# 自然言語処理に基づく4コマ漫画ストーリーデータセットの解析

Analysis of Four-scene Comics Story Dataset based on Natural Language Processing

岩崎 凌<sup>\*1</sup> 森 直樹<sup>\*1</sup> 上野 未貴<sup>\*2</sup> Ryo Iwasaki Naoki Mori Miki Ueno

\*<sup>1</sup>大阪府立大学 \*<sup>2</sup>豊橋技術科学大学 Osaka Prefecture University Toyohashi University of Technology

Comic computing is a branch of computing dealing with comics in engineering. Although comics are multi-modal data with natural languages and pictures, numerous studies in the field focus on images in comics rather than the comic content. To make models understand the contents in comics, we should deal with natural languages in comics in the form of character words. We used a dataset which was suitable for analyzing comic contents and previously proposed two tasks: sentiment analysis and variety analysis. However, two tasks did not go well because of the number of data. We demonstrated using data augmentation and analyzed the results to determine the feasibility of computers understanding comics.

# 1. はじめに

近年,人工知能による小説や漫画,アニメ,漫画といった創 作物を対象とした研究が大きな関心を集めている.創作物理解 や自動生成といった試みは工学的興味深く意義の大きい反面, そもそも人の創作物理解は高次の知的活動であり,どういった タスクであれば計算機が創作物を理解したといえるのかを定義 することさえ現状では難しい.

人の創作物の中で,特に漫画を工学的に扱う研究分野が近 年、コミック工学を中心に発展している.漫画を対象とする研 究では自然言語と画像を持つマルチモーダルなデータを扱う ため、人工知能研究の対象として適している.この分野では、 現在でも数多くの研究が報告されているが、多くは漫画の持つ 画像を対象としており、ストーリーといった漫画の意味を自然 言語から解析しようとする研究は少ないのが現状である.その 一因としては、上で述べたような意味理解のためにどのような タスクを設定すればよいかが明確ではないという点が挙げられ る.このため、表層的な情報だけではなく、ストーリーにまで 踏み込んだ研究は十分になされていない.そのような困難にも かかわらず、創作物を理解できる人工知能の構築への情動は尽 きることがない.

このような研究を実現するためには、まずデータセットが 必要となるが、人工知能と創作に関する研究分野では著作権 などの問題からデータセットを準備することが容易ではない 場合がある.しかしながら既存のデータセットには、研究に必 要な情報などのラベルを持っていないという問題点があった. この点を解決し、創作者視点を積極的に取り入れたデータセッ トとして4コマ漫画ストーリーデータセット[Ueno 18] が提 案されている.ただし、このデータセットはまだ規模が十分で はないといった別の問題が存在する.そこで、本稿では Data Augmentation を利用することで、4コマ漫画ストーリーデー タセットのデータ数の問題点を軽減し、これまで実現できな かった新しい実験手法について検討する.

これらを踏まえ、本実験では人工知能による意味理解のために感情分析と多様度分析の2つのタスクを用意し、Data Augmentation による結果を比較する.本実験はデータ量の関

大阪府堺市中区学園町1-1, iwasaki@ss.cs.osakafu-u.ac.jp

係から十分な結果は得られていないが,今後の人工知能による 創作物理解の可能性を示すという立場で結果を解析する.2章 では,参考研究を示し,3章で本実験で扱うデータセットにつ いて説明する.4章で実験手法や条件について説明し,5章で それらの考察を述べる.

# 関連研究

1章で述べたように,漫画を対象とした研究分野では画像認 識が用いられることが多い. Rigaud らの研究では漫画の吹き 出しに焦点を当て,吹き出し内のテキストを自動で抽出しよう と試みた [Rigaud 15, Rigaud 16]. 松井らの研究では,人の 描いたスケッチから似た絵を持つ漫画内の絵を自動で抽出する システムを提案した [Matsui 17].藤野らは4コマ漫画の構造 理解のために4コマ漫画の読解順序識別をした [Fujino 18]. これらのすべてが漫画内の画像のみ,あるいは絵に密接に関係 したラベルを用いた研究である.

漫画の自動生成に関する研究も報告されている [Ueno 14] [Fukuda 17]. 上述のような例と異なり,これらの研究は画像 だけでなく自然言語も対象としているが,これらの研究の目的 は漫画の解析という我々の目的とは異なっており,自然言語か ら漫画を解析しようとする我々の研究とは異なる立ち位置にあ る.現在は漫画の完全な自動生成というのは難しいことを踏ま えると,我々が試みるような解析が今後の人工知能による創作 研究の発展に非常に重要なものであることがわかる.

最後に漫画に関するデータセットについて述べる.漫画を 対象とした工学研究のためのデータセットはいくつか存在す る.日本の漫画を対象としたデータセットとしては Manga109 [Matsui 17, Ogawa 18] が挙げられる.Manga109 は 109 冊 の既存の漫画を含んでおり,漫画のテキストが文字に起こさ れていたり,漫画内のキャラクターや吹き出しについてのアノ テートがなされていたりする.工学的な漫画研究ではよく用 いられるデータセットであるが,自然言語から漫画を解析する 場合には漫画のストーリーに関するラベルが少ないことから Manga109 の持つラベルでは不十分であると考えられ,本稿 では Manga109 は用いず,4 コマ漫画ストーリーデータセッ トを用いることとする [Ueno 18].データセットの詳細は次章 で述べる.

連絡先: 岩崎 凌, 大阪府立大学 工学研究科,



図 1:4 コマ漫画ストーリーデータセットのデータの一例 (c) 作画: 鈴木市規(シナリオ:(株)スポマ 播村早紀/豊橋技 術科学大学 上野未貴)

(c) 作画: 浦田カズヒロ(シナリオ: (株) スポマ 播村早紀/豊 橋技術科学大学 上野未貴)

# 3. 研究用データセット

漫画研究用のデータセットがいくつか存在することは2章 で述べた通りであるが,漫画をデータとして使用する際は著作 権の問題から既存のデータセットを使用した方が良い. 我々が 研究に用いる4コマ漫画ストーリーデータセットは人工知能に よる創作研究の発展のために研究者が一から開発に関わった世 界初の研究用のデータセットである.以下に我々が研究でこの データセットを使う目的およびデータセットの特徴を述べる.

特徴の一つ目として、ストーリー解析に適した情報を多く含 んでいる点にある.市販された漫画をデータとした場合,著作 権などの問題に加え,計算機上で扱うためのデータが少なく, 漫画の意味理解を目的とした研究には適さないという問題があ る.例えば漫画に登場するキャラクターの感情は明示されてい ないため,読者によるアノテートによってラベルを付与する必 要があるが、アノテートされたラベルが漫画家の意図とは異な る可能性を否定できない.4コマストーリーデータセットでは 研究に関わっているため、そのような問題点を解決することが できる.

二つ目として、上野は4コマ漫画の構造を、

- 一般:標準的な起承転結をもつ
- 繰り返し:1,2コマ間の類似が3,4コマ間でも起きる
- 出オチ:1 コマ目におかしな絵が描かれてオチがある
- タイトルオチ:最後にタイトルを見返してオチがわかる
- 再帰:4コマ目から1コマ目に戻り話として成立する
- 参照:1つ以上前の話の続きの話となる
- 連続:連続した4コマを2話並べて8コマで話となる

と定義し、これに従ってデータセットを作成している点が挙げ られる.現在は、同一のストーリーを4コマ目がオチとなる

- t- t	-		24	Mil.
7.	1.	イー	14	21V
Ľ.	1.	)	$\sim$	<u>4</u> X

実験	ラベル	データ数	拡張後データ数			
感情分析	驚愕	18	2252			
	その他	147	19165			
多様度分析	多様度あり	674	482062			
	多様度なし	766	519826			

ー般と出オチの2つの構造から描いたものがデータとして存 在しており、本稿では計算機でこういった構造の把握も目的と している.

最後の特徴としては、Manga109のような市販漫画によっ て構成されたデータセットとは異なり、4 コマストーリーデー タセットのデータは本データセットのために5人の漫画家に よって描き下ろされている点である. ストーリーの解析をする 際にオリジナリティの観点から同一プロットを複数の漫画家が 描くことは稀有なため、そういったデータの収集に基づく研究 は困難であるため、これは非常に貴重な特徴である. なお5人 の漫画家の描く絵はそれぞれ絵のタッチが異なるため、少年漫 画タッチ, 少女漫画タッチ, 青年漫画タッチ, ギャグ漫画タッ チ, 萌え漫画タッチで描かれたデータがあると言い換えること ができる.この特徴は画像解析の方向からだけでなく、自然言 語から解析する場合にも非常に役立つものである. 作家がテキ ストから絵を起こすにあたってもとにしたテキストデータは同 一のものであるが、このテキストデータは各作家が自分の思う ようにセリフや感情ラベルを変更してもよかったため、作家ご とに差が生まれている.これによって作家ごとの考えの違いな ども把握することができる.

#### 4. 実験

本実験では,感情分析と多様度分析を行う.表1に各実験 で用いるデータ数を示す.各実験で使用するラベルに関する説 明は以下で述べる.

#### 4.1 Data Augmentation について

本実験では日本語 WordNet [Bond 11] のシソーラスを用い てデータを拡張する手法を用いる.データセットに含まれる状 況文,セリフを Juman++ [Morita 15] を用いて分かち書き をし,日本語 WordNet で類似語を持つ単語を類似語に置き換 えた.文の中に類似語を持つ単語が複数あっても,類似語に置 き換える単語は同時に1つまでとした.例えば,5つの単語か ら構成される文章があり,各単語が5つの類似語を持ってい た場合,その文からは新しく25文が生成されることとなる.

#### 4.2 感情分析

本実験で使用するデータセットは全 7 種類の感情ラベル (ニュートラル, 驚愕, 喜楽, 恐怖, 悲哀, 憤怒, 嫌悪)を持っ ているが, データ数と解析の難しさの問題から今回はニュー トラルを抜き, 感情ラベルを驚愕とそれ以外(喜楽, 恐怖, 悲 哀, 憤怒, 嫌悪)の2つとする.入力には感情ラベルが付与 されている状況文あるいはセリフー文を用い,上記の感情ラベ ルを識別する.図3に感情分析で使用する self-attention モ デル[Lin 17]を示す.

#### 4.3 多様度分析

多様度分析では、漫画家がある状況を絵に起こす際に多様度 があるかどうかを識別する.ここで多様度ラベルはデータセッ トに含まれておらず、我々が独自でアノテートしたものである



図 2: self-attention モデル

ことに注意する.本実験で使用するデータセットに含まれる 複数作家の描いた漫画を用い,2つのタッチの異なる同じエピ ソードで同じコマの絵のみを比較することでアノテートした. アノテート時には,

- 1. キャラクターの数と登場キャラクター
- 2. キャラクターの描き方
- 3. 背景の有無

#### に着目した.

1. が異なるタッチ同士は 2. や 3. の違いにかかわらず「多 様度あり」とし, 1. が等しい場合,「多様度なし」とした. そ のコマにキャラクターの頭部や手などの一部分のみしか登場し ていない場合, キャラクターとみなさなかった.

1. が等しく, 2. が異なる場合は「すこしの多様度あり」とした. キャラクターの描き方とはコマに登場するキャラクター の表現の仕方である. 漫画において, 登場人物を描く方法は 様々である. キャラクターをそのまま描くこともできれば, コ マ内の吹き出しに小さく顔を描くことで, 登場人物はコマ内に 登場していないがその吹き出しの横に小さくキャラクターを描 くことで話者を明示する手法もある. このようなキャラクター の描き方の違いが上記の 2. の一例である.

1. が等しく, 3. が異なる場合は「すこしの多様度あり」とした. ここでいう背景には,キャラクターのいる場所を表すような背景に加え,その話に関係してくるアイテム(例えば,弁当に関する話の際は弁当箱)なども含まれる.

実験時,入力には話者ラベル,感情ラベル,状況文,セリフ とタッチラベルを使用した.2つのタッチのデータを比較し, 多様度のないものか少しでも多様度のあるものかの2クラス の識別をするため,話者ラベル,感情ラベル,状況文,セリフ は各タッチごとに入力した.話者ラベルは,ラベルなし,メイ ンキャラクター A,メインキャラクター B,その他の4次元 の one-hot ベクトルであり,感情ラベルは、ラベルなしか前 述した7つの感情の8次元の one-hot ベクトルである.なお 入力に使うデータはもともとデータセットに含まれているもの である.図4に実験に使用したモデルを示す.

## 5. 実験結果

表 5. に実験結果を示す. 表 5. より, どちらの実験でも訓 練時の Accuracy が 1 に近くなっていることから, 訓練デー タにおける数の少ないラベルに対しても学習ができているこ とがわかる. 特に感情分析ではデータの偏りが大きいことか



図 3: 感情分析に用いるモデル (self-attention は図 2 参照)

ら, Data Augmentation がない状態ではデータ数の多いラ ベルだけを出力するような識別機となっていることが確認でき たが, Data Augmentation によってそういったデータに対し ても特徴を学習できるようになっていることがわかる. 多様度 識別では Data Augmentation によって結果が良くなってい ることがわかる. 一般的に, Data Augmentation はロバスト 性を高めるという目的に使用されることが多く,必ずしもモデ ルのパフォーマンスを改善するものではない. しかしながら本 実験で用いた Data Augmentation 前のデータの数は非常に 少なく,適切な特徴を学習するには心もとないデータ数であっ た. 全データを通して一度しか出てこない単語も多かったため に,各データに対して非常に過学習のしやすい状態であったの が, WordNet による Data Augmentation によって本来テス トデータにしか出現しないような単語も学習できるようにな り,性能が向上したものと考えられる.

# 6. まとめ

4 コマ漫画ストーリーデータセットを用いた実験の例として 感情分析と多様度分析の 2 つを示した. Data Augmentation を用い,感情分析と多様度分析での結果を比較した. Data Augmentation により結果の改善が見られた. 本実験ではシソーラ スを用いることでデータを拡張したが, Data Augmentation の方法は他にも考えられ,そちらも比較検討する.

#### 謝辞

本研究は一部,日本学術振興会科学研究補助金基盤研究(C) (課題番号 26330282)の補助を得て行われたものである.

本研究は一部, JST, ACT-I(グラント番号:JPMJPR17U4) の支援を受けたものである.

4 コマ漫画ストーリーデータセットの制作に対し,ご協力い ただいた漫画家の方々,株式会社スポマへ謝意を示します.

## 参考文献

[Bond 11] Bond, F.: Japanese SemCor : A Sense-tagged Corpus of Japanese (2011)

[Fujino 18] Fujino, S., Mori, N., and Matsumoto, K.: Recognizing the Order of Four-Scene Comics by Evolutionary Deep Learning, in *Distributed Computing and Artificial Intelligence*, 15th International Conference, DCAI



図 4: 多様度識別に用いるモデル (self-attention は図 2 参照)

	衣 2:	の結果)
--	------	------

実験	Train Accuracy	Test Accuracy	拡張時 Train Accuracy	拡張時 Test Accuracy
感情分析	0.8923	0.8941	0.9999	0.9000
多様度分析	0.6886	0.5181	0.9873	0.5475

2018, Toledo, Spain, 20-22 June 2018., pp. 136–144 (2018)

- [Fukuda 17] Fukuda, K., Fujino, S., Mori, N., and Matsumoto, K.: Semi-automatic Picture Book Generation Based on Story Model and Agent-Based Simulation, in Leu, G., Singh, H. K., and Elsayed, S. eds., *Intelligent* and Evolutionary Systems, pp. 117–132, Cham (2017), Springer International Publishing
- [Lin 17] Lin, Z., Feng, M., Santos, dos C. N., Yu, M., Xiang, B., Zhou, B., and Bengio, Y.: A Structured Selfattentive Sentence Embedding (2017)
- [Matsui 17] Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., and Aizawa, K.: Sketch-based Manga Retrieval using Manga109 Dataset, *Multimedia Tools and Applications*, Vol. 76, No. 20, pp. 21811–21838 (2017)
- [Morita 15] Morita, H., Kawahara, D., and Kurohashi, S.: Morphological Analysis for Unsegmented Languages using Recurrent Neural Network Language Model, in *EMNLP* (2015)
- [Ogawa 18] Ogawa, T., Otsubo, A., Narita, R., Matsui, Y., Yamasaki, T., and Aizawa, K.: Object Detection for Comics using Manga109 Annotations, *CoRR*, Vol. abs/1803.08670, (2018)

- [Rigaud 15] Rigaud, C., Thanh, N. L., Burie, J., Ogier, J., Iwata, M., Imazu, E., and Kise, K.: Speech balloon and speaker association for comics and manga understanding, in *ICDAR*, pp. 351–355, IEEE Computer Society (2015)
- [Rigaud 16] Rigaud, C., Pal, S., Burie, J.-C., and Ogier, J.-M.: Toward Speech Text Recognition for Comic Books, in *Proceedings of the 1st International Work*shop on coMics ANalysis, Processing and Understanding, MANPU '16, pp. 8:1–8:6, New York, NY, USA (2016), ACM
- [Ueno 14] Ueno, M., Mori, N., and Matsumoto, K.: 2-Scene Comic Creating System Based on the Distribution of Picture State Transition, in Omatu, S., Bersini, H., Corchado, J. M., Rodríguez, S., Pawlewski, P., and Bucciarelli, E. eds., *Distributed Computing and Artificial Intelligence, 11th International Conference*, pp. 459–467, Cham (2014), Springer International Publishing
- [Ueno 18] Ueno, M.: Four-Scene Comic Story Dataset for Softwares on Creative Process, in New Trends in Intelligent Software Methodologies, Tools and Techniques - Proceedings of the 17th International Conference SoMeT\_18, Granada, Spain, 26-28 September 2018, pp. 48–56 (2018)