

視覚的特徴と感情的特徴が比喩生成に与える影響についての検討

The Influence of Visual and Emotional Features on Metaphor Generation

長瀬 カイセイ[†], 寺井 あすか[†]
Kaisei Nagase, Asuka Terai

[†] 公立はこだて未来大学
Future University Hakodate
g2125058@fun.ac.jp, aterai@fun.ac.jp

概要

抽象画像に対する比喩生成において、視覚的特徴に加え情緒感覚的類似性が重要であると指摘されている。本研究では情緒感覚の一種である感情に焦点を当て、物体認識モデル ResNet と感情推定器 ArtEmis を利用し、視覚的類似性と感情的類似性を用い、視覚的・感情的特徴が比喩に与える影響を検討した。加工画像を用いたファインチューニングモデルを用いて比較した結果、形状を強調した視覚的類似度に感情的類似度を加味したスコアが最も人の比喩生成を反映できることが示された。

キーワード: 比喩生成, 物体認識, 感情推定

1. はじめに

比喩とは「A は B のようだ」といったように、主題と喩える語の間に存在する類似性を通じ、主題の理解を助ける修辞技法である。抽象画像に対する比喩生成においては形状特徴が重要な役割を示すことが指摘されている (Yamamura et al., 2024)。一方、比喩においては喩える語と喩えられる語の情緒感覚的イメージが似ているほど理解を助けると指摘されている (楠見 孝, 2005)。画像に対する比喩生成における視覚的特徴の影響に関する検討は行われているが、情緒感覚的類似性の影響に関する検討は未だ数少ない。本研究では以上を踏まえ、感情に焦点を当て、比喩として用いられる対象との視覚的類似性と感情的類似性を用い、抽象画像に対する比喩生成における視覚的特徴と感情的特徴が与える影響について検討する。

2. 方法

本研究では、25 枚の抽象画像を対象に、ILSVRC2012 (Russakovsky et al., 2015) の 1000 の物体クラス内の具象画像との類似度に基づき比喩候補推定を行った。視覚的特徴の抽出のため、学習済みの物体認識モデル ResNet(He et al., 2016) を用いた。また、感情的特徴の抽出には、芸術画像に対し感情と説明文をラベリングした ArtEmis データセット (Achlioptas et al., 2021) で

作成された感情推定器を利用した。このモデルによって、各画像に対し 9 種類の感情に対する確率分布を得る。(Yamamura et al., 2024) の研究を参考に、視覚的特徴が比喩生成に与える影響をより詳細に分析するため、物体認識モデルに対して画像加工を行った画像を用いたファインチューニングを実施した。特定の視覚的要素を強調することによって比喩生成に有効な要素を強調した。具体的には「Canny エッジ検出化 (形状強調)」「ブラー化 (形状抑制)」「グレースケール化 (色抑制)」の 3 種類の加工を行い、それぞれに特化したモデルを作成した (図 1)。ファインチューニングは ILSVRC2012 の学習用画像約 120 万枚を加工し実施した。モデルの精度を表 1 に示す。

図 1 画像加工例

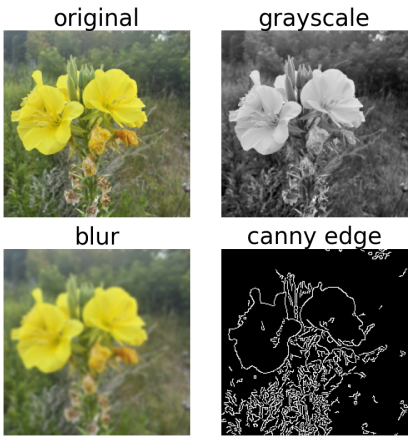


表 1 各ファインチューニング条件における物体認識精度

Top-5 Test Accuracy	
original	0.91
grayscale	0.87
blur	0.82
canny edge	0.76

本研究では比喩生成の候補集合として ImageNet の 1000 クラスを用い、入力画像と各クラスの具象画像

の特徴ベクトルとの類似性から比喩認識スコアを算出し、スコアにより各クラスが表す概念が抽象画像に対して生成される比喩として用いられる順位付けを行う。ILSVRC2012の検証用データセットから各クラス50枚を抽出し、各クラスごとに視覚的・感情的特徴ベクトルのセントロイドを算出した。入力画像と各クラスのセントロイドとのcos類似度を計算し、特徴類似度と感情類似度を取得した。クラス*i*について、視覚的類似度(sim_i^f)と感情類似度(sim_i^e)に対し重み(α)を付けて加算した値を比喩認識スコアとする。

$$sim_i = sim_i^f + \alpha * sim_i^e \quad (1)$$

重みの値を変化させ、視覚的特徴または感情的特徴が比喩生成に与える影響を検討した。順位比喩認識スコアを基に1000クラスの順位づけを行い、画像に対して生成される比喩に該当するクラスの順位の逆数(逆順位)を用いて比較を行った。画像に対する比喩は、富樫他(2024)による実験データから入力画像として用いた抽象画像に対する説明として人が実際に用いた比喩を利用した。参加者が抽象画像に対して自由に回答した比喩表現からILSVRC2012の1000物体クラスに該当するものを抽出した。比喩として用いられる概念がWordNet(Fellbaum, 1998)において物体クラスが表す概念の上位概念である場合、下位概念に該当するクラスも画像に対する比喩とみなした。例えば、「鳥」という比喩に対し、「ダチョウ」「めんどり」など下位概念のクラスも画像に対する比喩とみなした。画像に対して生成された比喩が複数存在する場合は、最も高い順位を採用した。

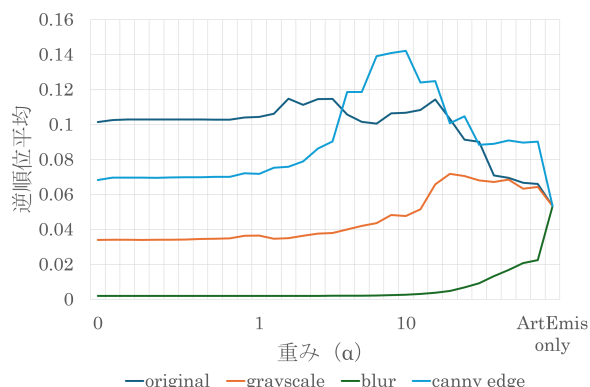
3. 結果

比較結果を図2に示す。縦軸は画像に対して生成された比喩に該当するクラスの逆順位平均であり、横軸は重み α である。Cannyエッジ検出化(形状強調)モデルを用い、視覚的類似度に感情的類似度を組み合わせた場合($\alpha=10$)が最も逆順位平均を向上させた。一方、感情的類似度を加味しない場合($\alpha=0$)、オリジナルモデルが最も高い値を示した。

4. 考察

以上の結果から、視覚的類似度に感情的類似度を加味することで最も人の比喩生成を反映することが示された。このことは、視覚的類似性だけでなく、感情的類似性も抽象画像に対する比喩生成に影響を及ぼす可能性があることを示唆している。一方、感情類似度を加味しない場合は、オリジナルモデルが最良であった。Yamamuraらの先行研究(2024)では形状強調モ

図2 重み α に対する逆順位平均



デルが最も比喩生成を反映していると指摘されていたが、本研究における画像類似度に基づいた評価指標の算出では、画像が持つ特徴の情報量を最も多く利用することができるオリジナルモデルが有利に働いたと考えられる。

5. まとめ

本研究ではILSVRC2012の1000クラスを対象とした検証を行った。今後の課題として、抽象画像と生成された比喩そのものの類似性に基づく検証が必要である。

謝辞

本研究はJSPS科研費基盤研究(C)24K15073の助成を受けたものです。

文献

- Achlioptas, P., Ovsjanikov, M., Haydarov, K., Elhoseiny, M., & Guibas, L. J. (2021). ArtEmis: Affective language for visual art. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11569–11579.
- Fellbaum, C. (Ed.) (1998). *WordNet: An electronic lexical database*. MIT press.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
- 楠見 孝 (2005). 心理学と文体論—比喩の修辞効果の認知— 表現と文体, 491–501.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252.
- 富樫 遥登・白川 琉叶・寺井 あすか (2024). 抽象的画像に対する比喩生成と創造性の関係: 脳波による検討 日本認知科学会第41回大会, pp. P-2–8.
- Yamamura, N., Chikazoe, J., Yoshimoto, T., Jimura, K., Sadato, N., & Terai, A. (2024). Perceptual features of abstract images for metaphor generation. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 28(1), 94–102.