

絵文字の量を制御可能な絵文字自動挿入

富田 龍平¹田村 晃裕¹加藤 恒夫¹藤田 杏樹²¹同志社大学大学院 理工学研究科 ²大阪大学大学院 生命機能研究科

{ctwh0181@mail4, aktamura@mail, tsukato@mail}.doshisha.ac.jp

u088870i@ecs.osaka-u.ac.jp

概要

本研究では、絵文字の量を制御可能な絵文字挿入手法を提案する。従来の絵文字挿入手法は、1つの絵文字を挿入することを前提としている。しかし実際には、複数の絵文字が使われることが多々ある。そこで本研究では、指定された絵文字割合（文中の文字数に対する絵文字数の割合）に基づき絵文字を自動挿入する手法を提案する。提案手法は、まず、BERTにより絵文字の挿入位置と各位置における絵文字数を推定する。その後、RNNにより挿入する絵文字系列を生成することで、文中の複数箇所に複数絵文字の挿入を可能とする。Xから作成したデータを用いた実験の結果、F値は低い値であったが、人手による主観評価では、人が行った絵文字挿入よりも好ましい挿入ができることを確認した。

1 はじめに

絵文字は、書き手の感情など、文字だけでは伝えきれない情報を伝えられたり [8, 11, 9]、読み手に親近感やポジティブな印象を与えられたり [10] できることから、文に対応する絵文字を予測する研究 [2, 1, 6] や絵文字を含まない文に絵文字を自動挿入する研究 [4] が行われている。これらの従来研究では、1文に対応する絵文字は1つであると仮定しており、従来の絵文字挿入手法では、1文に対して1つの絵文字しか挿入することができない。しかし実際には、書き手が強い感情を伝えたい時などは、複数の絵文字を文に含めることが多々ある。

そこで本研究では、絵文字挿入対象の文と共に、文中の文字数に対する絵文字数の割合を表す絵文字割合を入力として受け付け、指定された絵文字割合で絵文字を自動挿入する手法を提案する。提案手法は、BERT[3]による「挿入位置推定」とRNNによる「絵文字系列生成」により、文中の複数箇所に複

数絵文字の挿入を可能とする。具体的には、まず、BERTにより、挿入対象文の各単語に対して、直後に挿入される絵文字数を推定する分類問題を解くことで、絵文字の挿入位置と各位置における絵文字数を推定する。その後、1つ以上の絵文字が挿入されると推定された位置において、RNNにより挿入する絵文字系列を生成することで、文中の複数箇所に複数の絵文字を挿入する。

X (旧 Twitter) から作成したデータを用いて提案手法の自動評価と人手による主観評価を行った。主観評価では、絵文字挿入により文の意味が変わっていないか、文が不自然になっていないか、文の感情がより伝わるようになったか、文の印象が良くなったかを評価した。その結果、F値は16.4%と低い値となったが、主観評価では人手による絵文字挿入と比較して、提案手法は意味を変えることなくより自然に絵文字を挿入し、文の感情がより伝わり、印象が良い絵文字挿入ができることを確認できた。

2 関連研究

本節では、本研究に関連する従来研究を概観する。これまで、文に対応する絵文字を予測する研究が複数行われている。これらの研究は絵文字を文に挿入するものではないが、挿入される絵文字を決めるタスクとして捉えられるため、本研究と関連する。従来の絵文字予測は、1文に対応する絵文字は1つであると仮定し、文の分類問題として絵文字予測を解いている。Barbieriら [2] は、双方向LSTMを用いて、ポスト (旧ツイート) に対して絵文字を予測するモデルを提案している。このモデルは人手による予測よりも高い性能であることが実験により確認されている。そして、Nusratら [6] は、BERTを用いてポストに対する絵文字を予測するモデルを提案している。また、Barbieriら [1] は、Instagramの写真と説明文に対して絵文字の予測を行うモデルを提案

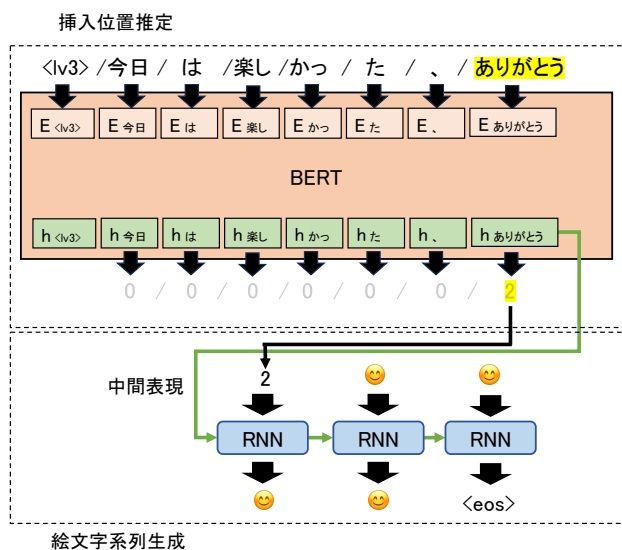


図1 提案手法の概要図

している。このモデルは、ResNets を用いて獲得した写真の特徴表現と FastText を用いて獲得した文の特徴表現を結合した特徴表現に基づき絵文字を予測するマルチモーダルなモデルである。

テキストに絵文字を自動挿入する研究として、これまで、Kwon ら [4] が絵文字を含まないポストへの絵文字挿入モデルを提案している。この研究では、絵文字の種類は出現位置と関連があることに着目し、絵文字の挿入位置の予測と挿入する絵文字の予測を同時学習するモデルを提案している。提案されたモデルは、双方向 LSTM でテキストをエンコードし、Bi-Affine 層で絵文字位置を予測する。そして、予測した絵文字位置に基づき Contextualized Dynamic-Smoothing により挿入する絵文字を予測する。従来の絵文字自動挿入は、従来の絵文字予測同様、1 文に対応する絵文字は 1 つであると仮定しており、1 文に対して 1 つの絵文字しか挿入することができない。

3 提案手法

本節では、指定された絵文字割合に基づき、文中の複数箇所に複数の絵文字を挿入可能な絵文字自動挿入手法を提案する。提案手法の概要図を図 1 に示す。提案手法は、BERT による「挿入位置推定」と RNN による「絵文字系列生成」により絵文字を自動挿入する。

提案手法では、まず、絵文字挿入対象の文 x と共に絵文字割合を入力として受け付ける。絵文字割合は、その度合いを示すタグ l (以降、「割合タグ」と

呼ぶ) で表す。図 1 中の $\langle lv3 \rangle$ は、絵文字割合が 6.5% 以上を示す割合タグである。そして、BERT により、指定された絵文字割合に基づき、絵文字の挿入位置と各位置における絵文字数を推定する。具体的には、挿入対象文の先頭に割合タグを結合したトークン列を BERT に入力し、割合タグ以外の各トークンに対して、直後に挿入される絵文字数を推定する。絵文字数が 0 の場合は絵文字が挿入されないことを表す。 t 番目のトークン x_t の直後の絵文字数 N_t が n である確率 $p(N_t = n | x, l)$ は式 1 の通り算出する。

$$p(N_t = n | x, l) = \text{softmax}(W h_t + B) \quad (1)$$

ここで、 h_t は BERT により算出された x_t の分散表現、 W と B はパラメータ行列である。

その後、1 つ以上の絵文字が挿入されると推定された位置 t 毎に、RNN により挿入する絵文字系列を生成する。 t 番目のトークン x_t の直後 (位置 t) の絵文字系列 y_t は式 2 の通りである。

$$y_t = \text{RNN}(h_t, N_t) \quad \text{for } N_t \geq 1 \quad (2)$$

y_t の生成では、RNN の隠れ状態の初期値を x_t の分散表現 h_t に設定し、生成する絵文字数 N_t を RNN に入力して、前方の絵文字から逐次生成する。そして、生成された各絵文字系列を各位置に挿入することで絵文字挿入を行う。

提案手法の絵文字挿入モデルの学習では、挿入位置推定の損失 ($\text{loss}_{\text{BERT}}$) と絵文字系列生成の損失 (loss_{RNN}) をあわせた損失 ($\text{loss}_{\text{BERT}} + \text{loss}_{\text{RNN}}$) に基づき、BERT と RNN を end-to-end で学習する。 $\text{loss}_{\text{BERT}}$ は BERT の出力と正解出力との交差エントロピー誤差であり、 loss_{RNN} は RNN の出力と正解出力との交差エントロピー誤差である。

4 実験

4.1 実験設定

本実験では、X から作成したデータを用いて提案の絵文字挿入手法を評価する。絵文字の自動挿入は、絵文字なしテキストから絵文字を含むテキストへの系列変換問題としても捉えることができる。そして近年では、特殊タグを変換前のテキストに結合させることで、テキスト生成を制御する手法が数多く提案されている [5]。そこで、提案手法の比較対象として、Transformer [7] により、絵文字挿入対象文の先頭に割合タグを結合したテキストを絵文字が挿入された文に変換することで絵文字挿入を行う手法

も評価する。なお、この Transformer による絵文字挿入では、絵文字挿入後の文において、絵文字以外の部分が絵文字挿入前の文から変わってしまう（入力文が保持されない）可能性があることに注意されたい。

実験データは Twitter API v2¹⁾を用いて抽出した 2020 年と 2021 年に投稿されたポストから作成した。具体的には、Unicode13.0 の絵文字を含む日本語文 3,451,965 文を収集し、URL やハッシュタグ、アカウント名などの不要な情報の削除と NFKC 形式による Unicode 正規化を行い、各文を BERT Tokenizer によりトークナイズした。ただし、絵文字は 1 個単位でトークナイズした。また、絵文字は肌色の情報を削除し、絵文字が一箇所に 6 個以上連続する場合は 6 個目以降の絵文字を削除した。そして、各文に対して絵文字割合を算出し、算出した絵文字割合に応じた割合タグを付与した。本実験では、絵文字割合が 6.5%以上の時は<lv3>、3%から 6.5%の時は<lv2>、3%未満の時は<lv1>の割合タグを用いた。また、emoji モジュール²⁾を用いて絵文字を除去³⁾することで、絵文字を含む各文に対して絵文字なし文を付与した。このデータの中から開発データと評価データとして 5,000 事例ずつを抽出し、残りの事例を学習データとした。

絵文字挿入の性能は、挿入位置と挿入した絵文字系列の両方が正しかった場合を正解として算出する適合率、再現率、F 値と、割合制御率及び原文保持率で評価した。割合制御率とは、絵文字挿入後の文における絵文字割合から求めた割合タグが入力として与えられた割合タグと一致した割合である。原文保持率とは、絵文字挿入後の文から絵文字を除去した文が入力の絵文字なし文と一致した割合である。

提案手法による絵文字挿入モデルでは、挿入位置推定で用いる BERT は、東北大学が公開している BERT-base⁴⁾を事前学習モデルとして使用した。ハイパーパラメータはデフォルト値を用いた。絵文字系列生成で用いる RNN は、1 層の LSTM モデルを用いた。隠れ層の次元は 756 次元とした。学習では、optimizer は Adam (学習率は $2e-5$, $\beta_1=0.9$, $\beta_2=0.999$)、バッチサイズは 32 とした。提案モデルは

表 1 自動評価指標による性能評価結果

	適合率	再現率	F 値	割合制御率	原文保持率
Transformer	0.053	0.149	0.078	0.836 0.923	0.671
提案手法	0.115 0.111	0.288 0.323	0.164 0.165	0.801 0.820	1.000

8 エポック学習し、開発データにおける F 値が最も高いエポックのモデルを評価した。

Transformer による絵文字挿入モデルは、OpenNMT-py⁵⁾を用いて実装した。エンコーダ層とデコーダ層はそれぞれ 6 個スタックし、注意機構のヘッド数は 8、埋め込み次元は 512、全結合層の次元は 2,048 とした。学習では、optimizer は Adam (学習率は $2e-3$, $\beta_1=0.9$, $\beta_2=0.998$)、ドロップアウト率は 0.1、バッチサイズは 4,069、学習ステップは 200,000 回、warmup ステップは 8,000 とし、損失関数は平滑化 ϵ が 0.1 のラベル平滑化交差エントロピーを用いた。評価では、開発データに対して原文保持率が最も高いモデルを評価した。

4.2 実験結果

実験結果を表 1 に示す。表 1 において、原文保持率以外は上段が評価データ全体に対する性能である。ただし、Transformer による手法では入力文が保持されない場合、適合率、再現率、F 値を求めることができない。そこで、Transformer による手法で入力文が保持された評価データに対する適合率、再現率、F 値、割合制御率を下段に示す。

表 1 より、提案手法は Transformer による手法と比較して、割合制御率は劣るものの、Transformer による手法で入力文が保持された評価データに対して、適合率、再現率、F 値が高いことが分かる。また、原文保持率の比較から、提案手法は原理的に入力文を必ず保持して絵文字を挿入可能（原文保持率が 100%）だが、Transformer による手法は入力文を保持できない場合が多いことが分かる。絵文字の自動挿入を実用場面で使用する場合、入力文が変わってしまうことは大きな欠点であることも考慮すると、提案手法の方が Transformer による手法よりも優れた絵文字自動挿入手法であると考えられる。

一方で、提案手法も Transformer による手法も適合率、再現率、F 値の値自体は非常に低かった。これは、異なる絵文字系列であっても同様の感情を表現できる場合が多いが、自動評価では絵文字系列が

1) <https://developer.twitter.com/ja/docs/twitter-api>

2) <https://github.com/carpedm20/emoji/>

3) emoji モジュールによる絵文字除去後に残る異体字セレクタも削除した。

4) <https://huggingface.co/cl-tohoku/bert-base-japanese-v3>

5) <https://opennmt.net/>

表2 絵文字の挿入位置推定性能

	適合率	再現率	F 値
提案手法	0.871	0.748	0.805

完全一致した場合のみ正解とみなすため、評価が過剰に厳しくなった可能性が高い。そこで、人手による主観評価実験も行った。詳細は 5.2 節に記す。

5 考察

5.1 挿入位置推定性能

提案手法は、挿入位置推定と絵文字系列生成により絵文字を自動挿入する。4 節では end-to-end の性能を評価したが、本節では、BERT による挿入位置推定の性能を評価する。なお、本評価では、絵文字の挿入位置のみに着目し、挿入する絵文字数の正誤は考慮しない。

表 2 に、挿入位置推定により絵文字が挿入されると判定された箇所の精度、再現率、F 値を示す。表 2 より、適合率の方が再現率よりも高いことが分かる。これは、絵文字挿入は書き手による個人差があり、特定の人に独特な挿入の箇所を学習するのが難しかったが、多くの人が挿入する一般的な箇所への挿入は適切に判断できた可能性がある。

5.2 主観評価

本節では、提案手法と Transformer による手法の主観評価を行う。また、人手により絵文字を挿入した文の評価も行い比較する。主観評価では、Transformer による手法で入力文が保持された評価データから、割合タグ毎に 35 文ずつランダムで抽出し、提案手法と Transformer による手法の絵文字挿入結果及び 4 節の定量評価において正解データとして使用した絵文字つき文を評価した。つまり、評価対象は各手法で 105 文ずつ合計 315 文である。評価者は 29 名（全員 20 代で男性 23 名、女性 6 名）で、各評価者は 315 文全てを評価した。評価は、絵文字が含まれない入力文と絵文字挿入後の文を比較し、以下の 4 つの項目を評価した。

意味 絵文字挿入により文の意味が変わったか？

自然さ 絵文字挿入により文が不自然になったか？

感情 絵文字挿入により文の感情がより伝わるようになったか？

印象 絵文字挿入により文の印象が良くなったか？

各項目は 1 から 5 の値で評価した。「意味」と「不自

表3 主観評価結果 (Avg: 平均, SD: 標準偏差)

		意味↓	不自然さ↓	感情↑	印象↑
人手	Avg	1.84	1.71	3.77	3.57
	SD	1.20	1.11	0.89	0.93
Trans-former	Avg	1.81	1.66	3.76	3.59
	SD	1.18	1.05	0.89	0.91
提案手法	Avg	1.60	1.43	3.94	3.76
	SD	1.03	0.83	0.77	0.82

然さ」は値が低いほど望ましく、「感情」と「印象」は、値が高いほど望ましい。

主観評価の結果を表 3 に示す。表 3 より、全ての評価項目において、提案手法は Transformer による手法や人手による絵文字挿入よりも良い結果であることが分かる。この結果から、提案手法はその他の方法よりも、意味を変えることなくより自然に絵文字を挿入し、絵文字挿入により、文の感情がより伝わるようになり印象良くすることができることが分かった。

6 おわりに

本研究では、BERT による絵文字位置推定と RNN による絵文字系列生成により、絵文字割合を制御可能な絵文字挿入手法を提案した。X から作成したデータを用いた実験の結果、F 値は低い値であったが、人手による主観評価では人が行った絵文字挿入よりも好ましい挿入ができることが確認でき、提案手法の有効性を実験的に確認できた。本研究では絵文字挿入において絵文字割合を制御したが、絵文字の挿入傾向は書き手の年代や性別によっても変わりうる。今後は絵文字割合以外の観点で制御可能な絵文字挿入手法にも取り組みたい。

参考文献

- [1] Francesco Barbieri, Miguel Ballesteros, Francesco Ronzano, and Horacio Saggion. Multimodal emoji prediction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pp. 679–686, 2018.
- [2] Francesco Barbieri, Miguel Ballesteros, and Horacio Saggion. Are emojis predictable? In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pp. 105–111, 2017.
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the*

Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186, 2019.

- [4] Jingun Kwon, Naoki Kobayashi, Hidetaka Kamigaito, Hiroya Takamura, and Manabu Okumura. Making your tweets more fancy: Emoji insertion to texts. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pp. 770–779, 2021.
- [5] Louis Martin, Éric de la Clergerie, Benoît Sagot, and Antoine Bordes. Controllable sentence simplification. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pp. 4689–4698, 2020.
- [6] Muhammad Osama Nusrat, Zeeshan Habib, Mehreen Alam, and Saad Ahmed Jamal. Emoji prediction in tweets using bert. *arXiv preprint arXiv:2307.02054*, 2023.
- [7] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pp. 5998–6008, 2017.
- [8] Joseph B Walther, Kyle PD’ addario. The impacts of emoticons on message interpretation in computer-mediated communication. *Social science computer review*, Vol. 19, No. 3, pp. 324–347, 2001.
- [9] 高橋直己, 上野舞夕, 浜田百合, 庄司裕子. 絵文字を用いた文章における感情伝達効果に関する研究. *日本感性工学会論文誌*, Vol. 21, No. 1, pp. 135–142, 2022.
- [10] 北村英哉, 佐藤重隆. 携帯メールへの絵文字付与が女子大学生の印象形成に与える効果. *感情心理学研究*, Vol. 17, No. 2, pp. 148–156, 2009.
- [11] 廣瀬信之, 牛島悠介, 森周司. 携帯電話メールによる感情の伝達に顔文字と絵文字が及ぼす影響. *感情心理学研究*, Vol. 22, No. 1, pp. 20–27, 2014.