

# マイクロブログにおける画像投稿の文脈に基づくオンデマンド検索手法

小泉 実加<sup>†</sup> 吉永 直樹<sup>††</sup> 豊田 正史<sup>††</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

<sup>††</sup> 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: †{mkoizumi, ynaga, toyoda}@tkl.iis.u-tokyo.ac.jp

**あらまし** マイクロブログに投稿される大量の画像の中には、実世界で起こっている事象の理解を助けるような有用な画像も多数存在しており、画像投稿をユーザの目的に応じて効率的に収集できればユーザにとって有益となる。マイクロブログにおける画像検索においては、たとえば「犬の写真」のように画像中に存在する具体物に基づいて検索を行いたい場合と、「地震の被害の写真」「心温まる画像投稿」のように画像投稿がなされた状況や意図、あるいはそれを見た人が返す反応や抱く感情のような、投稿が持つ文脈に基づいて検索を行いたい場合がある。しかし、文脈を表現する説明的な語が投稿文に含まれることは少ないため、その文章情報を用いて投稿を適切な投稿文脈に紐付けるのは困難である。画像特徴を用いたとしても、類似した画像が投稿文脈に応じて異なる意味を持つことが多いため、画像から投稿文脈を獲得するのは難しい。そこで本研究では、ユーザが求める投稿文脈を持つ少数の画像投稿をクエリとして、文脈に沿って画像投稿を収集するタスクに取り組む。投稿内の文章や画像の情報に加え、リプライや前後の投稿などを用いて画像投稿の投稿文脈をベクトルで表現し、入力クエリの投稿どうしのベクトル距離が小さくなるようオンデマンドで距離学習を行うことで、投稿文脈の類似した投稿を獲得する。実験では、人手で投稿文脈への分類を行った Twitter のデータを用いて、提案手法の有効性を評価した。

**キーワード** マイクロブログ、画像投稿検索、文脈獲得

## 1. はじめに

現在、多くのソーシャルメディアではテキストだけでなく画像の投稿が可能となっており、投稿された画像は情報発信やユーザ同士のコミュニケーションにおいて重要な役割を果たしている。画像投稿が可能なソーシャルメディアは多岐にわたり、Instagram や flickr など画像投稿が主体のものも存在するが、それぞれのソーシャルメディアごとに投稿内容の傾向は異なる。Twitter を始めとするマイクロブログは使い方の自由度が高く、その分投稿画像の内容も幅広い。実際に投稿される画像としては、日常の体験に関する画像、ニュース性の高い画像や情報の共有や拡散を目的とした画像、イラストをはじめとする創作物の画像などが挙げられる。

マイクロブログにおける画像投稿に顕著な特徴として以下の2点がある。

- (1) 災害や事故の瞬間を現場で捉えた写真のような、速報性のある画像が投稿されやすい。
- (2) 情報伝達を主眼とした画像が投稿されやすい。

(1) は、地震や火事などが起こった際に、たまたま現場に居合わせたユーザがその様子を撮影した画像を投稿するものである。報道機関が現場で撮影する映像より即時性が高いため、マイクロブログに投稿された画像がテレビや新聞の報道で利用されることもある。(2) に関しては、マイクロブログには知識や注意喚起などを広めることを目的とした投稿も多く、その伝達性の向上のために画像が利用されることがある。具体例としては、災害時の対処法を図解したイラストや、有用な情報の掲載され

ている Web サイトのスクリーンショットなどが挙げられる。

このように、マイクロブログに投稿された画像はそのとき起こっているできごとに関する理解や判断を助けるための有用な情報源となる。

Twitter で投稿画像を収集する際には、たとえば「火事の写真」のように画像に含まれる具体物に基づいて検索を行いたい場合もあるが、「地震の被害を写した写真」や「災害への備えに関する情報」、「心温まる画像投稿」といった、画像投稿がなされた状況や意図、あるいはそれを見た人が返す反応や抱く感情(以下、これを投稿文脈と呼ぶ)に基づいて検索を行いたい場合も多い。特に、マイクロブログにおける画像投稿の特徴として上述したような画像を収集する際には、投稿文脈に基づく検索が必要になることが多い。たとえば、「地震の被害を写した写真」というクエリに対して、スーパーの陳列棚の商品が落下している様子や、自宅の本棚が倒れている写真、あるいは家屋が倒壊している写真など、「被害」に当てはまる様々な画像投稿が提示されることが望ましい。このような文脈に基づく抽象的な検索を、画像に付随する文章情報や、画像情報のみを用いて行うことは困難である。以下でその理由を述べる。

まず、投稿文を対象として検索を行う場合、「地震の被害」や「心温まる画像」などの具体性に欠く検索語は投稿文中に出現しにくいという問題や、画像の内容を説明する十分な文章情報が付随しているとは限らないという問題が存在する。特に、マイクロブログにおける画像投稿の特徴(1)で挙げたような写真は急いで投稿されることが多く、さらに写真からある程度状況が理解できることから、説明が不十分であることが多い。一方

で画像を検索対象とした場合、画像が類似していても投稿文脈が異なる場合や、逆に画像が類似していても投稿文脈が同じ場合があり、画像特徴は投稿文脈の推定に十分ではない。たとえば、愛犬の日常を捉えた写真と保護された迷子の犬の写真は、画像自体は似ているものの投稿意図は異なる。

そこで投稿文と画像の両方を用いて検索を行うことを考えると、今度は検索の際にどの特徴量を重視するべきかはクエリの投稿文脈によって異なるという点が問題となる。例を挙げると、自然災害が発生した際に投稿されることが多い「災害の備えに関する情報」という文脈の画像投稿に関しては、災害時に道具がない状態で米を炊く方法の手順を示す画像から、避難時に持ち出すべきもののリストの画像まで様々な表層的特徴を持つ画像が存在する。そのため、画像情報よりも文章情報を重視する必要がある。一方で、「被災地への応援イラスト」という文脈の画像投稿では、同じキャラクターを扱ったイラストが多いなど画像に類似性があるため、画像特徴も重要となる。

そこで本研究では、マイクロブログとして Twitter を取り上げ、Twitter における画像投稿を投稿文脈に基いて収集することを目的とし、上で述べた課題に対処した手法を提案する。まず、投稿文脈に基づく検索では単語や画像をクエリとして用いることが困難であることから、投稿文脈を表現するクエリとして、目的とする投稿文脈を持つ既知の少数の画像投稿を用い、クエリ投稿と文脈が類似している画像投稿を獲得するというタスクを設定する。提案手法では、画像投稿内の文章と画像に加え、投稿内の文章に不足する情報を補うため、その前後に投稿されたツイートやその投稿に対する他ユーザの反応のツイートを用いて画像投稿の投稿文脈ベクトルを構築する。この投稿文脈ベクトルを利用して、検索対象の画像投稿の中から、クエリの投稿と投稿文脈ベクトル間距離の近いものを検索結果として取り出す。ここで、ベクトル間の距離を計算する際にクエリの性質に応じてその各次元に重み付けを行うことで、重視する特徴量の調整を行う。次元ごとの重みは、クエリとして与えられた投稿について、それらの投稿文脈ベクトル間の距離が近くなるように距離学習 [1] を行うことでオンデマンドで算出する。

実験では、多数の文脈に関連画像が投稿され、それらが多くの有用な情報を含んでいたこととして熊本地震を選び、熊本地震に関する複数の文脈について画像投稿の検索を行い、手法の性能を評価した。まず、熊本地震が発生した時期の画像投稿を約 1 万件収集し、その一部を投稿文脈で分類してデータセットを構築した。そして、各文脈について画像投稿の検索を行い、提案手法の有効性の確認と提案手法が有効な文脈の検証。クエリの投稿数を変化させたときの検索性能の評価を行った。

## 2. 関連研究

投稿文脈を考慮した Twitter における画像投稿検索という目的に対して、本研究では画像投稿をクエリとするタスクを設定している。同様のタスクに取り組んだ研究は筆者の知る限りでは存在しないため、本章ではまず、テキストをクエリとした画像検索および Twitter における投稿推薦に関する研究について述べ、その後、Twitter における投稿の内容を関連投稿を用い

て拡張している研究について述べる。

### 2.1 文章を用いた画像検索

画像や動画など多様なメディアを扱うソーシャルメディアの普及に伴い、テキストから画像を検索するようなクロスモーダル検索の必要性が高まり、多くの研究がなされている。

テキストを用いた画像検索手法に関しては、正準相関分析をはじめとする手法でテキスト特徴と画像特徴を同一空間にマッピングし、その空間内で近接するデータを検索結果とする手法 [2] [3] や、異なるモダリティのデータの結合分布を学習することで画像とテキスト間の相関を得る手法 [4] などがある。また、画像と文章のより複雑な対応関係を捉えるため、深層学習によってモダリティ間のマッピングを非線形に行う手法 [5] [6] も提案されている。これらの研究が対象としているのは、文章と画像の表現の間で表層的な対応がとれているデータであり、Twitter のように、文脈次第で表層が類似した画像に全く異なる文章が付与されたり、逆に全く異なる画像に似た文章が付くようなデータに対してそのまま適用するのは困難である。

### 2.2 マイクロブログの投稿推薦

本研究では、投稿文脈を収集するためのクエリとして、テキストではなく、目的とする投稿文脈を持つ既知の画像投稿を用いる。これはマイクロブログ上の投稿推薦に近いタスクであるといえる。しかし以下で挙げる既存研究はいずれもユーザの過去のリツイートの特徴からそのユーザの興味を引きそうな投稿を推薦するというものであり、本研究のように任意に選んだ投稿をクエリとするものではない。

投稿推薦タスクへの取り組みとしては、まず、ユーザとそのユーザが過去にリツイートした投稿の関係を利用して協調フィルタリングを適用するものが挙げられる [7] [8] [9]。投稿されたばかりでユーザの反応がまだ少ないツイートに関しては、どのようなユーザに好まれるかの情報が存在しないことから、ツイート内の単語とユーザ間の関係を用いた手法 [7] や、ハッシュタグとユーザ間の関係を用いた手法 [8] などが用いられている。

また、文脈を考慮した画像投稿推薦に関する研究としては、投稿文章に加え、画像を OCR したデータ、URL が投稿に含まれる場合はそのリンク先の内容、画像を Google 画像検索にかけた結果を画像の文脈として用いて画像投稿の推薦タスクに取り組んだものがある [10]。この手法では、特徴量の性質から、画像中に文章が含まれるような加工済み写真やスクリーンショット、ニュース記事などの外部ページの写真、既に検索エンジンで内容の推測が可能な内容を捉えることに特化しており、一般ユーザが投稿した写真や出現したばかりのトピックの文脈の考慮は難しい。

### 2.3 マイクロブログにおける文脈情報の利用

Twitter には、一投稿あたりの文章量が少ないために同一のトピックに関する投稿を連続した複数の投稿に分けることが多い、リツイート機能により他人の投稿に対するコメントが気軽に行える、といった特徴があり、対象となる投稿以外に、その内容を理解するのに役に立つ投稿が存在する。そのため、投稿が持つ文脈を表現するツイート資源として、同一ユーザが同じトピックについて行った前後の投稿や、投稿に対する他のユー

ザの反応を利用できる。ここでは、それらの文脈情報の収集や活用を行っている研究について述べる。

同一ユーザの連続した投稿におけるトピックの共有に着目した研究としては、LDA を用いてツイートのトピック分類を行う際に隣接する投稿が一定確率で連続していると仮定する手法 [11] や、時間的に連続するツイートが同文脈で投稿されたかを判別する研究 [12] が挙げられる。また、特定の投稿に対するユーザの反応を用いてその性質の分類を行っている研究としては、ある投稿が人を笑わせることを目的に投稿された「ネタ投稿」であるかどうかの判別を行ったもの [13] がある。[13] においてはユーザの反応として投稿へのリプライ、引用リツイート、コメント、リツイートした直後の投稿内容を利用して二値分類問題を解いている。

### 3. 提案手法

本章では、クエリとして与えられた画像投稿と類似する文脈を持つ投稿を収集するための手法について述べる。全体の流れとしては、

(1) クエリおよび検索対象の画像投稿を、投稿文脈を表すベクトル表現に変換する。

(2) (1) のクエリ投稿のベクトル表現を用いてオンデマンドで距離学習を行い、そのクエリに適した距離行列を学習する。

(3) (2) の距離行列を用いてクエリと画像投稿の距離を計算し、クエリと距離が近い投稿を類似文脈投稿とする。

の 3 ステップである。以下でそれぞれの手順の詳細を述べる。

#### 3.1 投稿文脈のベクトル表現

単体の画像投稿に含まれる情報は少ないため、その投稿を取り巻く数種類の投稿を併せて用いて、画像投稿の投稿文脈のベクトル表現を構築する。投稿文脈は、ユーザがどのような状況・意図でその投稿を行ったかという投稿経緯と、投稿に対して他のユーザがどのような反応をしているかという投稿の受け止められ方という 2 つの要素から構成されるとする。前者を表すものとして、画像投稿自体に加えて、投稿の時間的前後にそのユーザがどのような投稿を行ったかという情報を用いる。後者を表すものとして、他のユーザがその投稿にどのようなリプライを送っているか、また、その投稿をリツイートした時にどのような反応をしているかという情報を用いる。以下で各特徴量の詳細を述べる。

また、以下の各特徴量のベクトルは、それぞれ 0-1 の範囲で要素の正規化を行ったうえで連結することでその画像投稿の投稿文脈を表現するベクトルとする。

##### 3.1.1 画像投稿自体のベクトル表現

投稿文のベクトル表現として、文中の単語の分散表現 (CBOW に基づく word embedding [14] を採用) の平均を用いる。投稿文中の単語は、形態素解析器 MeCab<sup>(注1)</sup> および MeCab 用辞書の Neologd [15] を使用して投稿文の形態素解析を行い、単語分割を行うことで獲得した。そのうち、名詞 (一般名詞および固有名詞)、動詞、形容詞、形容動詞からストップワードを除い

たものの分散表現の平均をとった。分散表現は画像投稿を収集した期間のツイートで学習を行うことで、その期間に適応的な表現を獲得するものとする。

また、画像情報として、Convolutional Neural Network (CNN) を用いた画像分類のためのモデル AlexNet [16] に、投稿中の画像を入力した際の第 7 層における出力を用いる。これは 4096 次元のベクトルとなる。モデルとして、ImageNet データにおける学習済みモデル [17] を用いた。画像投稿中に複数の画像が含まれる場合は、1 枚目の画像のみを用いた。文章と画像両方の情報を併用する際には、PCA (Principal Component Analysis) を用いて次元圧縮を行い、文章情報に基づいて構築したベクトルと次元数を統一する。

##### 3.1.2 関連投稿のベクトル表現

以下で述べる 3 種類の関連投稿の文章をそれぞれベクトル表現する。ベクトルの構築方法は、投稿本文ベクトルの構築方法に準じる。各関連投稿の定義を以下に示す。

**同一ユーザの周辺投稿** [12] より、同一ユーザの投稿においては、時間的に近接した連続する投稿は文脈を共有している可能性が高い。そこで、対象となる画像投稿の前後について、投稿間隔が 3 分以内である一連の通常ツイート (リプライ、リツイート、引用リツイートを除いたツイート) をまとめて周辺投稿とし、投稿文脈として利用する。投稿間隔が 3 分以内である一連のツイートというのは、たとえば対象投稿の 2 分後と 4 分後に投稿があれば、その両方を指す。リプライやリツイートはそこで話題が切り替わる可能性が高いため、それらが現れたらそれ以前/以降のツイートの収集を止める。また、URL を含むツイートは、外部 Web ページの共有や何らかの Web サービスからの連携投稿である場合が多く、やはり話題が切り替わる可能性が高いため、URL を含むツイートが現れた段階で同様に収集を止める。

**リプライ** 対象となる画像投稿に対して投稿されたリプライを、その投稿に対する他ユーザの反応として扱う。

**リツイートに対する反応** Twitter においては、ユーザがリツイートの直後にその内容に関するコメントを投稿することが多い [18]。そのため、リツイートから 3 分以内に投稿された通常ツイートをリツイートに対する反応として利用する。周辺投稿と同様に、話題が切り替わっている可能性の高い、URL を含むツイートは除外する。リプライは、その内容がリプライ元のツイートに関連している確率は非常に高い。しかしながら、Twitter においてリプライの頻度は低く、それだけで十分なデータ量を確保することは難しい。一方で、リツイート後の投稿は、その内容がリツイートに関係している確率はリプライに比べて下がるものの、リプライより多くの投稿を安定して収集することができる。

##### 3.2 オンデマンド距離学習

投稿文脈のベクトル表現は、投稿の本文及び画像、関連投稿のベクトルを連結したものであり、クエリによって重視する特徴量が異なると考えられる。たとえば似たリプライをもらいやすい投稿文脈ならリプライ、画像が似ている投稿なら画像の次元に投稿間距離の計算時に重みをかけることで、それらの次元

(注1) : <http://taku910.github.io/mecab/>

の要素が類似している投稿との類似度をより高めることができる。

この次元の重みを学習する手法としては、距離学習 [1] を用いる。距離学習は主にクラスタリングにおいて用いられる半教師あり手法で、分類対象のデータの一部にラベル付けを行い、それらのうち、同じラベルを持つ要素同士の距離が近く、異なるラベルを持つ要素動詞の距離が遠くなるよう、距離行列を学習するものである。本研究では、距離学習の手法として Large Margin Nearest Neighbor (LMNN) [19] を用いた。LMNN では、異なるクラスに属する要素同士が大きなマージンを持つようにしつつ、 $K$  近傍の要素は同じクラスに属するように距離行列の学習を行う。

距離学習時には、与えられるクエリに含まれる投稿を全て同じクラスに属する要素として扱い、クエリとは投稿文脈の異なる投稿を異なるクラスに属する要素として扱う。クエリと投稿文脈の異なる投稿は、投稿文が 30 字以上の画像投稿の中から、その投稿文ベクトルとクエリの画像投稿の投稿文ベクトル間のユークリッド距離が小さい順に一定の数を取得して用いた。本文の長さに閾値を設けたのは、本文が長いほど内容に具体的な内容の説明が含まれる可能性が高くなり、投稿文ベクトルの信頼度が上がるためである。

## 4. 実験

実験では、まず多様な関連画像が投稿されるようなできごとが発生した前後の期間の画像投稿を収集し、そのできごとに関する画像投稿を投稿文脈に人手で分類することで画像投稿データセットを構築した後、それぞれの投稿文脈に基づく画像投稿検索を行い、その評価を行った。

### 4.1 実験設定

#### 4.1.1 評価用データセット

評価実験用に独自に構築した Twitter データセットについて述べる。以下で使用したデータはいずれも筆者が所属する研究室において継続的に収集・蓄積しているツイートデータから抽出したものである。このツイートデータは、約 278 万のツイート公開ユーザの投稿を継続的に収集したものであり、収集対象のユーザとしてまず、2011 年 3 月に 30 名程度の著名な日本人ユーザを選択し、それらのユーザがメンションやリツイートを行ったユーザをさらに再帰的に収集対象とし、順次拡大していったものである。

今回の実験では、多数の文脈で関連画像投稿が行われ、有用な投稿が多く観測されたできごととして熊本地震を選んだ。熊本地震は主に 2016 年 4 月 14 日から 16 日にかけて熊本県及び大分県で相次いで発生した最大震度 7 の一連の地震である。

地震が発生した日を含む、2016 年 4 月 11 日から 4 月 23 日までの 2 週間において、その期間中にリツイートされた回数が上位の 10,000 件の画像投稿を抽出し、そこから画像が取得できなかった投稿を除いた 9245 件を画像投稿データセットとした。リツイート回数上位の投稿を収集したのは、リツイート回数が多ければ、その投稿に価値を見出したユーザの数も多いと考えられるためである。次に、上記のリツイート回数上位の

表 1 リプライ数の分布

リプライ数	0	1	2	3	4	5	6-10	11-20	21-
投稿数	3166	2355	1267	792	458	325	637	205	40

表 2 周辺投稿数の分布

周辺投稿数	0	1	2	3	4	5	6-10	11-20	21-
投稿数	7558	920	309	142	72	36	93	72	43

表 3 リツイートへの反応投稿数の分布

反応投稿数	0-9	10-19	20-29	30-39	40-49	50-99	100-199	200-
投稿数	1507	2078	1567	1027	681	1360	668	357

画像投稿データに対して、それらの投稿へのリプライ、周辺投稿、リツイートへの反応投稿を収集した。リプライに関しては、リツイート数の多いツイートに自動でリプライを飛ばすような bot アカウントからの投稿は除外した。その結果、リプライが取得できたのは 6079 件、周辺投稿が取得できたのは 1687 件であった。独自に収集したツイートデータから抽出しているため、リプライやリツイートへの反応については特に、実際よりも少ない投稿しか獲得できていないものが多い。それぞれ収集した数の分布を表 1、表 2、表 3 にまとめた。表より、周辺投稿が取得できている投稿は全体の約 15 % と非常に少ないことが分かる。一方、リツイート直後の投稿は安定してある程度の数を集めることができている。

評価用に、上記データセットからランダムに選択した 2500 件を検索対象投稿とし、同じ投稿文脈を有する画像投稿の分類を筆者が手動で行うことで投稿文脈を付与した。これらの投稿のうち、熊本地震に関する投稿全てを投稿文脈によって分類し、その中からその文脈に属する画像投稿が多い代表的な投稿文脈について、それらに基づく画像投稿の検索実験を行うこととした。以下で、その 9 文脈について詳細を述べる。また、各文脈の代表的な画像投稿を図 1 に、投稿数の分布を表 4 に示す。各投稿文脈のラベルは筆者が分類した投稿文脈に対し、便宜的に付与したものである。

**災害時に役立つ情報** 災害時の安全や生活環境の改善のための知識を与える投稿。たとえば避難所にあるもので作れる簡易オムツの作り方の図解、避難時に持ち出すものリスト、地震における避難の際にペットをどう扱えばよいのかという情報をまとめた画像などを含む。

**被災者に役立つ情報** 給水や炊き出し、避難所の案内、被災地で空き巣が出現しているという注意喚起の文章のスクリーンショットなど、即時的に被災者に役に立つ投稿に加え、上記の災害時に役立つ情報を合わせたもの。

**地震の情報** 地震の速報や震度の情報、あるいは活断層の位置の解説など、地震に関する情報を伝える投稿。報道機関のアカウントが発信したものや、テレビや Web サイトのスクリーンショットを投稿したものなどがある。

**地震の被害** 地震被害に立ち会ったユーザが投稿した、地震の被害を写した写真。スーパーの棚やユーザ自身の家の本棚が崩れている写真や、熊本城が崩れている写真など。

**マスコミ批判** 被災地でのマスコミの振る舞いを批判するもの



図 1 熊本地震における代表的な投稿文脈の投稿例

表 4 代表的な文脈における画像投稿数

被災者に役立つ情報	65
災害時に役立つ情報	32
地震の情報	60
地震の被害	41
迷子のペット	25
自衛隊などの災害出動	31
いい話 (支援や応援への感謝・評価)	24
応援	82
マスコミ批判	15

や、記者の Twitter の投稿における配慮にかけた振る舞いを避難するものなど、テレビのスクリーンショットや文字ベースの画像が多い。

**応援** 熊本地震では、「#くまモン頑張れ絵」というハッシュタグで、熊本への応援のメッセージを伝えるくまモンのイラストが多く投稿された。そのような、熊本への応援メッセージを表現する投稿、イラストが大部分を占める。

**迷子のペットの情報** 地震によって逃げたペットを探している人や、逃げてきたペットを預かっている人による、その情報を伝えるための投稿。

**いい話 (支援や応援への感謝・評価など)** 被災地で活動する自衛隊や支援物資を送ってくれた人への感謝を述べる投稿など。

**自衛隊などの災害出動** 自衛隊や海上保安庁、在日米軍などによる災害出動に関する画像投稿。大部分がそれらの期間の公式アカウントから投稿されたものである。

#### 4.1.2 投稿文脈ベクトル

リツイート回数上位の画像投稿データと同期間の約 670 万件

のツイートをを用いて 100 次元の単語の分散表現を学習した。投稿文のベクトル表現と画像のベクトル表現を連結する際には、画像ベクトルを 100 次元に次元削減した。

#### 4.1.3 距離学習

クエリの投稿を正例とし、負例は全データセットから 3.2 で述べた方法で 100 件取得した。距離学習の実装は python の metric-learn ライブラリ<sup>(注2)</sup>を用いた。検索対象の画像投稿とクエリとの間の投稿文脈ベクトルの距離は、学習した距離行列を用いて計算したクエリ内の各投稿との距離を全て足し合わせたものとした。

### 4.2 実験結果

#### 4.2.1 文脈ごとの性能評価

以下で述べる各手法について、各文脈で画像投稿検索を行った。1つのクエリに用いる画像投稿の数は3件とし、各投稿文脈で4種類のクエリを用いて検索を行った。このときの、検索結果の上位30件における正答数の平均によって評価を行う。その結果を表5に示す。なお、その文脈の正答数と平均の差の最大値を±の後に記している。実際のTwitterにおける運用ではリツイートで見かけた投稿をクエリとする機会が多くなることが予測される。すなわち、リツイートの頻度が高い画像投稿ほどクエリになりやすいと考えられる。そこで、本研究においてクエリを選択する際も、その文脈の投稿をリツイートの多い順に並べ、上位から順に重複なく3つづつ、4パターン取得した。

**ベースライン** 投稿内の文章のみを利用し、距離学習を行わ

(注2) : <https://all-umass.github.io/metric-learn/index.html>

表5 文脈ごとの top30 正答数の平均

手法	投稿文脈								
	迷子のペット	災害時に役立つ情報	応援	いい話	被害	災害出動	地震の情報	マスコミ批判	被災者に役立つ情報
ベースライン	<b>17.0±2.0</b>	3.0±3.0	10.0±6.0	<b>5.25±2.25</b>	2.25±2.25	13.5±2.5	19.25±1.75	2.25±2.25	15.5±5.5
ベースライン+画像	12.75±5.25	2.75±1.75	6.25±3.75	1.5±1.5	3.0±3.0	12.0±4.0	18.25±2.25	1.25±1.25	9.75±4.25
本文	14.0±4.0	<b>8.75±1.75</b>	11.25±4.75	1.5±1.5	6.5±3.5	<b>16.0±3.0</b>	20.5±1.5	3.25±3.75	<b>15.75±6.75</b>
本文+画像	8.25±2.75	4.5±2.5	6.75±4.25	0.5±1.5	5.75±3.75	14.0±1.0	20.5±4.5	1.0±1.0	10.0±7.0
本文+関連投稿	11.0±1.0	6.5±3.5	<b>13.0±3.0</b>	1.75±1.25	<b>10.25±6.25</b>	14.5±1.5	21.25±2.25	<b>5.25±2.75</b>	10.75±5.75
本文+関連投稿+画像	8.25±2.25	4.25±0.75	7.0±2.0	1.75±1.75	8.0±7.0	13.5±1.5	<b>21.5±2.5</b>	2.5±2.5	8.5±5.5

ない

**ベースライン+画像** 投稿内の文章と画像を利用し、距離学習を行わない

**本文** 投稿内の文章のみを利用し、距離学習を行う

**本文+画像** 投稿内の文章と画像を利用し、距離学習を行う

**本文+関連投稿** 投稿内の文章に加え関連投稿を利用し、距離学習を行う

**本文+関連投稿+画像** 投稿内の文章と画像に加え関連投稿を利用し、距離学習を行う

表5より、「迷子のペット」および「いい話」文脈以外の7文脈で距離学習の効果が確認できた。また、「応援」「地震の被害」「地震の情報」「マスコミ批判」文脈において関連投稿の効果が確認できた。一方で「災害時に役に立つ情報」「被災者の役に立つ情報」「災害出動」といった文脈の投稿は、本文のみを用いたほうが正答数が多い。これらの文脈の投稿は、情報の伝達を主眼とし、本文中に多くの情報が含まれていることが多い。しかし、各特徴量の重みとして働く距離行列の要素の合計を見ると、これらの文脈ではリプライとリツイートへの反応に本文よりも大きな重みが付与されていた。特にリツイートへの反応は、リツイートに関係ない内容の投稿も多く含まれるため、ノイズとして働きうる。一方で、画像情報の利用はほぼ全ての文脈において精度を下げる結果となっている。今回用いた画像特徴はCNNの学習済みモデルの中間層であったが、このモデルは一般画像認識のためのImageNetデータセットで学習されたもので、学習データにはTwitterにアップロードされているようなイラストやスクリーンショットなどは含まれていないため、よりTwitterの画像に適した画像特徴を選ぶ必要があるだろう。

#### 4.2.2 クエリの投稿数の変化と正答数

クエリの投稿数を変化させた場合の正答数の変化を、図2に示す。これは、各文脈について、まずリツイート数上位の2投稿をクエリとし、その後リツイートが多い順に順次1投稿ずつクエリを拡張してそれぞれのクエリにおける正答数を数えたものである。

「応援」や「被害」は文章の少ない投稿が多く、文章ベースの検索が難しい文脈であるが、提案手法ではクエリの投稿数が少ない段階で有効性を示している。「被災者の役に立つ情報」「災害時に役に立つ情報」はクエリ数の増加に伴い関連投稿を利用した際の精度が上昇する。特徴量ごとの重みを確認すると、クエリの投稿数が少ない時点ではリツイートへの反応への重みが大きかったが、投稿数の増加に伴い本文への重みも増加している。また「地震の情報」の画像は、震度を示した地図や断層の

位置を示した地図など類似したものが多かったため、画像投稿が有効に働いている。

## 5. まとめと今後の課題

本稿では、マイクロブログにおける投稿文脈に基づく画像投稿獲得を目的とし、所望の投稿文脈を持つ複数の画像投稿自身をクエリとして、類似投稿文脈を持つ画像投稿を獲得するタスクを設定した。そして、多様な関連投稿を用いて投稿内容を拡張して投稿ベクトルを構築し、クエリ内の投稿ベクトル間の類似度が高くなるようオンデマンドに距離学習を行った後、検索対象の投稿ベクトルとの類似度を計算し類似度の高い投稿を集める手法を提案した。実験では、実際のTwitterの画像投稿に文脈のアノテーションを行ったうえで各文脈について投稿検索を行い、多くの文脈で距離学習および関連投稿が有効であることを確認した。

今後の課題としては、まず、よりTwitterに適した画像特徴の利用やリツイートへの反応のノイズ軽減、投稿を行ったユーザやリツイートをしたユーザの情報の利用など、利用する特徴量の改善・拡張が挙げられる。次に、クエリの投稿数が少ない場合や、クエリに含まれる投稿の表層的な特徴の分散が大きい時に過学習を避けるため、次元ごとではなく、特徴量ごとに重み付けをする距離学習を行うことを検討している。また、本文に十分な情報を含むような、本文の情報のみを用いた場合に非常に高い精度を示す文脈も存在したことから、距離学習に加え、クエリ内の投稿の単語の重複や文章長など、より表層的な文章特徴から特徴量への重み付けを行う手法も検討したい。

## 文 献

- [1] Liu Yang and Rong Jin. Distance metric learning: A comprehensive survey. *Michigan State University*, Vol. 2, No. 2, 2006.
- [2] Nikhil Rasiwasia, Jose Costa Pereira, Emanuele Coviello, Gabriel Doyle, Gert RG Lanckriet, Roger Levy, and Nuno Vasconcelos. A new approach to cross-modal multimedia retrieval. In *Proceedings of the 18th ACM international conference on Multimedia*, pp. 251–260. ACM, 2010.
- [3] Yunchao Gong, Qifa Ke, Michael Isard, and Svetlana Lazebnik. A multi-view embedding space for modeling internet images, tags, and their semantics. *International journal of computer vision*, Vol. 106, No. 2, pp. 210–233, 2014.
- [4] David M Blei and Michael I Jordan. Modeling annotated data. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 127–134. ACM, 2003.
- [5] Fei Yan and Krystian Mikolajczyk. Deep correlation for matching images and text. In *Proceedings of the IEEE Con-*

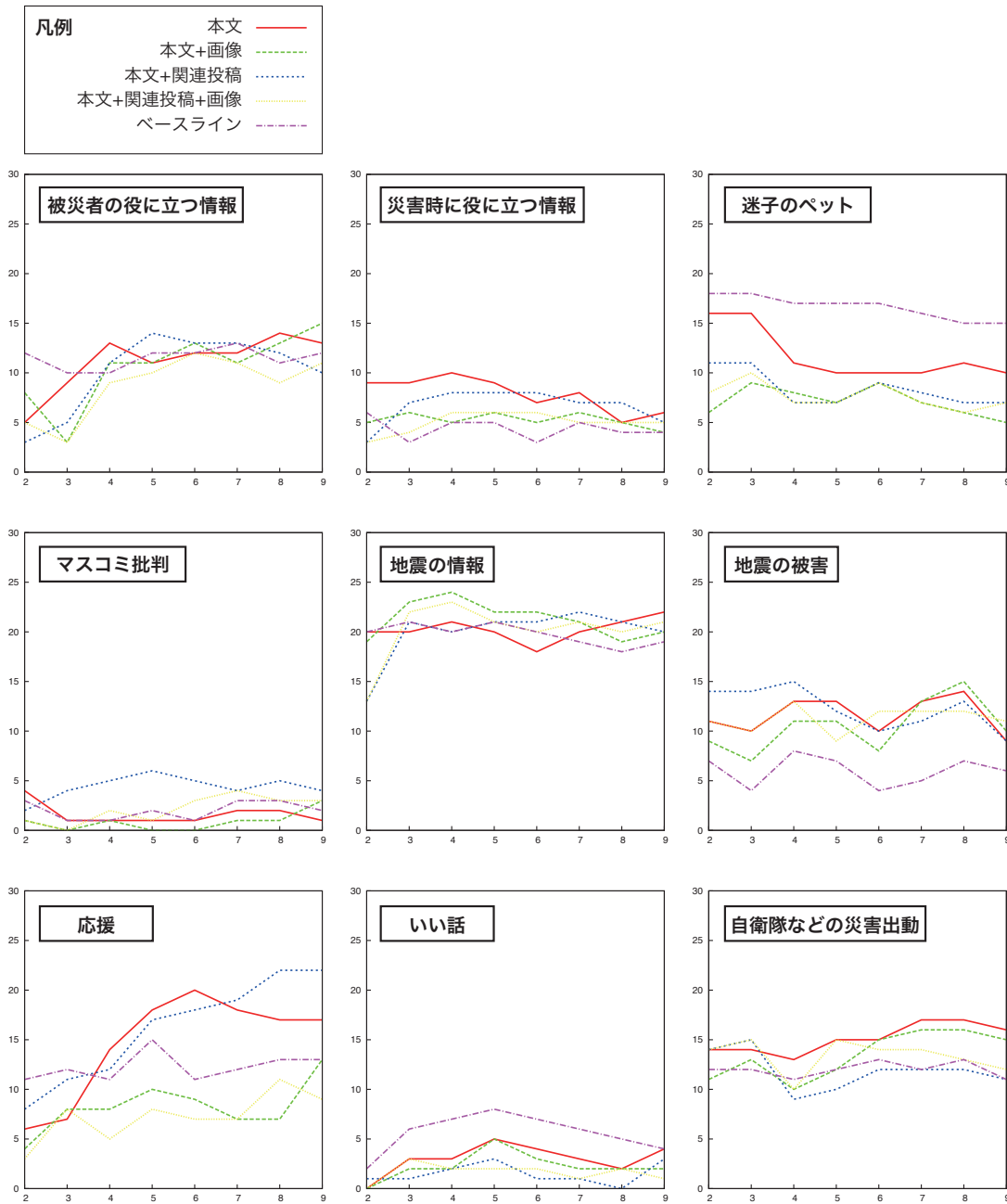


図2 クエリの投稿数と正答数 (縦軸: 正答数, 横軸: クエリの投稿数)

ference on Computer Vision and Pattern Recognition, pp. 3441–3450, 2015.

- [6] Andrej Karpathy, Armand Joulin, and Fei Fei F Li. Deep fragment embeddings for bidirectional image sentence mapping. In *Advances in neural information processing systems*, pp. 1889–1897, 2014.
- [7] Kailong Chen, Tianqi Chen, Guoqing Zheng, Ou Jin, Enpeng Yao, and Yong Yu. Collaborative personalized tweet recommendation. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pp. 661–670. ACM, 2012.
- [8] Wei Feng and Jianyong Wang. Retweet or not?: personalized tweet re-ranking. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 577–586. ACM, 2013.
- [9] Liangjie Hong, Aziz S Doumith, and Brian D Davison. Co-factorization machines: modeling user interests and pre-

dicting individual decisions in twitter. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 557–566. ACM, 2013.

- [10] Tao Chen, Xiangnan He, and Min-Yen Kan. Context-aware image tweet modelling and recommendation. In *Proceedings of the 2016 ACM on Multimedia Conference*, pp. 1018–1027. ACM, 2016.
- [11] 中村直哉, 菅野遼平, 高村大也, 奥村学ほか. 隣接するツイート間の関係を考慮したマイクロブログのトピック推定. 研究報告自然言語処理 (NL), Vol. 2012, No. 4, pp. 1–6, 2012.
- [12] 新谷歩生, 関洋平, 佐藤哲司. 投稿間隔に基づくマイクロブログからの話題チャンク抽出に関する一検討. In *DEIM Forum*, 第2011巻, 2011.
- [13] 林田宗一郎, 牛尼剛聡. ユーザの反応を利用したネタツイート自動分類手法. In *DEIM Forum*, 2014.
- [14] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space.

*arXiv preprint arXiv:1301.3781*, 2013.

- [15] Sato Toshinori. Neologism dictionary based on the language resources on the web for mecab, 2015.
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [17] BVLIC AlexNet Model. [http://dl.caffe.berkeleyvision.org/bvlc\\_alexnet.caffemodel](http://dl.caffe.berkeleyvision.org/bvlc_alexnet.caffemodel).
- [18] 塚本悠馬, 笹野遼平, 高村大也, 奥村学ほか. マイクロブログ上の告知投稿に対する非明示的な関連投稿の収集. 研究報告自然言語処理 (NL), Vol. 2013, No. 14, pp. 1–8, 2013.
- [19] Kilian Q Weinberger and Lawrence K Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, Vol. 10, No. Feb, pp. 207–244, 2009.