

小学校における NIE のための Web ニュースを補足する画像の分類

小栗太樹[†] 村田 真澄[‡] 安藤 一秋^{*}

[†]香川大学工学部 [‡]香川大学大学院工学研究科 ^{*}香川大学創造工学部

1. はじめに

近年、小学校において、新聞を教材として活用する教育 NIE (Newspaper in Education) が実施されている。新聞記事を選ぶこと、読むことなどを通じて、読解力の向上や自己判断力などを養うことができると報告されている[1, 2]。しかし、一般の新聞記事や Web ニュース記事は、子供向けに書かれておらず、また、内容を理解・補足するための図や写真などもほとんど付与されていないため、NIE 実践時には、小学生が読めない、理解できないなどの問題がある。NIE 実践指定校で NIE 担当教員にインタビューした結果、NIE を実践する教師は、新聞記事を探す労力に加え、記事内容を補足する資料の準備にもさらに時間を要する問題が生じていることを確認した。

そこで本研究では、教師が選択した Web ニュース記事に対して、記事内容を補足する画像コンテンツを Web 上から検索して、提示するシステムの構築を目的とする。先行研究[2]では、画像検索クエリの生成と画像周辺テキストを用いた画像のスコアリング手法を提案した。

画像には、写真、グラフ、図、表などのタイプがあり、記事内容や授業目的によって求められる画像のタイプは異なる。Web 上で検索して得られる画像は多数存在するため、タイプごとに分類して表示することで、必要な画像を見つけやすくなると考えられる。

そこで本稿では、画像をグラフ、図、表、その他の 4 タイプに分類する手法について検討する。画像を分類するための情報として、画像自身から得られる特徴と画像周辺テキストから得られる特徴が活用できる。本稿では、まず、周辺テキストのみを用いた手法について検討する。

2. 関連研究

本研究の先行研究として、村田ら[2]は、画像検索クエリ生成法と画像周辺テキストを用いた検索画像のスコアリング法を提案している。ニュース記事と Wikipedia 記事に対して、パターンマッチングによりクエリを生成する手法と、小学校の教科書に含まれる図表キャプションを利用した手法の 3 種類でクエリを自動生成する。そして、そのクエリで検索された画像に対し、画像周辺テキストとニュース記事の TF-IDF 値を基にした cosine 類似度でスコアリングする。

近藤ら[3]は、与えられたテキストから重要語を抽出し、外部 API を利用することで、対象テキストに関連する動画やブログ等のコンテンツを推薦する手法を提案している。一般ユーザに対して幅広い内容のコンテンツを網羅的に検索・推薦することを目指したものである。小学校の

NIE で利用することを考えた場合、多くのコンテンツを提示するより、質の高いコンテンツを提示する方が効率的である。

2. 補足画像提示システムの概要

提案システムの処理手順を以下に示す。

- STEP 1: 教師が選択した Web ニュース記事に対して、重要語を抽出、または、教師が重要語を入力
- STEP 2: 重要語をもとに、Wikipedia 記事や教科書キャプションからクエリを生成
- STEP 3: 画像検索 API を用いて画像を検索
- STEP 4: ニュース記事と画像周辺テキスト、画像特徴などを基に、画像タイプの分類とスコアリング
- STEP 5: 検索画像を、総合ランキングを提示、またはタイプ別にリランキングして提示

以降、本稿では、STEP 4 の画像周辺テキストを利用した画像のタイプ別分類法について検討する。

3. 画像のタイプ別分類法の検討

本稿では、画像をグラフ、図、表、その他の 4 タイプに分類する。各タイプを以下のように定義する。

- ・グラフ 棒グラフ、円グラフなどのグラフを含む画像
- ・図 ある事柄について、文章やイラスト、矢印などの記号を用いて説明している画像
- ・表 あるデータや複数の項目をまとめた表を含む画像
- ・その他 上記以外の画像。写真やイラストなど

3.1 データセット

本稿では、画像を分類するための情報として、画像周辺テキストを用いるため、Web 上の画像および画像周辺テキストをデータセットとして構築する必要がある。

データセットの構築について述べる。無作為に選択したニュース記事から人手で重要語を抽出し、重要語を検索クエリとして、画像検索 API を用いて画像および Web ページを収集する。そこから、HTML 構造を利用して周辺テキストを抽出する。事前調査の結果を基に、Web ページ内の日本語文章が含まれているタグのうち、画像 URL が記述されている箇所より、前後にそれぞれ 3 タグ、計 6 タグ分の日本語文章、TITLE タグ、対象画像の alt 属性を周辺テキストとして利用する。

最終的に、各タイプ 100 画像、計 400 画像を収集し、これらの周辺テキストと共にデータセットを構築した。周辺テキストは、1 画像あたり平均 421 文字となった。

3.2 素性・学習モデル

本稿では、画像周辺テキストから生成した bag-of-words 素性と 3 種類の学習モデルを組み合わせた手法と、word2vec 素性と SVM (Support Vector Machine) による手法で性能を比較する。

Classification of Image to Supplement the Web News for NIE in Elementary Schools

Taiki Oguri[†], Masumi Murata[‡], Kazuaki Ando^{*}

[†]Factory of Engineering, Kagawa University,

[‡]Graduate School of Engineering, Kagawa University

^{*}Factory of Engineering and Design, Kagawa University

(1) bag-of-words 素性

画像周辺テキスト中の名詞、動詞、形容詞の bag-of-words を素性として用いる。出現頻度は考慮せず、その単語が出現した場合に 1、出現しない場合に 0 とする。データセット 400 文章中で出現文章が 3 以下、360 以上の単語は、ノイズとして除去する。

多クラス分類タスクの学習モデルは多数存在するが、一般的に高い分類性能が報告されている、one-versus-reset 法の SVM, Random Forest, Gradient Boosting の 3 モデルを採用する。

(2) word2vec 素性

単語の分散表現である word2vec[4] を素性として用いる。word2vec においては、意味の近い単語から生成されたベクトルは類似したベクトルとなる特徴を持つことが期待される。周辺テキストのベクトルは、画像周辺テキストで出現する各単語（名詞、動詞、形容詞）の分散ベクトルを加算したもので構成し、加算値を単語数で割った値を素性値に利用する。word2vec の学習には、2018 年 10 月時点の日本語 Wikipedia 記事全文を用い、次元数は 50 とする。

学習モデルは one-versus-reset 法の SVM を採用する。

4. 実験

4.1 実験設定

分類手法の性能を確認するため、評価実験を行う。形態素解析には MeCab を用い、辞書は、NEologd を利用する。各学習機の実装には、scikit-learn を用い、ハイパーパラメータは、グリッドサーチによりチューニングする。適合率、再現率、F 値を評価尺度として、作成したデータセットを用いて、10 分割交差検証で各手法の性能を評価する。

4.2 実験結果

実験結果として、4 タイプの各評価指標の平均値を表 1 に示す。表 1 から bag-of-words を素性とした SVM (BOW-SVM) がどの指標においても最も高い値となることが確認できる。逆に、word2vec を素性とした SVM (W2V-SVM) が最も性能が低く、素性の改良が必要といえる。

表 1. 実験結果

	適合率	再現率	F 値
BOW-SVM	0.64	0.63	0.64
BOW-RF	0.60	0.56	0.58
BOW-GB	0.63	0.61	0.62
W2V-SVM	0.56	0.54	0.55

BOW-SVM のタイプごとの分類結果を表 2 に示す。F 値については、グラフが最も高い値となったが、適合率では図が高くなった。そこで、タイプごとに周辺テキストを分析した結果、他のタイプにはあまり出現しない特徴的な単語が出現していることが確認できた。

表 2. BOW-SVM のタイプ別分類結果

	適合率	再現率	F 値
グラフ	0.65	0.70	0.67
図	0.67	0.61	0.64
表	0.64	0.58	0.61
その他	0.58	0.64	0.61

4.3 特徴語への重み付けによる性能比較

4.2 節の結果を基に、画像タイプ別の特徴語に重み付け

することで性能向上を試みる。

対象は、4.2 の実験で最も性能が高い BOW-SVM とする。1 タイプにのみに頻出する特徴語として、データセットの画像周辺テキスト中で、タイプ別の出現確率が 10% 以上の名詞を計 52 単語（グラフ：24 語、図：5 語、表：16 語、その他：7 語）抽出した。その一部を表 3 に示す。これらの特徴語が周辺テキストに出現した場合、対応する素性値を 2 とする。その他の素性値は、0 か 1 である。

表 3 重み付けする特徴語の一部

グラフ	グラフ, 推移, 調査, 発表, 増加, 減少, 割合
図	図, 仕組み, 発生
表	表, 対象, 違い, 比較
その他	写真, 理由, 影響, 紹介

4.1 節と同じ設定で、実験した結果を表 4 に示す。特徴語に重み付けすることにより、BOW-SVM の平均値と比べて、適合率が 4 ポイント、再現率が 4 ポイント、F 値が 3 ポイント向上した。この結果から、画像タイプと関連のある単語への重み付けが有効であると考えられる。

表 4 重み付けの結果

	適合率	再現率	F 値
グラフ	0.75	0.65	0.70
図	0.66	0.68	0.67
表	0.65	0.59	0.62
その他	0.65	0.77	0.70
平均	0.68	0.67	0.67

5. おわりに

本研究では、小学校教師が NIE の授業準備の負担を軽減することを目的に、Web ニュース記事の内容を補足する画像コンテンツを提示するシステムの構築を進めている。本稿では、そのシステムで利用する画像タイプ判定手法について検討した。評価実験の結果、bag-of-words を素性とした SVM が最も高い性能を示した。次に、画像のタイプ別の特徴語に重み付けすることで、分類性能の向上を確認した。

今後は、新たな素性の追加や画像解析の導入などによる分類性能の向上を目指す。最終的には、システムとして実装し、総合評価を行う。

謝辞

本研究の一部は JSPS 科研費 16K00478 の助成を受けて実施した。

参考文献

- [1] NIE 教育に新聞を, <http://nie.jp/>
- [2] 村田他, “小学校における NIE のための Web ニュース記事を補足する画像コンテンツの検索”, IPSJ2018 講演論文集, pp.437-438, 2018.
- [3] 近藤他, “重要語抽出を用いた外部 API からの関連コンテンツ推薦”, JSAI2010 論文集, 1D2-1, pp.1-4, 2010.
- [4] Mikolov, T., et al., “Efficient Estimation of Word Representations in Vector Space”, Cornell University Library arXiv:1301.3781, 2013.