

## 検索画像を介在させた語から語の連想模擬法の検討

## Word association simulation by feeding the color-name detected in web-retrieved images to BERT

相馬 佑哉<sup>†</sup>      堀内 靖雄<sup>†</sup>      黒岩 眞吾<sup>†</sup>  
Yuya Soma      Yasuo Horiuchi      Shingo Kuroiwa

## 1. はじめに

我々は、人の語から語の連想（以下、語連想と表記）を、学習済み BERT[1]を用いて模擬する手法の検討を行っている[2]。これまでに表 1 の文に加え、人の視覚情報に基づく連想語を出力するために「<刺激語>から連想される色は MASK です」等の文による連想を試みたが、刺激語と関係の低い色も出力されるという問題が生じた[2]。本稿では、頭の中で対象物の色を想起しているという状況を模擬するために、刺激語で検索した画像から色を抽出しテキスト情報として加えることで、人の連想に近づくかを検討した。

## 2. 画像の色情報を用いた BERT による語連想手法

## 2.1 BERT を用いた語連想手法

本稿では、BERT の Masked Language Model を用いて MASK に連想語が出現する文（以下、連想文と表記）を使用して連想模擬を行う。表 1 に実験で使用する連想文を示す。MASK には鍵括弧「」を付与した。（以下、通常の連想文と表記）また、MASK に色の単語が入るような連想文 4 文（表 1 の 1 と 2 の言葉を色に、3 と 4 の MASK を MASK 色とした。以下、色の連想文と表記）を設定した[2]。

本稿の提案手法では、検索画像から得られた色情報（以下、画像色情報と表記）を連想文の前に連結する手法を検討した。連結した文は「<刺激語>は〇色です。」「これは〇色です。」「〇。」であり、〇には色名が入る。この 3 文と表 1 の通常の連想文 4 文を、以下の例のように各々連結した（3 文×4 文=12 文）。

赤。イチゴから連想される言葉は「MASK」です。

## 2.2 画像色情報の取得

画像色情報の取得方法を下記に示す。

1. 刺激語をキーワードとして Bing 画像検索を行い、上位 30 枚を取得する。（実施日:2021/05/05）
2. 画像 30 枚に対して Microsoft Azure Computer Vision API でオブジェクト検出[3]を行い、検出オブジェクト名によらず confidence の値が最も高い領域を取得する。オブジェクトが検出されなかった場合は画像全体を取得する。
3. 取得した領域に対して Microsoft Azure Computer Vision API で配色検出[4]を行い、色情報（前景）を出力する。色情報は英語表記（Red 等）で出力されるため、日本語表記（赤等）に変換する。
4. 3. で得られた色（画像 30 枚分）の中で最も多い色名を画像色情報とし、連想文に加える。最も多い色が複数ある場合は confidence の合計値が最も高い色名を採用する。

表 1 通常の連想文

1	<刺激語>から連想される言葉は「MASK」です。
2	「MASK」は<刺激語>から連想される言葉です。
3	<刺激語>は「MASK」を連想させる言葉です。
4	「MASK」を連想させる言葉は<刺激語>です。

<sup>†</sup> 千葉大学 Chiba University

## 3. 画像色情報を用いた連想実験

## 3.1 実験条件

正解とする人の連想語としては「連想語頻度表」[5]を用いた。刺激語として、人の連想は視覚情報に基づくとの推測を検証するために連想語上位 5 語に色（日本語 Wikipedia において出現頻度の高い、赤、青、黄、緑、白、黒、ピンク）が存在する 38 語（表 2 の“刺激語”の列に示す）を用いた。

評価基準としては、連想語頻度表の上位 5 語以内の連想語と一致した数を**連想語スコア**とした。BERT も上位 5 語を出力させるため、最大値は 5 となる。実験結果では 38 語の連想語スコアを平均した値を示す。また、連想色のみに着目した連想語スコアを**連想色スコア**とする。この場合も、BERT による連想語としては上位 5 語を出力するが、本実験の刺激語の連想色は一つなので最大値は 1 となる。

## 3.2 実験結果

表 2 に実験に用いた刺激語と ①「連想語頻度表」の連想色（表中、“連想色”）、②2.2 の手法で取得した画像色情報（“画像色情報”）、③提案手法による BERT の上位 5 語中で出現頻度の高かった色（“BERT”）を示す。

表 3 に平均連想色スコア、上位 5 語中で色を表す語の割合、色の中で正しい連想色が最上位に出力された割合を示す。このうち、上位 5 語中で色を示す語の割合は、手法ごとの色の連想のされやすさ（しすぎも含め）を表している。また、表 4 に平均連想語スコア、色を除いた連想語スコア、上位 5 語中で色を示す語の割合（再掲）を示す。

なお、取得した検索画像 1140 枚（刺激語 38 語×画像 30 枚）に対して、筆者による主観評価を実施した。結果としては、妥当な画像を取得できた割合は 88.5%（1140 枚中 1009 枚）、その中で妥当な領域を取得できた割合は 60.1%（1009 枚中 606 枚）、さらにその中で妥当な色情報を取得できた割合は 80.0%（606 枚中 485 枚）であった。

## 3.3 考察

表 3 に示したように、“提案手法”の連想色スコアが“色の連想文”に比べ向上（0.53→0.74）し、色の中で正しい連想色が 1 番目に出力される割合も増加（0.21→0.61）した。このことから、色だけが MASK に現れるような文制約を掛けなくても、提案手法（前文で視覚から得られた色をテキストとして提示する）により、色を連想させることができたといえる。また、正解色が与えられれば確実にその語が連想されることもわかった。一方で、画像色情報が誤っていた場合は、後段で BERT を適用しても改善はみられなかった。ただし、“ワイン、メロン、パンダ、とんぼ、にきび”では、下位ではあるが正解色が含まれている点から、色情報の与え方の工夫により改善できる可能性もある。

また、表 4 より“提案手法”の連想語スコアが“通常の連想文”に比べ向上（0.63→0.92 : +0.29）していることがわかる。しかし、上位 5 語に色が占める割合が約 4 語と、色以外の

表2 刺激語と色の関係

	刺激語	連想色	画像色	BERT
人の連想色と画像色情報が一致	りんご	赤	赤	赤, 赤色, 緑, 緑色, 青
	イチゴ	赤	赤	赤, 赤色, 黄色, オレンジ
	トマト	赤	赤	赤, 赤色, 緑, 緑色
	バナナ	黄	黄色	黄色, 赤, オレンジ, ブルー
	レモン	黄	黄色	黄色, 赤, 黄, 青, オレンジ
	わさび	緑	緑	緑, みどり, 緑色
	カエル	緑	緑	緑, 緑色, 青, 黄色
	自然	緑	緑	緑, 緑色, 青, グリーン
	野菜	緑	緑	緑, 緑色, 青
	うどん	白	白	白, 黒, 白色, 青, 緑色
	ごはん	白	白	白, 白色, 黒, ピンク
	とうふ	白	白	白, 白色, 黄色
	シャツ	白	白	白, 白色, 黒, 赤
	シルク	白	白	白, 白色, 黒
	タオル	白	白	白, 白色, 黒, 赤
	マスク	白	白	白, 白色, 黒, ホワイト, 赤
	ミルク	白	白	白, 白色, ホワイト
	からす	黒	黒	黒, 黒色, 赤, 黄色
	インク	黒	黒	黒, 黒色, 赤, ブラック
	ピアノ	黒	黒	黒, 白, ブラック, 青, 赤
悪魔	黒	黒	黒, 赤, 黒色, ブラック	
電話	黒	黒	黒, 赤, 黒色, 白, ブラック	
さくら	ピンク	ピンク	ピンク, 青, ブルー	
不一致	キムチ	赤	橙	橙, 黄色, オレンジ
	ワイン	赤	黒	黒, 赤, 白, 黒色, 青
	ひよこ	黄	茶	茶色, 灰色, 茶, 緑
	カレー	黄	茶	茶, 茶色, 緑, 青, 赤
	チーズ	黄	茶	茶色, 赤, 青, 茶
	メロン	緑	白	白, 白色, 緑色, ピンク, 桃
	パンダ	白	黒	黒, 黒色, 白
	だるま	赤	白	白, 白色, 黒, 黒色
	とんぼ	赤	白	白, 白色, 桃, 赤
	にきび	赤	茶	茶色, 茶, 赤, 黄色
	サンゴ	赤	白	白, 白色, 黒
	マグマ	赤	灰	灰色, 灰
	バケツ	青	白	白, 黒, 白色, 赤
	ベンチ	青	白	白, 赤, 黒
	タイヤ	黒	灰	灰色, 赤, 青, 白, 灰

連想を大きく阻害していることもわかる。実際，“提案手法”の連想語スコアの向上は色の連想語による向上である。具体的には、色による向上分が 28/38 語(+0.74)で色以外の語による減少分が 18/38 語(-0.47)となっていた（これ以外に，“インク”に対する連想語“色”で+1/38）。

この副作用を確認するために、「連想語頻度表」から人が色を連想しない刺激語 9 語（アニメや育児等）を選出し、追加実験を行った。画像色情報を付与しない場合の連想語スコアは 0.44 で、上位 5 語中に色は出現しなかった。これに対し、画像色情報を与えた提案手法では色が平均 3.78 語出力され、連想語スコアは 0.00 であった。つまり、提案手法は色を無理やり連想する手法となってしまう。こ

表3 上位5語の平均連想色スコア

手法	連想色スコア	上位5語色の割合	連想色1位率
通常の連想文	0.00	0.03	0.00
色の連想文	0.53	3.82	0.21
提案手法(全体)	0.74	3.95	0.61
(一致)	1.00	4.00	1.00
(不一致)	0.33	3.87	0.00

表4 上位5語の平均連想語スコア

手法	連想語スコア	色を除く連想語スコア	上位5語色の割合(再掲)
通常の連想文	0.63	0.63	0.03
色の連想文	0.66	0.13	3.82
提案手法(全体)	0.92	0.18	3.95
(一致)	1.17	0.17	4.00
(不一致)	0.53	0.20	3.87

これらのことから、提案手法を単独で用いるのではなく、色以外の連想語との統合が必須であると結論できる。また、色情報が物を特定する情報となりにくい場合は、画像からの色情報の伝播を抑制する機構の検討も必要である。

また、表2中、不一致だったものには多色で構成されている物があり、その中で人が注目する色が何であるかを反映させる必要がある。もう一つの課題として、カレーやひよこの画像から得られた色が茶であった点があげられる。これらは画像から得られた色が色情報としては正しいにもかかわらず、人の連想では黄色となる事例であり、人の連想の興味深い点と考えられる。

#### 4. おわりに

本稿では、BERT を用いた語連想において、刺激語で検索した画像群から色情報を抽出し、テキスト情報に変換した上で MASK タスクの前文入力として与える手法の検討を行った。その結果、色を連想する語に対しては人の連想に近づけることができた。しかし、副作用も大きく、単独での語連想手法としては不相当であることも明らかとなった。今後は、BERT への補助入力として画像特徴量を用いる手法について検討する。

#### 謝辞

本研究を進めるに当たり、乾・鈴木研究室の訓練済み日本語 BERT モデル[6]を利用させていただきました。本研究は JSPS 科研費 JP20K11860 の助成を受けたものです。

#### 参考文献

- [1] Jacob Devlin, et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, arXiv:1810.04805, 2018.
- [2] 相馬佑哉, ほか. “人間と BERT の語から語らの連想の比較”, 言語処理学会 第27回年次大会, pp.794-798, Mar. 2021.
- [3] Microsoft Azure. “Detect common objects in images”, <https://docs.microsoft.com/ja-jp/azure/cognitive-services/computer-vision/concept-object-detection>, 引用日期: 2021/05/31.
- [4] Microsoft Azure. “Detect color schemes in images”, <https://docs.microsoft.com/ja-jp/azure/cognitive-services/computer-vision/concept-detecting-color-schemes>, 引用日期: 2021/05/31.
- [5] 水野りか, ほか. “連想語頻度表—3 モーラの漢字・ひらがな・カタカナ表記語—”, ナカニシヤ出版, 2011.
- [6] 東北大学乾・鈴木研究室. “Pretrained Japanese BERT models”, <https://github.com/cl-tohoku/bert-japanese>, 引用日期: 2021/05/31.