

Flow-base モデルを用いた文のスタイル変換

江畑 拓哉[†] 堀江 和正^{††} 北川 博之^{††}

[†] 筑波大学 大学院 システム情報工学研究科 〒 305-8573 茨城県つくば市天王台 1-1-1

[†] 筑波大学 計算科学研究センター 〒 305-8577 茨城県つくば市天王台 1-1-1

E-mail: [†]ebata@kde.cs.tsukuba.ac.jp, ^{††}{horie,kitagawa}@cs.tsukuba.ac.jp

あらまし 一般に、文はその内容（コンテンツ）と表現方法（スタイル）の2つから構成される。近年、文章構成や対話システムの「個性付け」などを旨とした、文のコンテンツを維持したままスタイルのみを変更するスタイル変換課題が注目を集めている。これまでに提案されたスタイル変換手法は、変換前後のスタイル文に対する大規模データセットを必要とするため、英語や中国語といった一部の言語にしか適用されていなかった。本研究では、それらの学習サンプルが少なくても利用できるスタイル変換手法の開発・提案を目的とし、具体的には、言語モデルと Flow-base モデルを組み合わせたスタイル変換手法を提案した。また、日本語文の著者変換タスクを通じて、本手法及び、Flow-base モデルの有効性について検討した。

キーワード 自然言語処理, スタイル変換課題, 深層学習, 言語モデル, 潜在表現

1 はじめに

一般に、文は伝えようとする内容（コンテンツ）とその表現方法（スタイル）から決定される。例えば「今日は晴れである。」と「今日は晴れていますね。」の2つの文は、『今日が晴れている。』というコンテンツは一致しているものの、語句に一部異なるものが用いられている。特に日本語では、これらのスタイルは著者や性別といったドメインから強く影響を受ける [1, 2] (表 2) ことが知られており、対話システムの個性付けなどにスタイルを利用できると考えられる。本研究では、ある文をそのコンテンツを変化させずに異なるスタイルを持つ文へ変換する課題（以下、スタイル変換課題と表記）について考える。

近年では、主に深層学習を用いたスタイル変換モデルが複数提案されている。特に、Jonas Mueller らの提案する手法 [3] では、(1) 深層学習モデルを用いて文を適当な潜在表現空間に写像、(2) ある評価関数を基づいて潜在表現空間上のベクトル（以下、潜在表現ベクトルと表記）を操作、(3) 操作後の潜在表現ベクトルを文空間に写像、というスタイル変換の一手順を提案している。

これらの従来手法は、文を、「言語規則を含んだ潜在表現空間」へ写像すること、その変換を獲得することを前提としている。そして、この変換の獲得には大量の文サンプルが必要である。しかし、これらの手法で用いた目的関数は、この潜在変数空間への写像とスタイル同時に最適化するものになっている。そのため、変換前後のスタイルの文を大量に用意する必要があった。このため、これらスタイル文のデータセット（以下、スタイルデータセットと表記）を十分に用意できる、一部の言語（英語、中国語など）の特定のスタイル（極性、時制など）の変換にしか適用できなかった。以上を踏まえ、本研究では従来手法よりも少ないスタイルデータセットからでも、適切な潜在表現空間

が獲得できるスタイル変換手法の開発・提案を目的とし、適用可能な言語やスタイルの幅を広げることを目指す。

提案手法では、まず言語モデルを用いた潜在表現空間の獲得を考える。言語モデルは、大量の文（以下、言語モデルデータセットと表記）によって訓練された汎用特徴量埋め込みであり、一般に、特定の課題についての学習サンプルが少ないケースの転移学習元として利用される。本研究においても言語モデルを用いることで、文の言語規則を含んだ潜在表現空間の獲得できると思われる。しかし、さらに「スタイル変換に適した潜在変数空間」を言語モデルに直接獲得させようとした場合、文の言語規則の学習が不十分になったり、目的関数が複雑になったりすることが知られている [3, 4]。

そこで、本研究では言語モデルを用いて潜在表現空間を獲得した後に、Flow-base モデルを用いてコンテンツとスタイルが分離されるような別の潜在表現空間へ写像する手法を提案する。Flow-base モデルは写像前後の潜在表現空間を可逆な関数で結びつけるモデルであり、言語モデルで獲得した潜在表現空間を損なうことなく、スタイル変換に適した潜在変数空間を獲得、相互の空間を可逆に行き来することができる。

本発表時点では、実際に本モデルでスタイルを変換した具体例を紹介するとともに、Flow-base モデル導入の効果について検討する。

2 先行研究

近年、深層学習の発展に伴い、これを用いたスタイル変換手法が複数提案されている [3, 5, 6]。これらの手法では、スタイル変換前後の文ペアが得られないことから、教師なし学習のアプローチが用いられる [3, 5, 6]。

先行研究では、文を潜在表現空間に写像し、潜在表現空間から文へ復元するという機械翻訳のアプローチを応用したモデル

	変換対象	スタイルデータセット (スタイルごとの文数)
STC	ポジティブ ↔ ネガティブ	250k~
S2BS	シェークスピア ↔ 現代作家	400k~
本研究	芥川龍之介 ↔ 小川未明	15k~

表 1 先行研究と本研究で扱うスタイルデータセットサイズの比較. 先行研究と比較して本研究で扱うデータセットサイズは 10% ~ 5% しかない. STC(Style Transfer from Non-Parallel Text by Cross-Alignment [5]), S2BS (Sequence to Better Sequence [3]),

が提案されている. Tianxiao Shen らの提案した手法 [5] では, まず各スタイルの文からスタイル情報を除いたコンテンツのみを抽出, 潜在表現空間へ写像する. そして潜在表現空間上の潜在表現ベクトルから文へ復元する際に, スタイル情報を別途追加することで, 特定のスタイルの文を生成する. この手法は, 生成された文の「スタイル反映度」を目的関数に直接組み込んでおり, スタイル情報を文へ反映させやすいという特徴を持っている. 一方, この目的関数の設計上, 大規模なスタイルデータセットが学習に必要となる. このため, 本研究のように小規模なスタイルデータセットしか使うことができないという前提 (表 1) では, 適用が困難である.

他には, 文を多変量正規分布空間に従う潜在表現空間へ写像し, 潜在表現空間上でスタイル情報を操作する Jonas Mueller らの提案する手法 [3] がある. この手法では, スタイル情報の操作に勾配法を用いており, 表現が大きく変化しづらい著者などのスタイル変換に適している. また, その妥当性が理論的に証明されている点 [3] も, この手法の優れた点である. しかしこの手法は言語規則とスタイル変換の両方に携わる, 複雑な目的関数の最適化を行うことから, 安定した学習を行うには高いノウハウが要求される. また, 文の潜在表現空間が多変量正規分布になるように学習を行うと, 言語規則が潜在表現空間に対して十分に反映されないという指摘がある [4].

本研究では, これらの手法と比べて, 各部分の学習の難易度を下げるとともに, スタイルデータセット以外の学習データを文の言語規則獲得に利用できることを目指す. 具体的には, 文の言語規則に関わる目的関数とスタイルに関わる目的関数とを別々に分けることで, 学習段階やデータセットを分けることを試みる.

2.1 言語モデル

言語モデル [7, 8, 4] とは, 文の言語規則や文の意味 (スタイルとコンテンツを含む) を潜在表現空間へ写像するモデルである. それぞれの文は文体をはじめとする言語規則やコンテンツやスタイルなどで類似度が異なっている. 文をそれらの類似度を反映した潜在表現空間に写像することが言語モデルの目的である.

一般に自然言語処理の課題を深層学習を用いて取り組む際には, この潜在表現空間の学習が暗黙的に必要となる. このため, 言語モデルは, 十分な学習サンプルが得られない場合に, 転移

学習元として利用することで, 潜在表現空間の獲得を補助することが出来る.

一般に言語モデルは, 文を潜在表現空間へ写像する役割を持つ Encoder と, 潜在表現空間から文へ逆写像する役割をもつ Decoder の 2 種類のモデルで構成される. また訓練は大量の文から構成される言語モデルデータセットを用いて, Encoder と Decoder を介して復元される文が元の文と一致することに関する目的関数を最適化することで行われる [4, 7].

本研究は小規模なスタイルデータセットしか使うことができないため, 言語モデルを用いて文の言語規則に関わる潜在表現空間の獲得を目指す.

2.2 Flow-base モデル

Flow-base モデル [9, 10, 11] とは, サンプルしたデータの分布と異なる分布とを可逆に変換するモデルである. 具体的には, 可逆な変換関数 f_i を用いて, $f = f_L \circ \dots \circ f_1$ という関数合成を行うことで, 分布間を可逆に変換する. 全体の変換関数 f が可逆である条件はそれぞれの関数 f_i が逆関数を持つことであり, それぞれの関数が逆関数が存在する条件は逆関数定理より, ヤコビ行列が正則行列の性質を持つことである. 学習は, 式 1 で示される対数尤度の最大化を用いて行われる.

本手法では, 言語モデルによって得られる潜在表現空間と多変量正規分布空間を Flow-base モデルを用いて結びつける. 多変量正規分布を変換先の空間として選択した理由は, スタイルやコンテンツといった独立な潜在要素を潜在表現空間上でほぐす (Disentangle) ことが知られており [12], 異なる 2 文の潜在表現ベクトルの線分上にある文が自然な文を保ちながら意味を連続的に変化させられる [4] ためである.

$$\log p_X(X = x) = \log p_H(H = f(x)) + \log \left| \det \frac{\partial f(x)}{\partial x} \right| \quad (1)$$

3 提案手法

本手法の構造は, 言語規則に関するモデルである言語モデルと, スタイル変換に関わるモデルである Flow-base モデルとクラス分類モデルに二分される. 本手法の学習は (1) 言語規則に関わる LM Training, (2) スタイル変換に関わる Style Training の二段階に分かれており, それぞれ別の目的関数を用いて最適化する. これにより, 各学習を容易にし, それぞれに別のデータセットを利用可能にすることが本提案手法のコンセプトである.

スタイル変換を行う際には, スタイル変換を行いたい文を言語モデルの Encoder と Flow-base モデルによって多変量正規分布空間に写像, クラス分類モデルを用いて潜在表現ベクトルのスタイル情報を操作, Flow-base モデルの逆関数と言語モデルの Decoder を用いてスタイルが変換された文を生成する, という手順を取る.

3.1 構造

図 1 に提案手法の全体像を示す. 提案手法は, (1) 言語モデ

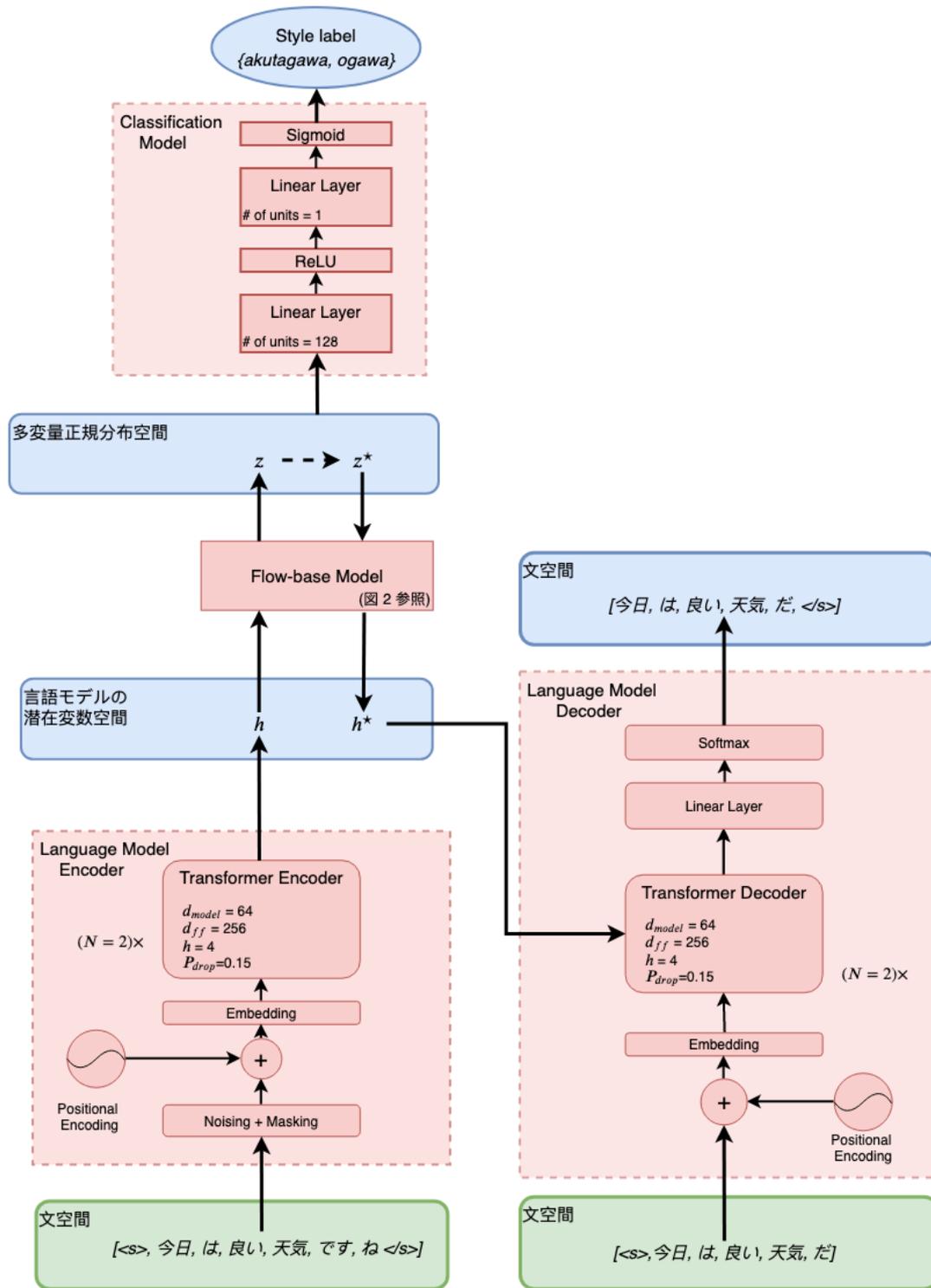


図1 提案手法の概要図. 緑の部分が入力, 赤い部分がモデル, 青い部分が出力となっている. 推論時はまず言語モデルの Encoder と Flow-base モデルを介して得られる z について Style label を目標 label に近づくように最適化し z^* を得る. 次に z^* Flow-base モデルの逆関数と言語モデルの Decoder を経由することでスタイル変換された文を獲得する.

ル (2)Flow-base モデル (3) クラス分類モデル の 3 つを組み合わせた構造である.

3.1.1 言語モデル

言語モデルでは, 文から文に関する潜在表現空間への写像を行

う. 本モデルは図1の Language Model Encoder と Language Model Decoder によって構成される. 本モデルの Encoder への入力は元の文をトークナイズしたものの先頭と末尾に開始記号 $<s>$ と終了記号 $</s>$ を加えたものであり, 出力は潜在表

芥川龍之介	小川未明
僕等はちょうど京橋の擬宝珠の前に佇んでいた。	子供は、黙って、うつむいていました。
——お母さんにも困りましたね。	姉は、あきれて、困ってしまいました。

表 2 著者による表現手法の差。芥川龍之介が比較的に漢字の割合が多いのに対して、小川未明はひらがなの割合が多く句点を多く用いていることが観察できる。

現 h である。Decoder への入力元文に開始記号 $\langle s \rangle$ を加えたものと潜在表現 h であり、出力は元文のそれぞれのトークンに対する確率値である。

今回使用した言語モデルは Transformer [13] と同様のものであるが、本課題に合わせて一部修正を行なった。具体的には、まず Input に汎化性能を向上させる Noising [7], Masking [7, 4] を追加している。さらに文を表す潜在表現ベクトル h を Transformer の Encoder から抽出している。

3.1.2 Flow-base モデル

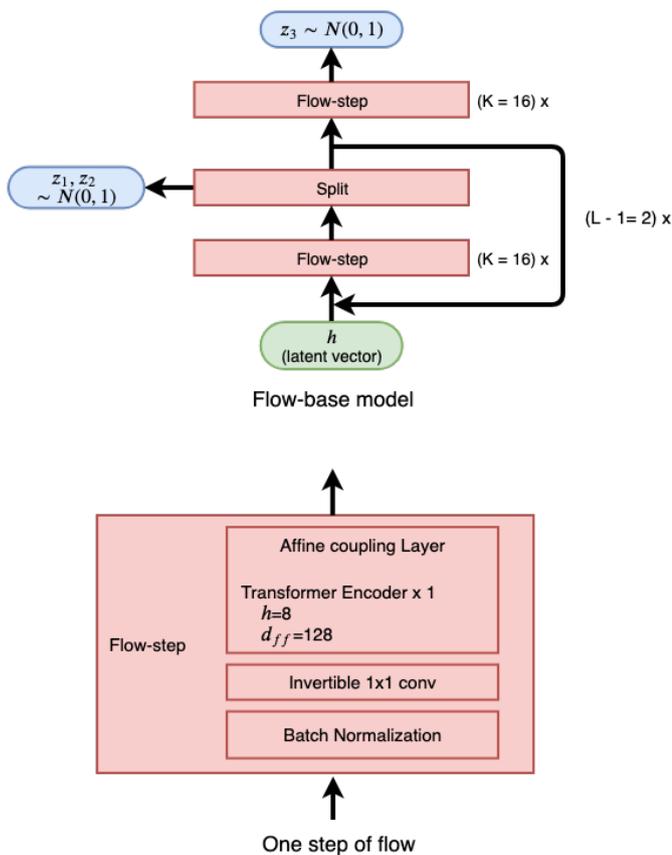


図 2 Flow-base モデルの概要図。

Flow-base モデルでは、言語モデルから獲得できる潜在表現を多変量正規分布空間へ写像する。本モデルの概要を図 2 に示す。本モデルへの入力は言語モデルから得られる潜在表現ベクトル h であり、出力は多変量正規分布に従う潜在表現 $z = (z_1, z_2, z_3) \sim N(0, 1)$ である。

本モデルは、Glow [9] をベースに構築されているが、Affine Coupling レイヤーには FlowSeq [14] で用いられているレイヤーを、Actnorm レイヤーには Batch Normalization レイ

ヤーをそれぞれ用いた。

3.1.3 クラス分類モデル

クラス分類モデルは、スタイルをクラスとみなした分類を行う。本モデルは図 1 の Classification Model によって構成される。本モデルへの入力は Flow-base モデルを通して得られた多変量正規分布に従う潜在表現ベクトルであり、出力はスタイルのラベル $\{0(\text{akutagawa}), 1(\text{ogawa})\}$ の確率値である。

本モデルは、先行研究である Sequence to Better Sequence [3] を元に設計されており、2 層の線形レイヤーとそれぞれの活性化関数に ReLU (Rectified Linear Units) [15], Sigmoid 関数をそれぞれ適用した非常に小さなモデルとなっている。

3.2 訓練手法

訓練は、LM Training 後に、訓練済みの言語モデルを元に Style Training を行う、という手順で行う。Flow-base モデルとクラス分類モデルは Style Training で同時に訓練されることから、Flow-base モデルの写像先が多変量正規分布空間はスタイルとコンテンツがより分離するような制約のかかった空間になることが期待される。

なお、提案手法では、言語モデルの潜在変数空間をスタイル変換のための潜在変数空間に近づける Fine-Tuning を行っていない。

3.2.1 LM Training

LM Training では、言語モデルデータセットを用いて、式 2 で示す損失関数について言語モデルを訓練する。損失関数は入力文と Encoder と Decoder によって得られる生成文との再構成誤差でのみ構成されている。 $B(\tilde{X}|X)$ は入力文 x に対する Noising, Masking であり、 $E(H|\tilde{X})$ は Encoder, そして $D(X|H)$ は Decoder を表している。 H は潜在表現空間を指しており、LM Training によって入力文から潜在表現ベクトル h が獲得できることを期待する。

$$L_{rec}(x) = \frac{1}{2} \|x - D(X|H = E(H|\tilde{X} = B(\tilde{X}|X = x)))\|_2 \quad (2)$$

3.2.2 Style Training

Style Training では、スタイルデータセットを用いて Flow-base モデルとクラス分類モデルを同時に訓練する。このときの Flow-base モデル、クラス分類モデルの損失関数は、式 3 を用いた。

この式は、Flow-base モデルに関する損失関数とクラス分類モデルに関する損失関数とを、重み β を用いて和を取ったものである。この損失関数を用いることで、Flow-base モデルが写像する潜在表現空間は、多変量正規分布空間かつスタイルとコンテンツがより分離した空間になることを期待している。なお、

β はハイパーパラメータ調整の結果 1.0 とした。

Flow-base モデルに関する損失関数 (式 4) は負の対数尤度であり, p_H, p_Z はいずれも多変量正規分布 $N(\mathbf{0}, \mathbf{1})$ に従う確率分布である. クラス分類モデル $C(Z = f(x))$ に関する損失関数 (式 5) はクラスラベル l とのバイナリクロスエントロピー $H_B(P, Q)$ である.

$$L_{flow}(h) + \beta L_{class}(h) \quad (3)$$

$$\begin{aligned} L_{flow}(h) &= -\log p_H(H = h) \\ &= -(\log p_Z(Z = f(h)) \\ &\quad + \log |\det \frac{\partial f(h)}{\partial h}|) \end{aligned} \quad (4)$$

$$\begin{aligned} L_{class}(h) &= H_B(P = l, Q = C(Z = f(h))) \\ &= -l \log C(Z = f(h)) \\ &\quad + (1 - l) \log(1 - C(Z = f(h))) \end{aligned} \quad (5)$$

3.3 スタイル変換手順

実際の文のスタイル変換では, まず文を言語モデル, Flow-base モデルを用いて多変量正規分布に従う潜在表現空間へ写像する. 次に得られた潜在表現ベクトル z を, クラス分類モデルへ入力し, 分類結果を目標のクラスに近づくよう最急降下法を用いてベクトルを操作する. ただし, 文のコンテンツを維持できるように, 最急降下法には元の潜在変数ベクトルから出来るだけ離さない制約を加える.

具体的なアルゴリズムを Algorithm 1 に示す. iteration は最急降下法の反復回数, lr は最急降下法の学習率であり, L_{grad} は式 6 に示される. L_{grad} の第一項は, 目標クラスと現在の潜在表現のクラス分類結果との最小二乗誤差であり, 目標スタイル y へ近づけるための項である. 第二項は, 初期潜在表現と現在の潜在表現との最小二乗誤差であり, 元の文の潜在表現 z_{init} の近傍内でスタイルの最適化が行われるための正則化項である. そして $\alpha \in \mathbb{R}^+$ は 2 つの項の重みである. なお本実験では iteration を 1000, lr を 0.2 とした.

Algorithm 1 Style Transfer

Input: sentence x , keep_orig $\in \mathbb{R}^+$,
iteration $\in \mathbb{N}^+$, lr $\in \mathbb{R}^+$

Output: transferred sentence x^*

```

1:  $z = f(E(x))$ 
2:  $z_{init} = z$ 
3: for  $i = 0$  to iteration do
4:   loss =  $L_{grad}(z)$ 
5:    $z = z - lr \nabla \text{loss}$ 
6: end for
7:  $x^* = D(f^{-1}(z))$ 
8: return  $x^*$ 

```

$$\begin{aligned} L_{grad}(h) &= \|y - C(Z = f(h))\|_2 \\ &\quad + \alpha \|z - z_{init}\|_2 \end{aligned} \quad (6)$$

4 実験

本実験では, Flow-base モデルを導入したことの有効性検証を目的に, クラス分類精度の観点からスタイル変換結果を評価した.

4.1 データセット

本研究のデータセットは, 青空文庫から著作の多い著者である, 芥川龍之介, 小川未明, 夏目漱石らの作品 (付録 8.1) の文の中から, 下記の条件を満たす文を用いた.

- 40 文字以上, 1 文字以下である文
- 『から始まる文
- 』で終わる文

なおスタイルデータセットの訓練データは, 言語モデルデータ内の芥川 / 小川の文サンプルから構成した. 各データセットの内訳を表 3 に示す.

一般に自然言語処理のデータセットはそのプライバシー保護や著作権のために公開が困難である. そこで, 本研究ではすべてのデータを青空文庫¹ から収集することで対処した.

4.2 実験設定

スタイルデータセットのデータは割合として学習データ:評価データ:検証データ=6:2:2 へ分割し, 言語モデルデータセットについては文数として訓練データ:評価データ:検証データ=66k:2k:3k へ分割した. 文のトークナイズには SentencePiece [16] を用い, 語彙数を 8000 とした. トークナイズの結果, データセットの最大トークン長は 40 となった.

評価手法としては, Flow-base モデルを用いた場合 (提案手法) とそうでない場合 (提案手法 w/o Flow-base) との, クラス分類モデルの分類精度の比較を用いた.

4.3 結果

まず, 提案手法の芥川龍之介, 小川未明へのスタイル変換の例を, 表 4, 5 にそれぞれ示す. 全体として句点の割合と漢字の使用率について変化があることがわかる. 実際にすべてのデータセット内における芥川龍之介, 小川未明の文中では, 句点の頻度は 48.0% : 82.3%, ひらがなと漢字の使用割合の差は 24.9 : 50.8 であり, お互いの文に対する特徴を反映できていると考えることができる.

次に, クラス分類精度の比較結果を表 8 に示す. 提案手法は提案手法 w/o Flow-base に比較して分類精度が 91.3% まで向上したことを確認できた. この結果から, 提案手法によって, スタイルとコンテンツがよりほぐされた潜在表現空間を獲得することができたと考えられる.

さらに, Flow-base モデルの導入の有無によるスタイル変換の差を, スタイルデータセットの検証データ内の芥川龍之介の文「自分も, すぐに死ぬであろう。」のスタイル変換を用いて観察した. 結果を表 6, 7 に示す. このスタイル変換結果は,

1: <https://www.lpaozora.gr.jp/>

L_{grad} の α を 0.020 から減少させながらサンプルし、生成文が元の文と異なってから文法的に誤った生成文が得られるまでを取り出したものである。提案手法は、提案手法 w/o Flow-base に比べて自然言語として意味の通る文をサンプルできていることが確認できた。また、特に「です」という表現は芥川龍之介の文中よりも小川未明の文中の方が出現割合が 3.9 % 多く、Flow-base モデルを適用したことで芥川龍之介と比較した小川未明の特徴を捉えることができていると考えられる。

	文数	備考
言語モデルデータセット	665k	うち 芥川 32k / 小川未明 11k
スタイルデータセット	30k	芥川 15k / 小川 15k
合計	680k	

表 3 データセットの内訳

	Acc.
Flow-base モデルあり	91.3%
Flow-base モデルなし	88.6%

表 8 Flow-base モデルの有無とクラス分類モデルの検証データセットにおける精度

5 考察

本実験では、クラス分類精度という評価軸において Flow-base モデルの導入が有効であることは確認できた。これはスタイルをより独立に近い要素として取り扱う、スタイル変換課題に適した潜在変数空間を獲得できたと考えられることによる。

しかし、Flow-base モデルがスタイル変換について有効であることを確認するに十分な結果は得られていない。この原因としては、Flow-base モデルに与えている学習データがスタイルデータセットのみとなっていたことで、潜在表現空間同士の結びつきが不十分であった可能性がある。そこで、Flow-base モデルの事前学習として言語モデルデータセットを用いた学習を行うことで、適切な文空間を獲得できるのではないかと考えている。

また本手法は、スタイル変換手順におけるハイパーパラメータに対する依存性が強いということがわかった。例えば、表 6 をみると、ハイパーパラメータ α の違いによって、変換結果に大きく影響を与えていることがわかる。また、最適な変換結果となるハイパーパラメータの組み合わせは、文やスタイルごとに異なることも確認できた。これは、実際の問題に適用する際には大きなデメリットとなる可能性がある。

また、本実験では、先行研究で取り扱う最大トークン長が 15 ~ 30 [5, 7] であることに対して、本実験ではスタイルデータセットの不足を補うため、最大 40 トークンの文を各データセットに含めた。一般にトークン長の長い文は、文の言語規則学習の観点から学習が困難である。実際に言語モデルの再構成誤差は、主に機械翻訳で用いられる生成文と元の文の類似度を測る指標である、BLEU Score [17] では 54.5 (Upper is Better),

完全一致文の割合では 31.9 % という低い値に留まっている。このため、本研究目的の達成のためにはより長い文に対応できるような言語モデルの設計が必要であると考えられる。

また本研究が対象とした著者というスタイルは、極性についてのスタイルに比べ、「すぎ」と「きらい」のような明らかなスタイル差を決定づけるような単語がないため、明確にスタイルの差を示すことが難しい。このため、先行研究で用いられる定性評価を導入することが困難であると考えられる。

BLEU Score	完全一致文の割合
54.5	31.9%

表 9 言語モデルの再構成誤差

6 まとめ

本研究では、少数のスタイルデータしか得られないという設定においてもスタイルの変換方法を獲得できる深層学習モデルを開発・提案した。本手法は、言語モデルを用いて文の言語規則を含んだ潜在表現空間を獲得、その潜在表現空間を Flow-base モデルを用いてスタイル変換に適した別の潜在表現空間と可逆に結びつけることで、文の言語規則とスタイル変換両方を鑑みた潜在表現空間を獲得する。そして本手法によって獲得できる潜在表現空間が、スタイルを独立した要素として扱えることを、スタイル分類という課題から評価した。

今後の課題としてはデータの収集方法の見直しや、スタイル変換そのものの客観的評価の導入が挙げられる。また本手法は先行研究とは問題設定が大きく異なりそれらのモデルサイズを参考にすることが難しいため、スタイル変換・スタイルのクラス分類の精度の両面から最適なモデルのハイパーパラメータを探索する必要がある。

7 謝辞

本研究は、2019 年度共同研究 (SKY 株式会社) (CPE01017) 「データエンジニアリングの知見の応用による SKYSEA Client View のログ及び資産情報の処理の高速化・軽量化・高度化」の助成を受けたものです。また、本研究は JSPS 科研費 JP 19109834 の助成を受けたものです。

References

- [1] J. Ueno, “Gender differences in japanese conversation,” 2008.
- [2] 金. 絵利子 and 岡. 剛, “カーネル法による構文に着目した作家の文体の特徴づけと類似性分析,” 人工知能学会論文誌, vol. 32, no. 3, 2017. DOI: 10.1527/tjsai.F-G94.

原文	変換結果
この時、奥から三十五六の女房が出て来た。 やっぱり、お化けかね 愛なんて、そんなもの。 黙れ！馬鹿者 と、巡査は答えました。 御縁談とか、外に何か恋愛という様なことはなかったのですか	この時奥の三十五六の女房が、出て来た。 やっぱりお化けかね 愛なんてそんなものだ。 馬鹿者黙れ！ と巡査は答えました。 御縁談という様なことは、恋愛とか何か外になかったのですか

表 4 芥川龍之介へのスタイル変換例

原文	変換結果
今日は晴れている。 細君は、お勝手のカーテンから顔を出して笑った。 負けですよ！ そこをジッと睨んで腕を組んだ。 これは残念だが仕方がない。 ——すぐわたしが行って見てやります。	晴れているね。 細君は、お勝手から顔を出して、笑ったの。 だめですよ！ 腕を組んで、ジッと睨んだ。 これはおしまいだが仕方がない。 ——わたしが行って見てやりますよ。

表 5 小川未明へのスタイル変換例

芥川龍之介への変換			小川未明への変換		
α	style	Styled Sentence	α	style	Styled Sentence
0.014	0.0	自分も、すぐに死ぬであろう。	0.011	1.0	自分も、すぐに死ぬであろう。
0.013	0.0	自分も、すぐに死ぬであろう。	0.01	1.0	すぐに、死ぬであろう。
0.012	0.0	すぐに、自分も死ぬであろう。	0.009	1.0	、死ぬであろう。
0.011	0.0	己もすぐに死ぬ。	0.008	1.0	、もう、死ぬにちがいないと思ったのです。
0.01	0.0	これも亦すぐには云はない。	0.007	1.0	、、きみをなめたのです。

表 6 提案手法

芥川龍之介への変換			小川未明への変換		
α	style	Styled Sentence	α	style	Styled Sentence
0.013	0.0	自分も、すぐに死ぬであろう。	0.015	1.0	自分も、すぐに死ぬであろう。
0.012	0.0	自分は、自分もすぐに死ぬ程。	0.014	1.0	、、すぐに死ぬであろう。
0.011	0.0	自分は が、旨之旨。	0.013	1.0	、、さびしい明智探偵であろう。

表 7 提案手法 w/o Flow-base モデル

- [3] J. Mueller, D. Gifford, and T. Jaakkola, “Sequence to better sequence: Continuous revision of combinatorial structures,” in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, International Convention Centre, Sydney, Australia: PMLR, Aug. 2017, pp. 2536–2544. [Online]. Available: <http://proceedings.mlr.press/v70/mueller17a.html>.
- [4] S. R. Bowman, L. Vilnis, O. Vinyals, A. M. Dai, R. Józefowicz, and S. Bengio, “Generating sentences from a continuous space,” in *CoNLL*, 2015.
- [5] T. Shen, T. Lei, R. Barzilay, and T. S. Jaakkola, “Style transfer from non-parallel text by cross-alignment,” in *NIPS*, 2017.
- [6] Z. Hu, Z. Yang, X. Liang, R. Salakhutdinov, and E. P. Xing, “Toward controlled generation of text,” in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., ser. Proceedings of Machine Learning Research, vol. 70, International Convention Centre, Sydney, Australia: PMLR, Aug. 2017, pp. 1587–1596. [Online]. Available: <http://proceedings.mlr.press/v70/hu17e.html>.
- [7] T. Shen, J. Mueller, R. Barzilay, and T. S. Jaakkola, “Latent space secrets of denoising text-autoencoders,” *ArXiv*, vol. abs/1905.12777, 2019.
- [8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” *CoRR*, vol. abs/1810.04805, 2018.
- [9] D. P. Kingma and P. Dhariwal, “Glow: Generative flow with invertible 1x1 convolutions,” in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., Curran Associates, Inc., 2018, pp. 10 215–10 224. [Online]. Available: <http://papers.nips.cc/paper/8224-glow-generative-flow-with-invertible-1x1-convolutions.pdf>.
- [10] L. Dinh, D. Krueger, and Y. Bengio, “Nice: Non-linear independent components estimation,” *CoRR*, vol. abs/1410.8516, 2014.
- [11] L. Dinh, J. Sohl-Dickstein, and S. Bengio, “Density estimation using real nvp,” *ArXiv*, vol. abs/1605.08803, 2016.
- [12] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *CoRR*, vol. abs/1312.6114, 2013. arXiv: 1312 . 6114. [Online]. Available: <http://arxiv.org/abs/1312.6114>.
- [13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” in *NIPS*, 2017.
- [14] X. Ma, C. Zhou, X. Li, G. Neubig, and E. H. Hovy, “Flowseq: Non-autoregressive conditional sequence generation with generative flow,” in *EMNLP/IJCNLP*, 2019.
- [15] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ser. ICML’10, USA: Omnipress, 2010, pp. 807–814, ISBN: 978-1-60558-907-7. [Online]. Available: <http://dl.acm.org/citation.cfm?id=3104322.3104425>.
- [16] T. Kudo and J. Richardson, “Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing,” in *EMNLP*, 2018.
- [17] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “Bleu: A method for automatic evaluation of machine translation,” in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ser. ACL ’02, Philadelphia, Pennsylvania: Association for Computational Linguistics, 2002, pp. 311–318. DOI: 10.3115/1073083.1073135. [Online]. Available: <https://doi.org/10.3115/1073083.1073135>.

8 付 録

8.1 データセットに含まれる文の著者一覧

- 小川未明
- 夏目漱石
- 芥川龍之介
- 森鷗外
- 坂口安吾
- 太宰治
- 海野 十三
- 江戸川乱歩
- 岡本 綺堂
- 北大路魯山人
- 楠山正雄
- 小酒井不木
- 佐々木味津三
- 高村光雲
- 中里介山
- 中谷宇吉郎
- 北条民雄
- 夢野久作