

## LSTM を用いた俳句自動生成器の開発

## Development of Automatic Haiku Generator Using LSTM

米田 航紀 \*1

Koki Yoneda

横山 想一郎 \*2

Soichiro Yokoyama

山下 倫央 \*2

Tomohisa Yamashita

川村 秀憲 \*2

Hidenori Kawamura

\*1北海道大学 工学部

School of Engineering Hokkaido University

\*2北海道大学 大学院情報科学研究科

Graduate School of Information Science and Technology, Hokkaido University

The creation of art using deep learning has been paid attention to in recent years.

Also, there is a haiku as an art that has long been popular in Japan.

In this research, we demonstrate the usefulness of deep learning as art creation by making haiku from motifs, which is a general method of creating haikus, using deep learning. First, we train LSTM based on a large amount of past haiku, let it generate a string.

Second, we extract the ones that satisfy the condition as a haiku from the generated character string and calculate the evaluation value as to whether it fits the motif image or not.

If the evaluation value is high, it is assumed that the generated haiku matches the motif image.

In this process, we conducted an experiment to confirm whether LSTM was able to learn rules as a haiku.

## 1. 序論

### 1.1 研究背景

近年、深層学習の研究の発展により文章の翻訳や、画像のキャプション生成など様々な場面で使用されている。

また、中でも注目を集めている研究として絵などの芸術の生成がある。この分野では DCGAN を使用し、絵を作成する研究 [1] 等がある。

また、日本で古くから親しまれている芸術として俳句がある。俳句は世界最短の定型詩とされ、日本の伝統芸術である。俳句の特徴として、17 音である、季語を含んでいる、切れ字を 1 つ以下含むといったルールが明確に決められている。そのため、他の芸術よりも俳句はルールを使用することにより生成された句の評価が比較的行い易い。数値的に評価することがたやすく工学的な分析が容易であるため、深層学習に使用される損失関数の計算を定量的に行うことができる。また、俳句は長い歴史があり日本で広く親しまれているため、深層学習で使用される教師データが数多く存在している。従って、本研究では芸術の生成に深層学習を使用する一例として俳句を使用し、その有用性を検証する。俳句や和歌を自動生成する研究は昔からされてきているが、多くは単語を組み合わせ、何らかの方法で評価値を出して評価が高いものを出力するという仕組みになっている。[2] 中にはディープラーニングを使用したものもあるが、画像に基づいた俳句を生成するという試みは少ない。

### 1.2 本研究の目的

本研究では、深層学習の芸術作成での有用性を示すとともに、日本の伝統芸術である俳句の作成支援システムの作成を行う。具体的な俳句の作成方法として、吟行をはじめとして俳句作成において一般的な方法である「モチーフから俳句をつくる」ことに着目する。中でもモチーフとして写真を取り上げ、写真から俳句を作成することを俳句作成過程として、それを深層学習を用いて再現する。深層学習が芸術作成に有用であるかを検証するため、作成した俳句に独創性があるか、俳句のルールを学

習できているか、画像をもとに作成できているかといった面から検証を行う。

## 2. データ収集

本研究に必要なデータは大きく分けて 2 種類である。

- 俳句データ
- 俳句と画像のマッチングデータ

以下、それぞれについて解説する。

### 2.1 俳句データ

俳句を学習する際に使用するデータである。データを含む俳人は小林一茶、正岡子規、高浜虚子の 3 人と、現代の俳人多数及び大塚凱とした。

まず web 上のデータベースのスクレイピングで収集した。大塚氏の俳句は本人に提供して頂いた。次に読みがっていない句に対して手動で俳句の読みを付加した。最後に 17 音になっていない俳句や季語が含まれない俳句、「(」等の記号を含む俳句は学習データとして不適であるため除外した。最終的には計 38,506 句となった。

### 2.2 俳句と画像のマッチングデータ

俳句と画像との適合度を学習する際に使用するデータである。俳句と画像を 1 組として扱うデータであり、このデータで使用する俳句は前述の俳句データで収集した俳句となる。使用する画像は画像素材販売サイト「imagenavi」に提供していた画像である。

画像のアノテーションに俳句の季語が含まれていた場合、その画像と俳句を 1 組として扱うことで、機械的に収集した。また、手動で加えることも行った。同様にして俳句と画像の組み合わせの候補を集め、全国のボランティアやアルバイトの方々に、どの画像が俳句に適合しているかを判定して頂き、マッチングデータとして加えた。最終的には 369,754 組収集した。

連絡先: 米田航紀, 札幌市北区北 1 4 条西 9 丁目北海道大学  
大学院情報科学研究科 9 階 調和系工学研究室, E-mail:  
yoneda@complex.ist.hokudai.ac.jp

### 3. システム

本研究が構築したシステムは、大きく分けて3つの構成に分けることができる。

- 俳句生成
- 俳句フィルタ
- 俳句と画像の適合判定

それぞれの部分について解説する。

#### 3.1 俳句生成

2章で作成した俳句のデータを入力して学習し、学習したモデルを使用して文字列を生成する。ここでは学習部分と生成部分に分けて説明する。

##### 3.1.1 俳句学習

本研究では文章生成において成果を上げている LSTM[3] を使用して俳句を生成する。

始めに、学習データをシャッフルし、学習データに使用された文字を出現順に若い番号から ID を与える。次に、学習データの文字をそれぞれの ID に変換し、バッチに分ける。バッチ内のサンプルは、予め設定した BPTT の長さ分の ID の列となる。改行文字にあたる ID が出現した場合でも、そこで区切ることせず続行する。最後にサンプルを先頭から順に 1-of-K 符号化し projection layer に入力してベクトルに変換する。そうして得られたベクトルを LSTM に入力する。

LSTM の出力は次に出現する文字の確率であり、正解データは次の文字の ID 番目が 1 のワンホットベクトルである。サンプルを最後まで入力し終えた段階で LSTM の状態をリセットし、誤差を計算する。

パラメータは LSTM 層を 3、LSTM ユニット数をそれぞれ 1024、最適化手法を Adam、学習率を 0.02、減衰率を 0.99、エポック数を 300、バッチ数を 50、サンプル長を 100、目的関数をクロスエントロピーとした。

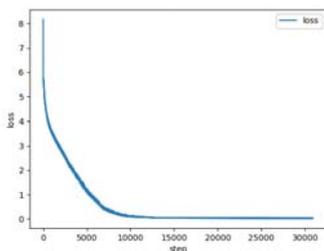


図 1: LSTM の学習の様子

##### 3.1.2 俳句出力

学習した LSTM を使用し、文字列を出力する。

入力を改行文字とし、次の文字の確率を計算してルーレット選択をし、選ばれた文字が次の入力となる。改行文字出現してから次の改行文字が出現するまでに選択された文字列が俳句候補となる。改行文字が指定した回数出現した際、出力を停止する。

#### 3.2 俳句フィルタ

俳句生成部で出力した文字列は俳句としての条件を満たしていない文字列を含んでいるため、満たしている文字列のみを取り出す。ここでは俳句としての条件を 17 音になっている、季語を 1 つのみ含む、切れ字を 1 つ以下含む、この 3 つとする。

#### 3.3 俳句とモチーフ画像の適合判定

ここでは、フィルタを通して俳句と判定された文字列の中から、モチーフ画像に適合する句を抜き出す。そのための方法として、深層学習を使用し画像と俳句がどの程度適合するかを学習する。適合度が高ければその俳句をモチーフ画像に基づいて生成した俳句であると定める。

前準備として、使用する学習データに使われている漢字それぞれに出現順に若い番号から ID を与える。適合判定では、2章で作成した画像と俳句の組を学習データとして使用する。負例として、同数の画像と俳句の実在しない組を学習データと同数だけランダムに生成する。

まず学習データの画像を Inception-v3 を使用して特徴量ベクトルを得る。次に学習データの俳句に使用されている漢字の使用回数を数え、その漢字の ID 番目の要素が使用回数となるベクトルを得る。ここまでで得られた 2 つのベクトルを結合し、入力ベクトルとする。適合判定では、全結合層のニューラルネットワークに入力し、入力された組が実在の組がどうかを判定する。学習パラメータは中間層が 3 層でそれぞれのユニット数が 1024, 512, 256 ユニット、エポック数を 1000、最適化手法を Adam、学習率を 0.00001、目的関数をクロスエントロピーとして学習した。

図 3 に評価例を示す。評価値が高い俳句と評価値を示す。

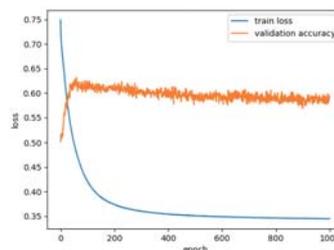


図 2: 画像とのマッチング学習

### 4. 実験

本研究では 3 つの実験を行う。パラメータを変化させた場合に、LSTM が俳句の性質を学習する能力がどのように変化するかを調べるとともに、本研究で生成された俳句がモチーフ画像に適合していることを示す。

- 生成された文字列と学習データの類似度に関する実験
- 生成された文字列の俳句フィルタにかけた際の判定の割合に関する実験
- 俳句とモチーフ画像の適合度判定の妥当性に関する実験

以下ではそれぞれの実験について解説する。



日本の桜も見えず一つ星,1.0  
 ものなしはうしろく見えて桜かな,1.0  
 花見でも肩事もなしに蘆の花,1.0  
 桜垣濃くらははれと成にけり,1.0  
 花の春天窓の窓もあなたかな,1.0  
 心見や桜がさして子のわられ,1.0  
 死そはる花のはぐれも花見哉,1.0  
 小坊主が寝に来て桜咲にけり,1.0  
 足音の心と並ぶ桜哉,1.0

図 3: 画像とのマッチング学習結果

#### 4.1 生成された文字列と学習データの類似度に関する実験

この実験では、生成された文字列と学習に使用した俳句の類似度の最小値を測る実験を行う。

この実験で使用する類似度として、Levenshtein 距離を使用する。

以下のパラメータを組み合わせて学習して文字列を 10,000 行生成し、生成した文字列を学習データとの Levenshtein 距離の最小値をそれぞれ計算した。

- LSTM 層数:2, 3
- LSTM ユニット数:256, 512, 1024

他のパラメータは 4.1.1 章で示したものと同様にした。

図 4, 図 5 に結果を示す。x 軸は最小 Levenshtein 距離、y 軸は該当する俳句の割合であり、l は層数、u はユニット数を表す。

図 4 はどちらもユニット数を 1,024 と設定したものであり、最小 Levenshtein 割合が高い、つまり学習データと同一の文字列を出力することが多いことが分かる。

図 5 はいずれもユニット数が少ないものであり、最小 Levenshtein 割合が高く、学習データにあまり類似していない文字列を出力することが多いことが分かる。

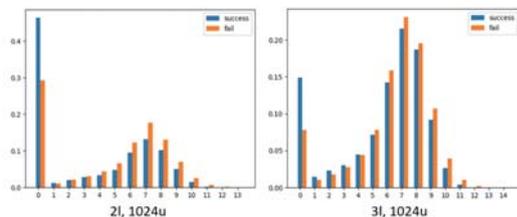


図 4: 学習データに類似した分布

#### 4.2 生成された文字列の俳句フィルタにかけた際の判定の割合に関する実験

この実験では生成された文字列と入力データを俳句フィルタにかけて、それぞれの判定の分布から LSTM が俳句として重要である季語や切れ字といった要素を学習できているかを調べ

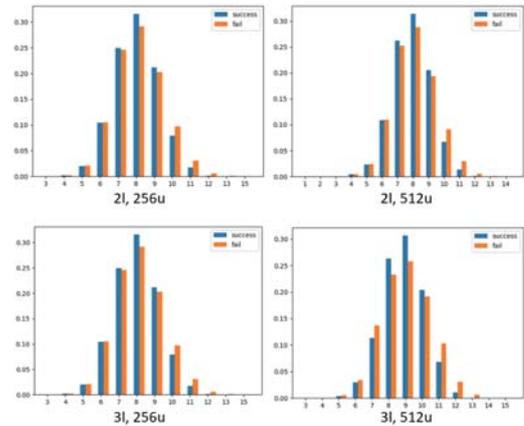


図 5: 学習データに類似していない分布

る。4.1 章と同様のパラメータで学習した複数のモデルでそれぞれ文字列を学習データと回数出力し、フィルタにかけて判定の内訳を見る。図 6,7,8 に結果を示す。x 軸はそれぞれのモデルであり、左端は学習データをフィルタにかけた際の判定結果である。y 軸はその条件で該当した俳句数を表す。図 6 の timeout の項目は、処理時間削減のために、音数を探索する処理に 1 秒以上かかる場合に処理を終了した句の数を表す。図 8 の cannot read の項目は、切れ字判定に必要な読みを得られなかった句の数を表す。

3 つの図全てにおいてユニット数が 1024 のとき、学習データとの分布に近く、学習データの俳句に含まれる季語や切れ字といったルールを学習できたといえる。

また、4.1 章と比較すると、フィルタの判定の分布が学習データに近くなるパラメータは学習データの句を出力する割合が高く、トレードオフの関係であることが分かる。

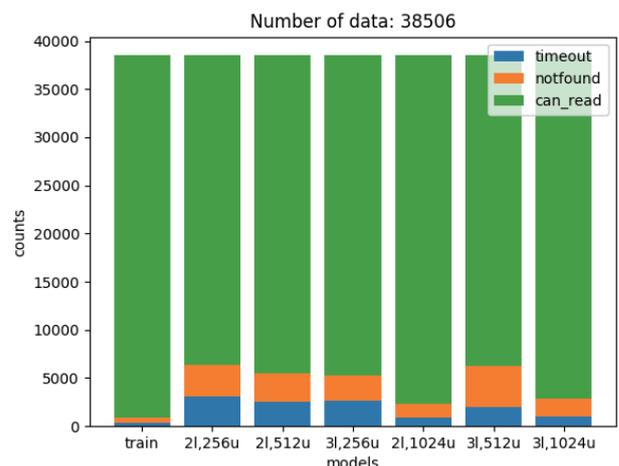


図 6: フィルタ判定分布:音数

#### 4.3 俳句とモチーフ画像の適合度判定の妥当性に関する実験

ここでは、俳句と画像の適合度判定が妥当であるかを実験する。フィルタを通して得られた俳句と任意の画像を入力していき、適合度を計算する。任意の画像に写っているキーワード

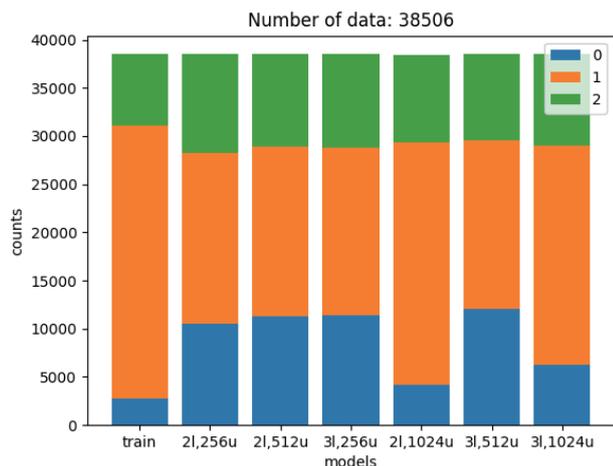


図 7: フィルタ判定分布:季語

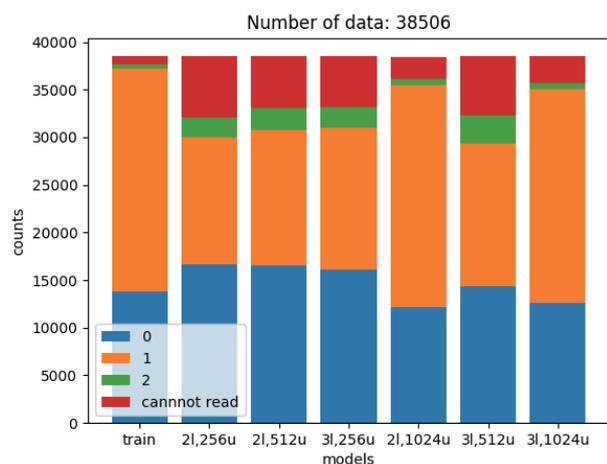


図 8: フィルタ判定分布:切れ字

を設定し、キーワードを含んでいる俳句がそれぞれの評価値でどの程度出現するか確認する。例えば、桜の画像なら桜というキーワードである。結果を図 9,10,11 に示す。

図 9 はキーワードを含む句が多く高い評価をされているが、図 10,11 は高い評価をされている句が少ない。

従って、画像に適合する俳句を選ぶことができる画像もあるが、精度は高くないことが分かる。

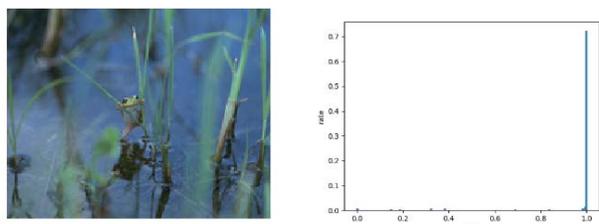


図 9: マッチング評価実験 (キーワード:蛙)

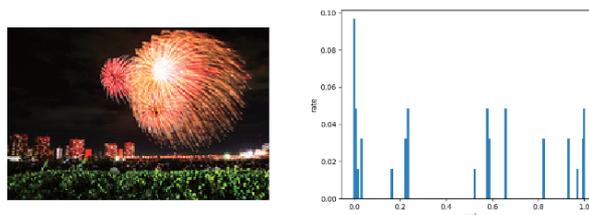


図 10: マッチング評価実験 (キーワード:花火)

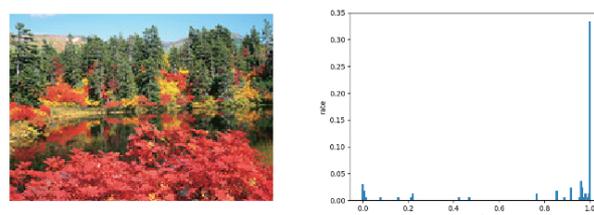


図 11: マッチング評価実験 (キーワード:紅葉)

## 5. 結論

本研究では一茶など過去の有名な俳人と現在の俳人の句を LSTM に大量に入力することで俳句の音数や季語などのルールを学習した。その学習済み LSTM を使用して文字列を大量に生成し、俳句としての条件を満たすものを抜き出し、モチーフ画像と適合するかどうかを算出した。

実験結果より、LSTM を使用することで、俳句のルールを学習できることを確認した。また、モチーフ画像と適合する俳句を抜き出すことができるが、適合する句を高く評価できないことも多いため、今後の課題として俳句とモチーフ画像の適合評価器の改善が挙げられる。

## 謝辞

本研究は札幌 AI ラボにご支援いただき実施されました。

本研究のデータ作成として、データ提供を快く引き受けてくださった大塚凱様、データ作成をお手伝いしていただいた松山市役所及び愛媛大学俳句研究会をはじめとする全国のボランティアの皆様とアルバイトの皆様には心より感謝申し上げます。

## 参考文献

- [1] Alec Radford, Luke Metz, Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", 2015, arXiv.
- [2] Tosa, Naoko, Hideto Obara, and Michihiko Minoh. "Hitch haiku: An interactive supporting system for composing haiku poem." International Conference on Entertainment Computing. Springer, Berlin, Heidelberg, 2008.
- [3] Sundermeyer, Martin, Ralf Schlter, and Hermann Ney. "LSTM neural networks for language modeling." Thirteenth Annual Conference of the International Speech Communication Association. 2012.