

指導教員（主査）：山本祐輔 准教授
副査：杉山岳弘 教授，大木哲史 准教授

2021 年度 静岡大学大学院総合科学技術研究科 修士論文

芸術作品に興味を促すストーリーテリング

静岡大学 大学院総合科学技術研究科 情報学専攻 所属
学籍番号 71030046

永野 里佳奈

2022 年 2 月

概要

本稿では芸術作品を敬遠している人々が興味を持てるよう、絵画画像およびそれに紐づくキャプションを並び替え、鑑賞者に提示する手法を提案する。提案手法では、すべてのキャプションを通して1つのストーリーを作ることで、鑑賞者の芸術に対する抵抗感を減らし、興味を引きやすくすることで、作品を印象付けることを狙う。具体的には、まず2種類のキャプションモデルから作品画像にキャプションを生成する。キャプションの種類を複数用いることで、ストーリーの流れに変化を持たせる。次に、接続判定によってキャプションを選択し、順序づける。接続判定は、ストーリーテリング用のデータセットを用いて、起承転結の各ストーリー構成要素を反映できるようにした。また、結部分に詩的な表現を取り入れることによって、最後にオチが生まれるようにし、興味の持てるストーリーを作成できることを狙った。評価実験の結果から、提案手法のビジュアルストーリーと比較手法の解説文では大きな違いが得られなかった。一方、評価者の芸術に対する関心の高さの違いから、評価結果に差があることが明らかになった。この要因として、ビジュアルストーリーをきっかけとして、ストーリーを想像せず評価していたことが考えられる。今後は、想像力を補えるようキャプションやストーリー生成方法について改善を行う必要がある。

目次

第 1 章	はじめに	5
第 2 章	関連研究	8
2.1	美術教育	8
2.2	ICT 技術を利用した美術鑑賞支援	9
2.3	ビジュアルストーリーテリング	9
第 3 章	提案手法	11
3.1	芸術作品データセットの作成	12
3.2	利用するキャプションモデル	13
3.3	起・転・結の要素を並び替えに反映する接続判定器の構築	14
第 4 章	評価実験	18
4.1	タスクの流れ	18
4.2	評価用ビジュアルストーリーの選択	19
4.3	評価用画像および解説文の用意	19
4.4	提案手法および比較手法による芸術への興味・関心	19
4.5	事後アンケート	20
4.6	実験協力者	20
4.7	実験協力者に示す出力例	20
第 5 章	結果	24
5.1	各手法の評価結果	24
5.2	実験協力者が持つ芸術への関心の強さごとの評価結果	26
5.3	提案手法の出力結果例	26

第 6 章 考察	32
6.1 提案手法の有効性	32
6.2 提案手法の問題点	33
第 7 章 まとめ	36
参考文献	38

図目次

1.1	提案手法の出力例	7
3.1	提案手法の概要図	12
3.2	一般画像キャプションモデルと 詩的画像キャプションモデルの出力例	14
3.3	Visual Storytelling Dataset (ViST) の例	15
3.4	「起—結」接続判定器の学習データ	16
4.1	画像群_1 の提案手法（上）と比較手法（下）の出力結果	21
4.2	画像群_2 の提案手法（上）と比較手法（下）の出力結果	22
4.3	画像群_3 の提案手法（上）と比較手法（下）の出力結果	23
5.1	提案手法の出力結果の中で、ストーリー性の評価値が高かった例（画像 群_2, 画像群_4, 画像群_10）	28
5.2	提案手法の出力結果の中で、面白さの評価値が高かった例 (画像群_6, 画像群_10)	29
5.3	提案手法の出力結果の中で、興味喚起度の評価値が高かった例 (画像群_1, 画像群_2)	30
5.4	提案手法の出力結果の中で、興味喚起度の評価値が低かった例 (画像群_11, 画像群_17)	31

表目次

3.1	各接続判定器作成の正解ペア、不正解ペアの組み合わせ	16
3.2	各接続判定器作成の分類性能の評価結果	17
5.1	各手法の評価結果	24
5.2	各ビジュアルストーリーの評価結果	25
5.3	実験協力者が持つ芸術への関心の強さごとの評価結果	26

第1章

はじめに

芸術は、人々の生活や文化、教育面など様々な面で重要な役割を持っている [1]。芸術鑑賞は、人々の感性や創造力を豊かにし、精神的満足を高められることが期待されている。また、アーティストの創造力に基づいた思考や問題提起は、アート思考としてビジネス面にも生かすことができると注目され始めている [2]。

2020年から現在に至るまで、コロナ禍の影響を受けて、博物館や美術館等の施設への来館頻度が減少している [3]。その一方で、情報技術の発達により、芸術に触れる新しい機会が生まれている。例えば、Google Arts & Culture や各美術館が提供するストリートビューなど、オンライン鑑賞が盛んに行われるようになった。

しかし、このような機会があるものの、文化庁の調査 [4] によれば、過去1年間で文化芸術を「まったく・ほとんど鑑賞していない」と回答した人が46.1%であった。その回答をした理由として、「関心がないから」(35.4%)、「特にない・分からない」(22.8%)が上位に挙げられた。このことから、人々が芸術を鑑賞しない理由は、不満による具体的なものではなく、芸術にそもそも興味・関心を持っていないことが原因であると考えられる。

本稿では、これまで芸術作品を敬遠してきた人々が、興味を持てるようになることを目的とする。そのためには、絵画に焦点を当て、作品からキャプションを自動生成し、複数の作品群を読みたくなるようなストーリー立てで並び替え、鑑賞者に提示する手法を提案する。4コマ漫画のようにストーリー化することによって、鑑賞者が芸術作品へ親近感や共感できることが期待される。このように芸術への抵抗感を減らすことで、興味を持つようになることを狙う。

一般的な芸術鑑賞の場とされる美術館では、作品を理解するための手がかりとして解説文がつけられる。例えば、1枚目の絵画作品には「悲嘆に暮れる表情や大胆な仕草が、峻厳な金地に浮かび上がり、キリストの受難の感動をドラマチックに伝えてくれる精緻な絵

付けのパネルである。^{*1}」という解説文が作品につけられ、展示される。このような解説文は、作品の歴史や背景などが中心であり、専門用語も多く含まれるため、身近に感じにくい内容といえる。

美術館のストリートビューを含め、館内での体験向上を狙ったアプローチ [5][6][7][8] は、様々に提案されている。これらの対象は美術館訪問者であり、芸術にある程度興味がある人に向けたものである。芸術に興味がない人は、必然的に鑑賞機会や作品の見方、背景知識について興味がある人よりも少なく、既存のアプローチでは芸術への興味に繋がりづらい。そのため、芸術に興味がない人に向けた新しいアプローチが必要であると考えられる。

提案手法では、まず入力画像からキャプションを自動生成するモデルを作成する。MS COCO[9] の画像とアノテーションで学習をした画像キャプションモデルと、詩を出力する画像キャプションモデルの 2 種類を用いる。芸術作品画像および出力された 2 種類のキャプションを、提案手法で利用するデータセットとする。出力に面白みを持たせるため、すべて同じキャプションモデルではなく、異なるものを使用する。RoBERTa[10] の隣接文予測を利用した接続判定器を構築することにより、出力される 3 つのキャプションの組み合わせが、ストーリーとして成立するようになる。

具体的な提案手法の出力例を図 1.1 に示す。1 番目から 3 番目まで通じたストーリーを考えると、窓から 2 人組が、もう人が帰ることのない家に飾られたクリスマスツリーを眺めている、という流れをイメージすることができる。芸術作品画像の情報を用いてストーリー作成することで、作品に関する知識を持たない人でも理解しやすく、身近に感じられる見方を提示する。

提案手法による出力を美術館のポスターや、パンフレットなどのコンテンツとして利用する可能性が考えられる。ポスターやパンフレットは、不特定多数の人の目に触れる場所に掲示されるため、これまで芸術に興味がない人の目にも入る。このとき、提案手法の出力によって展示作品を分かりやすく印象づけることで、芸術への興味に繋がりやすくなると考えられる。

^{*1} <https://www.metmuseum.org/art/collection/search/471905>

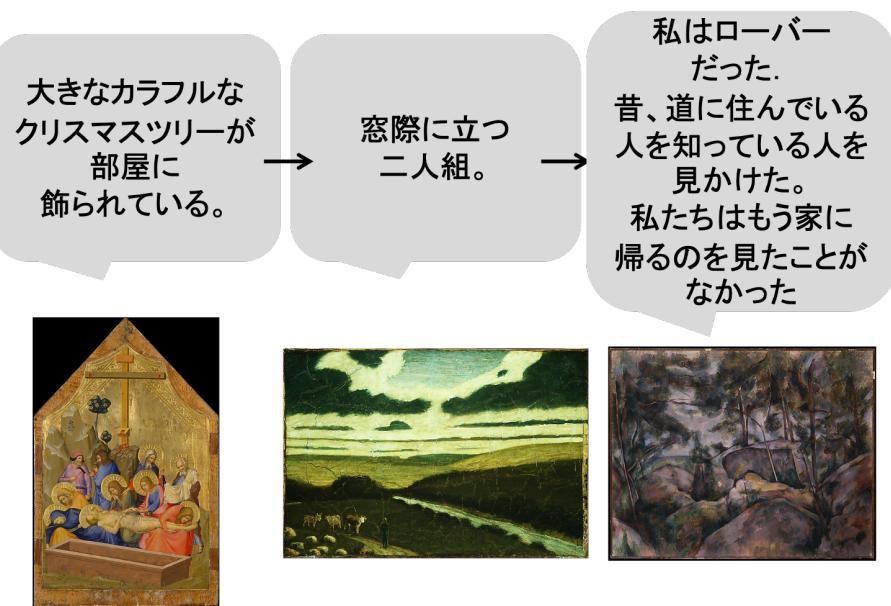


図 1.1 提案手法の出力例

第2章

関連研究

2.1 美術教育

まず、一般的な人々が芸術を敬遠してしまう要因として、芸術に対するネガティブなイメージがあると考えられる。縣・岡田[11]は、芸術への動機づけを考察するため、美術に関わりのない人に対し、アートのイメージについて調査した。その結果、アートの必要性は認識されているものの、難解なイメージや、自分とは関係のない才能や技術のある人のものというイメージを持たれていることが分かった。本稿では、このようなネガティブなイメージを低減できるようにし、芸術鑑賞の動機づけを高める。

知識のない学習者は鑑賞の方向性が曖昧になるため、対話を汲み取りつつ、適切に作品理解へ導くことが必要である[12]。しかし、実際の美術教育は、学習者の裁量によるものが多い状態である。Zimmerman[13]は、美術教育による創造性育成の課題について述べている。美術教育の現場では、創造性は生まれつきのものであるとされ、教師は生徒の創作活動に直接介入するものではなかった。さらに金子[14]は、日本の美術教育が、学習者の自己完結で行われている点を指摘した。学習を受ける側の創作活動が、教育そのものに直結するとされ、教師や教育内容の存在が失われている。このように、現代の美術教育は、学習者個人に委ねられていることが問題として挙げられる。そのため、外部からの学習支援が適切に得られず、学習者の理解や興味・関心に個人差が生まれやすくなっていると考えられる。

本稿では、美術教育を受けていたものの、これまで芸術に理解や興味を持ちにくかった人にも楽しめられるよう、専門的な知識を必要としないストーリー形式の鑑賞支援を行う。

2.2 ICT 技術を利用した美術鑑賞支援

近年, ICT 技術の発展により, 新しい鑑賞支援が提案されている。例えば, 作品や美術館での鑑賞体験を自分自身や近しい人と関連づけることで, 新しい価値を示したものがある。Spence ら [5] は, 美術館の作品を親しい人にギフトとして贈るアプリケーションを提案した。アプリケーション利用者に, 相手に贈るために作品を選ぶという, 作品鑑賞の新しい視点を提供している。

Fosh ら [6] は, 訪問者が同伴するパートナーに向けて, 展示作品の鑑賞方法をデザインする手法を提案した。評価実験の結果, パートナーは訪問者がデザインした鑑賞方法を肯定的に受け入れた。大衆に向けた通常の鑑賞とは異なる, パートナー個人に向けた鑑賞を体験できることが明らかになった。Muntean ら [7] は無形の民族文化を理解し, 体験する展示設計とシステムを提案した。提案システムは, 民族文化の展示物がどのような使われ方をしたかについて, 現代の物との関連性から訪問者に探索させることを企図している。これらの関連研究は, 自分と作品との関連性を考えさせることによって, 作品に個人的な意味を持たせている。

また, 訪問者を魅了しつつ, 教育面の充実も考慮した展示内容も考案されている。Mallavarapu ら [8] は, 訪問者の動きによって映像が変化する体験型展示に対して, 視覚的なフィードバックを行う展示方法を提案した。フィードバックによって, 訪問者は展示のコンセプトを理解し, 映像の変化の原因を考えるようになった。

このように, 訪問者にとって美術館の体験がより近しいものになり, 学習効果が高くなるよう ICT を活用した様々なアプローチが提案されている。しかし, これらの提案は美術館訪問者に向けたものであり, ある程度芸術に興味・関心のある人が対象となっている。本稿では, 美術館や博物館への来館意欲を高めるために, 芸術に関心のない人に芸術に対する興味を持たせる方法を提案する。

2.3 ビジュアルストーリーテリング

ビジュアルストーリーテリングは, 複数の画像の並びからストーリーを自動生成する手法である。Li ら [15] は, ストーリーテリングの新たな要素として, 画像から抽出した感情の特徴をストーリー生成に取り入れるモデルを作成した。これまでモデル作成には, 画像の視覚情報のみを利用してきていたが, 人間が行うように同じ画像でも異なるストーリーが生まれるよう, 感情に着目した。Yu ら [16] は, フォトアルバムが与えられたときアルバム

の中で代表的な写真を選択し、ストーリーを作成するモデルを提案した。3種のRNNで構成されたモデルを用いて、アルバム内写真のベクトル化・代表的な写真の選択・ストーリーの作成を行った。Krauseら[17]は、1枚の画像からストーリーとして捉えられる複数の文章を自動生成するモデルを提案した。通常の画像キャプショニングでは、入力画像から1文を生成するため、限られた画像情報しか表現できなかった。Krauseらはこの問題を解決するため、文章ベクトル生成用と単語生成用で2種のRNNを作成し、複数の文章を生成した。

興味を駆り立てる画像キャプショニングについて様々な研究がされている[18][19]が、本稿では事実を述べるよりも見た人を惹きつけ、記憶に残りやすいと考えられる[20]ストーリーテリングを参考にする。実際に、ストーリーテリングはマーケティングや広告にも活用され、商品価値を高める効果がある。Kangら[21]は、ラジオCMにおいて、ストーリーテリングの効果を調査した。結果、ストーリーテリングは、広告に対する好意的な反応と正の相関があることが明らかになった。また、複数の画像を使うため、画像キャプショニングよりも一回で知ることができる作品数が増える利点もある。本稿では、複数の画像キャプションを物語形式に並べることで、鑑賞者に芸術作品を印象づけ、興味を持たせることを狙う。

第3章

提案手法

本章では、鑑賞者の興味をひく絵画作品群のビジュアルストーリー生成の方法について述べる。ここでビジュアルストーリーとは、提案手法により生成された3つの絵画作品と文章の組み合わせと定義する。提案手法の概要図は以下の通りである（図3.1）。提案手法は、画像キャプショニングと文接続判定を組み合わせ、キャプションとその絵画を並び替えることでストーリーを生成する。ストーリーは、3つのキャプションと絵画で構成される。これらは、ストーリー構成の種類である「起承転結」の起・転・結部分を表している。承は起の補足内容になるため、本稿では省略し、ストーリー構成に最低限必要な3つに絞った。また、鑑賞者が芸術作品に興味を持てるよう、提案手法による出力は、次の条件を満たすようにする。

- 内容に面白さを含んでいる
- キャプションの並びがストーリーとして成り立つ

具体的には、次の手順でストーリーの生成を行う。

1. 事前に起・転用と結用の画像キャプションモデルを構築する。
2. 事前に芸術作品画像を(1)のキャプションモデルに入力した結果を、芸術作品データセットとしてデータベースに格納する。
3. 事前にRoBERTa[10]の隣接文予測に対して、各ストーリー構成にふさわしい前後関係かを予測するようにファインチューニングを行い、接続判定器を構築する。
4. 芸術作品データセット内のキャプションの全組み合わせを接続判定する。
5. 「起一結」接続判定器、「起一転」接続判定器、「転一結」接続判定器が最も高い組み合わせとなるキャプションと、それに紐づく芸術作品画像を合わせて出力する。

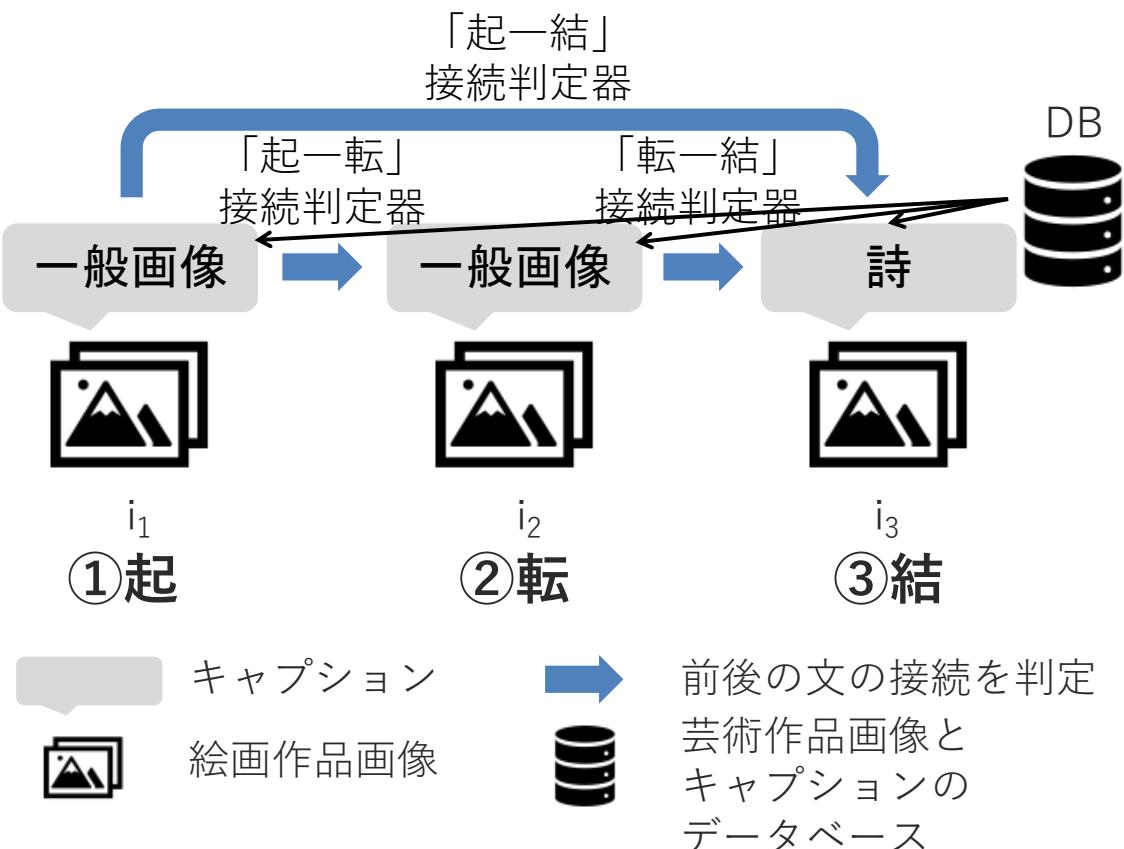


図 3.1 提案手法の概要図

以下、提案手法を構成する各要素の詳細について述べる。

3.1 芸術作品データセットの作成

本稿では、ビジュアルストーリーの題材としてメトロポリタン美術館^{*1}のパブリック・ドメイン画像の、Painting カテゴリに属する 3,080 枚の中から、ランダムに 100 枚を抽出して用いる。この画像および画像を 2 種類のキャプションモデルに入力した結果を、データベースに格納しビジュアルストーリー生成の素材とする。

^{*1} <https://www.metmuseum.org/about-the-met/policies-and-documents/open-access>

3.2 利用するキャプションモデル

提案手法のキャプションモデルは、次の2種類を使用する。

- 一般画像キャプションモデル
- 詩的画像キャプションモデル

1番目の起、2番目の転には一般画像キャプション用い、3番目の結には詩的画像キャプションを用いる。

2種類用いるのは、1つ目の条件である「内容の面白さを含んでいる」を満たすためである。一般画像キャプションは、画像の内容を直接説明する文章である。一方の詩的画像キャプションは、画像内容をシンボルとして、情景や感情などを加味した間接的な表現となっている。提案手法では、最後の「結」のみに「起」・「転」の説明文とは異なる意外な表現の詩的画像キャプションを持ってくることで、ストーリーにオチを作ることを狙う。ストーリー全体の流れに変化を持たせることで、面白さを加えることができると考えられる。

一般画像キャプションモデルは、一般的な画像とそのキャプションを用いて構築する。一般画像キャプションモデルの構築には、画像認識に Convolutional Neural Networks (CNN)，文章生成に Long short-time memory (LSTM) を利用した [22]。データセットは、MS COCO[9] の各画像とそれに紐づくキャプション 617,394 件を用いて学習を行った。

詩的画像キャプションモデルは、Liu ら [23] が提案する自由形式詩自動生成モデル img2poem を利用した。img2poem は、詩を自動生成するため、画像から詩のインスピレーションを得られるシンボルを抽出する。そして、詩と画像情報それぞれを物体・感情・情景の側面から分類し、関連づけて学習を行っている。

例えば、作品画像「アルプスを越えるボナパルト」を入力した場合、各キャプションモデルの出力は次の通りである（図 3.2）。一般画像キャプションモデルでは、「a man and woman are riding on a horse. (男性と女性が馬に乗っている。)」という文章が出力される。詩的画像キャプションモデルでは、「i am a coal-truck waiting for a time time i will go to the chalk water and look back down the way (私は時を待つ石炭トラック 私は白亜の水に行き、道を振り返るだろう)」という文章が出力される。画像内容を直接的に説明する一般画像キャプションと比べて、詩的画像キャプションは、心情などの画像内容を元に想像される要素を加えた文章である。

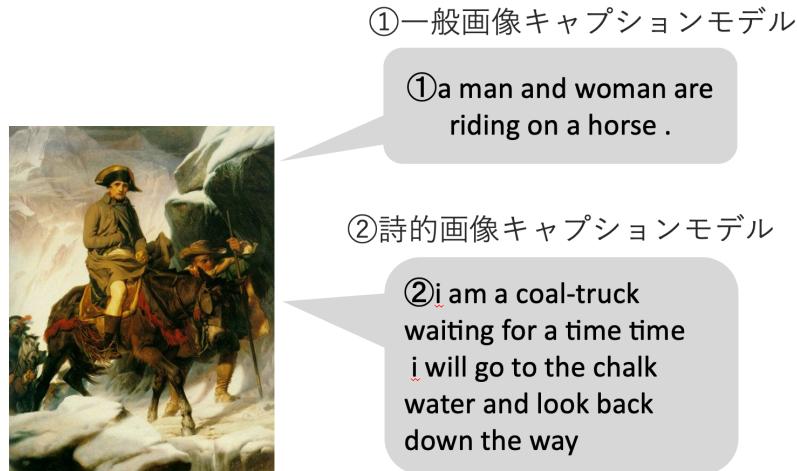


図 3.2 一般画像キャプションモデルと
詩的画像キャプションモデルの出力例

3.3 起・転・結の要素を並び替えに反映する接続判定器の構築

2つ目の条件である「キャプションの並びがストーリーとして成り立つ」を満たすための接続判定器について述べる。本稿では、3つのキャプションが1つのストーリーとして成立するよう、各ストーリー構成(起—転、転—結、起—結)の前後関係に着目する。そのために、RoBERTa の事前学習済みモデルの隣接文予測を用いた接続判定器を構築する。

接続判定器では、ストーリーを構成する上での「起・転・結」の関係性を学習させる。そのための学習データとして、ストーリーテリングのデータセット Visual Storytelling Dataset (以下、VIST と表記する) [24] を用いる。VIST は、5枚一組の画像セットがストーリーになるよう、各画像に対しアノテーションが付けられたものである (40,155組)。本稿では、5文の中で1文目が「起」、3文目が「転」、5文目が「結」を構成しているものとして捉える。提案手法では、以下の3種類のテキストペアを VIST の各ストーリーから抽出し、学習データとして用いる。

- 接続判定器で扱いたい関係性のペア（正解ペア）

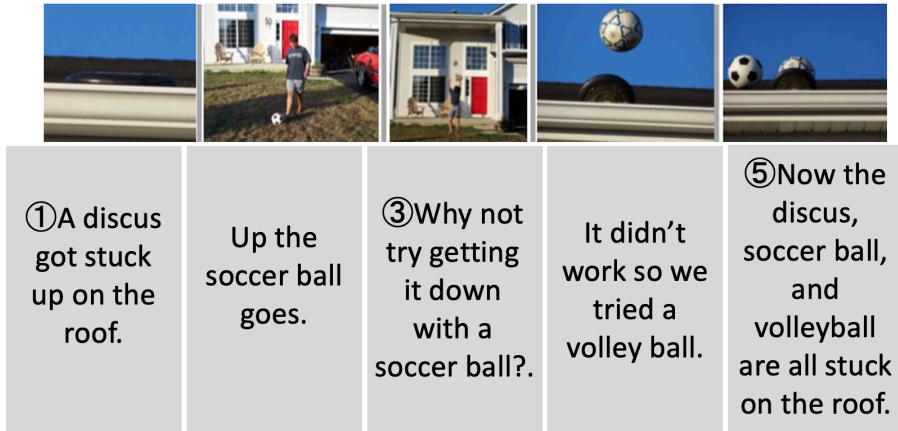


図 3.3 Visual Storytelling Dataset (VIST) の例

2. (1) の順序を逆にしたペア (不正解ペア)
3. (1) と始まり, もしくは終わりのどちらか一方が異なるペア (不正解ペア)

(1), (2) によって, テキストの組み合わせだけでなく前後関係を判断することを狙う. さらに (1), (3) によって, 前後関係の判定精度を上げることを狙う. 例えば, 「起一結」接続判定器の場合の学習データは次のようになる (図 3.4).

1. 1 文目・5 文目
2. 5 文目・1 文目
3. 1 文目・3 文目

実際に用いる各接続判定器の正解データおよび, 不正解データの組み合わせを以下に示す (表 3.1). また, 今回構築した各接続判定器の分類性能の評価結果について, 表 3.2 に示す.

「起一結」接続判定器の場合

正解ペア

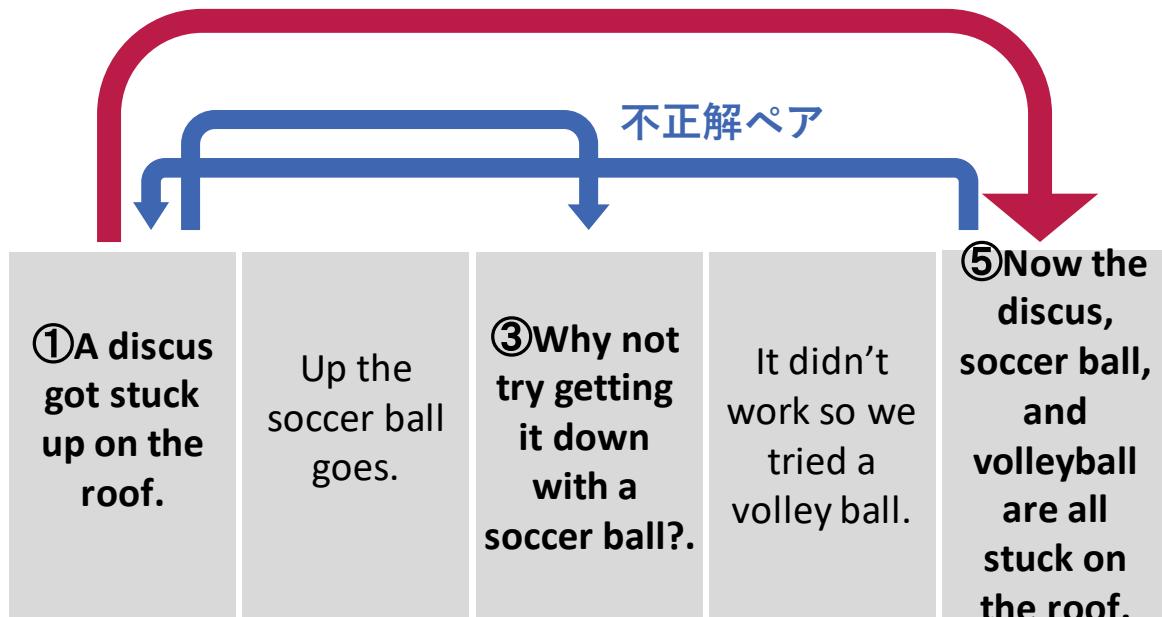


図 3.4 「起一結」接続判定器の学習データ

表 3.1 各接続判定器作成の正解ペア、不正解ペアの組み合わせ

	正解ペア (40,155 件)	不正解ペア (80,310 件)
「起一結」接続判定器	1 文目・5 文目	5 文目・1 文目 1 文目・3 文目
「起一転」接続判定器	1 文目・3 文目	3 文目・1 文目 1 文目・5 文目
「転一結」接続判定器	3 文目・5 文目	5 文目・3 文目 1 文目・5 文目

表 3.2 各接続判定器作成の分類性能の評価結果

	正解率	適合率	再現率	MCC (マシユーズ相関係数)	balanced ACC (平均正解率)	f1 (F1スコア)
「起—結」接続判定器	0.79	0.74	0.56	0.50	0.73	0.64
「起—転」接続判定器	0.76	0.65	0.64	0.44	0.72	0.63
「転—結」接続判定器	0.77	0.71	0.54	0.46	0.71	0.61

第4章

評価実験

本章では、提案手法の有効性を確かめるための評価実験について述べる。評価実験ではこのビジュアルストーリーおよび、比較手法となる絵画作品と解説文の組み合わせを複数の評価者に評価してもらった。

4.1 タスクの流れ

実験協力者に行ってもらったタスクは、以下の流れで行った。

1. 提案手法で生成された 20 個のビジュアルストーリーに対する評価
2. 比較手法で用意した 20 個の画像および解説文に対する評価
3. 事後アンケート

ステップ 1 とステップ 2 では、「出力された文章と絵画全体の面白さ」「出力された文章の組み合わせのストーリー整合性」「絵画作品への興味喚起」の 3 つの観点からビジュアルストーリーの評価を行った。ステップ 3 では、事後アンケートとして実験協力者の性別、学年、現在持っている芸術への興味関心、美術館への来館頻度等を聞いた。

実験協力者に提示した説明文は、以下の通りである。

このアンケートでは、絵画作品につけられた文章が、どの程度作品に興味を持たせるかどうかを評価していただきます。

アンケートでは、3 つの絵画作品とそれを説明した文章の組み合わせを提示します。これらを閲覧していただき、

- 文章と絵画作品全体の面白さ

- 提示文の組み合わせがストーリーとして成り立っているか
 - 絵画作品をどの程度知りたくなったか
- について、最も当てはまるものを選択肢から回答してください。
質問は全部で 40 問あります。回答時間の目安は、15 分です。

4.2 評価用ビジュアルストーリーの選択

提案手法として、実験協力者に 20 個のビジュアルストーリーを提示した。評価用のビジュアルストーリーは、以下の方法で 20 個に絞り込んだ。

1. 各接続判定における接続していると予想される確率の積を計算し、これをランクイングスコアとする。
2. ランキングスコア順にすべてのビジュアルストーリーを並び替える。
3. 起にあたる 1 枚目の画像が異なるよう、ビジュアルストーリー上位 20 件を選択し、これを評価用ビジュアルストーリーとした。

4.3 評価用画像および解説文の用意

比較手法として、4.2 における評価用ビジュアルストーリーと同様の 3 つの絵画作品に対して、各解説文を付与したもの用意した。なお、解説文はメトロポリタン美術館のウェブサイト^{*1}から取得した。

4.4 提案手法および比較手法による芸術への興味・関心

提案手法および比較手法の解説文それぞれについて、どの程度芸術に興味を持つことができたかを 3 つの観点から調べた。

1. 出力された文章と絵画全体の面白さ
2. 出力された文章の組み合わせのストーリー整合性
3. 絵画作品への興味喚起

そのために実験協力者に以下の質問をした。

^{*1} <https://www.metmuseum.org/art/collection>

1. 「提示された文章と絵画作品の全体を見て、どの程度面白いと感じたか」
2. 「提示された文章の組み合わせがストーリーとして成り立っているか」
3. 「提示された文章と絵画作品の全体を見て、提示された 3 つの絵画作品の作者や背景などについて、インターネットや書籍を利用し、どの程度知りたくなったか」

アンケート結果は、5段階のリッカート尺度 [25] で測定する。例えば 1 つ目の質問であれば、5=とても面白い、4=やや面白い、3=どちらともいえない、2=あまり面白くない、1=全く面白くない、の選択肢から回答してもらった。

また、提案手法と比較手法で提示する順番が与える影響を考慮し、「提案手法・比較手法」の順と、「比較手法・提案手法」の順に見せる場合の 2 つに実験協力者を 3 名ずつ分けた。

4.5 事後アンケート

実験協力者がどの程度芸術に興味を持っているかを、以下 3 つの質問で調査した。

「これまでの平均的な 1 年間の美術館の来館頻度」と、「美術館以外（オンライン上など）での絵画作品の鑑賞頻度」の 2 つの質問については、6 段階で回答してもらった(6=月数回以上、5=月 1 回程度、4=2~3 ヶ月に 1 回程度、3=半年に 1 回程度、2=年に 1 回程度、1=全く・ほとんどない)。3 つ目の「現在持っている芸術への興味・関心度合い」という質問には、5 段階で回答してもらった(5=ともある、4=ややある、3=どちらともいえない、2=あまりない、1=全くない)。

また、被験者の属性として性別、学年を聞いた。

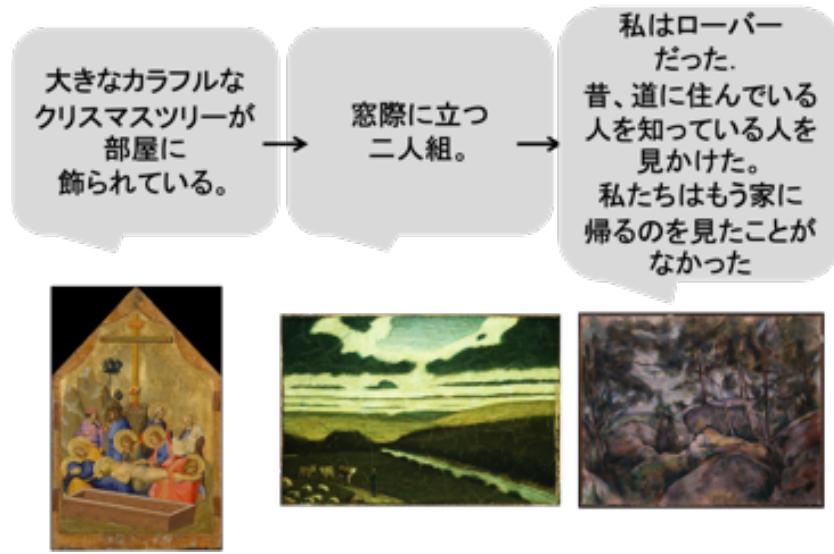
4.6 実験協力者

本稿では、情報学を専攻している大学生・大学院生 6 名を募集し、アンケート評価を実施した。

4.7 実験協力者に示す出力例

実験協力者に評価してもらう比較手法と提案手法の出力例を表 4.1、表 4.2、表 4.3 に示す。

提案 手法



比較 手法

悲嘆に暮れる表情や大胆な仕草が、峻厳な金地に浮かび上がり、キリストの受難の感動をドラマチックに伝えてくれる精緻な絵付けのパネルである。この作品は、バチカン市国アボリティカ図書館に所蔵されているミサ書を彩色したことで知られる「聖ゲオルギオス写本師」の現存する数少ない作品であり、この作品と同名の作品（登録番号 61.200.1）がある。フィレンツェで修行した後、1309年から1377年までローマ教皇庁があったアヴィニヨンで過ごし、シエナ派やフランス絵画の特徴を取り入れた。



この風景は、簡略化されたディテール、平坦な形態、パターン化された構図など、驚くほど現代的である。バルビゾンの影響を受けたライダーの初期の作品に登場した羊飼いや動物たちが、ここでは、厳格な対角線に沿って流れる小川、中景を埋める曲がりくねった丘、対照的な帶で上昇し、光の輪郭でうねる深い青の形で最高潮に達する雲などの定型的要素に圧倒されます。このような作品から、ライダーより30年後に生まれた前衛画家マースデン・ハートリーは、彼を「アラベスクの巨匠」と呼ぶに至ったのである。

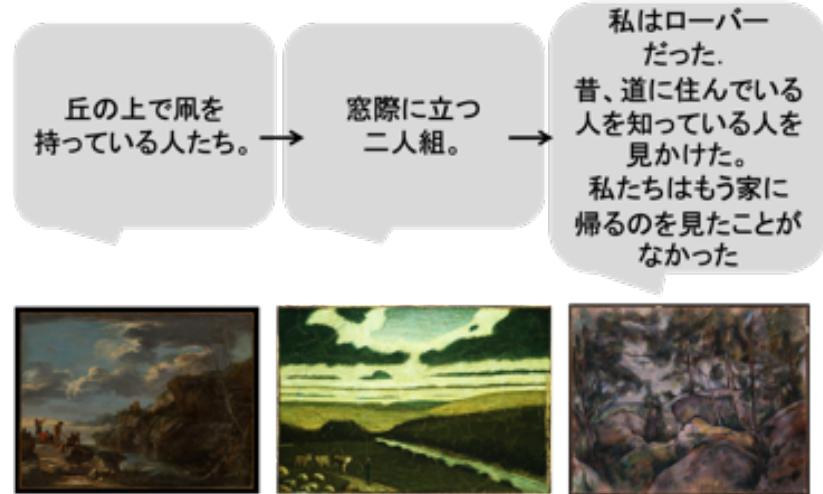


セザンヌはこの構図で、静物画の果物と同じように岩を扱い、微妙に変化する色のメッセージで形を表現している。緑、青、紫の色調と、中央の金色の陽光がアクセントとなり、石にきらめくような躍动感を与えている。1890年代半ばのセザンヌの油彩画によく見られる、水彩画のように薄く塗られた顔料が、その繊細さをより一層引き立てる。舞台となったのは、パリ南東部に位置するフォンテーヌブローの森で、目を見張るような景観、岩、古木の檻で知られ、世紀半ばから野外画家たちを惹きつけってきた。



図 4.1 画像群_1 の提案手法（上）と比較手法（下）の出力結果

提案 手法



悲嘆に暮れる表情や大胆な仕草が、峻厳な金地に浮かび上がり、キリストの受難の感動をドラマチックに伝えてくれる精緻な絵付けのパネルである。この作品は、パリカン市国アボリスティカ図書館に所蔵されているミサ書を彩色したことで知られる「聖ゲオルギオス写本師」の現存する数少ない作品であり、この作品と同名の作品（登録番号 61.200.1）がある。フィレンツェで修行した後、1309年から1377年までローマ教皇庁があつたアヴィニヨンで過ごし、シエナ派やフランス絵画の特徴を取り入れた。

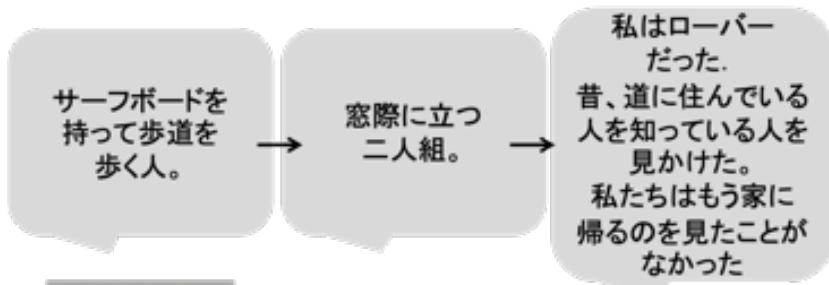
この風景は、簡略化されたディテール、平坦な形態、パターン化された構図など、驚くほど現代的である。バルビゾンの影響を受けたライダーの初期の作品に登場した羊飼いや動物たちが、ここでは、厳格な対角線に沿って流れる小川、中景を埋める曲がりくねった丘、対照的な帯で上昇し、光の輪郭でうねる深い青の形で最高潮に達する雲などの定型的要素に圧倒されます。このような作品から、ライダーより30年後に生まれた前衛画家マースデン・ハートリーは、彼を「アラベスクの巨匠」と呼ぶに至ったのである。

セザンヌはこの構図で、静物画の果物と同じように岩を扱い、微妙に変化する色のメッセージで形を表現している。緑、青、紫の色調と、中央の金色の陽光がアクセントとなり、石にきらめくような躍動感を与えている。1890年代半ばのセザンヌの油彩画によく見られる、水彩画のように薄く塗られた顔料が、その繊細さをより一層引き立てている。舞台となったのは、パリ南東部に位置するフォンテーヌブローの森で、目を見張るような景観、岩、古木の徑で知られ、世紀半ばから野外画家たちを惹きつけってきた。



図 4.2 画像群_2 の提案手法（上）と比較手法（下）の出力結果

提案 手法



比較 手法

ラスクはパンを二度焼きした硬いピスケットで、賞味期限が長いことから人気のおやつだった。パン屋はラスク、ピスケット、ジンジャーブレッド、茶菓子などを作って、少年や青年に街頭で売り歩かせることが多かった。少なくとも一度は、少年たちの乱暴な振る舞いを理由に、市から行商人の雇用許可を取り消されたことがある。パン職人たちは、「田舎者、船乗り、その他一過性の人々」に携帯菓子が人気であることを指摘し、議会に執行猶予を願い出たのである。



この風景は、簡略化されたディテール、平坦な形態、パターン化された構図など、驚くほど現代的である。バルビゾンの影響を受けたライダーの初期の作品に登場した羊飼いや動物たちが、ここでは、厳格な対角線に沿って流れる小川、中景を埋める曲がりくねった丘、対照的な帶で上昇し、光の輪郭でうねる深い青の形で最高潮に達する雲などの定型的要素に圧倒されます。このような作品から、ライダーより30年後に生まれた前衛画家マースデン・ハートリーは、彼を「アラベスクの巨匠」と呼ぶに至ったのである。



セザンヌはこの構図で、静物画の果物と同じように岩を扱い、微妙に変化する色のメッセージで形を表現している。緑、青、紫の色調と、中央の金色の陽光がアクセントとなり、石にきらめくような躍動感を与えている。1890年代半ばのセザンヌの油彩画によく見られる、水彩画のように薄く塗られた顔料が、その繊細さをより一層引き立てている。舞台となったのは、パリ南東部に位置するフォンテンヌブローの森で、目を見張るような景観、岩、古木の樺で知られ、世紀半ばから野外画家たちを惹きつけってきた。



図 4.3 画像群_3 の提案手法（上）と比較手法（下）の出力結果

第 5 章

結果

本章では、第 4 章で述べた評価実験の結果について示す。以下、3 つの観点で提案手法、比較手法の評価を行った。

1. 出力された文章と絵画全体の面白さ（以下、面白さと表記）
2. 出力された文章の組み合わせのストーリー整合性（以下、ストーリー性と表記）
3. 絵画作品への興味喚起（以下、興味喚起度と表記）

なお、比較手法は提案手法で作られたビジュアルストーリーを構成する絵画の解説文を表示したものであるため、ストーリー性は評価していない。

5.1 各手法の評価結果

表 5.1 に 3 つの観点における各手法の評価結果の平均値を示す。表 5.1 からすべての観点において、提案手法よりも比較手法の方が評価が高かった。しかし、提案手法の「面白さ」、「ストーリー性」、「興味喚起度」の Fleiss Kappa 値は -0.05, -0.09, -0.05 であった。また、比較手法の「面白さ」、「興味喚起度」の Fleiss Kappa 値は 0.02, -0.05 であった。このことから、評価者間の評価はあまり一致していないことを示唆している。

表 5.1 各手法の評価結果

	面白さ	ストーリー性	興味喚起度
提案手法	2.48	2.48	2.29
比較手法	2.71	-	2.61

表 5.2 各ビジュアルストーリーの評価結果

	面白さ		興味喚起度		ストーリー性
	提案手法	比較手法	提案手法	比較手法	提案手法
画像群_1	2.83	2.17	2.33	2.00	2.50
画像群_2	2.67	2.17	2.33	2.00	3.17
画像群_3	2.50	3.00	2.50	2.83	2.67
画像群_4	2.50	3.00	2.50	2.50	3.00
画像群_5	1.83	2.33	2.00	2.33	2.17
画像群_6	3.00	2.67	1.67	2.33	2.33
画像群_7	2.00	2.83	2.33	2.83	2.83
画像群_8	2.33	2.17	2.00	2.00	2.33
画像群_9	2.33	3.00	2.33	3.00	2.50
画像群_10	3.00	3.50	2.83	3.17	3.17
画像群_11	2.67	3.33	2.33	3.33	2.50
画像群_12	2.17	1.83	2.00	1.83	1.83
画像群_13	2.50	2.67	2.00	2.67	2.50
画像群_14	2.17	3.33	2.33	2.83	2.00
画像群_15	2.33	2.50	2.00	2.17	1.83
画像群_16	2.50	2.67	2.00	3.17	2.83
画像群_17	2.00	3.00	2.00	3.17	1.83
画像群_18	2.83	2.83	2.83	3.60	2.67
画像群_19	2.83	2.50	2.33	2.50	2.33
画像群_20	2.50	2.50	2.33	2.17	2.50

表 5.2 に提示した各絵画作品画像群の各観点における評価の平均値を示す。順位は、評価用ビジュアルストーリーを選択する際（4.2 節）に用いたランクingsコア順になっている。概ね各ビジュアルストーリーにおいても、比較手法の方が優れていることから、提案手法によって芸術作品に興味を促せたとは主張できない。

表 5.3 実験協力者が持つ芸術への関心の強さごとの評価結果

	提案手法		比較手法	
	関心が高い人	関心が低い人	関心が高い人	関心が低い人
面白さ	2.60	2.41	2.54	2.76
ストーリー性	2.95	2.24	-	-
興味喚起度	2.60	2.14	2.43	2.70

5.2 実験協力者が持つ芸術への関心の強さごとの評価結果

事後アンケートでの「これまでの平均的な 1 年間の美術館の来館頻度」、「美術館以外（オンライン上など）での絵画作品の鑑賞頻度」という実際の芸術鑑賞頻度を問う質問的回答から、実験協力者を「芸術への関心が高い」群と、「芸術への関心が低い」群の 2 つに分けた。具体的にはこれらの質問のいずれかに対して、6=月数回以上、5=月 1 回程度、4=2~3 ヶ月に 1 回程度、3=半年に 1 回程度と回答した人を、「芸術への関心が高い」群とした（2 人）。一方、2=年に 1 回程度、1=全く・ほとんどないと回答した実験協力者を、「芸術への関心が低い」群とした（4 人）。

表 5.3 に各群の評価の平均値を示す。表 5.3 より、関心が高い人にとって提案手法と比較手法に大きな差はなかった。一方、関心が低い人にとっては提案手法より比較手法の方が評価が高かった。この結果から、普段芸術を鑑賞しない人に対して、提案手法は芸術作品に興味を促せたとは主張できない。

5.3 提案手法の出力結果例

5.3.1 ストーリー性が高かった出力結果例

図 5.1 に提案手法の出力結果の中で、ストーリー性が 3 以上の例（画像群_2, 画像群_4, 画像群_10）を示す。これらを見ると、2 枚目の「窓際に立つ 2 人組。」というキャプションと関連性があるといえる。2 枚目のキャプションから、家の中の状態もしくは窓から外の様子を眺めていることが想像できる。画像群_4 では部屋にいる男女 2 人を表しており、画像群_2 では外の廻を持っている様子、画像群_10 では山の様子を眺めている、という想像に繋がったといえる。

5.3.2 面白さが比較手法より高かった出力結果例

図 5.2 に提案手法の出力結果の中で、比較手法と比べたときに面白さの評価が高かった例（画像群_6, 画像群_10）を示す。画像群_6 では、1 枚目に「スケートボードで坂道を走っている男」という、画像の内容とは異なるものの、思わず男性の姿勢がそのように見えてくるようなキャプションが付けられている。画像群_10 では、1 枚目に「大きな山を背にした、大きな山並み。」というキャプションが付けられている。画像の内容に概ね沿っているが、キャプションが日本語として不自然であること、5.3.1 節からストーリー性はある程度保たれていることから、面白さに繋がっているといえる。

5.3.3 興味喚起度が比較手法より高かった出力結果例

図 5.3 に提案手法の出力結果の中で、比較手法と比べたときに興味喚起度の評価が高かった例（画像群_1, 画像群_2）を示す。また、図 5.4 に評価が低かった例（画像群_11, 画像群_17）を示す。評価が高かった画像群_1, 画像群_2 を見ると、起部分の文章が転や結部分に繋がっている。一方、評価が低かった画像群_11 はストーリーとしては成り立っていると考えられるが、すべてを通して読むとやや面白みに欠ける。画像群_17 は、起部分の「犬が瓶を持った牛を見ている。」という文章が画像と関連性がなく、また転や結部分へ文章が繋がっていないといえる。

ストーリー性の評価が高かった提案手法出力例

画像群_2



画像群_4



画像群_10



図 5.1 提案手法の出力結果の中で、ストーリー性の評価値が高かった例（画像群_2, 画像群_4, 画像群_10）

面白さの評価が高かった提案手法出力例

画像群_6



画像群_10



図 5.2 提案手法の出力結果の中で、面白さの評価値が高かった例

(画像群_6, 画像群_10)

評価が高かった提案手法出力例

画像群_1



大きなカラフルなクリスマスツリーが部屋に飾られている。

窓際に立つ二人組。

私はローバーだった。
昔、道に住んでいる人を知っている人を見かけた。
私たちはもう家に帰るのを見たことがなかった

画像群_2



丘の上で凧を持っている人たち。

窓際に立つ二人組。

私はローバーだった。
昔、道に住んでいる人を知っている人を見かけた。
私たちはもう家に帰るのを見たことがなかった

図 5.3 提案手法の出力結果の中で、興味喚起度の評価値が高かった例
(画像群_1, 画像群_2)

評価が悪かった提案手法出力例

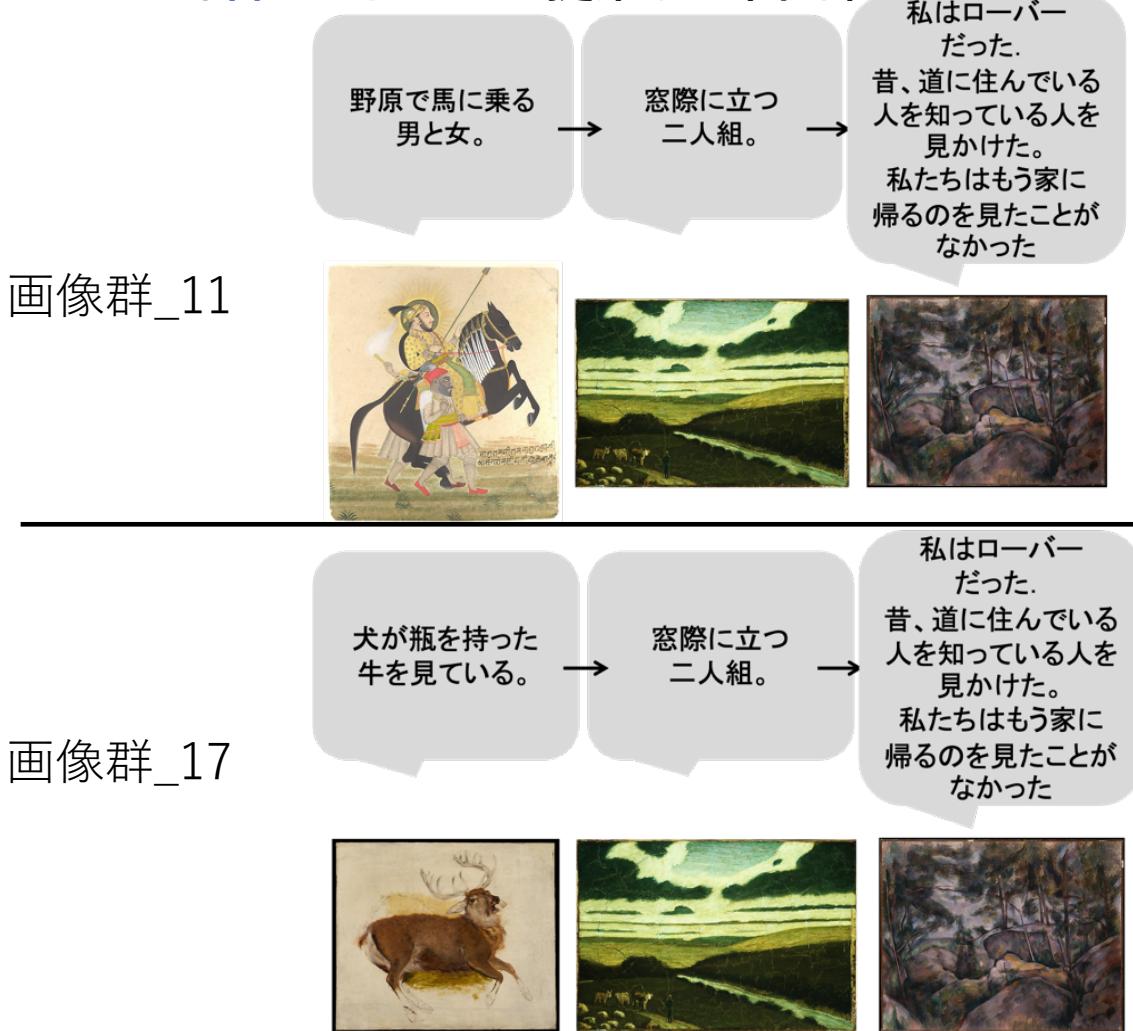


図 5.4 提案手法の出力結果の中で、興味喚起度の評価値が低かった例
(画像群_11, 画像群_17)

第 6 章

考察

本章では第 5 章の評価実験の結果をもとに、提案手法の考察と改善方法について述べる。

6.1 提案手法の有効性

表 5.1, 表 5.2 の結果から、絵画作品への興味の喚起度合いは、平均的には提案手法によって生成されたビジュアルストーリーと一般的な解説文と比較して差はなかった。一方、表 5.3 が示しているように、もともとの芸術へ関心度別に評価者の評価を分析したところ、評価値に傾向の差が見られた。芸術への関心が高い人には提案手法の方が興味喚起度スコアが高く、提案手法によるビジュアルストーリーは一般的な解説文よりも興味を喚起できる可能性がある。一方、今回対象としている芸術への関心が低い人にとっては、提案手法のほうが興味喚起度スコアは低く、一般的な絵画の解説文の方がよいという結果が得られた。

提案手法によって生成されるビジュアルストーリーは一般的な絵画の解説文ほどはとつつきにくいものではないが、3 コマ分の絵画や生成されたキャプションから行間を補ってストーリーを想像する必要性がある。その種の想像行為は向き不向きがあり、芸術への関心が高い人の方が馴染みがあると考えられることから、芸術への関心の違いによって評価傾向にも差が現れたと考えられる。

6.1.1 興味を促せるビジュアルストーリーの要素

表 5.2 の評価結果を見ると、ビジュアルストーリーによっては提案手法の方が面白さや興味を促せた可能性がある。図 5.2 の出力結果から、付けられたキャプションが画像内容とやや差異があることが面白さに繋がる要因としてと考えられる。次に、図 5.1 の出力結果から、2 枚目の転部分に繋がりがあることがストーリーを成り立たせるために必要であると考えられる。図 5.3 からも、転部分を除いて起と結の組み合わせからもストーリーが想像できることが、評価の高い要因であると考えられる。一方図 5.4 から、単純に文章として成立していても面白みに欠けると興味を促すことに繋がらず、評価が低かったと考えられる。特に、ストーリー性および面白さの両方の評価が高い画像群_10 は、提案手法の興味喚起度も各ビジュアルストーリー中で高く、ストーリー性と面白さは興味喚起に繋がる要素となる可能性がみられる。

6.2 提案手法の問題点

6.2.1 評価の不一致

5.1 節における、提案手法および比較手法の Fleiss Kappa 値の結果から、実験協力者の評価が一致していないことがいえる。表 5.1 の評価結果はすべての実験協力者の回答の平均値をとっており、サンプルサイズも 6 人と少ないため、回答のばらつきの影響を受けていると考えられる。この原因として、同じビジュアルストーリーを提示したとしても、その中のどの部分を評価するかが実験協力者によって異なることが考えられる。ビジュアルストーリー全体を通しての文章の組み合わせや画像群か、1 作品の画像やキャプションに着目して評価を行ったのか今回の評価実験では明らかでない。また、絵画やキャプションのどの部分に关心を持つか、あるいはビジュアルストーリーから何を感じるかは、見る人の感性や好みに依存すると考えられる。そのため、評価が評価者間で一致しないという事象が生じたと考えられる。

また、評価アプローチにも被験者間で解釈の違いがあった可能性がある。今回の提案手法では、ビジュアルストーリーを見た時にストーリーを考えてしまうきっかけを作り、芸術作品に興味を持つことを期待していた。一方、評価実験を行う際には、評価者にそのような想像をしてストーリーを考えることを説明せずにビジュアルストーリーの評価を行ってもらった。そのため実験協力者はビジュアルストーリーを見ても、そこからストーリーを想像することはせずに評価をしていた可能性がある。特に、ストーリー性を評価するた

めの「提示された文章の組み合わせがストーリーとして成り立っているか」という質問では、3つの文章が論理的に繋がっているかどうかという意味に解釈されたことが考えられる。

そのほかの点では、提案手法・比較手法の個々で点数付けを行う方法よりも、2つを比較をする2点比較法を用いた方が、評価者の負担を減らし、結果も明確になりやすいといえる。また、ビジュアルストーリーにすることで芸術に興味を持つというアイデアと、実際に提案手法がそれを達成できたかという実装の部分が、今回の評価実験では混ざっている問題も挙げられる。このように、今回正確に提案手法を評価できなかった可能性があるため、評価方法について改善が必要と考えられる。

6.2.2 ビジュアルストーリー全体への注目度

本稿では、提案手法と比較手法とともに、絵画作品と文章の組み合ったビジュアルストーリーを評価者に提示した。しかし画像と文章を提示した際に、画像の方が印象に残るため、評価対象であるビジュアルストーリーではなく、絵画作品のみで判断したという意見が見られた。このことから、本来目的としていたビジュアルストーリー全体を評価するよりも、絵画作品のみの情報から評価を行ってしまった可能性が考えられる。評価者および評価対象によってどこに着目するかが変わってきてしまうことは、6.2.1節で述べた評価の不一致の要因にも繋がるといえる。

6.2.3 提案手法と比較手法の情報量の違い

今回提案手法よりも解説文である比較手法が優れていた要因として、比較手法の方が情報量が多かったことが考えられる。提案手法によるビジュアルストーリーは、絵画作品に対して短い3文のみがつけられている。一方、比較手法である解説文はある程度の長さのある文章となっており、評価者に対して提供している情報量に差があることがみられる。また、提案手法は文章から想像力を働かせてストーリーを考える必要がある。比較手法はそのような必要がなく、詳細に作品の情報が述べられているため、提案手法よりも興味を持ちやすくなったと考えられる。

6.2.4 日本語への自動翻訳による文章の不自然さ

本稿では、提案手法に必要な画像キャプション生成モデルや、文接続判定器の学習には英語のデータセットを用いている。そのため、提案手法の元の出力結果は英語の文章に

なっている。一方、評価実験の際は実験協力者の母国語に合わせて、出力された文章を日本語に自動翻訳して提示している。日本語に翻訳すると、このような原文が持つ表現が完全に再現できない問題が生じる。特に結部分の詩的画像キャプションでは、英単語の韻を踏むことで詩的な表現の工夫がなされている。そのため、今後英語を母国語とする人を対象として、提案手法を評価してもらうことで、今回行った評価実験とは異なる結果が得られると考えられる。

6.2.5 キャプションとストーリー生成方法の改善

今後改善する方法として、まずキャプション生成モデルの改善が考えられる。本稿で利用したキャプション生成モデルは、一般的な写真画像およびそのキャプションをデータセットとして学習をしたものである。今回入力画像として絵画作品を使用していることから、画像に適したキャプションが生成されないことがあった。一方で、画像に適さないキャプションによって、ストーリーに面白さを与える要因になることも考えられる。そのため、生成された各キャプションについて人間が評価をし、その評価値をストーリー生成の要素として取り入れることが改善として挙げられる。

次に、接続判定器構築の際に用いた学習データの改善が考えられる。今回学習データについては、5文1ストーリーのVISTに対して、1文目・3文目・5文目が起・転・結を構成していると仮定した。しかし、ストーリーによっては必ずしも抽出した文が起・転・結を構成しているとは限らない。この点については文の分類を人間に評価してもらうことや、4コマ漫画のデータセットを用いるなど、起・転・結の構成を正確に取得する改善が考えられる。また、単純にデータ数を増やすための方法として、より大規模なストーリーデータセット[26][27]や、ショートストーリーに適したROCStories Dataset[28]などを用いることが挙げられる。

第7章

まとめ

本稿では、芸術に関心がない人が興味を持てるよう、複数の画像キャプションの並び替えすることで、芸術作品をストーリーとして鑑賞できる方法を提案した。具体的には、画像キャプショニングモデルと、前後のキャプションの繋がりを見る接続判定器を構築した。画像キャプショニングモデルには、一般的なモデルと詩を自動生成するモデルの2種類を使い、文章全体に変化を持たせることで面白みを与えるよう設計した。接続判定は、ストーリーテリング用のデータセットを用いて、起承転結の各ストーリー構成要素を考慮できるようにした。また、結部分に詩的な表現を取り入れることによって、最後にオチが生まれるようにし、興味の持てるストーリーを作成できることを狙った。

絵画作品への興味の喚起度合いは、平均的には提案手法によって生成されたビジュアルストーリーと一般的な解説文と比較して差はなかった。一方、もともとの芸術へ関心度別に評価者の評価を分析したところ、芸術への関心が高い人には提案手法で生成したビジュアルストーリーの方が一般的な解説文よりも興味が喚起されるという傾向が見られた。逆に芸術への関心が低い人にとっては、一般的な絵画の解説文の方がより高く評価されるという傾向が見られた。

本稿では、芸術作品を用いてストーリー形式にすることで、通常の鑑賞とは異なる、専門的な知識を必要としない見方を提供したり、印象付けやすくしたりする手法を提案した。これにより、鑑賞者が芸術に対して親近感や共感を得られることを期待する。

研究業績

学術報告

1. 永野里佳奈, 山本祐輔. 芸術作品に興味を促すキャプションの自動生成. 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021.
2. 永野里佳奈, 山本祐輔. 芸術作品に興味を促すストーリーテリング. 東海関西データベースワークショップ (DBWS2021), 2021.
3. 永野里佳奈, 山本祐輔. 芸術作品に興味を促すビジュアルストーリーの自動生成. 第 14 回データ工学と情報マネジメントに関するフォーラム (DEIM2022), 2022 (予定) .

賞罰

1. 東海関西データベース 優秀学生プレゼンテーション賞, 2021 年 9 月.

参考文献

- [1] 文化審議会. 文化芸術立国の実現を加速する文化政策（答申）－「新・文化庁」を目指す機能強化と 2020 年以降への遺産（レガシー）創出に向けた緊急提言－, 2016.
- [2] 文化審議会文化政策部会. アート市場活性化を通じた文化と経済の好循環による「文化芸術立国」の実現に向けて, 2021.
- [3] みずほ総合研究所株式会社. 令和 2 年度「博物館ネットワークによる未来へのレガシー継承・発信事業」における「博物館の機能強化に関する調査」事業実績報告書, 2021.
- [4] 文化庁. 文化に関する世論調査 報告書, 2019.
- [5] Jocelyn Spence, Benjamin Bedwell, Michelle Coleman, Steve Benford, Boriana N. Koleva, Matt Adams, and Ju Row Farr. Seeing with new eyes: Designing for in-the-wild museum gifting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, No. 5, pp. 1–13, 2019.
- [6] Lesley Fosh, Steve Benford, Stuart Reeves, and Boriana Koleva. Gifting personal interpretations in galleries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 625–634, 2014.
- [7] Reese Muntean, Alissa N. Antle, Brendan Matkin, Kate Hennessy, and Jordan Wilson. Designing cultural values into interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, No. 623, pp. 6062–6074, 2017.
- [8] Aditi Mallavarapu, Leilah Lyons, Stephen Uzzo, Wren Thompson, Rinat Levy-Cohen, and Brian Slattery. Connect-to-connected worlds: Piloting a mobile, data-driven reflection tool for an open-ended simulation at a museum. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, No. 7, pp. 1–14, 2019.

- [9] Xinlei Chen, Hao Fang, Tsung-Yi Lin, Ramakrishna Vedantam, Saurabh Gupta, Piotr Dollár, and C. Lawrence Zitnick. Microsoft coco captions: Data collection and evaluation server. arXiv:1504.00325, 2015.
- [10] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. arXiv:1907.11692, 2019.
- [11] 縣拓充, 岡田猛. 美術の創作活動に対するイメージが表現・鑑賞に及ぼす影響. 教育心理学研究, Vol. 4, No. 58, pp. 438–451, 2010.
- [12] 和田咲子, 山田芳明. 美術作品鑑賞における対話と作品理解の関係についての一考察. 美術教育学, No. 29, pp. 645–655, 2008.
- [13] Enid Zimmerman. Reconceptualizing the role of creativity in art education theory and practice. *Studies in Art Education*, Vol. 4, No. 50, pp. 382–399, 2009.
- [14] 金子一夫. 現代美術教育学研究の問題点とその解決—贈与交換論による美術教育の再定義を通して—. 美術教育学, No. 38, pp. 179–191, 2017.
- [15] Nanxing Li, Bei Liu, Zhizhong Ha, Yu-Shen Liu, and Jianlong Fu. Emotion reinforced visual storytelling. In *Proceedings of the 2019 on International Conference on Multimedia Retrieval*, ICMR ’19, pp. 297–305, 2019.
- [16] Licheng Yu, Mohit Bansal, and Tamara L. Berg. Hierarchically-attentive rnn for album summarization and storytelling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, EMNLP ’17, pp. 966–971, 2017.
- [17] Jonathan Krause, Justin Johnson, Ranjay Krishna, and Li Fei-Fei. A hierarchical approach for generating descriptive image paragraphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR ’17, pp. 317–325, 2017.
- [18] Kurt Shuster, Samuel Humeau, Hexiang Hu, Antoine Bordes, and Jason Weston. Engaging image captioning via personality. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, CVPR’ 19, pp. 12516–12526, 2019.
- [19] Kota Yoshida, Munetaka Minoguchi, Kenichiro Wani, Akio Nakamura, and Hirokatsu Kataoka. Neural joking machine: Humorous image captioning. arXiv:1805.11850, 2018.
- [20] Anna Lundqvist, Veronica Liljander, Johanna Gummerus, and Allard van Riel.

The impact of storytelling on the consumer brand experience: the case of a firm-originated story. *Journal of Brand Management*, Vol. 20, No. 4, pp. 283–297, 2013.

- [21] Jin-Ae Kang, Sookyeong Hong, and Glenn T. Hubbard. The role of storytelling in advertising: Consumer emotion, narrative engagement level, and word-of-mouth intention. *Journal of Consumer Behavior*, Vol. 19, pp. 47–56, 2020.
- [22] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR ’15, 2015.
- [23] Bei Liu, Jianlong Fu, Makoto P. Kato, and Masatoshi Yoshikawa. Beyond narrative description: Generating poetry from images by multi-adversarial training. In *Proceedings of the 26th ACM international conference on Multimedia*, MM’18, pp. 783–791, 2018.
- [24] Ting-Hao K. Huang, Francis Ferraro, Nasrin Mostafazadeh, Ishan Misra, Jacob Devlin, Aishwarya Agrawal, Ross Girshick, Xiaodong He, Pushmeet Kohli, Dhruv Batra, et al. Visual storytelling. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL ’16, 2016.
- [25] Rensis Likert. A technique for the measurement of attitudes. *Archives of Psychology*, No. 140, pp. 44–53, 1932.
- [26] Nader Akoury, Shufan Wang, Josh Whiting, Stephen Hood, Nanyun Peng, and Mohit Iyyer. Storium: A dataset and evaluation platform for machine-in-the-loop story generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, EMNLP’ 20, pp. 6470–6484, 2020.
- [27] Angela Fan, Mike Lewis, and Yann Dauphin. Hierarchical neural story generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, EMNLP’ 20, pp. 889–898, 2018.
- [28] Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James AllenIn. A corpus and evaluation framework for deeper understanding of commonsense stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL HLT’ 16, pp. 839–849, 2016.

謝辞

本研究の遂行ならびに論文の作成にあたり，ご指導賜りました静岡大学情報学部准教授の山本祐輔先生に謹んで深謝の意を表します。

本論文をまとめるにあたり，副査として有益な御助言と御教示を賜りました静岡大学情報学部教授の杉山岳弘先生，静岡大学情報学部准教授の大木哲史先生にこころより感謝の意を表します。

本研究の遂行ならびに論文の作成にあたり御協力いただいた，静岡大学情報学部山本研究室の皆様に感謝いたします。

最後に，これまで温かく見守ってくれた家族に感謝します。

2022年3月 永野 里佳奈