

# 深層学習によるマイクロブログの投稿位置推定手法の検討

河野 慎<sup>1,a)</sup> 植田 一博<sup>2,b)</sup> 徳田 英幸<sup>1,c)</sup>

**概要:** インターネットにある情報からその投稿された場所を推定することは、実世界の状況を把握し、災害やイベントなどに関して知る上で、非常に重要である。本稿では、マイクロブログを対象とし、深層学習を適用した位置推定手法を提案する。提案手法では、マイクロブログの投稿内容とその投稿に付与された画像からその投稿された場所を推定する。実験では、位置情報が付与されたマイクロブログを収集し、提案手法とベースライン手法による位置推定の精度を比較し、その結果を報告する。

## Microblog Geolocation with Deep Neural Networks

MAKOTO KAWANO<sup>1,a)</sup> KAZUHIRO UEDA<sup>2,b)</sup> HIDEYUKI TOKUDA<sup>1,c)</sup>

### 1. はじめに

インターネットには、世の中の情勢や商品など様々な情報が公開されている。人々は、これらの情報を収集し、実世界で起きていることを推論し、理解することができる。インターネットのサービスであるマイクロブログは人々が経験したり、体験したりしたことについて投稿するプラットフォームとして重要である。特に位置情報が付与されたマイクロブログは、実世界のどこで起きているのか、詳細に理解する上で非常に有用である。一方、プライバシーの問題から位置情報が付与されているマイクロブログの数は、付与されていないものに比べて、非常に少ない。しかし、人々は投稿内容などからそれが示す位置を推論することが可能である。

そこでマイクロブログの投稿内容から位置推定を行うことで、実世界で起きていることをコンピュータで理解しようと試みる研究は数多く存在する [24][8][1][11][7]。これらの研究では、位置推定タスクを自然言語処理を用いたタスクとしてみなし、地名など場所に関する情報や人々が持つ



図 1 画像付きマイクロブログと位置推定の概要。画像付きマイクロブログが与えられた時、Galaxy モデルは各日本の都道府県で投稿されたと推定される確率分布を出力する。

事前知識を用いて特徴量を設計している。しかし、場所に関する情報などは日々増えていくため、これらを全て特徴量に含めて、自然言語処理を行うのは難しい。一方で、位置推定以外の自然言語処理タスクでは、深層学習を適用することで従来の人によって特徴量を設計するアプローチよりもその精度向上を達成している [4][12][19][20]。そこで、本稿では位置推定タスクに対して深層学習を適用し、従来のように特徴量の設計をせずに推定の精度向上を目的とする。

また投稿内容からではなく画像からその位置を推定する研究も存在する [21]。本稿で扱うマイクロブログにも画像を付与する機能が実装されており、画像からマイクロブ

<sup>1</sup> 慶應義塾大学大学院政策・メディア研究科  
Keio University, Fujisawa, Kanagawa, 101-0062, Japan

<sup>2</sup> 東京大学大学院総合文化研究科  
The University of Tokyo

a) makora@ht.sfc.keio.ac.jp

b) ueda@gregorio.c.u-tokyo.ac.jp

c) hxt@ht.sfc.keio.ac.jp

グの位置推定を行うことも可能となる。特に付与されている画像が写真である場合、その写真に投稿された場所に関する情報が写っていることが考えられ、投稿内容のテキストと組み合わせることで位置推定の精度向上が期待される。したがって本稿では、画像が付与されているマイクロブログの位置推定を行う。

本稿では、画像付きマイクロブログの位置推定タスクを多クラス分類問題として扱い、日本の都道府県を分類クラスとする。次にマイクロブログのテキストに recurrent neural network (RNN) を、また画像には convolutional neural network (CNN) を適用し、学習させる。位置の推定時は、入力された画像付きマイクロブログが投稿された可能性が高い都道府県に割り当てるように、各モデル及び、二つのモデルを組み合わせたモデルが離散確率分布を出力する。この二つのモデルを組み合わせたモデルを本稿では、Galaxy モデルと呼ぶ。

実験の結果、Galaxy モデルを用いることで、従来のようにテキストのみや画像のみによる位置推定に比べ、高い精度で推定をすることが可能であることを確認した。マイクロブログによっては、投稿内容に位置推定を行うには十分な情報がない場合などがある。実際、テキストのみを用いた位置推定では達成できていないものが多い。一方でテキストと画像を組み合わせることでその位置推定の精度が向上していることから、Galaxy モデルがテキストと画像それぞれから位置推定に有効な特徴を学習していることがわかった。

本稿における貢献は次の通りである。

- マイクロブログの投稿位置推定への深層学習の適用
- マイクロブログの投稿位置推定にテキストと画像の両方を利用
- 投稿位置推定のための Galaxy モデルの提案

## 2. 関連研究

本節では、マイクロブログの投稿位置を推定する関連研究について説明する。マイクロブログの投稿位置を推定するためのアプローチとして、マイクロブログの投稿内容、つまりテキストを用いる方法が挙げられる。従来の手法では、人々が位置推定する際に用いている知識や係り受けなどの文法的な知識を表現できるような特徴量を設計し、機械学習手法に入力していた [22][13]。しかし、マイクロブログの投稿内容がもつ情報量は、その文字数の少なさや投稿する頻度の問題からニュース記事や通常のブログなどに比べて少ないため、特徴量設計が難しい。一方で、自然言語処理の研究において、RNN が様々なタスクに適用されている [4][12][19][20]。2つの文章が与えられた時に、その2つの文章の関連性スコアを算出する意味関連性タスクにおいて、Kai ら [19] は、long-short term memory (LSTM) と Tree-LSTM を適用し、既存手法を上回った。その他の深層

学習手法も、マイクロブログのような短文の評価極性を推定するタスクにおいてその性能を発揮している [2][14][16]。したがって、マイクロブログの位置推定に対して RNN を適用することで、その精度の向上が期待される。

また、マイクロブログの情報量を増やすアプローチとして、一つのマイクロブログの投稿内容以外の情報を追加する方法が挙げられる。伊川ら [8] や鈴木ら [24] はユーザの過去の投稿内容を集約することでマイクロブログがもつ情報量を増やし、位置を推定しようと試みている。同様に Hong ら [7] や Kinsella ら [11] はユーザの投稿履歴からユーザの出現位置の分布を確率モデルで表現し、推定を行っている。しかし、ユーザが過去に投稿した内容が現在の位置に関係しているとは限らずノイズとなることが考えられるため、過去どれだけ遡る必要があるかユーザごとに考慮する必要がある、その扱いは難しい。過去の投稿による情報量追加以外では、Cheng ら [1] のマイクロブログを投稿したユーザのプロフィールや投稿した際の IP アドレスの追加が挙げられる。

そしてマイクロブログは画像を付与することができ、この画像を投稿内容に加えて位置推定に利用することが考えられる。マイクロブログではなく、Flickr に投稿された画像が撮られた位置を推定する PlaNet というシステムが提案されている [21]。PlaNet には、CNN モデルに Inception モデル [18] を適用されており、世界中の都市の粒度で 24.5 % の精度で推定することが可能となっている。この研究から、位置推定において画像が有効であることがわかる。

以上のことから、多くの位置推定に取り組んでいる研究では、特徴量の設計や情報の追加などに尽力している一方で、深層学習を様々な自然言語処理に関するタスクへの適用することで、これらを必要とせずに目的を達成できている。先行研究から、マイクロブログにおいて深層学習が有効であることがわかり、特に RNNs は自然言語処理タスクにおいて非常に有効であることがわかるが [4][12][19][20]、本稿のようにマイクロブログに RNNs を用いて位置推定タスクを多クラス分類問題として取り組む研究はない。また、マイクロブログに付与された画像と投稿内容を組み合わせる位置推定に取り組む研究もない。

## 3. 提案手法

本稿では、画像付きマイクロブログの位置推定タスクを多クラス分類問題としてみなす。そこで、マイクロブログの投稿内容 (テキスト) から位置推定を行う LSTM モデルと、マイクロブログに付与された画像から位置推定を行う CNN モデルを適用し、テキストと画像の二つから位置推定を行う Galaxy モデルを提案する。本節では、それぞれのモデルについて説明をし、モデルの学習方法について述べる。

### 3.1 テキスト学習のための LSTM

RNN は、文法的な情報など入力される連続した単語の情報を内部で保持することができるため、様々な自然言語処理タスクに問題に適用されてきた手法である。RNN への入力データはマイクロブログの投稿内容となり、任意の  $T$  個の単語  $w_t$  で構成される投稿内容  $tweet$  は  $tweet = (w_1, w_2, \dots, w_T)$  と表される。この単語  $w_t$  は *one-of-k* ベクトルで表現されており、予め用意された語彙数  $d$  次元で  $i_{w_t}$  番目の要素が 1 となっており、それ以外の要素は 0 となっている。RNN は、

$$x_t = W^{xw} w_t \quad (1)$$

$$h_t = \sigma(W^{hx} x_t + W^{hh} h_{t-1} + b^h) \quad (2)$$

$$y_t = \psi(W^{yh} h_t) \quad (3)$$

となっている。隠れ層  $h_t$  は  $m$  次元ベクトルとし、重み  $W^{xw} \in \mathbf{R}^{n \times d}$ ,  $W^{hx} \in \mathbf{R}^{m \times n}$ ,  $W^{hh} \in \mathbf{R}^{m \times m}$ , バイアス  $b^h \in \mathbf{R}^m$  とし、 $\sigma(\cdot)$  と  $\psi(\cdot)$  は活性化関数を表す。

本稿で適用する LSTM ネットワーク (LSTM) [3][6] は、隠れ層を除いて RNN と同じ構造になっている。LSTM の隠れ層は LSTM ユニットとも呼ばれており、内部状態を保持するゲート  $c_t$  と隠れ層への入力と出力、そして初期化を制御するために三種類のゲート  $i_t$  と  $o_t$ ,  $f_t$  を含んだ構造をしており、

$$h_t = o_t \odot \sigma(c_t) \quad (4)$$

$$o_t = \sigma(W^{ox} x_t + W^{oh} h_{t-1} + W^{oc} c_t + b^o) \quad (5)$$

$$c_t = f_i \odot c_{t-1} + i_t \odot \sigma(W^{hx} x_t + W^{ch} h_{t-1} + b^h) \quad (6)$$

$$i_t = \sigma(W^{ix} x_t + W^{ih} h_{t-1} + W^{ic} c_{t-1} + b^i) \quad (7)$$

$$f_i = \sigma(W^{fx} x_t + W^{fh} h_{t-1} + W^{fc} c_{t-1} + b^f) \quad (8)$$

で定義される。ただし、 $\sigma(\cdot)$  は活性化関数を表し、一般的に *sigmoid* 関数が用いられることが多い。LSTM 内部の活性化関数  $\sigma(\cdot)$  には *hard sigmoid*  $= \max(0, \min(1, 0.2x + 0.5))$  関数を用いる。*hard sigmoid* 関数は *sigmoid* 関数を近似したものであり、*sigmoid* 関数に比べて計算量が少なく済む。マイクロブログは構造的な文章で書かれていることが少ないため、LSTM は位置推定において重要な手がかりとなる単語を蓄積し、不必要な単語は破棄する必要がある。ゆえに、マイクロブログへの LSTM の適用は、有効であると考えられる。本稿では、LSTM の出力に対し、*dropout*[17] と非線形関数  $\text{ReLU}(x) = \max(0, x)$  を持つ *batch normalization*[9] を適用した全結合ニューラルネットワーク (Fully Connect, FC) を追加する。*Dropout* は層に含まれるユニット同士の結合を確率的になくす技術であり、これによりモデルの汎化性能が向上する。また、*batch normalization* は、*interl covariate shift* と呼ばれる学習により特徴量の分布が変化してしまい、収束が遅くなってしまう現象を抑えるための技術である。*Batch normalization* は

$$BN(h; \gamma, \beta) = \beta + \gamma \frac{h - \widehat{E}(h)}{\sqrt{\widehat{Var}(h) + \epsilon}} \quad (9)$$

と表され、特徴量  $h$  の平均を 0, 分散を 1 にする。これらの技術を用いることで、マイクロブログの非構造的な文章に対応することが可能となる。LSTM の詳しい構造を図 2(a) に示す。

### 3.2 画像学習のための CNN

本稿では VGG-16[15] をもとに構造を決めた。CNN は、画像を 2 次元のまま入力として扱い、畳み込みむことで、画像の特徴量を有効的に抽出する手法である。畳み込みは

$$r_i = \phi(W * h + b)_i = \phi \left( \sum_{j,k} (W \otimes x)_{jk} + b \right) \quad (10)$$

で計算される。ただし、畳み込みの重みは  $W \in \mathbf{R}^{d \times d \times m \times n}$ , バイアスは  $b \in \mathbf{R}^n$  となっており、 $d$  はフィルタの大きさ、 $m$  は入力される画像データ  $h$  のチャンネル数、 $n$  は出力される画像データのチャンネル数を表す。VGG-16 モデルの場合、入力される画像サイズは  $224 \times 224$  ピクセルとなっているが、本稿では  $32 \times 32$  ピクセルにし、チャンネル数を順に 32, 32, 64, 64, 128, 128, 256, 256 とした。各層の非線形関数には *ReLU* を採用し、LSTM と同様に *batch normalization* を適用した。

### 3.3 Galaxy モデル

Galaxy モデルは、LSTM と CNN を組み合わせたモデルとなっており、画像付きマイクロブログの投稿内容と画像から投稿された位置の推定を行う。LSTM と CNN それぞれの最終の出力層を取り除き、直前の層からの出力を連結し、FC に入力させる。Galaxy モデルは、

$$h_i = FC([LSTM(\text{text}_i); CNN(\text{image}_i)]) \quad (11)$$

$$y_i = \psi(W h_i + b) \quad (12)$$

であり、 $LSTM(\text{text}_i)$  は  $i$  番目のマイクロブログの投稿内容の LSTM による出力を、 $CNN(\text{image}_i)$  は  $i$  番目のマイクロブログの画像の CNN に出力を表し、 $[a; b]$  はベクトル  $a$  と  $b$  の連結を意味する。LSTM, CNN の重みはすでに学習したものを適用し、学習時に更新はしないように固定し、学習時は式 (11) と式 (12) のパラメータのみを更新する。Galaxy モデルの構造を図 2(c) に示す。

### 3.4 モデルの学習

各モデルの出力はマイクロブログが各次元に対応する位置ラベルである確率を表す多項分布となるため、 $\psi(\cdot)$  をソフトマックス関数

$$\text{softmax} = \frac{\exp(a_i)}{\sum_{j \neq i} \exp(a_j)} \quad (13)$$

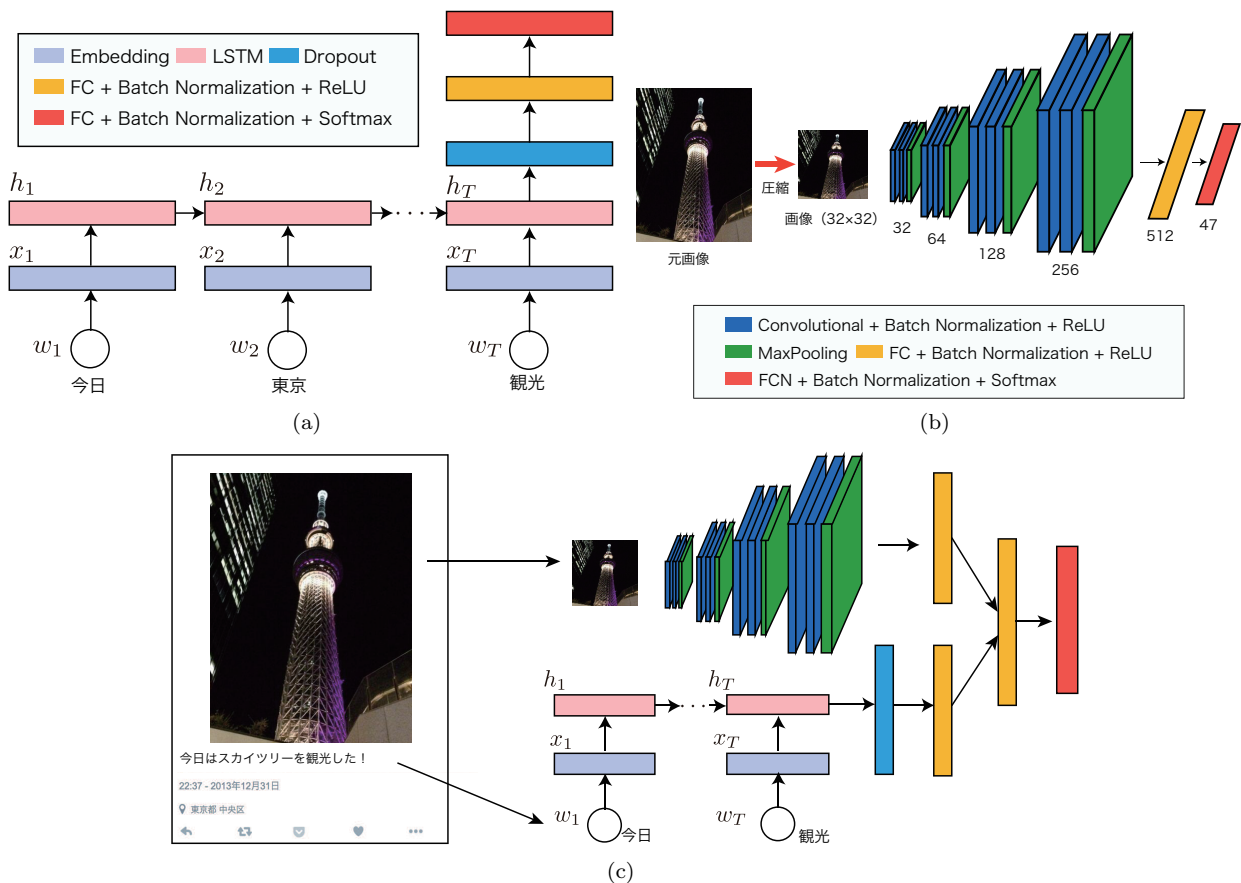


図 2 マイクロブログ投稿位置推定モデルの構造. (a) テキストを学習する LSTM, (b) 画像を学習する CNN, (c) Galaxy モデル.

とする. そして, モデルが最小化する誤差関数は  $n$  番目の位置情報付き発言の  $k$  クラス分類におけるクロスエントロピー誤差となり,

$$E_n = - \sum_k t_{nk} \ln y_{nk} \quad (14)$$

で表される. ただし,  $y_{nk}$  は  $n$  番目のマイクロブログの投稿に対してモデルが推定したラベルの  $k$  次元目の値を,  $t_{nk}$  は  $n$  番目の位置情報付き発言の正解ラベルの  $k$  次元目の値を表す. またパラメータを最適化するアルゴリズムに *adaptive moment estimation* (Adam) [10] を採用する. Adam の特徴として, 計算の効率が良く, 省メモリで済み, スケールの影響も受けにくく, 大規模なデータやパラメータに対して対応が可能なのが挙げられる. なお学習率は全てのモデルにおいて 0.0001 とした.

## 4. 実験

### 4.1 データセット

本稿では, Twitter においてユーザの現在位置情報が付与されたマイクロブログを用いる. その際に収集する対象は, 日本国内の位置情報が付加されたものとし, 収集期間は 2013 年 11 月とした. まず先行研究 [5] に従い, 1 ヶ月の間にマイクロブログを 10 回以上投稿したユーザを抽出

し, また bot 等のユーザを除くため投稿回数が月 500 回以上であるユーザを削除した.

#### 4.1.1 前処理

本稿で用いるマイクロブログは日本語中心に書かれており, あらかじめ前処理を行う必要がある. まず日本語以外の記号及び英数字を取り除く. 次に, 形態素解析エンジンである mecab\*<sup>1</sup> を用い, 名詞のみを抽出する. 抽出した名詞のうち, 収集した全ての位置情報付き発言において出現頻度が 5 回未満, および 50% 以上に出現する単語を取り除く. その結果, 総単語数は 102,053 個となった. 前処理の流れを図 3 に示す.

#### 4.1.2 正解ラベルの付与

分類をする際に, 位置情報付き発言の正解ラベルを付与する必要がある. 本稿では, 国土交通省による平成 26 年度大字・町丁目位置参照情報\*<sup>2</sup>にある日本の 47 都道府県を利用した. 各都道府県にはその中心地となる位置情報が付与されている. そして各マイクロブログの正解クラスは, 各都道府県の緯度経度情報の中で最も近いものに割り当てるという方法を用いた.

\*<sup>1</sup> <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

\*<sup>2</sup> <http://nlftp.mlit.go.jp/isj/>

表 1 収集した位置情報付き発言について

年月	発言数	ユーザ数	10 未満	500 以上	除去後		画像付き
					ユーザ数	発言数	
2013/11	22,938,345	225,845	179,558	121	46,166	17,426,655	120,577

東京八重洲バスターミナル。綺麗に改良されてる\(^o^)/

↓日本語抽出

東京八重洲バスターミナル綺麗に改良されてる

↓品詞分解

東京 / 八重洲 / バスターミナル / 綺麗 / に / 改良 / さ / れ / てる  
名詞 / 名詞 / 名詞 / 名詞 / 助詞 / 名詞 / 動詞 / 動詞 / 動詞

↓名詞のみ抽出

東京 / 八重洲 / バスターミナル / 綺麗 / に / 改良 / さ / れ / てる

↓半角スペースで区切る

東京 八重洲 バスターミナル 綺麗 改良

図 3 テキストの前処理の流れ。マイクロブログの投稿内容から英数字や記号を除き、日本語のみにする。次に品詞で分解し、名詞と分類された単語のみを抽出する。そして抽出した単語を半角スペースで区切る。

表 2 実験結果

手法	分類精度 (%)
	Acc
ランダム	2.15
最頻地域区分	17.12
LSTM (テキスト)	35.34
CNN (画像)	29.91
Galaxy (両方)	<b>49.22</b>

#### 4.2 投稿位置推定精度の比較

位置推定における提案手法の有効性の検証のため、本実験では都道府県の推定精度に関する指標を用いた。

都道府県の推定精度に関する指標には、マイクロブログの正解都道府県と N 個の推定した都道府県候補に正解が含まれていた場合の精度 (Acc) のマイクロ平均を用いた。

ベースラインには、ランダムと最頻地域区分を用いた。ランダムは、ランダムで都道府県を推定都道府県とする。最頻地域区分では、訓練データに最も多く含まれている都道府県を推定都道府県とする。データは訓練、テストデータに 8:2 にわけ、実験に用いた。

画像付きマイクロブログの位置推定を行った結果を表 2 に示す。表 2 より、ランダムの場合  $\frac{1}{47} \approx 2.128\%$  となっている。次に最頻出区分では、もっとも多かったのは東京都であり、テストデータ全体の 17% を占めていた。テキストのみを用いた LSTM や画像のみを用いた CNN に比べ、2 つを組み合わせた Galaxy モデルの位置精度が最も良かった。このことから、Galaxy モデルがテキストと画像それぞれから位置推定に有効である特徴量を学習していることがわかる。

表 3 予備実験結果 (都道府県名が含まれていないマイクロブログの場合)

手法	分類精度 (%)
	Acc
SVM	13.634
LSTM	<b>17.261</b>

#### 4.3 考察・今後の課題

都道府県を推定する先行研究として鈴木ら [24] の研究が挙げられる。鈴木らは *opal*<sup>\*3</sup> を用いているが、*opal* を提案している Yoshinaga ら [23] の実験によって *support vector machine* (SVM) が同等の推定精度と報告されている。本稿では、先行研究手法との比較を行っていないため、Galaxy モデルが有意であることは示せていない。予備実験にて、都道府県名がテキストに含まれていないマイクロブログの位置推定を行った結果を表 3 に示す。この結果から、本稿で用いたデータとは一部異なるため、直接の比較はできないが、SVM よりも LSTM の方が位置推定において有効であることがわかるため、Galaxy モデルが先行研究よりも有効であることが期待される。

本稿はメモリの制約上などから画像を最初に  $32 \times 32$  にリサイズしたが、この結果かなりの情報が失われていると考えられる。本稿で参考にした VGG-16 は  $224 \times 224$  ピクセルであり、今後この画像のリサイズを大きくする、もしくは任意のサイズでも入力可能なモデルを提案することで、リサイズに伴う情報量の損失を抑え、位置推定精度の向上に期待できると考えられる。

PlaNet[21] では、画像から位置推定を行う作業を実際に人にやらせようゲームを Web 上で公開し、その結果 PlaNet が人よりも高い精度で推定できることを報告している。今後、本研究においても同様に人々に実際に画像付きマイクロブログから位置推定を行ってもらい、Galaxy モデルとの比較なども行っていく必要があると考えている。

#### 5. 結論

本稿では、画像付与されたマイクロブログの投稿位置推定のため、Galaxy モデルを検討した。位置推定タスクを多クラス分類問題としてみなし、Galaxy モデルに確率分布を出力させた。これによって、マイクロブログが投稿された位置を表現し、日本の各都道府県で投稿された確率を割り当てることが可能となる。従来の手法では、マイクロブログの投稿内容を用いて特徴量を人手で設計し、位置推定を行っていた。Galaxy モデルはマイクロブログの投稿

\*3 <http://www.tkl.iis.u-tokyo.ac.jp/ynaga/opal/>

のテキストをRNNによって学習し、マイクロブログの画像をCNNによって学習することで位置推定を可能とする。実験において、ベースラインと比較した結果、テキストと画像の両方を組み合わせたGalaxyモデルが最も精度よく位置を推定することができた。

本稿では、日本で投稿されたマイクロブログの都道府県、及び付与された位置情報との誤差距離の推定に取り組んだが、その他の行政区分や地域メッシュ<sup>\*4</sup>などの粒度や海外の国々にも提案手法は適用が可能であり、今後様々なアプリケーションに適用することができ、実世界を理解するための手がかりとして役立つと期待している。

## 参考文献

- [1] Zhiyuan Cheng, James Caverlee, and Kyumin Lee. You are where you tweet: a content-based approach to geolocating twitter users. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pp. 759–768. ACM, 2010.
- [2] Cícero Nogueira dos Santos and Maira Gatti. Deep convolutional neural networks for sentiment analysis of short texts. In *COLING*, pp. 69–78, 2014.
- [3] Felix Gers. Long short-term memory in recurrent neural networks. *Unpublished PhD dissertation, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*, 2001.
- [4] Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.
- [5] Bo Han, Paul Cook, and Timothy Baldwin. Text-based twitter user geolocation prediction. *Journal of Artificial Intelligence Research*, pp. 451–500, 2014.
- [6] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, Vol. 9, No. 8, pp. 1735–1780, 1997.
- [7] Liangjie Hong, Amr Ahmed, Siva Gurumurthy, Alexander J Smola, and Kostas Tsioutsoulouklis. Discovering geographical topics in the twitter stream. In *Proceedings of the 21st international conference on World Wide Web*, pp. 769–778. ACM, 2012.
- [8] Yohei Ikawa, Miki Enoki, and Michiaki Tatsubori. Location inference using microblog messages. In *Proceedings of the 21st international conference companion on World Wide Web*, pp. 687–690. ACM, 2012.
- [9] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- [10] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [11] Sheila Kinsella, Vanessa Murdock, and Neil O’Hare. I’m eating a sandwich in glasgow: modeling locations with tweets. In *Proceedings of the 3rd international workshop on Search and mining user-generated contents*, pp. 61–68. ACM, 2011.
- [12] Ryan Kiros, Yukun Zhu, Ruslan R Salakhutdinov, Richard Zemel, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. Skip-thought vectors. In *Advances in Neural Information Processing Systems*, pp. 3276–3284, 2015.
- [13] Stephen Roller, Michael Speriosu, Sarat Rallapalli, Benjamin Wing, and Jason Baldridge. Supervised text-based geolocation using language models on an adaptive grid. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1500–1510. Association for Computational Linguistics, 2012.
- [14] Aliaksei Severyn and Alessandro Moschitti. Twitter sentiment analysis with deep convolutional neural networks. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 959–962. ACM, 2015.
- [15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [16] Richard Socher, Jeffrey Pennington, Eric H Huang, Andrew Y Ng, and Christopher D Manning. Semi-supervised recursive autoencoders for predicting sentiment distributions. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 151–161. Association for Computational Linguistics, 2011.
- [17] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1929–1958, 2014.
- [18] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.
- [19] Kai Sheng Tai, Richard Socher, and Christopher D Manning. Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*, 2015.
- [20] Oriol Vinyals and Quoc Le. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.
- [21] Tobias Weyand, Ilya Kostrikov, and James Philbin. Planet-photo geolocation with convolutional neural networks. *arXiv preprint arXiv:1602.05314*, 2016.
- [22] Benjamin P Wing and Jason Baldridge. Simple supervised document geolocation with geodesic grids. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pp. 955–964. Association for Computational Linguistics, 2011.
- [23] Naoki Yoshinaga and Masaru Kitsuregawa. Kernel slicing: Scalable online training with conjunctive features. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 1245–1253. Association for Computational Linguistics, 2010.
- [24] 鈴木有, 鍛冶伸裕, 吉永直樹, 豊田正史. 過去の投稿を活用したマイクロブログユーザの現在位置推定. In *DEIM Forum*, 2015.

<sup>\*4</sup> <http://www.stat.go.jp/data/mesh/>