要がある。本方式では国立国語研究所コーパス開発センターが提供している分類語彙表[19]を利用し、表2で示した分類項目における単語を抽出しシーンオブジェクト単語抽出機能に用いる。抽出した単語に関してすべての語彙をカバーできている訳ではないが、単語数が3816語と膨大であるため実用に耐えうる網羅性を持つと判断できる。

3.4.3. シーンオブジェクト単語抽出機能

3.4.2節で抽出した単語と形態素解析機能を利用して抽出した名詞のみのテキストデータにおける単語が

表2 分類語彙表から抽出したシーンに関する単語

分類項目	見出し本体例1	見出し本体例2
空間・場所	冷暗所	行楽地
都会・田舎	都会	村里
固有地名	日本	オセアニア
社会・世界	人間界	桃源郷
現場	仕事場	漁場
社寺・学校	神社	大学
事務所・市場・駅など	銀行	発電所
店・旅館・病院・劇場など	露店	居酒屋
教室・兵営など	刑務所	選手村
住居	集合住宅	城
家屋・建物	駅ビル	仏閣
部屋・床・廊下・階段など	居間	図書室
地類(土地活用)	山林	馬場
道路・橋	線路	農道
その他土木施設	防空壕	釣り堀
山野	山脈	砂山
川・湖	運河	滝
海・島	深海	列島
地相	樹海	竹やぶ
地帯	北極	熱帯

一致している場合その単語を本方式におけるシーンオブジェクト単語と定義し抽出する。

またシーンごとのセグメンテーションに関しては本方式では Python の日本語のテキストに対してルールベースによる文区切りを行うライブラリである `ja_sentence_segmenter`[20]を利用し区切られた文 40 文ごとにシーン区切りを置いている。

3.4.4. シーンオブジェクト取得機能

定めたシーンオブジェクト単語に対応した画像を取得する事で各シーンにあったシーンオブジェクトを取得することができる。

画像の取得には Python のライブラリである `google-images-download`[21]を利用し、Web 画像検索を行い取得する。

また漢字のみで構成された単語を画像検索した際に中国語と誤判定され正確な表示が出ない場合がある。本方式では検索する単語に対して全角スペースおよび「がぞう」の文字を追加することで解決をしている。

3.5. 画風変換機能

取得したシーンオブジェクトに対して認識したノベルスタイルを用いてノベルスタイルにあったシーンオブジェクトに変換する事で挿絵を生成する。

この変換には画風変換と呼ばれる技術が用いられる。画風変換とはベースとなる画像を与えられた画像の画風に変換することである。現在ではディープラーニングを利用した画風変換[22]や CycleGAN[23]を利用した手法が考えられている。

本方式においてはベース画像となるシーンオブジェクトに CycleGAN の手法を用いてノベルスタイルを学習したモデルを作成し、そのモデルを利用して画風変換を行う。

CycleGAN の特徴として教師データが要らない点が上げられる。従来の画風変換の手法としてあげられる Pix2Pix[24]においては画像の学習の際に教師データとして変換前と想定される変換後の画像のペアを用いて学習させる必要がある。しかし CycleGAN においてはペアを作製する必要は無く、与えられた画像群のスタイルの特徴を学習することで画風変換を可能としている。本方式の目的である画風変換においては画像のペアを作製することが困難だと思われるため CycleGAN を利用する。実装に当たり CycleGAN の実装には `pytorch-CycleGAN-and-pix2pix` [25]を用いた。

以上により本方式の目的であったノベルスタイルに合ったシーンオブジェクトである挿絵を作成することが出来る。

4. 評価実験

4.1. 実験環境

ノベルスタイル知識ベースの生成には前述の通り小説投稿サイトである小説家になろうより表 1 に記載した小説を利用している。またシーンオブジェクト生成のための場所の情報の抽出も、表 2 に示すように分類語彙表を用いて取得している。

4.2. 実験1 ノベルスタイルの認識の検証

4.2.1. 実験目的

本実験では本方式によって定めたノベルスタイル認識手法に関して、ユーザが入力したテキストとノベルスタイル知識ベースとの間で正しい類似性が現れているか判断する。

小説投稿サイト小説家になろう[1]より著者が設定できる項目であるジャンルに関して、恋愛、SF、ファンタジー、ホラー、ヒューマンドラマ、歴史、推理に該当する作品を 10 万文字以上の条件のもと総合順位順にそれぞれ 100 作品選択し本方式における提案手法を利用しそれら小説のノベルスタイルを認識する。認識の際に類似されたノベルスタイルを類似度の高いものから各 2 つ取得し、正解ジャンルが本実験の際に定めたジャンルと等しいか評価を行う。

4.2.2. 実験結果

実験結果は表 3 に記載する。該当作品数とはそのジャンルにおいて 10 万文字以上という条件に合致している作品数である。

4.2.3. 実験考察

一致率の高いジャンルとして恋愛、SF、ファンタジー、歴史、推理があげられる。特に歴史のジャンルにおいては 97%と非常に高い一致率を示している。要因として歴史のジャンルにおいて利用される単語に関して、歴史以外のジャンルで利用されることが少なく、Doc2Vec を用いてベクトル化した際にそれら単語により強く類似されているのではないかと考えられる。11

表 3 実験 1 の実験結果

正解ジャンル	該当作品数 (件)	一致率 (%)
恋愛	12,093	68
SF	4,398	69
ファンタジー	35,044	81
ホラー	590	2
ヒューマンドラマ	3,293	10
歴史	1,293	97
推理	1,005	69

種類と少ないノベルスタイル知識ベースではあるが、高い一致率が得られた。

また一致率の低いジャンルとしてホラー、ヒューマンドラマがあげられる。ホラーにおいて本稿において定めたホラー作品は日本における田舎の少年が繰り広げるホラー作品である。しかし小説家になろうにおいてホラーのジャンルは幅広くゾンビやウイルス感染などのパニックホラー、圧倒的強者による生殺与奪の権利を奪われ強制的にゲームをさせられるデスゲーム、主人公がシリアルキラーやサイコパスである小説など様々な要素が存在しており、本稿において定めた作品のみでは正確に類似が出来ないのではないかと考えられる。またホラー作品では設定として学校を軸においた作品が挙げられており、それら要素が他のジャンルの作品と類似されているのではと考えられる。ヒューマンドラマにおいては、ヒューマンドラマの定義として人間らしさを主題とした作品とあげられているが、それぞれの作品により恋愛要素が強いものや時代背景としてファンタジーのジャンルの要素が用いられている物など他のジャンルの作品と比較的類似がされやすいためこのような結果となっていると考えられる。これらジャンルでは少ないノベルスタイル知識ベースの影響によりノベルスタイルを正しく認識出来ていない。

以上より本方式におけるノベルスタイル認識手法はノベルスタイル知識ベースにおいて種類が少ないため一致率が低いと思われる箇所があるが、本稿におけるノベルスタイル認識機能として妥当であると考えられる。

4.3. 実験2 シーンオブジェクト単語抽出の検証

4.3.1. 実験目的

本実験では本稿によって定めたシーンオブジェクト生成手法に関して、有効な抽出が行われているか判断する。

小説家になろうより[1]作品タイトル「隻眼・隻腕・隻脚の魔術師～森の小屋に籠っていたら早 2000 年。気づけば魔神と呼ばれていた。僕はただ魔術の探求をしただけなのに～」のテキストデータをより先頭 15,386 文字取得し、3.5.2 節の機能を用いて区切られたシーン 10 個に対して、本稿によって定めたシーンオブジェクト抽出機能を利用しシーンオブジェクト単語の抽出を行う。また同様の区切られたシーンごとの小説に対して被験者 1 名が読書を行い、シーンを想像する際に役に立ったテキストを抽出する。

4.3.2. 実験結果

本稿によって定めたシーンオブジェクト抽出機能を利用したシーンオブジェクト単語の抽出実験結果は表 4 に記載する。被験者が読書を行い、同様のシーンオ

表 4 実験 2 の実験結果 1

シーン	本稿におけるシーンオブジェクト単語抽出機能
1	['村', '地帯', '村', '村', '村']
2	['家']
3	['間', '洗面所', '水流', '玄関', '森', '朝', 'てっぺん', '朝', '村']
4	['山', '部屋', '寝室', 'リビング', '家', '玄関', '手洗い', '場', '床', '床', '村']
5	['村', '村', '前線', '朝', '村']
6	['場', '場', '村', '家', '玄関', '家', '村']
7	['村', '村', '村', '場所', '場', '村', '村']
8	['麦畑', '家', '山', '山', '山', '中', '村']
9	['地獄', '現世', '間', '村', '地獄', '現世', '地獄', '地獄', '出会い', '場', '村', '家']
10	['部屋', '山', '道場', '道場', '朝', 'キッチン', 'お湯', 'リビング', '階段', '寝室']

表 5 実験 2 の実験結果 2

シーン	被験者によるシーンオブジェクト単語抽出
1	['村', '白い草原', '村', '柵で囲われた村']
2	['ひらけた場所', '家']
3	['家', '玄関', '洗面所', 'リビング', 'リビング']
4	['家', '寝室', '燃えている村']
5	['戦場', '火災']
6	['崩れ落ちる建物']
7	['燃えてる建物', '村', '廃村']
8	['焼け崩れた家', '瓦礫の山']
9	['瓦礫の山', '村', '家', '寝室']
10	['リビング', '家', '寝室']

表 6 実験 2 の混合行列

		シーンオブジェクト単語抽出機能	
		抽出された単語	抽出されていない単語
被験者によるシーンオブジェクト単語抽出	抽出された単語	村, 家, 玄関, 洗面所, リビング, 寝室	白い草原, 柵で囲われた村, 開けた場所, 燃えている村, 戦場, 火災, 崩れ落ちる建物, 廃村, やけ崩れた家, 瓦礫の山
	抽出されていない単語	地帯, 間, 水流, 朝, てっぺん, 山, 手洗い, 場, 床, 前線, 場所, 麦畑, 地獄, 現世, 出会い, 道場, キッチン, お湯, 階段	

ブジェクト単語の抽出実験結果は表 5 に記載する。これら結果を用いて作製した混合行列を表 6 に記載する。

4.3.3. 実験考察

本方式における機能では抽出されたが被験者には抽出されていない単語として「麦畑，地獄」があげられる。これらは「麦畑のような金色の髪」「地獄から救い出す」といった比喻表現において使われており，シーンを表す単語ではないが抽出されている。また「地帯」は寒冷地帯といった単語が分割されて形態素解析されたことにより正しく抽出出来ていない。

逆に本方式における機能では抽出されていないが被験者には抽出されている単語では形容詞を含んだシーンオブジェクト単語があげられる。

以上のことより，比喻表現を見分けることが出来ない課題などが挙げられるが，シーンオブジェクト単語を抽出することが出来ている。また今後の展望として，本方式における機能において抽出される単語を増やす事が重要だと考えられる。本方式における機能では抽出されていないが被験者には抽出されている単語に関してはシーンオブジェクト単語に付随する形容詞を取得するなどの機能を追加することで解決出来ると考えられる。

4.4. 実験 3 画風変換機能の検証

4.4.1. 実験目的

本稿によって定めた画風変換機能に関して，有効な変換が行われているか判断するために実験を行う。

前述した実験 2 の作品を用いて実験を行う。実験 2 において取得したシーンオブジェクト単語である村，森，キッチンを中心にシーンオブジェクトと定める。そのシーンオブジェクトに対して 3.3 節で定めた機能を用いて取得したノベルスタイルを適用させ画風変換を行う。今回取得されたノベルスタイルは「真の仲間じゃないと勇者のパーティーを追い出されたので、辺境でスロウライフすることにしました」である。

4.4.2. 実験結果

CycleGAN の学習において利用したノベルスタイルを表 7 に示す。本実験により生成されたシーンオブジェクトおよび，生成された挿絵を表 8 に示す。

4.4.3. 実験考察

結果より変換前の挿絵と比べて変換後の挿絵では比較的明度が高い画像に変換されていることが解る。この結果はノベルスタイルとして与えた画像は明度が高いことに起因していると思われる。学習に利用したノベルスタイルの枚数が 9 枚と少ないが妥当な画風変換が行われていると考察出来る。

5. まとめ

本稿ではノベルスタイルとシーンオブジェクトを

表 7 実験 3 に利用したノベルスタイル



表 8 実験 3 の実験結果

	変換前	変換後
村		
森		
キッチン		

用いた挿絵の自動生成方式について述べた。本方式における挿絵は小説文章からシーンオブジェクトを生成し，それに対してノベルスタイルを用いて画風変換することで生成される。

また，本方式を検証するためにノベルスタイルを定める際に利用した類似度の有効性，シーンオブジェクトの生成における分類語彙表を利用した抽出機能の有効性，画風変換機能の有効性を検証する実験を行った。

本方式を用いることで小説の著者が挿絵を生成するのを容易することができると考えられる。

今後の課題として，各シーンにおける感情を考慮し

たスタイルの生成やシーンオブジェクトの生成の際に時間帯や登場人物にあった挿絵の生成などがあげられる。またより実用的な機能として実際にテキストデータに対して挿絵を適切な位置に自動挿入することも考えられる。

参 考 文 献

- [1] 小説家になろう <https://syosetu.com/>
- [2] カクヨム <https://kakuyomu.jp/>
- [3] 田中麻巳. (2016). 小説の読みに与える挿絵の影響. 読書科学, 58(2), 74-86.
- [4] 和田裕一. (2019). 挿絵が物語文の読解における状況モデルの構築に及ぼす影響. 心理学研究, 90(4), 368-377.
- [5] Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49, 294–303
- [6] Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995a). The construction of situation models in narrative comprehension: An Event-Indexing Model. *Psychological Science*, 6, 292–297.
- [7] Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185.
- [8] 菅生健介, & 萩原将文. (2014). 感性を考慮した自然言語文からの風景画像生成システム. 日本感性工学会論文誌, 13(2), 371-379.
- [9] 西山晴彦, 大久保達真, & 松下温. (1997). Picnyck: 風景描写文から風景画像の創造. 情報処理学会論文誌, 38(5), 997-1007.
- [10] 村山貴志, 入江英嗣, & 坂井修一. (2020) Word2Vecを用いた挿絵自動挿入手法の提案. 情報処理学会インタラクシオン 2020 論文集, 377-380
- [11] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
- [12] Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., & He, X. (2018). Attngan: Fine-grained text to image generation with attentional generative adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1316-1324).
- [13] Qiao, T., Zhang, J., Xu, D., & Tao, D. (2019). Mirrorgan: Learning text-to-image generation by redescription. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1505-1514).
- [14] 五十嵐晃, & 上岡英史. (2014). シナリオの内容を反映した挿絵の生成手法. 研究報告モバイルコンピューティングとユビキタス通信 (MBL), 2014(5), 1-6.
- [15] MeCab: Yet Another Part-of-Speech and Morphological Analyzer <https://taku910.github.io/mecab/>
- [16] Neologd <https://github.com/neologd/mecab-ipadic-neologd>
- [17] Gensim <https://radimrehurek.com/gensim/>
- [18] Le, Q., & Mikolov, T. (2014, June). Distributed representations of sentences and documents. In International conference on machine learning (pp. 1188-1196). PMLR.
- [19] 国立国語研究所コーパス開発センター | 分類語彙表 <https://ccd.ninjal.ac.jp/goihyo.html>
- [20] ja_sentence_segmenter <https://pypi.org/project/ja-sentence-segmenter/>
- [21] google-images-download <https://github.com/hardikvasa/google-images-download>
- [22] Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.
- [23] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision (pp. 2223-2232).
- [24] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).
- [25] Jun-Yan Zhu and Taesung Park pytorch-CycleGAN-and-pix2pix | GitHub <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>