

# ハッシュタグ付与精度向上のためのSNS投稿内のメディア選択手法の提案

関谷虎汰郎<sup>†</sup> 鈴木 優<sup>†</sup>

<sup>†</sup> 岐阜大学知能情報学領域知能理工学専攻 〒501-1193 岐阜県岐阜市柳戸1番1

E-mail: <sup>†</sup>a4525044@edu-u.ac.jp, <sup>††</sup>ysuzuki@gifu-u.ac.jp

**あらまし** SNSの投稿に対して、機械学習を用いて投稿の内容にあうハッシュタグを付与する研究がいくつか行われている。これらの研究では、投稿内に含まれている画像やテキストといったメディアのなかから事前に入力に用いるものを決めて行われている。しかし、対象とする話題によっては、投稿ごとにハッシュタグを付与しやすいメディアが異なるのではなかと考えた。そこで、ハッシュタグ付与に用いるメディアを投稿ごとに選択することで精度が向上すると考え、研究を行った。本稿では、あらかじめ決めた話題の投稿を収集し、ハッシュタグ付与に用いるメディアを選択することで精度の向上を試みる。それぞれのメディアからハッシュタグを予測する。予測したハッシュタグの正誤を正解ラベルとすることで、ハッシュタグ付与に用いるメディアを選択するモデルを構築する。メディアを選択しない手法を用いたハッシュタグ付与の精度と比較することで、メディアを選択することにより精度が向上する話題があることを示す。

**キーワード** SNS, ソーシャルメディア, ハッシュタグ, 機械学習, Twitter

## 1 はじめに

Twitter<sup>1</sup>やInstagram<sup>2</sup>といったソーシャル・ネットワーキング・サービス(SNS)にはハッシュタグという機能が存在する。ハッシュタグとは、ユーザーが自分の投稿に付けるハッシュ記号#を先頭とした文字列である。ハッシュタグは投稿内の意見や話題を明示的に示している。また、ハッシュタグを用いることで、ユーザーは同じハッシュタグが付与された投稿の一覧を表示させるなど様々な利用方法が考えられている。

しかし、ハッシュタグの付与はユーザーの自主性に依存しており、ハッシュタグの付与されていない投稿も多い。そこで、機械学習を用いて自動的にハッシュタグを付与することを考える。これにより、ハッシュタグの付与されていない投稿を少なくすることで、ハッシュタグの利便性を高めることができる。

SNSの投稿内には文章や画像といった複数のメディアが含まれている。また、メディアにはユーザーの意見が記載された文章や、話題の中心となる人物の画像など投稿の内容に関連する情報が含まれていることが多い。そこで、メディアを入力とし、投稿の内容にあったハッシュタグを機械学習を用いて付与する。しかし、SNSの投稿には、内容と関連の少ないメディアが含まれていることもある。

例として、観光へ行った際の投稿を考える。観光地の名前をハッシュタグとして付与することで、旅行の行先を決めたいユーザーの手助けになると考えられる。テキストでは「きれいな景色をたくさん見れて楽しかった」といった旅行へ行った感想のみが記述されているとする。また、画像として観光地で撮



図1 Tweetに添付されている画像例

影した写真が添付されているとする。テキストに注目した場合、感想だけでは観光地に関する情報が乏しく、観光地の名前とは別のハッシュタグが付与されることが考えられる。画像に注目した場合、写真の景色を基に、観光地の名前をハッシュタグとして付与することが可能であると考えられる。また、観光に関連する投稿では、観光地に関連する情報は画像が多く持っていることが多いと考えられる。そのため、観光に関する投稿に対し内容にあったハッシュタグを付与する場合、あらかじめ画像に注目することで高い精度でハッシュタグを付与できると考えられる。

しかし、あらかじめ注目するメディアを決めることで精度が下がる場合が考えられる。例として、図1が添付されている投稿を考える。この投稿は、文章では「今の政権では行先が不安」といった現政権に対する意見が記述されており、ユーザーの感情を表すために画像が添付されているものとする。この投稿に対し、内容に適したハッシュタグとして「#与党不支持」を付与する。文章に注目すると政治に関連する単語が存在し、政権に対する不満も読み取れる。したがって、機械学習を用いてハッシュタグを付与する際、「#与党不支持」を付与することができ

1: <https://twitter.com/>

2: <https://www.instagram.com/>

ると考えられる。一方で、画像に注目した場合、政治に関する情報が入っておらず、政治とは関連のないハッシュタグが付与されることが考えられる。例のような投稿の場合、文章に注目する、あるいは画像と文章の両方に注目してハッシュタグを付与する必要があると考えられる。

別の例として、政権を批判している記事の写真が添付されており、文章では「もっといろんな人に知ってほしい」といった記述がされている投稿を考える。このような投稿の場合、文章からのハッシュタグ付与は難しく、画像に注目してハッシュタグを付与する必要があると考えられる。

上記の2つの例のように、投稿の内容と関連の少ないメディアが存在し、投稿ごとに注目したいメディアが異なることがある。このような投稿にハッシュタグを付与する際、事前に用いるメディアを決めていた場合、ハッシュタグ付与の精度が下がる要因となる。そこで、投稿の内容にあったハッシュタグを付与する際、投稿ごとに適切なメディアを選択することでハッシュタグ付与の精度が向上すると考えた。

本稿では、特定の話題において、ハッシュタグ付与に用いるメディアを選択することで精度が向上することを示す。適切なメディアを機械学習を用いて自動的に選択する手法について考える。適切なメディアとは、ハッシュタグ付与のタスクにおいて、入力に用いた際、投稿の内容にあったハッシュタグを付与できる可能性の高いメディアとする。各メディアからハッシュタグを付与した事例を訓練セットとすることで投稿内のメディアの中から適切なメディアを推定することができると考えた。このとき、投稿内のメディアを結合し、1つの入力データとする。これにより、画像とテキストの両方の特徴を考慮して推定することが可能である。

政治に関連する投稿を Twitter から収集し、投稿ごとに入力に用いるメディアを選択することにより、ハッシュタグ付与の精度が向上することを示すため実験を行った。実験の結果から、政治に関する Tweet を対象としたとき、大部分の Tweet で画像を入力として選択することとなった。そのため、メディア選択のハッシュタグ付与精度への影響は少なかったが、精度の向上が見られた。これにより、メディア選択によりハッシュタグ付与精度が向上する話題が存在することを示した。

## 2 関連研究

ハッシュタグを自動的に付与する研究は既にいくつか行われている。Qi ら [1] は Tweet のハッシュタグ付与タスクを分類問題として扱い、テキストを CNN モデルに入力することで分類を行った。さらに、テキスト内からハッシュタグ付与の際に注目すべき部分を算出し、CNN モデルと組み合わせる手法を提案した。また、Qi ら [2] はハッシュタグの付与を翻訳タスクとして扱い、テキストからハッシュタグ付与を行った。テキストとハッシュタグを異なる言語であるが同じ意味を持っている文字列として扱い、翻訳モデルを適用した。さらに、テキスト内のトピックはハッシュタグと関わりがあるとして、トピック抽出手法と翻訳モデルを組み合わせる手法を提案している。ま

た、Qi ら [3] はハッシュタグの付与を分類問題として扱い、テキストだけでなく画像も用いて分類を行っている。テキストと画像の特徴量をそれぞれ抽出する。テキストの特徴量を画像の特徴量を用いて補正し、画像の特徴量を補正後のテキストの特徴量を用いて補正する。その後、補正した特徴量を用いて分類を行うことで、異なるメディアを用いた分類を行っている。

上記の研究では、あらかじめ用いるメディアをテキスト、あるいはテキストと画像の両方と決め、ハッシュタグの付与を行っている。本研究では、SNS の投稿ごとに用いるメディアを選択することで、ハッシュタグ付与の精度が向上する話題が存在することを示す。

異なるメディアを入力に用いて総合的に学習を行うことで精度向上を図る研究がいくつか行われている。その中で、テキストと画像を結合し、分類に用いる手法がいくつか提案されている。Ignazio [4] らは商品の画像に商品の説明文を画像に結合して商品の分類を行っている。商品の説明文から特徴量を抽出し、得られた特徴量からテキストを画像に変換する、商品の画像にテキストを変換した画像を埋め込むことで1枚の画像とし、任意の CNN モデルに適応できる手法を提案している。三條ら [5] はクックパットにおける料理レシピの人気予測を料理の画像と料理の説明文を結合したデータを入力とし、予測を行っている。AlexNet や LSTM を用いて画像と説明文から特徴量を抽出し、得られた特徴量を結合する。結合した特徴量を全結合層からなるネットワークに入力することで学習を行っている。

本稿では、ハッシュタグ付与に用いるメディアを機械学習により選択する際、テキストと画像を結合したデータを入力として学習を行っている。異なるメディアを結合する手法と構築するモデルとして三條らの論文に記載されているモデルを用いて実験を行った。

## 3 提案手法

SNS の投稿が与えられたとき、内容にあったハッシュタグを付与するためにどのメディアを用いるかの推定を分類問題として扱う。提案手法の概要を図2に示す。SNS の投稿として Twitter の投稿である Tweet を用いる。また、メディアとして Tweet 内の文章からハッシュタグを除いたテキストと画像を対象とする。提案手法によるメディアの選択と選択したメディアを用いたハッシュタグの付与を以下の流れで行う。

- (1) ハッシュタグの付いている Tweet を用意する。
- (2) Tweet 内のテキストを用いてハッシュタグを推定する。
- (3) Tweet 内の画像を用いてハッシュタグを推定する。
- (4) ハッシュタグの推定結果と Tweet 内のハッシュタグを比較して各 Tweet に四種類のラベルを付与する。
- (5) ラベル付けした Tweet を用いてモデルを構築する。
- (6) 構築したモデルを用いて Tweet を四種類に分類する。
- (7) 分類結果を基に、ハッシュタグの付与に用いるメディアを選択する。
- (8) 選択したメディアを用いてハッシュタグを付与する。

### 3.1 データセットの作成

ハッシュタグの付与されている Tweet を用意する。このとき、Tweet に付いているハッシュタグを内容にあったハッシュタグとして扱う。Tweet 内のメディアを用いて内容にあったハッシュタグを機械学習を用いて推定する。推定結果と Tweet に付いているハッシュタグを比較し、同じハッシュタグであれば正しいハッシュタグを推定出来たものとして扱う。各 Tweet に対し、推定結果を基にラベルを付与することでデータセットを作成する。

#### 3.1.1 Tweet 内のテキストを用いたハッシュタグ推定

本稿では内容にあったハッシュタグの推定を分類問題として扱う。Tweet 内のテキストのみを用いてハッシュタグを分類することを考える。ハッシュタグの分類を BERT [6] を用いて行う。BERT とは自然言語処理タスクにおいて、汎用性の高いモデルである。転移学習を用いることで様々な自然言語処理タスクで高い精度を発揮している。データセットに用いる Tweet とは別にテキストを用意し、Tweet 内のハッシュタグを正解ラベルとしてファインチューニングを行う。データセットに用いる Tweet 内のテキストを BERT に入力し、ハッシュタグを分類する。分類したハッシュタグが Tweet に含まれている場合、正しいハッシュタグが分類できたものとする。

#### 3.1.2 Tweet 内の画像を用いたハッシュタグ推定

Tweet 内の画像のみを用いてハッシュタグを分類することを考える。ハッシュタグの分類を EfficientNet [7] を用いて分類する。EfficientNet とは画像処理タスクにおいて、汎用性の高いモデルである。BERT と同様に、転移学習を用いることで画像処理タスクで高い精度を発揮している。データセットに用いる Tweet とは別に画像を用意し、Tweet 内のハッシュタグを正解ラベルとしてファインチューニングを行う。データセットに用いる Tweet 内の画像を EfficientNet に入力し、ハッシュタグを分類する。

#### 3.1.3 Tweet に対するラベル付け

3.1.1 節と 3.1.2 節の分類結果を基にラベル付けを行う。3.1.1 節の分類結果は正しいが、3.1.2 節の分類結果が間違っている場合、“テキストのみ”のラベルを付与する。3.1.1 節の分類結果は間違っているが、3.1.2 節の分類結果が正しい場合、“画像のみ”のラベルを付与する。3.1.1 節と 3.1.2 節の分類結果が共に正しいハッシュタグを分類できていれば“両方”のラベルを付与する。3.1.1 節と 3.1.2 節の分類結果が共に間違っている場合、“どちらでもない”のラベルを付与する。

### 3.2 モデル構築

3.1 節で作成したデータセットを用いてモデルを構築する。構築する際、Tweet 内のテキストと画像を結合し、1 つの入力データ  $S$  を形成する。このとき、以下の二つの手法で結合する。

- 各メディアをベクトルに変換し、結合する。
- 各メディアから特徴量を抽出し、結合する。

構築するモデルは全結合層を隠れ層として持つ。活性化関数として式 1 のように定義される Relu 関数を適用した後に Dropout

を適用する。ここで、 $x_i$  は全結合層のユニットの 1 つから出力された値である。

$$Relu(x_i) = \begin{cases} 0, & x_i \leq 0 \\ 1, & x_i > 0 \end{cases} \quad (1)$$

出力結果として 4 次元のベクトルである出力  $O = (o_0, o_1, o_2, o_3)$  を得る。この出力  $O$  に対して式 2 のように定義される Softmax 関数を適用することで  $O' = (o'_0, o'_1, o'_2, o'_3)$  を得る。 $O'$  の各要素  $o'_i$  は入力データ  $S$  がラベル ( $i$ ) に属する確率を表している。よって、 $O'$  の要素内で最大値をとる要素に対応したラベルが分類結果となる。入力データ  $S$  からラベルを分類するタスクを解き、重みを学習する。

$$o'_i = Softmax(o'_i) = \frac{\exp(o'_i)}{\sum_{k=0}^n \exp(o'_k)} \quad (k = 1, 2, 3, 4) \quad (2)$$

#### 3.2.1 メディアをベクトルに変換し、結合する手法

テキストと画像をベクトル化し、連結することで 1 つ入力データ  $S$  を形成することを考える。Tweet 内のテキストを形態素解析することで、単語ごとに分割する。形態素解析には、辞書として mecab-ipadic-NEologd を適用させた MeCab を用いる。Tweet 内の単語毎に fastText [8] の分散表現を用いて単語ベクトル  $W_i$  にする。Tweet 内の単語ベクトルの平均  $\bar{W}$  を式 3 のように計算する。

$$\bar{W} = \frac{\sum_{i=0}^N W_i}{N} \quad (i = 1, 2, 3...N) \quad (3)$$

ただし、 $N$  は Tweet 内の単語の総数とする。 $\bar{W}$  を全結合層に入力し、得られた出力をテキストのベクトルとする。Tweet 内の画像を RGB の値を基に数値化し、平坦化したものを全結合層に入力する。全結合層の出力結果を画像のベクトルとする。テキストのベクトルと画像のベクトルを連結したデータを入力データ  $S$  とする。

#### 3.2.2 抽出した特徴量を結合する手法

テキストと画像からそれぞれ特徴量を抽出したのちに連結することで 1 つの入力データ  $S$  を形成することを考える。Tweet 内のテキストを学習済みの BERT に入力する。BERT の出力層への入力を取得し、全結合層に入力する。全結合層の出力結果をテキストの特徴量として扱う。Tweet 内の画像を学習済みの EfficientNet に入力する。EfficientNet の出力層への入力を取得し、全結合層に入力する。全結合層の出力結果を画像特徴量として扱う。テキストの特徴量と画像の特徴量を連結したデータを入力データ  $S$  として扱う。

### 3.3 構築したモデルによる分類とメディアの選択

与えられた Tweet 内のテキストと画像をモデル構築の際に用いた手法で結合する。その後、構築したモデルに入力することで 4 クラスに分類される。“テキストのみ”に分類された場合、テキストからハッシュタグを推定することで正しいハッシュタグを付与できると考えられる。したがって、テキストを選択する。“画像のみ”に分類された場合、画像を選択する。“両方”に分類された場合、3.1.1 節と 3.1.2 節の推定結果から精度の高

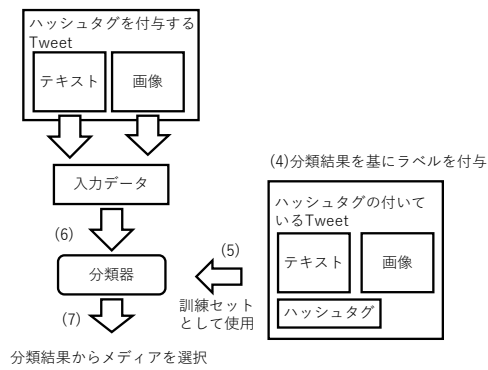


図2 メディア選択の流れ

い結果を得られたメディアを選択する。”どちらでもない”に分類された場合、ハッシュタグを推定するには特徴量が不十分であると考えられる。したがって、画像とテキストの両方を選択する。

### 3.4 ハッシュタグの付与

3.2 節で選択されたメディアを用いてハッシュタグを付与する。テキストを選択した場合、モデルはデータセットのテキストでファインチューニングを BERT を用いる。画像を選択した場合、モデルはデータセットの画像でファインチューニングをした EfficientNet を用いる。両方を選択した場合、メディア選択に用いたモデルと同様のネットワークを使用する。ただし、データセットのテキストと画像を 3.2 節でモデルを構築する際に用いて手法で結合し、ハッシュタグを正解ラベルとして学習したモデルを用いる。

## 4 評価実験

提案手法に従いハッシュタグ付与を行う。メディア選択を行わない手法と精度を比較することで、メディアを選択することで精度が向上する話題が存在することを示す。

### 4.1 手順

Twitter API を用いて Tweet の収集を行う。このとき、収集期間を 2021 年 12 月 16 日以前とし、Retweet を除いた Tweet をキーワード検索を用いて収集した。Twitter において、日本の政治に関連する話題の Tweet が多く見られる。政治に対する意見を表現する方法として、テキストに記載する他、画像を加工することで表現するなど Tweet ごとに異なる。そのため、メディア選択により精度が向上する話題であると考えた。対象とする話題を政治として実験を行う。Tweet 収集の検索条件として、政治に関する考えが分かる”#自民党支持”，”#自民党政治を終わらせよう”をキーワードとした。Tweet を 738 件収集し、369 件を 3.1 節にしたがって第一著者がラベルを付与し、データセットを作成した。このとき、URL の削除やリプライ情報の削除といった下処理を行った。また、画像が複数ある Tweet の場合、そのうちの一枚をランダムに選択し、使用した。また、ハッシュタグが複数付与されている場合、キーワード検索で用

いたハッシュタグのみを使用し、その他のハッシュタグはハッシュ記号を削除し、テキストとして扱った。

提案手法のモデル構築の際、leave-one-out 交差検証を行い、最も精度の高かったハイパーパラメータのモデルを使用した。変更したハイパーパラメータの一覧を表 1 に示す。データセットのうち、329 件のデータを訓練セットとして学習を行い、40 件のデータを用いて学習の収束を確認した。1 件のデータをテストセットとして推定を行った。訓練セットとテストセットの組み合わせを全通り行い、その平均をモデルの精度とした。このとき、学習の収束の確認に使用したデータは常に同じデータを用いた。データセットとは別に 369 件を 3.1.1 の BERT モデル、3.1.2 の EfficientNet モデルのファインチューニングに用いた。学習の際、各クラス数が同数になるようにデータ数の調整を行った。データ数の少ないクラスのデータをランダムにクラス間のデータ数の差だけ選択し、複製することでデータ数の調整を行った。

実験において、学習済み BERT モデルとして、東北大学の乾・鈴木研究室の訓練済み日本語 BERT モデル<sup>3</sup>を使用した。訓練済み BERT モデルは日本語版 Wikipedia を用いて事前学習が行われており、語彙数は 32,000 となっている。ファインチューニングの際、BERT モデルのパラメータは、最終層のみ更新するように設定する。また、学習済み EfficientNet モデルとして、github 上で公開されているモデル<sup>4</sup>を用いた。ImageNet<sup>5</sup>を用いて事前学習が行われており、モデルの 1 つである efficientnet-b0 を用いた。テキストをベクトル化する際に用いた fastText において、単語ベクトルの次元数を 300 次元、前後 10 単語内に出現する単語を予測するタスクにより学習を行った。また、訓練セットとして Twitter から 2020 年 1 月上旬から 11 月上旬までの期間で収集した 175,767,614 件の Tweet の文章を使用した。

メディアの選択が分類精度へ与える影響を調べるため、精度の比較を行う。メディア選択を用いない手法として、分類に用いるメディアを事前に決めた際の分類精度を比較対象とする。提案手法とは別に以下の 3 つの手法でハッシュタグ分類を行った。

- データセット内のテキストのみを用いた際の精度
- データセット内の画像のみを用いて分類を行った際の精度
- データセット内のテキストと画像を結合したデータを用いて分類を行った際の精度

テキストのみを用いた分類には BERT を、画像のみを用いた分類には EfficientNet をそれぞれモデルとして用いた。テキストと画像を結合したデータを用いる分類には提案手法と同じモデルを使用し、入力結合する手法として 3.2.2 節の手法を用いた。提案手法のモデル構築の際に用いたデータと同じ訓練セットと検証セットを用いてファインチューニングと学習を行い、テストセットを分類した。ただし、正解ラベルとして付与

3 : <https://github.com/cl-tohoku/bert-japanese>

4 : <https://github.com/lukemelas/EfficientNet-PyTorch>

5 : <https://www.image-net.org/index.php>

したラベルは用いず、各 Tweet に付与されていたハッシュタグを用いた。

## 4.2 結果と考察

### 4.2.1 データセット作成

まず、3.1.1 節と 3.1.2 節の手順に従いハッシュタグの分類を行った。分類結果の混同行列を表 2 と 3 に示す。＃自民党政治を終わらせようを”不支持”，＃自民党支持を”支持”と記載する。また、Tweet に付与されているハッシュタグを正解とした。精度を比較したとき、3.1.1 節の分類結果は  $accuracy = 0.62$ 、3.1.2 節の分類結果は  $accuracy = 0.81$  であった。これにより、メディア選択の際、”両方”と判定された Tweet は画像を用いてハッシュタグ付与を行う。また、ラベルを付与した際のそれぞれのデータ数を表 4 に示す。＃両方のデータ数が最も多く、＃テキストのみと＃どちらでもないのデータ数が少なかった。＃画像のみのデータ数がラベルのデータ数より多いことから、今回のタスクにおいて、画像を用いたハッシュタグ付与はテキストを用いるよりも容易であることが分かる。

＃テキストのみが割り振られた Tweet の画像を確認したところ、ピントのずれた写真や他政党の人物が映っている写真が＃画像のみの Tweet の画像よりも多かった。また、政治とは関連のない画像を政治上の出来事に例えている画像が存在した。例として、図 3 のような画像を示す。画像は男性が本を捨てながら片付けをしているイラストである。本を行政上の書類に見立て、重要な書類を捨てている状況を表している。このような画像は政治とは直接的な関連が乏しく、テキストが選択されるべき Tweet であると考えられる。テキストを確認したところ、自民党に対する記載が多く、さらに、「支持する」や「政権交代に行こう」といった自民党に対する考えが分かりやすい単語を含む Tweet が多かった。また、＃テキストのみが割り振られた Tweet 群にハッシュタグの偏りは少なかった。＃画像のみの Tweet の画像を確認したところ、政治家の写真が多く、また、画像内にユーザーの考えが分かる文章が埋め込まれているものが多かった。また、テキストを確認したところ、具体的な人名を挙げているものが多かった。＃画像のみの Tweet 群は＃不支持の Tweet が多く、偏りが見られた。＃両方が割り振られた Tweet の画像は＃画像のみの Tweet と同様に政治家の写真や、文章が埋め込まれた画像が多かった。文章は＃テキストのみの Tweet と同様に自民党に対する考えが分かる単語を含む Tweet が多かった。＃どちらでもないが割り振られた Tweet は数が少なく、大きな特徴は見られなかった。

### 4.2.2 メディア選択

次に、3.2 節の手順に従いモデルを構築し、分類を行った。モデルによる分類結果の混同行列を表 5 と 6 に示す。表中では、＃テキストのみを T、＃画像のみを I、＃両方を C、＃どちらでもないを N と表記する。テストセットの Tweet に付与されていたラベルを正解とした。データ数を最も数の多い＃両方の Tweet に合わせてデータ数の調整を行い、学習を行った。また、3.2.1 節の手順に従い入力データを形成した際の分類結果



図 3 「テキストのみ」が割り振られた Tweet の画像例

を 3.2.1 節の分類結果、3.2.2 節の手順に従い入力データを形成した際の分類結果を 3.2.2 節の分類結果と記載する。3.2.1 節の手法は＃両方のみ出力する結果となった。そのため、うまく学習が行えていないと考えられる。これは各メディアをベクトル化した後の処理として用いた平均化や平坦化の際に、考慮すべき特徴量が抽出できず、学習が行えなかったのではないかと考えられる。以降の結果において、3.2.2 節の手法のみ記載する。

メディア選択の結果、＃両方のラベルが最も多く、＃どちらでもないのラベルは出現しなかった。精度は  $accuracy = 0.67$  となった。テキストを用いたハッシュタグの分類は 55 件となり、残りの 274 件は画像を用いたハッシュタグの分類を行った。

### 4.2.3 ハッシュタグ分類と精度比較

最後に、ハッシュタグの分類を行い、精度の比較を行った。Tweet に付与されているハッシュタグを正解とした。それぞれの手法の混同行列を表 7-10 に示す。画像を用いた分類の結果は  $accuracy = 0.77$ 、テキストを用いた分類の結果は  $accuracy = 0.66$ 、テキストと画像の両方を用いた分類の結果は  $accuracy = 0.67$  となり、事前にハッシュタグ付与に用いるメディアを決める場合、画像を用いたハッシュタグを付与することで高い精度が得られる。これに対し、提案手法による分類結果は  $accuracy = 0.80$  となり、提案手法による分類が最も精度が高かった。

画像を用いた分類と提案手法を比較すると、大きな差はなく、メディア選択が精度へ及ぼす影響が小さいことが考えられる。そのため、メディア選択の結果が偏りの少ない話題を対象として実験を行う必要がある。また、＃テキストのみと割り振られた Tweet に注目すると、画像よりもテキストも用いることでハッシュタグ付与の精度が向上している。これは、メディア選択により精度が向上した実例であると考えられる。これにより、ハッシュタグ付与に用いるメディアを選択することで精度が向上する話題が存在することを示すことが出来たと考えられる。

テキストを用いたハッシュタグ分類と画像を用いたハッシュタグ付与分類の結果を比較する。各ラベルの recall を比較すると、不支持のラベルの recall は画像を用いたハッシュタグ分類のほうが高く、支持のラベルの recall はテキストを用いたハッシュタグ分類のほうが高い結果となっている。この結果は、用いるメディアによって、付与しやすい内容にあったハッシュタグが異なると考えられる。次に、提案手法により＃テキストのみの Tweet において、支持の Tweet が 33 件と

不支持より多く、画像を用いるのと比較し、テキストを用いたハッシュタグ分類が精度が高かった。メディア選択により、付与しやすいメディアを選択できたため、提案手法によるハッシュタグ分類が最も精度の高い結果となったと考えられる。また、用いるメディアによって recall に差があるタスクに対して、メディアを選択する手法は有用であると考えた。この仮説を実証するために、異なる話題で同様の実験を行い、結果を比較する必要がある。

5 おわりに

本稿では、SNS のハッシュタグ付与をメディア選択することで精度が向上する話題が存在すると考え、研究を行った結果を報告した。

SNS 投稿内のメディアを入力として、ハッシュタグを付与することを考える。このとき、投稿の内容と関連が少ないメディアであった場合、内容に合ったハッシュタグを付与することは困難である。さらに、投稿の内容と関連が少ないメディアが投稿ごとに異なる場合、事前に決めたメディアのみを用いたハッシュタグ付与は精度が下がると考えた。そこで、ハッシュタグ付与に用いるメディアを入力に用いる Tweet ごとに選択することを考えた。

投稿内の各メディアからハッシュタグを推定する。推定結果を投稿に付与されているハッシュタグと比較することでラベル

付けを行い、データセットを作成した。作成したデータセットを訓練セットとして用いることで、どのメディアを用いることで内容にあったハッシュタグ付与できるかを推定するモデルを構築した。構築したモデルの推定結果から入力に用いるメディアを選択する。選択したメディアを用いてハッシュタグ付与を行った。付与結果をメディア選択を用いない手法として画像を用いた付与、テキストを用いた付与、画像とテキスト両方を用いた付与の精度と比較した。これにより、ハッシュタグ付与に用いるメディアを選択することで精度が向上する話題が存在することを示す実験を行った。

実験では、作成したデータセットのうち、大部分で画像を選択することとなり、画像を用いた分類と大きな差は生じなかったが、提案手法により精度の向上が見られた。また、画像を用いた付与とテキストを用いた付与で比較すると、recall に差が生じていた。この結果から、用いるメディアごとに付与しやすいハッシュタグが存在しており、ハッシュタグを付与しやすいメディアを選択できたため、精度が向上したと考えられる。よって、メディアを選択することで精度が向上する話題が存在することを示すことができた。

実験の結果より、複数の改善点が考えられる。メディア選択の結果が大きく偏る結果となった。これにより、メディア選択

表 1 ハイパーパラメータの一覧

ハイパーパラメータ	取りうる値
各全結合層の次元	100 次元, 300 次元, 600 次元, 1000 次元
隠れ層の層数	1, 2
ドロップアウト率	0.0, 0.1, 0.2

表 2 3.1.1 節の分類結果

		分類結果	
		不支持	支持
正解	不支持	71	116
	支持	6	136

表 3 3.1.2 節の分類結果

		分類結果	
		不支持	支持
正解	不支持	157	32
	支持	30	110

表 4 付与したラベルごとのデータ数

ラベル	テキストのみ	画像のみ	両方	どちらでもない
データ数	71	112	171	16

表 5 3.2.1 節の分類結果

		分類結果			
		T	I	C	N
正解	T	0	0	55	0
	I	0	0	106	0
	C	0	0	190	0
	N	0	0	18	0

表 6 3.2.2 節の分類結果

		分類結果			
		T	I	C	N
正解	T	31	14	10	0
	I	14	63	29	0
	C	10	25	155	0
	N	0	2	16	0

表 7 提案手法によるハッシュタグ分類の結果

		分類結果	
		不支持	支持
正解	不支持	159	22
	支持	43	105

表 8 テキストを用いたハッシュタグ分類の結果

		分類結果	
		不支持	支持
正解	不支持	76	105
	支持	6	142

表 9 画像を用いたハッシュタグ分類の結果

		分類結果	
		不支持	支持
正解	不支持	173	8
	支持	68	80

表 10 テキストと画像の両方を用いたハッシュタグ分類の結果

		分類結果	
		不支持	支持
正解	不支持	82	99
	支持	10	138



がハッシュタグ付与の精度へ与える影響が小さいと考えられる。メディア選択の結果が大きく偏っていない話題を対象として実験を行う必要がある。また、対象とする話題を複数用意し、同様の実験を行い、結果を比較することでメディア選択が精度へ及ぼす影響を調べる必要がある。次に、テキストと画像を含み、ハッシュタグが付与されている Tweet を収集することが難しく、十分なデータ量で実験を行うことが出来なかった。そのため、データの収集方法を変える、生成したデータを入れるなど、データセットの作成方法を改めて考える必要がある。そのほかに、対象としたハッシュタグが2種類のみと少なく、分類が容易であったと考えられる。多値分類問題として改めて問題を設定し、実験を行う必要がある。

作成したモデルについて、三條らの研究で用いられたモデルを参考に作成したが、異なるメディアを結合する手法がいくつか提案されている。それらの手法を実装し、比較することで精度への影響を確かめる。また、メディア選択のためのモデルとハッシュタグ付与のモデルがそれぞれ独立したモデルとして作成している。Tweet を入力、ハッシュタグが出力となる end-to-end モデルで作成することで粒度の高い特徴量を考慮してハッシュタグ付与が行うことができる。本稿の実験では考慮出来ていなかった特徴量を考慮することで、精度が向上する可能性がある。

そのほかに、本稿で扱ったテキストや画像以外のメディアとして動画などが考えられる。データを収集することが困難であるがハッシュタグ付与に与える影響を確かめる必要がある。また、ハッシュタグ付与の研究は数多く存在するため、様々な手法を取り入れることで精度へ与える影響を確かめる必要がある。また、本稿で提案した手法をハッシュタグ付与以外に役立てることが出来ないか模索したいと考えている。

**謝辞** 本研究の一部は JSPS 科研費 18H03342, 19H04221, 19H04218 の助成を受けたものです。

## 文 献

- [1] Qi Zhang Yuyun Gong. Hashtag recommendation using attention-based convolutional neural network. "Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI-16, pp. 2782–2788, 2016.
- [2] Qi Zhan Zhuoye Ding, Xipeng Qiu. Learning topical translation model for microblog hash-tag suggestion. "Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, IJCAI-13, pp. 2078–2084, 2013.
- [3] Qi Zhang, Jiawen Wang, Haoran Huang, Xuanjing Huang, and Yeyun Gong. Hashtag recommendation for multimodal microblog using co-attention network. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pp. 3420–3426, 2017.
- [4] Ignazio Gallo, Alessandro Calefati, Shah Nawaz, and Muhammad Kamran Janjua. Image and encoded text fusion for multi-modal classification. *CoRR*, Vol. abs/1810.02001, , 2018.
- [5] 桂井麻里衣三條 智史. Deim forum 2018 p1-4 画像特徴とテキスト特徴に基づく料理レシピの人気予測,.
- [6] 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, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

- [7] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. *CoRR*, Vol. abs/1905.11946, , 2019.
- [8] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomáš Mikolov. Bag of tricks for efficient text classification. *CoRR*, Vol. abs/1607.01759, , 2016.