

# 漫画画像理解のためのキャラクター顔分類

小松 俊太<sup>1</sup> 古田 諒佑<sup>2</sup> 谷口 行信<sup>1</sup> 日並 遼太<sup>3</sup> 石渡 祥之佑<sup>3</sup>

東京理科大学<sup>1</sup> 東京大学<sup>2</sup> Mantra 株式会社<sup>3</sup>

## 1 はじめに

近年電子書籍が普及し、国内外での日本の漫画の人気が高騰している。これに伴い、セリフの自動翻訳や漫画の推薦システム、イラストの自動着色技術等の需要が高まっている。しかし、セリフの自動翻訳について、一般的な翻訳手法では作風やキャラクター固有の表現を扱うことが困難である。これを解決するためには、翻訳を行う前にセリフをキャラクターごとに分類する必要がある。そこで、セリフを発話しているキャラクターに対応する画像を推定し、その画像情報を用いてキャラクターを分類し、その結果を翻訳タスクに適用する方法が考えられる。これを実現するために、本研究では、任意の漫画タイトルに適用でき、高精度なキャラクター分類を可能とする画像クラスタリング手法の提案を目指す。

## 2 関連研究

未知の漫画キャラクターの顔画像をクラスタリングする手法が複数提案されている。

Tsubota ら [1] は CNN による距離学習と k-means によるクラスタリングを用いた手法を提案した。この手法では、既知の漫画画像データを用いて、漫画画像に特化した特徴抽出器を学習し、さらに対象タイトルごとに CNN を距離学習することで、クラスタリング精度の向上を示した。しかし、この手法では新しいタイトルに対して逐一モデルの学習が必要であり、汎用的ではない。また、k-means によってクラスタリングを行うため、タイトル内に出現するクラス数 (キャラクター数) が既知である必要がある。

柳澤ら [2] は Tsubota らの手法の問題点を解決し、キャラクター数が不明なタイトルに対応するために、HDB-SCAN を用いたクラスタリング手法を提案した。しかし、テストデータの一部を事前にアノテーションする必要があるなど、実用上で問題点が残っている。

## 3 実験

本実験では、キャラクター固有の特徴量のみを用いて未知のキャラクターをクラスタリングする手法として、一般的な顔画像で用いられる距離学習手法を適用して精度を確認し、従来手法との比較を行う。

### 3.1 実験条件

本実験では 109 タイトルの漫画画像から構成される Manga109 [3] データセットを使用した。訓練データとテストデータは従来手法 [1, 2] と同様に、複数収録されているタイトルの 2 冊目以降を除き、残った 104 冊のうちそれぞれ 93 冊と 11 冊に分割した。比較手法として、特徴抽出ネットワーク (ResNet18, ResNet50,

表 1 実験結果

手法	AMI ↑	NMI ↑
Tsubota ら [1]	-	0.71
柳澤ら [2]	0.43	-
ResNet18 + Linear	0.49	0.61
ResNet50 + Linear	0.50	<b>0.62</b>
ResNetFace18 + Linear	0.36	0.51
ResNet18 + ArcFace	<b>0.51</b>	<b>0.62</b>
ResNet50+ ArcFace	0.42	0.56
ResNetFace18+ ArcFace	0.46	0.58

ResNetFace18) と分類および距離学習のためのネットワーク (Linear, ArcFace) を用いて、図 1 に示すフローで実験を行った。ここで、ResNet18 および ResNet50 は ImageNet データセットで、ResNetFace18 は ResNet18 の入力チャネルを 1 チャネルとした上で LFW データセットで事前に学習したネットワークである。また、評価指標にはクラスタリングタスクの一般的な評価指標である AMI および NMI を用いる。

### 3.2 実験結果

実験結果を表 1 に示す。表 1 から、柳澤ら [2] の手法と比較すると複数の手法で AMI を上回っているが、Tsubota ら [1] の手法と比べて NMI が下回っている。また、ArcFace による距離学習および人の顔情報の適用が効果的ではないことが伺える。

## 4 今後の課題

以上の実験から、キャラクター固有の特徴のみでは不十分であることが示唆されており、タイトルまたは著者固有の作風等の情報を加えて学習する手法が必要である。今後は、これらの情報を用いて、未知の漫画キャラクターであっても再学習の必要がなく、クラス数に依存しないクラスタリング手法の提案を試みる。

## 参考文献

- [1] Koki Tsubota, Toru Ogawa, Toshihiko Yamasaki, and Kiyoharu Aizawa. Adaptation of Manga Face Representation for Accurate Clustering. In *SIG-GRAPH Asia 2018 Posters*, 2018.
- [2] 柳澤秀彰, 京極健悟, Ravi Jian, 渡辺裕. 漫画キャラクター顔画像のクラスタリングにおける特徴表現の改良. *FIT*, 2019.
- [3] Yusuke Matsui, Kota Ito, Yuji Aramaki, Azuma Fujimoto, Toru Ogawa, Toshihiko Yamasaki, and Kiyoharu Aizawa. Sketch-based Manga Retrieval using Manga109 Dataset. *Multimedia Tools and Applications*, 2017.