

# 主要キャラクターの抽出を目的とした漫画キャラクター画像のクラスタリング

Clustering of Comic Character Images for Extraction of Major Characters

学生会員 柳澤秀彰<sup>†</sup>, 山下拓朗<sup>†</sup>, 正会員 渡辺 裕<sup>†</sup>  
Hideaki Yanagisawa<sup>†</sup>, Takuro Yamashita<sup>†</sup> and Hiroshi Watanabe<sup>†</sup>

あらまし 電子コミックの市場規模が拡大する中で、漫画画像の内容を理解してメタデータを生成する技術の需要が高まっている。特に登場キャラクターはストーリーを理解する上で重要な要素の一つである。キャラクターの認識にはキャラクターごとの特徴を学習した機械学習が用いられるが、複数の作品についてデータセットを作成する作業には膨大なコストがかかる。したがって、データセットを効率的に構築するために教師なし学習によってキャラクター画像を分類する技術が必要となる。しかし、キャラクター顔画像は一般画像よりも類似度の表現が困難であるほか、多数のノイズデータが存在することから従来の画像クラスタリングを適用できない。本論文では、CNNの出力をDBSCANでクラスタリングすることによって類似度の高い画像のみを抽出する手法を提案する。実験結果より、提案手法が複数の主要キャラクターをクラスタとして抽出可能であることを確認した。

キーワード：漫画，クラスタリング，ニューラルネットワーク，次元削減

## 1. ま え が き

日本における電子コミックの売上は2016年の時点で1617億円にのぼり、電子書籍市場全体の売上の約81.8%を占めている<sup>1)</sup>。このような背景から、電子コミックを対象とした自動翻訳や検索、着色といったサービスの需要が高まっている。これらの実装のために、漫画の内容を理解し、自動的にメタデータを生成する技術が研究されている。特にキャラクターは漫画の内容を理解するために重要な要素の一つである。したがって、漫画に登場するキャラクターをリスト化することによって、電子コミックサービスをより効率的に利用できると思われる。従来研究では、教師なし学習によるキャラクターの認識手法として、画像クラスタリングを用いたキャラクターの分類が提案されている。しかし、一作品に登場するキャラクターの数が不定であることや、登場頻度の少ないキャラクターがノイズとなるといった問題がある。

われわれは、登場頻度の少ないキャラクターをクラスタリング対象から除外し、主要なキャラクターのみを分類することによるキャラクターリストの生成を目的として、CNNの出力をDBSCANでクラスタリングする手法を提案する。実験では、CNNのモデル、特徴量の次元削減方法、DBSCANのパラメータ決定方法についてそれぞれ比較を行う

ことで、クラスタリング手法の最適な構成を求め、主要キャラクター抽出における精度を検証する。

## 2. 関連研究

### 2.1 漫画を対象とした画像処理

漫画画像からの物体認識について、小川らは深層学習を用いた物体検出手法であるSSDによる複数の漫画物体の並列検出を行っている<sup>2)</sup>。また、成田らは漫画キャラクターについて事前学習したCNNを特徴抽出器として利用し、漫画からのキャラクター顔画像の検索システムを提案している<sup>3)</sup>。登場キャラクターのクラスタリングに関して、長尾らはSURF特徴量とk-meansを用いた顔画像のクラスタリングを行っている<sup>4)</sup>。坪田らは深層距離学習を用いて、個別の漫画作品に適応した画像特徴量を生成することにより、k-meansのクラスタリング精度を向上させることに成功している<sup>5)</sup>。ただし、この研究は特徴量の有効性の検証を目的としており、正解クラスタ数を事前に設定しているほか、登場頻度の少ない「その他」のキャラクターは対象から除外している。

### 2.2 一般画像のクラスタリング

画像データセットを対象としたクラスタリングについて、従来研究では、混合ガウスモデルを画像に適用し、情報ボトルネックによってクラスタリングする手法<sup>6)</sup>や、SIFTやSURFなどの局所特徴量をBags of Featuresに変換し、凝集クラスタリングを行う手法<sup>7)</sup>などが提案されている。近年では、深層学習より得られる画像特徴量を利用した手法が提案されている。Liuらはディープオートエンコーダと

2018年5月30日受付，2018年9月20日再受付，2018年12月13日採録  
<sup>†</sup>早稲田大学 大学院基幹理工学研究科 情報理工情報通信専攻  
(〒169-0072 新宿区大久保3-14-9 早大シルマンホール401号室, TEL 03-5286-3488)

クラスタアンサンブルを組み合わせることで画像クラスタリングに適した特徴表現を生成する手法を提案している<sup>8)</sup>。Guérinらは、ImageNetについて事前学習を行ったCNNの出力を特徴量として、既存のクラスタリングアルゴリズムを適用する手法を提案している<sup>9)</sup>。この研究では、大規模データセットを学習したCNNが一般物体のクラスタリングに対して有効であることを示している。本研究はこの研究に基づき、一般物体について学習したCNNをキャラクター顔画像の特徴抽出器として使用する。

### 3. 提案手法

主要キャラクターの抽出に特化したクラスタリング手法として、CNNモデルとDBSCANによる顔画像のクラスタリング手法を提案する。提案手法は主に、CNN特徴量の計算、特徴量の次元削減、DBSCANのパラメタ決定とクラスタリングの実行、の三つの工程に分けられる。以下の項では各工程の詳細について述べる。

#### 3.1 画像特徴量

キャラクター顔画像の特徴量としてCNNの出力する特徴ベクトルを得る。本研究では、CNNのモデルとしてVGG16、VGG19<sup>10)</sup>、InceptionV3<sup>11)</sup>、Xception<sup>12)</sup>、ResNet50<sup>13)</sup>の5種類について検討を行う。各モデルについて、一般物体データセットのImageNetを事前学習した初期重みを適用する。抽出する特徴ベクトルは文献<sup>9)</sup>を参考に、最終層の一つ前の層の出力とする。

#### 3.2 次元削減

CNNの出力する特徴ベクトルは4096次元または2048次元のデータである。しかし、高次元空間ではあるデータ点から見た他の点群が同様に遠くに位置するようになるため、データ間の距離差を捉えにくくなる「球面集中現象」と呼ばれる現象が発生する。このことから、高次元データを対象としたクラスタリングは困難である。したがって、一般的には、高次元空間における特徴を維持したまま低次元にデータを再編する次元削減を行うことで、クラスタリングしやすいデータを得る。次元削減には線形変換と非線形変換が存在し、前者の例としてPCA、後者の例としてKernel-PCAやt-SNEが存在する。

##### 3.2.1 主成分分析 (PCA)

PCAは、教師なしの次元削減として一般的な手法である。データの分散共分散行列を固有値分解することで固有ベクトルを計算し、固有ベクトルの中でデータの分散を最大化するものを第一主成分と定める。次に、第一主成分と直交するベクトルのうち、分散を最大化するものを第二主成分として、第三以降の主成分もそれまでに決定した主成分と直交するという条件の下で分散を最大化するベクトルを選んでいく。これらの主成分軸を基底として高次元データの射影を繰り返すことで任意の次元のデータに圧縮することが可能である。

##### 3.2.2 カーネル主成分分析 (Kernel PCA)

PCAは線形の関係性を基にした分析手法であるため、データがガウス分布に従う場合は有効であるが、非線形な構造を持つデータ集合に対しては効果的な変換ができないという問題がある。この問題に対して、Kernel PCA<sup>14)</sup>は、カーネル法を用いてPCAを非線形に拡張し、非線形の変換を行う手法である。はじめに、データをカーネル法によって元よりも大きな次元へ写像し、高次元空間においてPCAを実行することで非線形の次元削減を行う。本研究では、カーネルの関数としてコサイン類似度を使用する。

##### 3.2.3 t-SNE

t-SNE<sup>15)</sup>は主に高次元データの可視化に用いられるアルゴリズムである。高次元空間上の2点 $x_i, x_j$ の類似度と低次元空間上の2点 $y_i, y_j$ の類似度をそれぞれ確率分布 $p_{ij}$ と $q_{ij}$ で表現し、二つの確立分布の差が最小となるように確率分布のパラメタを最適化する。これによって局所的なデータ間の特性に基づいて低次元への変換が行われる。PCAが「類似しないデータを遠くへ配置する」ことを優先することに対して、t-SNEは「類似するデータを近くに配置する」ことを優先するアルゴリズムとなっている。一方でt-SNEの問題点として、2~3次元より次元数の高いデータへの圧縮が保証されていないことや、密なクラスタを広げ、疎なクラスタを縮めるという圧縮の特性のために相対的なクラスタのサイズ情報が失われるといった点がある。本研究では、t-SNEの計算処理の反復回数を1000回とし、パラメタ perplexity を30に設定する。

#### 3.3 クラスタリング

##### 3.3.1 DBSCAN

DBSCAN (Density-based spatial clustering of applications with noise)<sup>16)</sup>はある空間に点集合が与えられたとき、互いに多くの隣接点を持つ点の集合をグループにまとめ、最近接点が遠い点を外れ値とするクラスタリングアルゴリズムである。DBSCANは与えられたデータ点を、コア点、到達可能点、外れ値の三つに分類する。コア点は自身の半径 $\epsilon$ 以内にMinPts個以上のデータ点が存在する点、到達可能点は半径 $\epsilon$ 以内にMinPts個未満のデータ点しか存在しないが、半径 $\epsilon$ 以内にコア点が存在する点、外れ値は、半径 $\epsilon$ 以内にコア点が存在しない点である。次に、コア点の集まりからクラスタを作成し、隣接している到達可能点をそれぞれのクラスタに割り当てることでクラスタリングを行う。DBSCANはk-meansのようなクラスタリング手法と異なり、事前にクラスタ数を決定する必要がないほか、外れ値に対してロバストであるという利点がある。

##### 3.3.2 DBSCANのパラメタの決定

DBSCANの実行にはminPtsと $\epsilon$ の二つのパラメタを正しく設定することが重要である。minPtsは望まれる最小のクラスタサイズであり、クラスタリングを行うためには3以上に設定する必要があるが、大抵の場合より高い値に設定する方がノイズデータの分離において有効である。本研究では、20回以上登場するキャラクターを主要キャラクターと

定義し、 $\text{minPts} = 20$ に設定する。 $\epsilon$ はデータ点の結びつきを決定するパラメタである。 $\epsilon$ が非常に小さい場合にはデータの大部分は外れ値としてクラスタリングされない。また、大きな値の場合にはクラスタは併合され、データの大多数は同一のクラスタに存在する。一般的な $\epsilon$ の決定方法として、 $k = \text{minPts}$ の最近傍への距離をプロットした $k$ 距離グラフを解析することで最適な値を得る方法が用いられる。Soniらは点群の密度が一様でないデータセットに対して有効な $\epsilon$ を求める手法として、AGED (Automatic Generation of Eps for DBSCAN)を提案した<sup>17)</sup>。まず点 $x$ について1~ $k$ 番目の近傍点までのユークリッド距離を計算し、その平均値からADEN (average local density function)を求める。ここで、 $d$ はデータの次元数である。

$$INF(x, y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}$$

$$DEN(x, y_1, \dots, y_k) = \sum_{i=1}^k INF(x, y_i)$$

$$ADEN(x, y_1, \dots, y_k) = \frac{\sum_{i=1}^k INF(x, y_i)}{k}$$

次に、各点 $p_i$ について求めたADENの値を大きき順にソートして10分割したヒストグラムを作成する。最後に、ヒストグラムの中でデータを個以上含むバケットについて平均値を計算し、これらを $\epsilon$ の候補としてクラスタリング結果を検証することで最適な値を決定する。

本研究では、クラスタリングの自動化を目的として、AGEDによって求められた候補から一意な $\epsilon$ を決定する方法を検討する。漫画一冊に登場する主要キャラクタの割合はそれぞれ異なることから、それぞれの漫画作品におけるデータ分布に適した $\epsilon$ の値を求める必要がある。顔画像の特徴量から求めたADENグラフおよびAGEDによって提案される $\epsilon$ の候補を図1に示す。ここで主要キャラクタとノイズデータではデータ間の距離差が大きくなるため、グラフの勾配変化が起きる点に相当すると推測した。したがって、AGEDによって求められる $\epsilon$ の候補について、隣接している候補同士の間の傾きを求め、この傾きがすべての候補間の傾きの平均値より大きく変化する箇所を $\epsilon$ に設定する手法を提案する。

## 4. 実験

### 4.1 テストセット

本実験に使用する漫画画像は、Manga109<sup>18)</sup> <sup>19)</sup>で公開されている作品のうち作者の異なる11冊を使用する。用意されたアノテーションに従い、 $30 \times 30$  pixel以上のサイズの顔領域を切り出してテストセットを作成した。このとき、アノテーションはキャラクタの髪や耳の領域を含まないため、顔領域の矩形の横幅、縦幅をそれぞれ均等に2倍に拡張し、ページからはみ出した場合にはページに合わせてク

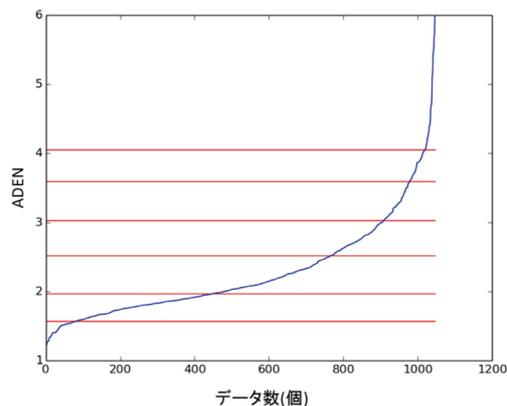


図1 漫画一冊の顔画像より算出した特徴量のADENグラフ  
赤線はAGEDによって求められた $\epsilon$ の候補を示す。

表1 テストセットにおけるクラス数と画像数  
クラス数はそれぞれ1個の「その他」のクラスを含み、括弧内は「その他」のクラスに属する画像数を示す。

タイトル	クラス数	画像数
ARMS	8	319 (122)
愛さずにはいられない	8	972 (141)
あっけら貫刃帖	10	704 (155)
あくはむ	8	1047 (64)
青すぎる春	10	552 (80)
天晴れ! カッポーレ	8	864 (75)
ありさ?	8	960 (156)
BEMADER・P	14	1226 (123)
爆裂! かんふー娘	11	1036 (172)
ベルモンド	12	847 (44)
ラブひな 14巻	11	1168 (43)

リップングするように切り出しを行った。次に顔画像をキャラクタ別に分類し、画像枚数が20枚以上のものを主要キャラクタとして個別のクラスに割り当てた。また、画像枚数が20枚未満のキャラクタはすべて「その他」のクラスとして割り当てた。作成されたテストセットの内容を表1に示す。CNNに入力する際には各画像のサイズを $224 \times 224$ にリサイズした。

### 4.2 クラスタリング評価基準

本研究は、主要キャラクタの抽出によるリストの生成を目的として、多数の主要キャラクタがそれぞれ異なるクラスとして抽出されているかを評価する。このとき、すべての画像を漏れなく抽出することは重視しないこととする。クラスタリングの評価指標として、DBSCANによってクラスとして抽出されたデータを対象に、以下のpurityとinverse purityを求める。

$$purity(\Omega, \mathbb{C}) = \frac{1}{N} \sum_k \max_j |\omega_k \cap c_j|$$

$$inverse\ purity(\Omega, \mathbb{C}) = \frac{1}{N} \sum_j \max_k |\omega_k \cap c_j|$$

ここで、 $N$ は得られたクラスタの総データ数、 $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$ は生成されたクラスタの集合、 $\mathbb{C} = \{c_1, c_2, \dots, c_j\}$ は正解クラスの集合を表す。このとき、「その他」に属する

画像はノイズデータとして、Cから除外している。purityは「異なるキャラクタが同じクラスタに入らない」ことを評価する基準、inverse purityは「同じキャラクタが異なるクラスタに入らない」こと評価する基準である。クラスタの正確さの評価指標として、purityとinverse purityの調和平均であるF値を求める。

$$F\text{値} = \frac{2}{\frac{1}{\text{purity}} + \frac{1}{\text{inverse purity}}}$$

このとき、抽出されるクラスタの数が少ないほどF値は大きくなる傾向が見られる。本研究では正確なクラスタを得ると同時に、可能な限り多数のキャラクタを抽出することを目的とすることから、purityとinverse purityに加えてクラスタ数を考慮した独自の評価基準 $\alpha$ を定め、本実験において最も重視する基準とする。

$$\alpha = \text{クラスタ数} \times \text{purity} \times \text{inverse purity}$$

### 4.3 実験1 画像特徴量と次元削減

特徴抽出に使用するCNNおよび次元削減の方法について、顔画像クラスタリングに最適な組み合わせを求める。CNNはVGG16, VGG19, InceptionV3, Xception, ResNet50

の5種類を使用する。はじめに、次元数とクラスタリング結果の関係を検証するため、Kernel PCAを用いて2次元から64次元まで次元数を変化させた結果を比較した。それぞれDBSCANのパラメタ $\alpha$ をk距離グラフの最小値から最大値に到達するまで1/100の間隔で変化させて、 $\alpha$ の値が最大となるときのクラスタリング結果を求めた。実験結果を図2に示す。このグラフより、次元数が大きいほどF値は上昇し、反対に抽出されるクラスタ数の平均値は少くなる傾向が確認できた。このような結果が得られた理由として、高次元のデータはデータ間の差異をより詳細に表現できる反面、球面集中現象の影響により、類似したデータをまとめることは難しくなるためだと推測できる。この結果、 $\alpha$ の値は16次元の辺りでピークをとることがわかった。

次に、異なる次元削減を行った場合のクラスタリング結果の変化を比較した。次元削減を行わない元データ、PCAによって64次元、16次元、2次元に圧縮したデータ、Kernel PCAによって64次元、16次元、2次元に圧縮したデータ、t-SNEによって2次元に圧縮したデータの8種類について検証した。このとき乱数の影響を考慮して、t-SNEは10回のクラスタリング結果の平均値を求めた。

実験結果を表2に示す。次元削減の方法について、t-SNE

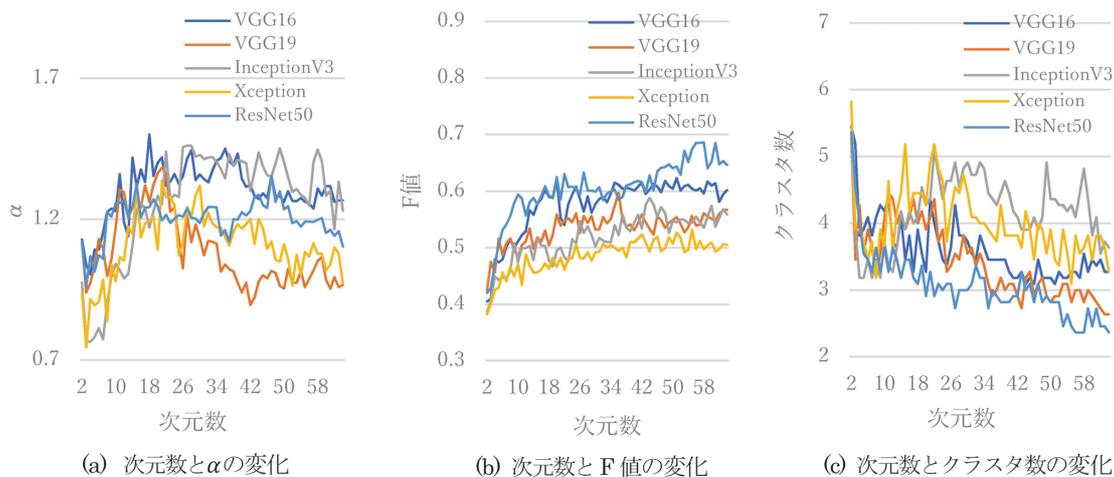


図2 データの次元数によるクラスタリング結果の変化

表2 特徴抽出器として使用するCNNと次元削減の手法によるクラスタリング結果の変化

		次元削減なし	PCA 64次元	PCA 16次元	PCA 2次元	KPCA 64次元	KPCA 16次元	KPCA 2次元	t-SNE 2次元
		VGG16	$\alpha$	0.711	0.814	0.911	0.900	1.267	1.376
	F値	0.634	0.585	0.503	0.427	0.602	0.586	0.405	0.494
	クラスタ数	1.46	2.18	3.00	3.82	3.273	3.82	5.455	7.41
VGG19	$\alpha$	0.665	0.837	0.939	0.868	0.966	1.242	1.114	1.848
	F値	0.580	0.561	0.514	0.411	0.567	0.511	0.424	0.483
	クラスタ数	1.82	2.36	3.27	4.55	2.636	4.09	5.273	7.54
InceptionV3	$\alpha$	0.429	0.536	0.634	0.718	1.230	1.255	0.975	1.560
	F値	0.519	0.544	0.503	0.368	0.559	0.504	0.386	0.471
	クラスタ数	1.27	1.36	2.00	4.09	3.636	4.18	5.273	6.33
Xception	$\alpha$	0.519	0.612	0.711	0.925	0.973	1.135	0.939	1.506
	F値	0.555	0.536	0.505	0.400	0.505	0.461	0.382	0.436
	クラスタ数	1.36	1.73	2.46	4.64	3.27	4.18	5.82	7.24
ResNet50	$\alpha$	0.828	0.961	1.145	1.135	1.103	1.206	1.120	2.149
	F値	0.646	0.596	0.560	0.440	0.646	0.592	0.420	0.499
	クラスタ数	1.82	2.36	3.35	5.82	2.36	3.45	5.36	7.98

表3  $\epsilon$ の決定方法によるクラスタリング結果の変化

	理想値	AGED 理想値	AGED 平均値	AGED 中央値	提案手法	シルエット
purity	<b>0.543</b>	0.533	0.492	0.507	0.498	0.319
inverse purity	0.482	0.510	0.517	0.537	0.472	<b>0.757</b>
クラスタ数	<b>7.98</b>	6.64	6.12	5.96	6.29	2.15
$\alpha$	<b>2.149</b>	1.821	1.634	1.621	1.743	0.519
F 値	0.499	0.506	0.481	<b>0.501</b>	0.468	0.435

© 石岡 ショウエイ



(a) 髪型が類似した画像を抽出したクラスタの例



(b) 背景が類似した画像を抽出したクラスタの例

図3 同じクラスタとして抽出された顔画像の例

(a) について左と中央は同一キのキャラクタだが、右は異なる。(b) はそれぞれ異なるキャラクタである。

を用いたとき他の手法で2次元に圧縮した場合と同等以上のF値でより多くのクラスタが抽出されることを確認した。このことから、DBSCANの適用において、データ全体の分散に着目した圧縮より局所的な類似度に着目した圧縮の方が有効であることが示された。また、CNNモデルの比較では、ResNet50を用いたとき他のモデルよりF値が高くなることを確認した。ResNet50によってクラスタとして抽出される顔画像の例を図3に示す。深層特徴量によるクラスタリングに共通した傾向として、キャラクタの髪や背景のテクスチャの類似性に基づいてクラスタリングされることが確認できた。このことから、異なるキャラクタの画像が類似した髪型や背景を持つときに、キャラクタの誤分類が発生しやすいことがうかがえる。また、ResNet50とそれ以外のCNNによる結果を比較すると、ResNet50では髪に着目したクラスタが抽出される割合が他のCNNよりも高いため、より良いクラスタリング結果が得られたと考えられる。

#### 4.4 実験2 DBSCANのパラメタ決定

AGEDによって求められる $\epsilon$ の候補から最適な値を決定する方法を検討する。画像特徴量として、実験1において最も高い $\alpha$ の値が得られたResNet50とt-SNEの組み合わせを使用して検証を行った。実験1で求められた結果を「理想値」、AGEDの候補より、クラスタリング結果の $\alpha$ の値が最大となるように選択した結果を「AGED理想値」とし

て、候補から平均値を $\epsilon$ に設定する方法(AGED平均値)、中央値を $\epsilon$ に設定する方法(AGED中央値)、勾配変化が起こる箇所を選択する方法(提案手法)の3種類について検討した。また文献<sup>20)</sup>で示されたシルエットスコアからを決定する手法を「シルエット」として合わせて評価した。それぞれ実験1と同様に各漫画作品について10回のクラスタリングを行い、平均値を求めた結果を表3に示す。提案手法を平均値や中央値に設定した場合と比較すると、inverse purityが低下する半面、ほぼ同等のpurityでより多くのクラスタを抽出することが可能であり、 $\alpha$ の値は高くなった。このことから、データ間距離の勾配変化から $\epsilon$ を決定する方法が有効性を持つことを確認した。また、シルエットスコアを使用した場合では少数のクラスタのみが抽出され、purityの値も低くなることから今回のクラスタリングには適さないことがわかった。

#### 5. むすび

本研究では、教師なし学習で漫画画像からキャラクタを分類する方法として、CNNの出力をDBSCANでクラスタリングする手法を提案し、その精度を評価した。実験結果より、DBSCANの適用について、ResNet50による特徴抽出とt-SNEを用いた次元削減が有効であることを確認し、DBSCANのパラメタの自動決定においてデータ間距離の勾配変化に基づく手法が有効性を持つという知見が得られた。今後の課題としてクラスタリングの精度を向上するために、fine-tuningを行ったCNNの利用や、画像からの背景除去処理が考えられる。

謝辞 本研究はJSPS科研費17K00511の助成を受けたものである。

#### 〔文 献〕

- 1) インプレス総合研究所：“電子書籍ビジネス調査報告書2017”，インプレス総合研究所（2017）
- 2) 小川，山崎，相澤：“並列化された検出器による高精度漫画物体検出”，映像学技報，42，4，pp.293-298（Feb. 2018）
- 3) R. Narita, K. Tsubota, T. Yamasaki and K. Aizawa: "Sketch-based Manga Retrieval using Deep Features", 2017 14th IAPR International Conference on Document Analysis and Recognition, pp.49-53 (Nov. 2017)
- 4) 長尾，渡辺：“コミックにおける主要キャラクタ同定の検討”，信学総大，D-21-3（Mar. 2016）
- 5) 坪田，小山，山崎，相澤：“キャラクタ顔特徴量の個別漫画への適応手法”，人工知能学全国大（32回），1K2-OS-2b-04（June 2018）
- 6) J. Goldberger, S. Gordon and H. Greenspan: "Unsupervised Image-set Clustering using an Information Theoretic Framework", IEEE transactions on image processing, 15, 2, pp.449-458（Feb. 2006）

- 7) T. Fukui and T. Wada: "Commonality Preserving Image-set Clustering Based on Diverse Density", in International Symposium on Visual Computing, pp.258-269 Springer (2014)
- 8) H. Liu, M. Shao, S. Li and Y. Fu: "Infinite Ensemble for Image Clustering", in Proceeding of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.1745-1754 (Aug. 2016)
- 9) J. Giéin, O. Gibaru, E. Nyiri and S. Thiery: "CNN Features are also Great at Unsupervised Classification", arXiv:1707.01700 (July 2017)
- 10) K. Simonyan and A. Zisserman: "Very Deep Convolutional Networks for Large-scale Image Recognition", arXiv:1409.1556 (Sep. 2014)
- 11) C. Szegedy, V. Vanhoucke, S. Ioffe and J. Shlens: "Rethinking the Inception Architecture for Computer Vision", in Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, pp.2818-2826 (June 2016)
- 12) F. Chollet: "Xception: Deep Learning with Depthwise Separable Convolutions", arXiv:1610.02357 (Oct. 2016)
- 13) K. He, X. Zhang, S. Ren and J. Sun: "Deep Residual Learning for Image Recognition", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778 (June 2016)
- 14) B. Schölkopf, A. Smola and K.R. Müller: "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", Neural Computation, 10, 5, pp.1299-1319 (July 1998)
- 15) L. van der Maaten and G. Hinton: "Visualization Data using t-SNE", Journal of Machine Learning 9, pp.2579-2605 (Nov 2008)
- 16) M. Ester, H.P. Kriegel, J. Sander and X. Xu: "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", in Proceeding of 2nd International Conference on Knowledge Discovery and Data Mining, pp.226-231 (Aug. 1996)
- 17) N. Soni and A. Ganatra: "AGED (Automatic Generation of Eps for DBSCAN) ", International Journal of Computer Science and Information Security, 14, 5 (May . 2016)
- 18) Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamasaki and K. Aizawa: "Sketch-based Manga Retrieval using Manga109 Dataset", Multimedia Tools and Applications, Springer (Nov. 2016)
- 19) T. Ogawa, A. Otsubo, R. Narita, Y. Matsui, T. Yamasaki and K. Aizawa: "Object Detection for Comics using Manga109 Annotations", arXiv:1803.08670 (Mar 2018)
- 20) E. Taskesen and M.J. T. Reinders: "2D Representation of Transcriptomes by t-SNE Exposes Relatedness between Human Tissues", PLoS ONE 11 (2) : e0149853, <https://doi.org/10.1371/journal.pone.0149853> (Feb. 2016)



柳澤 秀彰 2016年，早稲田大学大学院基幹理工学研究科情報理工情報通信学科卒業。現在，同大学大学院博士課程在学。漫画の物体検出，キャラクターの認識の研究に従事。学生会員。



山下 拓朗 2018年，早稲田大学基幹理工学部情報理工学科卒業。現在，同大学院基幹理工学研究科情報理工・情報通信専攻在学。深層学習を用いた画像処理。マンガ画像におけるキャラクターの認識に関する研究に従事。



渡辺 裕 1985年，北海道大学大学院工学研究科博士課程修了。同年，NTT入社。NTT研究所において画像の効能率符号化およびMPEG標準化に従事。2000年より，早稲田大学教授。映像認識・深層学習に興味を持つ。工学博士。正会員。