

異種の特徴量フィルタを用いた「かわいい」画像の分類方式

古宮大暉[†] 古渡翔太[†] 秋吉政徳[†]
 神奈川大学[†]

1. はじめに

人工知能による画像分類は、幅広い分野で用いられ、研究がなされている。分類精度は近年大きく向上し、顔認証による空港のセキュリティシステムや、画像を入力とした検索エンジンなど、身の回りのシステムにも用いられている。しかし、顔や動物の種類分けなど、特定の形状や概念を持つ画像分類が十分な精度で行われている一方で、曖昧な形容詞である「感性語」によって表現される画像の分類は困難である。感性語とは人間が得た刺激に対する心理的感情を表現する言葉である。感性語によって表現される画像には、定量的な指標や定義がなく、また画像ごとに特定の形状や概念をもたない。この分類問題は人間にできて人工知能にはできないという、人工知能における課題の一つでもある。

そこで本研究では、感性語であり、その表現が特に多岐に渡る「かわいい」という言葉で表現される画像に対して、ジャンルの分類の方式を提案する。

2. 提案方式

2.1 アプローチ

人工知能を用い、「かわいい」画像のジャンルを分類した研究がある[1]。研究では、画像の持つ色や局所的な特徴を、特徴量フィルタを用いて抽出し、抽出した3つの特徴量を用いてNN(Neural Network)による分類をおこなった。しかし、分類精度は35%程度にとどまり十分とはいえない結果であった。

そこで、新たな特徴量フィルタを作成し、従来研究の特徴量を補う特徴量を抽出した上で、異なる複数の機械学習分類器を用いた精度比較を行う。その結果から、「かわいい」画像の分類において有効な分類方式を提案する。

2.2 構成

本研究で提案する方式の構成を図1に示す。画像から4つの異なる特徴量を抽出し、その特徴量を合成した「全特徴量」を作成する。その全特徴量を入力に、異なる4つの分類器を用いて分類を行い、結果の比較を行う。

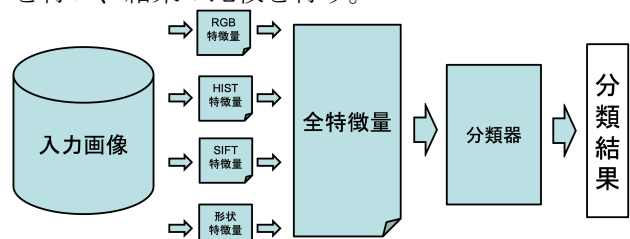


図1: 提案方式の構成

2.3 特徴量フィルタ

異なる4つの特徴量フィルタを作成し、特徴量を抽出する。

1つ目はRGBヒストグラム特徴量フィルタである。これは、画像の各ピクセルが持つRGBの各値の出現頻度をヒストグラムとして抽出した、RGB情報の特徴量である。

2つ目はRGB畳み込み特徴量フィルタである。画像に対して畳み込み処理をすることで、圧縮し特徴づけたRGB情報の特徴量である。

3つ目はSIFT特徴量フィルタである。画像の持つ特徴点とその特徴量を抽出するSIFT(Scale-Invariant Feature Transform)アルゴリズムを用いて得た特徴量情報を元に擬似画像を生成し、その擬似画像に畳み込み処理を行うことで、情報を圧縮し特徴づけた特徴量とした。

以上の3つは従来研究[1]にて提案された手法である。しかし、これらの特徴量では画像全体の形状を抽出できていないため、新たに画像を線画に変換し、畳み込み処理によって特徴づける「形状特徴量フィルタ」を作成する。線画は輪郭画像とは異なり、輪郭内部の線情報を得られるほか、グレースケール画像との比較において、CNN(Convolutional Neural Network)による分類精度が高かったという研究結果[2]がある。

以上の4つの特徴量フィルタによる特徴量を用いて分類を行う。

A classification method for “kawaii”
 images using heterogeneous image filters

[†]Daiki Komiya, Shota Kowatari, Masanori Akiyoshi,
 Kanagawa University

2.4 分類器

分類器には、4つの異なる機械学習分類器を用いる。NN、Random Forests、Ada Boost、SVM(Support Vector Machine)である。

また、本研究のベンチマークとして、画像分類において主流に用いられるCNNによる同一画像の分類実験を用いる。

3. 実験・結果

実験に用いた「かわいい」画像のジャンルとデータ数を表1に示す。

表1: 「かわいい」画像のデータセット

ジャンル名	データ数(枚)
きもかわ	570
ぶさかわ	464
ゆめかわ	702
ゆるかわ	830
病みかわ	559

このデータセットをジャンルごとに8割を学習データ、2割をテストデータに設定し、5分割の交差検証を行った。分類精度は5回の実験の精度の平均をもって評価する。

抽出した特徴量を入力に、異なる4つの分類器による分類結果とベンチマークのCNNの結果を、以下の図2に示す。

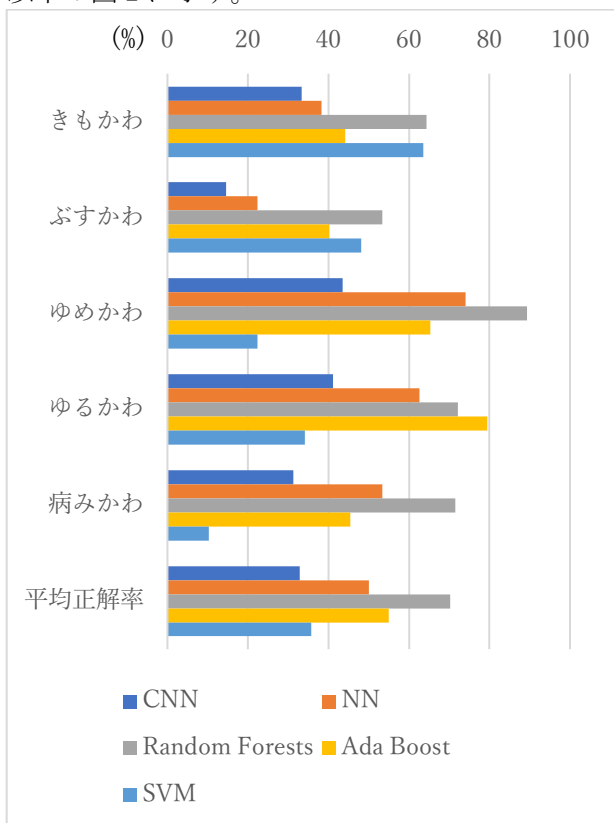


図2: 分類精度の比較

平均分類精度では、提案した4つの分類器全て

がベンチマークのCNN以上の精度となった。また、NNによる分類精度は、従来研究[1]の精度36.1%を上回る50.1%となった。異なる分類器ごとの平均分類精度の比較ではRandom Forestsが最も高く、次いでAda Boost、NNといった結果となった。

4. 考察

分類器ごとの比較ではRandom Forestsが最も高い精度で分類を行えた。

実験では決定木の深さを小さくしたときに分類精度が向上した。ここから、NNのような単体で強力なアルゴリズムよりも、複数の弱学習器による分類の方が「かわいい」画像のような曖昧な画像の分類には適していると考えられる。

また、Random Forestsの特徴量部分抽出は、特徴量全体としては捉えられなかった部分的な共通点を抽出することで、分類精度の向上に有効に働いたと考える。

Random Forestsによるアンサンブル学習は、本実験の入力画像で参考にした、一般の人々の感性を反映しているとも考える。

5. おわりに

本研究では、感性語によって表現され、定量的な指標を持たない「かわいい」画像に対して、従来研究[1]が用いた特徴量を補い、また異なる4つの分類器を用いた画像分類の比較実験を行うことで、「かわいい」画像の分類に適した方式の提案を試みた。実験結果から、新たに提案した形状特徴量フィルタは、従来研究で用いた特徴量を補い、分類精度向上に有効に作用した。また、分類器においてはRandom Forestsによる分類精度が最も高く、これはアンサンブル学習(バギング)の仕様や、特徴量の部分抽出が適していたと考えられる。

以上から、「かわいい」画像の分類においては提案方式に述べた4つの特徴量を用いた、Random Forestsによる分類方式が有効であると考える。

参考文献

- [1] 古渡 翔太・秋吉 政徳:特徴量フィルタを用いた「かわいい」の要素を含む画像の分類方法, 電気学会, 第87回情報システム研究会, IS-22-004, pp. 15-18(2022)
- [2] 旭 峻佑・成松 宏美・大和 淳司・平 博順: イラスト認識のための画像タイプ分類, 人工知能学会, 第35回全国大会, 2Yin5-20(2021)