

先端科学技術研究科 修士論文要旨

| | | | |
|--|--------------------------------|-----|--------------|
| 所属研究室 (主指導教員) | 数理情報学 (池田 和司 (教授)) | | |
| 学籍番号 | 2311079 | 提出日 | 令和 7年 1月 20日 |
| 学生氏名 | 小澤 海斗 | | |
| 論文題目 | 画像中のイヌに対するゼロショット・フューショットでの感情分類 | | |
| 要旨 | | | |
| <p>本研究の目的は、アノテーションデータを大量に用意することが難しいイヌの感情認識タスクにおいて、ゼロショット学習やフューショット学習がどの程度有効かを検証することである。具体的には、イヌの画像から「怒り」「悲しみ」「幸せ」「リラックス」の4つの感情を、学習データをほとんどあるいは全く使わずに正確に分類できるかという問いに取り組んだ。</p> <p>本研究では、CLIP (Contrastive Language-Image Pre-training) モデルを利用し、ゼロショットおよびフューショット学習による感情分類を行った。これには、イヌの感情状態(「幸せ」「悲しみ」「怒り」「リラックス」)にラベル付けされた画像データを使用し、ゼロショットでは事前学習済みモデルに直接タスクを実行させ、フューショットでは限られたデータを用いた再学習を実施した。また、EfficientNetを従来手法として選択し、モデル間の性能比較を行った。さらに、次元削減手法 (PCA, ICA, t-SNE) および類似性計算 (コサイン類似度, ユークリッド距離) を用いて、埋め込み表現の構造と感情分類能力を評価した。</p> <p>ゼロショット学習を行ったCLIP モデルは、アノテーションデータなしにもかかわらず、4クラスのイヌの感情を約61%で分類できることが確認された。一方で、フューショット学習では従来法であるEfficientNetが最大75.47%となり、相対的に優位な成績を収めた。中間層の解析では、埋め込み表現がイヌとヒトの感情のドメイン間の類似性を反映しており、CLIPモデルが感情状態のドメインを活用してイヌの感情をある程度特徴的に捉えていることが示唆された。</p> <p>以上の結果から、アノテーションデータを十分に用意することが困難な状況下では、ゼロショット学習可能なCLIP モデルが実用的な手法となり得ることが示唆された。一方で、少量でもデータを用意できる場合には、EfficientNet を含むフューショット学習がより高い精度を期待できる可能性がある。本研究の成果は、イヌの感情分類を効率化し、その応用範囲を拡大する上で重要なステップとなるものである。</p> | | | |