

アスキーアート分類手法の比較検討

Comparison of Classification Methods for Ascii Art

松本 和幸^{*1} 藤澤 日明^{*2} 吉田 稔^{*1} 北 研二^{*1}
Kazuyuki Matsumoto Akira Fujisawa Minoru Yoshida Kenji Kita

^{*1}徳島大学 ^{*2}青森大学
Tokushima University Aomori University

In recent years, a lot of non-verbal expressions have been used on social media. Ascii art (AA) is an expression by visual technique using characters. In this paper, we set up an experiment to classify AA pictures by using character features and image features. We try to clear which feature is more effective for the method to classify AA pictures. We proposed four methods; 1) character frequency based method, 2) character importance value based method and 3) image feature based method, 4) character's image feature based method. We trained the neural networks by using these four features. As the experimental result, the best classification accuracy was obtained with the feed forward neural networks using character's image feature.

1. はじめに

非言語表現の一つとして、電子掲示板などでアスキーアートという表現が用いられる。アスキーアートは、1行内に収まる顔文字や絵文字といった非言語表現とは異なり、複数行にわたって表現されるため、文字ベースでの解析が困難である。しかしながら、1枚のアスキーアートが表現できる意味内容は非常に多様であり、Web上のデータ解析において無視することができない表現といえる。アスキーアートは文字を点や線の代わりに用いて絵を視覚的に表すため、使用される一つ一つの文字は、アスキーアート中の人物の台詞やキャプションを除けば、ほとんど意味を持たない。したがって、アスキーアートは文字集合として扱うよりも、画像として扱うほうが適していると考えられる。本研究では、アスキーアートのカテゴリ分類に画像特徴量が有効かどうかを検証するため、アスキーアートの文字特徴量（文字出現頻度、文字重要度）と、アスキーアートを画像化して得た画像特徴量をそれぞれ機械学習により学習させてアスキーアートのカテゴリ分類器を作成する。これらの分類器を実験により評価することで、有効な特徴量について考察する。

2. 関連研究

谷岡ら [Tanioka2005] は、バイトパターンと形態素解析による品詞解析結果を用いてサポートベクターマシンによるアスキーアートを含むテキストとそうでないテキストとに分類する手法を提案し、90%の分類精度を得ている。林ら [Hayashi2009] は、アスキーアートを言語に依存せず抽出する手法として、同じ文字が連続して2回現れる回数に着目し、さらに、テキストから LZ77 圧縮による圧縮率、RLE 圧縮による圧縮率を素性として用いることにより、アスキーアートの抽出実験の結果、90%を超える高い F 値を得ている。鈴木 [Suzuki2011] は、林らと同様に圧縮率に着目しており、機械学習アルゴリズム C4.5 を用いて特定の言語に依存しない抽出を実現している。いずれの研究においても、アスキーアートかそうでないかの2値分類タスクを対象としたものであり、アスキーアートの意味理解については考慮されていない点で本研究とは異なる。

藤澤ら [Fujisawa2016] は、アスキーアートの持つ形状情報を活用するための手法として、アスキーアートを画像に変換することで、画像特徴量を用いたアスキーアートの類似度評価手法を提案した。画像特徴量を用いることで、顔文字をもとに作成されたアスキーアートに対して、使用されている文字やサイズの違いに影響されず画像特徴量の類似性を確認することに成功している。藤澤ら [Fujisawa2017] は、顔文字と、その顔文字をもとに作られた大型アスキーアートに対し、画像特徴量を抽出して類似度比較を行っている。この研究において、藤澤らは、形状特徴量である HOG (Histograms Oriented Gradients) が画像特徴量として有効であると述べている。本研究では、アスキーアートのカテゴリを文字特徴量と画像特徴量の2通りの特徴量を用いた分類器を機械学習により構築することで、従来はほとんど対象とされなかったアスキーアートのカテゴリ分類手法について比較検討する。

3. アスキーアートからの特徴抽出

3.1 画像特徴量を用いたアスキーアートの分類

1枚のアスキーアート全体から画像特徴量を抽出し、それを機械学習させることでアスキーアート分類器を作成する。アスキーアートの画像への変換には、AAtoImage[AAtoImage]を用いる。画像に変換されたアスキーアートは、N×Nピクセルのサイズの画像に正規化した後、二値化処理をおこないエッジ画像に変換する。本論文では、N=128と設定した。

前述の手順でカテゴリごとに準備したアスキーアートのエッジ画像を、畳み込みニューラルネットワーク (CNN) により学習し、アスキーアート分類器を作成する。画像の枚数による学習の偏りの影響を避けるため、あらかじめカテゴリ間の偏りが出ないように訓練データ数を揃えておく。

また、文字単位で画像特徴量 (HOG) を抽出し、その特徴量を k-means 法によりクラスタリングすることで得た文字画像のクラスの出現頻度ベクトル (文字画像ベクトル) をフィードフォワードニューラルネットワーク (FFNN) に学習させてアスキーアート分類器を作成する手法も試す。クラスタリングのクラス数を 1000 とした。

連絡先: 徳島大学理工学部, 南常三島町 2-1,
matumoto@is.tokushima-u.ac.jp

3.2 文字特徴量を用いたアスキーアートの分類

アスキーアートから、使用されている文字の頻度を用いる文字頻度ベクトル特徴を FFNN に学習させる方法と、単語重要度計算に用いられる TF*IDF の考え方に基づく文字重要度を特徴とした行列を CNN により学習させる方法を検討する。文字重要度 (CF*IAF) は、式 1,2,3 により計算できる。文字重要度を文字単位で計算することで、画像におけるピクセルの画素値と同じ効果を期待する。

$$CF * IAF = CF \times IAF \quad (1)$$

$$CF = \text{freq}_c^i \quad (2)$$

$$IAF = \log\left(\frac{N}{AF} + 1.0\right) \quad (3)$$

CF は、あるアスキーアート AA_i における文字 c の出現頻度を示す。N は、重要度の計算をおこなうアスキーアートの総数を示し、AF は、文字 c が幾つのアスキーアートに出現したかを示す。図 1 に、アスキーアートから文字重要度行列を抽出し、CNN により学習させる過程を簡単に示す。アスキーアートによって 1 行あたりの文字数や行数が異なることから、1 行における最大文字数を 100 文字、最大行数を 100 行と設定し、このサイズに満たない場合は、空白文字によりパディング処理を施した。また、このサイズよりも大きいものは、各行の先頭から 100 文字、1 行目～100 行目までを使用して特徴抽出をおこなった。

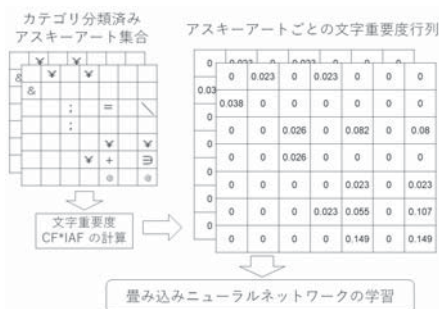


図 1: アスキーアートから文字重要度行列の抽出

同じ作品に登場するキャラクターをモチーフとした AA は、同じ作家が作成している場合が多く、同一キャラクターを描く際には差分のみを変更して作成することが多い。このため、文字頻度ベクトル特徴を用いるだけでも、同一作品カテゴリの判定が可能であると予想できる。しかし、同じ作品においても、デフォルメされたキャラクターであったり、キャラクターの人気や複雑度などから作られるアスキーアート数が少ないキャラクターも存在する。このことから、アスキーアートの種類によっては、文字特徴のみでのカテゴリ分類が困難なことが予想される。

4. 評価実験

4.1 実験設定

画像特徴による AA 分類器 (2 種類) と文字特徴による AA 分類器 (2 種類) がどの程度アスキーアートのカテゴリ分類に適しているかを確認するため、同じアスキーアートを基に作成したデータに対する評価実験をおこなう。AA 分類器の学習に用いたニューラルネットワーク (NN) の各層のパラメータは、

表 1, 2, 3 に示すとおりである。学習試行回数はいずれも 50 回とし、ミニバッチサイズを 256 と設定して学習をおこなった。畳み込み層におけるフィルタは、 3×3 のサイズでフィルタ数を 20, 32 とする。中間層の活性化関数に ReLU、最適化アルゴリズムに Adam を使用し、学習率を 0.0001 とする。出力層 (全結合層) の直前のドロップアウト層でのドロップアウト率を 0.5 とし、出力層の活性化関数に Softmax を用いる。

表 1: 文字頻度/文字画像ベクトルを学習させる FFNN の構造

Layer	出力	パラメータ数
全結合層-1	(500)	3374000
全結合層-2	(128)	64128
全結合層-3	(44)	5676

表 2: 文字重要度を学習させる CNN の構造

Layer	出力	パラメータ数
畳み込み層-1	(98, 98, 20)	200
畳み込み層-2	(96, 96, 20)	3620
プーリング層-1(Max pooling)	(48, 48, 20)	0
畳み込み層-3	(46, 46, 20)	18496
プーリング層-2(Max pooling)	(23, 23, 20)	0
Flatten 層	(10580)	0
全結合層-1	(1024)	58983424
全結合層-2	(44)	45100

表 3: 画像特徴量を学習させる CNN の構造

Layer	出力	パラメータ数
畳み込み層-1	(126, 126, 32)	320
畳み込み層-2	(124, 124, 32)	9248
プーリング層-1(Max pooling)	(62, 62, 32)	0
畳み込み層-3	(60, 60, 64)	18496
プーリング層-2(Max pooling)	(30, 30, 64)	0
Flatten 層	(57600)	0
全結合層-1	(1024)	58983424
全結合層-2	(44)	45100

4.2 実験データ

本研究では、Web サイト「AAMZ Viewer^{*1}」からダウンロード可能なまとめ zip ファイルに含まれる漫画やアニメ、ゲームなどの作品のキャラクターをモチーフに作成されたアスキーアートを対象とし、作品のシリーズ名をカテゴリとみなす。「匿名掲示板 2ちゃんねる」のオリジナルアスキーアートや「その他作品」に分類されるアスキーアートは実験対象から除いた。この理由として、これらをカテゴリの分類に含めると、様々な作品やキャラクターをモチーフにしたものが同一カテゴリに含まれることになってしまい、それらが訓練の際のノイズとなる可能性があると考えたためである。一方で、「原哲夫作品」などといった、シリーズを通して作風・絵柄が類似するキャラクターが登場する作品群に関しては、一つのカテゴリとして扱う。各アスキーアートの数が 900 以上のカテゴリ (計 44 カテゴリ) を実験対象とする。訓練データと検証データを、カテゴリごとに均等になるように 8:1 の比率で分け、残りのアスキーアートを評価データとする。データの内訳を表 4 に示す。

以上のように準備したデータをもとに、文字特徴量に基づく分類器 (2 種類) と、画像特徴量に基づく分類器 (2 種類) をそれぞれ作成する。

4.3 評価方法

ニューラルネットワークにより出力された分類カテゴリごとの確率値の上位 $k=3,5,10$ 件以内に、正解カテゴリが出力されれば分類成功とする場合の分類成功率を求める。また、上位 1 件に出力された場合のみを成功とした場合の完全一致率 (精

*1 <http://aa.yaruyomi.com/>

表 4: データの内訳

訓練データ	検証データ	評価データ
35200	4400	20309

度)も求める。さらに、 $k=10$ のときに分類成功した事例の平均順位について求めることで、分類器の性能を比較する。

4.4 実験結果

表 5 に、各分類器における分類成功率、完全一致率および平均順位を示す。「文字頻度」は、文字頻度ベクトルを FFNN により学習した分類器、「文字重要度」は、文字重要度行列を CNN により学習した分類器、「画像特徴」は、エッジ画像に変換して CNN により学習した分類器を示している。また、「文字画像」は、文字単位の HOG 特徴量をクラスタリングして得た文字画像ベクトルを FFNN により学習した分類器を示している。

表 5: 実験結果

	文字頻度	文字重要度	画像特徴	文字画像
分類成功率 $k=3$ (%)	41.19	21.61	32.96	45.22
分類成功率 $k=5$ (%)	49.57	28.92	39.18	52.65
分類成功率 $k=10$ (%)	64.30	44.16	51.72	65.96
完全一致率 (%)	27.59	11.12	25.64	31.97
平均順位	3.307	4.307	3.299	3.046
作品別精度平均 (%)	31.41	13.13	45.50	35.33

実験の結果、文字画像ベクトルによる手法が、他の手法よりも高い分類成功率、完全一致率および平均順位が得られた。また、画像特徴に基づく手法は、作品別精度平均で最も高い結果が得られた。

5. 考察

精度上位および下位作品の例を示しながら、各手法ごとの違いについて考察する。表 6, 7, 8, 9 は、各手法における精度が上位 10 件、下位 10 件の作品とその精度を示したものである。上位 10 件だけをみると、画像特徴を用いた手法はすべての作品において 70%以上の分類精度を示しており、一方で、文字重要度に基づく手法では、60%以上の成功率の作品は存在しないことがわかる。

図 2 に、4 手法を通して最も高い精度を得た作品「BLEACH」のアスキーアートを示す。アスキーアートからはある程度複雑な絵柄であることがわかる。また、このカテゴリに属する他のアスキーアートをみると、ほぼ同じアスキーアートで、少しの差分があるだけのものも多くみられた。この作品のアスキーアートの種類は全部で 913 種類であり、すべてのデータを学習および評価に用いた。このため、絵柄や作風が独特であるということよりも、同じようなアスキーアートが同一カテゴリに含まれることが結果に大きく影響したといえる。しかし、画像特徴を用いた手法は、文字頻度や文字重要度を用いた手法とは

表 7: 完全一致率 (文字重要度)

順位	作品名	精度	順位	作品名	精度
1	恋姫無双	52.27	35	コードギアス	6.13
2	逆転裁判	42.86	36	魔法少女リリカルなのは	6.09
3	プリキュア	30.33	37	Rozen Maiden	5.53
4	原哲夫作品	29.92	38	涼宮ハルヒの憂鬱	5.29
5	ドラえもん	29.5	39	咲-Saki-	4.86
6	銀河英雄伝説	27.1	40	東方 project	4.68
7	ポケットモンスター	22.23	41	機動戦士ガンダム	3.84
8	魔人探偵脳噛ネウロ	21.97	42	特撮	3.75
9	機動戦艦ナデシコ	17.56	43	魔法少女まどか☆マギカ	3.54
10	スクライド	17.39	44	遊戯王	3.07

表 8: 完全一致率 (画像特徴)

順位	作品名	精度	順位	作品名	精度
1	BLEACH	92.63	35	魔法少女リリカルなのは	17.34
2	逆転裁判	91.84	36	機動戦士ガンダム	17.15
3	銀魂	87.1	37	魔法少女まどか☆マギカ	17.14
4	化物語	80.81	38	遊戯王	16.67
5	キン肉マン	80.73	39	特撮	15.27
6	原哲夫作品	74.8	40	プリキュア	14.37
7	恋姫無双	73.3	41	TYPE-MOON 作品	13.46
8	さよなら絶望先生	71.32	42	涼宮ハルヒの憂鬱	13.25
9	銀河英雄伝説	70.56	43	東方 project	12.86
10	スクライド	70.29	44	Rozen Maiden	11.91

異なる作品が上位に多く出現していることから、文字の種類に大きく影響を受けなかったと考えられ、汎用性の高い手法であるといえる。

4 手法に共通して高い精度が得られた作品として、「原哲夫作品」「銀河英雄伝説」などがあげられる。これについて、各作品カテゴリに属する小分類の数を比較すると、「原哲夫作品」が 19 種類、「銀河英雄伝説」が 24 種類であり、大きな差はなかった。さらに、キャラクターの表情だけが異なるものや、全身・頭部だけの違い、一つのアスキーアートにキャラクターが複数、単独で描かれているといった違い、台詞の有無の違いなど、類似するアスキーアートが同一カテゴリに多量に含まれている点で共通している。いずれの場合においても、アスキーアートのサイズが異なっても、用いられる文字の種類は同じであるため、文字を特徴とした場合にも高い精度が得られたと考えられる。

一方で、精度が低かった作品について、画像特徴量、文字画像ベクトルを用いた手法は他の 2 手法と比較して平均して 2 倍以上の精度が得られている。画像特徴量から得られる形状特徴により作品の絵柄の特徴を上手くとらえることが出来た結果であると考えられる。これらのことから、全体的な精度が低くなった原因は、評価データに多く含まれているカテゴリの作品が、類似するアスキーアートが多いタイプであったため、文字頻度特徴が有利になった結果であると考えられる。

また、どの手法においても下位 10 件に含まれた「機動戦士ガンダム」のように、登場人物だけでなく、登場するメカ類が含まれる場合に、同じカテゴリとして分類することが困難であったと考えられる。これは、アスキーアートのように基本的に線のみで構成された絵の場合、まったく異なる対象(例えばメカ類と人間)を描いた作品に対し、作風や絵柄で共通点を見

表 6: 完全一致率 (文字頻度)

順位	作品名	精度	順位	作品名	精度
1	機動戦艦ナデシコ	83.59	35	Rozen Maiden	8.29
2	銀河英雄伝説	73.36	36	コードギアス	6.51
3	魔人探偵脳噛ネウロ	66.67	37	銀魂	6.45
4	恋姫無双	65.91	38	ストライクウィッチーズ	5.57
5	プリキュア	64.25	39	BLEACH	5.26
6	原哲夫作品	62.99	40	VOCALOID・UTAU	3.44
7	さよなら絶望先生	56.59	41	機動戦士ガンダム	3.43
8	ゾイド	55.36	42	女神転生シリーズ	3.42
9	ジョジョの奇妙な冒険	51.3	43	TYPE-MOON 作品	2.96
10	福本伸行作品	47.11	44	魔法少女まどか☆マギカ	2.27

表 9: 完全一致率 (文字画像)

順位	作品名	精度	順位	作品名	精度
1	機動戦艦ナデシコ	83.97	35	TYPE-MOON 作品	19.31
2	銀河英雄伝説	71.03	36	とある魔術の禁書目録	18
3	プリキュア	70.87	37	ストライクウィッチーズ	16.04
4	魔人探偵脳噛ネウロ	68.94	38	BLEACH	15.79
5	恋姫無双	67.05	39	化物語	12.12
6	原哲夫作品	66.14	40	Rozen Maiden	11.78
7	さよなら絶望先生	64.34	41	VOCALOID・UTAU	10.64
8	ゾイド	63.1	42	コードギアス	9.58
9	ドラえもん	50.36	43	女神転生シリーズ	4.79
10	ジョジョの奇妙な冒険	48.1	44	魔法少女まどか☆マギカ	4.25



図 2: 分類成功したアスキーアート (画像特徴)

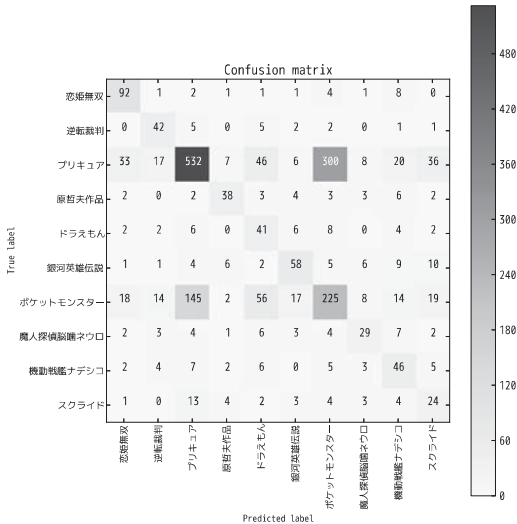


図 3: 完全一致率上位 10 件間の混同行列 (文字重要度)

出すことは人間であっても難しいことから容易に推測できる。アスキーアートのように、収集が比較的困難なものを訓練データとする場合には、人工的なノイズ付加によるデータの拡張や、線画からソースとなる画像を推測し、カテゴリ分類するなどといった工夫が必要と考えられる。

図 3 に、最も精度が低かった文字重要度による手法における完全一致率上位 10 件の作品間での混同行列、図 4 に、最も作品別精度平均が高かった画像特徴量による手法における完全一致率下位 10 件の作品間での混同行列を示す。図からわかる通り、文字重要度を用いた手法は、誤分類が多くみられる。とくに、「ポケットモンスター」、「プリキュア」、「ドラえもん」が混同されやすい傾向がみとれる。これらの作品はすべて子供向けアニメ作品であることが共通しているため、絵柄に類似点があったと考えられる。一方で、画像特徴量を用いた手法の場合、偏った誤分類傾向はあまり見られず、どの作品にも同程度に誤分類しやすい結果となった。

6. おわりに

本論文では、アスキーアートの分類手法として、文字特徴による手法と、画像特徴による手法を比較検討した。アスキーアートのカテゴリとして、そのアスキーアートのモチーフとなった漫画やアニメ、ゲーム作品の作品名またはシリーズ名を設定し、ニューラルネットワークを用いて分類器の学習をおこなった。実験の結果、文字単位の HOG 特徴量による文字画像ベクトルを FFNN により学習した分類器を用いた場合に全体的な成功率が最も高くなり、画像特徴量を CNN により学習した分類器を用いた場合に、作品ごとでみたときに高い精度（最

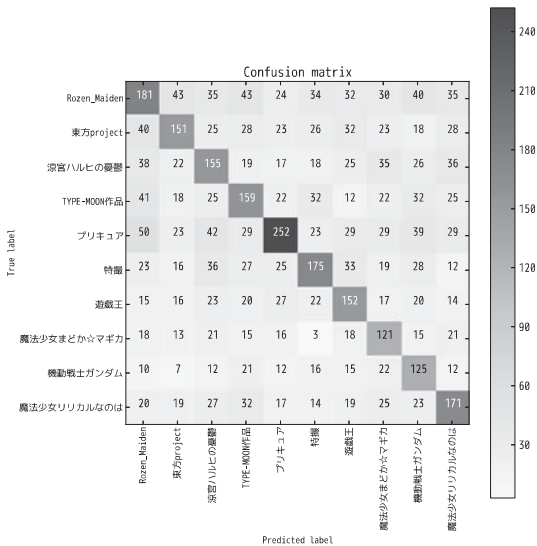


図 4: 完全一致率下位 10 件間の混同行列 (画像特徴)

大 92.63%) を得ることが出来た。
今後の課題として、文字頻度ベクトルを用いた手法では不利と考えられる異なる文字で同じキャラクターを表現しているような場合における画像特徴量の優位性を、評価実験をおこなうことで実証したい。また、作品による分類だけでなく、「デフォルメ」か「非デフォルメ」の分類や、モチーフとなっている物体のカテゴリ認識についても検討したい。さらに、写真やイラスト画像をエッジ画像とすることにより、非アスキーアートを訓練データとして用いてアスキーアートのカテゴリ分類器を作成し、その評価をおこないたい。

謝辞

本研究の一部は、科学研究費補助金 15K16077, 15K00425, 15K00309 によりおこなわれた。

参考文献

[Tanioka2005] 谷岡 広樹, 丸山 稔: 形態素解析に基づく SVM を用いたアスキーアートの識別, 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, 2005-02-25, Vol.104, No. 668, pp. 25-3, 2005.

[Hayashi2009] 林 和幸, 鈴木 徹也: テキスト圧縮を用いた言語に依存しないテキストアート抽出法, 研究報告デジタルドキュメント (DD), 情報処理学会, 2009-09-18, Vol.2009, No.3, pp.1-6, 2009.

[Suzuki2011] 鈴木 徹也: データ圧縮率を用いるテキストアート抽出法におけるテキストの正規化手法, 言語処理学会第 17 回年次大会発表論文集, pp.842-845, 2011.

[Fujisawa2016] 藤澤 日明, 松本 和幸, 吉田 稔, 北 研二: 画像特徴量を用いたアスキーアートからの顔文字検出, 人工知能学会全国大会, 3H4-OS-17b-1, 2016.

[Fujisawa2017] 藤澤 日明, 松本 和幸, 吉田 稔, 北 研二: 画像特徴量を用いた大型アスキーアートの分類手法の提案, 人工知能学会全国大会, 2O3-OS-22b-3, 2017.

[AAtoImage] <http://www.nicovideo.jp/watch/sm20296302>, 最終アクセス日 2016-3-28.