

マンガ画像の分類手法に関する検討

小城 凱^{1,a)} 伊東 栄典²

概要: 近年、海賊版サイトによるマンガ画像の公開が問題である。マンガは印刷された紙媒体としても発行されているため、マンガのページ画像のスキャンは簡単で、そのため DRM(デジタル権利管理) による海賊版対策は有効ではない。本研究では海賊版マンガの抑制を行うために、マンガのページ画像から、マンガのタイトルや巻数、話数を推定する分類手法を検討している。今回、分類手法の枠組みと、Manga109 データを用いた予備実験を行った。本発表では検討内容と予備実験の結果を報告する。

キーワード: 海賊版マンガ, フィルタリング, 画像認識, パターン認識, 機械学習

A study on classification method of manga Images

KAI KOJO^{1,a)} EISUKE ITO²

Abstract: Pirate edition of Manga is a serious problem. Manga is published as a printed paper book, it is easy to make digital image of manga pages by scanner, so DRM (Digital Rights Management) are not effective for prevention of pirate edition. In this study, in order to suppress pirated manga, we study some classification methods which estimates the title, number of volumes, and number of episodes of manga from manga page images. In this paper, we conducted a preliminary experiment using the framework of classification methods. We also report the details of the study and the results of preliminary experiments.

Keywords: Japanese manga, pirate edition, bootleg, filtering, pattern recognition, machine learning

1. はじめに

近年、海賊版サイトによるマンガ画像の公開が問題となっている。海賊版とは著作物を、著作権者や発行者に無断で複製し提供するものを言う。スキャナー機器やネットの普及に伴い、海賊版マンガがネット上で流通し始めた。2018年には海賊版サイトの「漫画村」が爆発的に利用された [1]。政府は2019年4月13日に知的財産戦略本部会合・犯罪対策閣僚会議を開き、ネット上の海賊版サイトに対する緊急対策を行い、日本国内のISPで侵害コンテンツの通信をブロック可能にした。しかしながら2023年1月現在でも海賊版サイトは存在している。

電子書籍ではDRM(デジタル権利管理)による利用制限が有効である。一方、海賊版の対策はDRMだけでは困難

である。印刷されたマンガのスキャンは簡単であるため、DRM範囲外からマンガのページ画像を作成できてしまう。

本研究では海賊版マンガの抑制を行うために、マンガのページ画像から、マンガのタイトルや巻数、話数を推定する分類手法を検討する。高精度のマンガ画像分類器が実現できれば、迷惑メールフィルタのように海賊版マンガのフィルタとして利用できる。海賊版サイトはクラウド事業者のサーバ・ストレージにデータを保持しているため、クラウド事業者による海賊版マンガ画像検索にも利用できる。

本論文では、マンガページ画像の分類器の構造を述べる。また分類器が用いるマンガの特徴をについても議論する。また手動で取得した小規模のマンガページ画像とManga109データを用いた試験的な分類システムによる予備実験の結果を述べる。

¹ 九州大学工学部電気情報工学科

² 九州大学情報基盤研究開発センター

^{a)} kojo.kai.183@s.kyushu-u.ac.jp

2. 用いたデータ

マンガ分類器が使うパターン認識や機械学習にはマンガの画像データが必要である。入手手段は、以下のものが挙げられる。

- (1) 紙マンガのスキャン
- (2) 電子マンガのスクリーンショット
- (3) 既存マンガデータセット利用

それぞれの利点、問題点を挙げる。紙マンガスキャンの長所は任意の漫画のデータの作成ができることで、短所は裁断の手間および劣化状態による画像品質にばらつきである。電子マンガからの画像取得の場合、利点は任意のマンガで高品質画像を取得可能なことで、短所は手作業の場合のデータ取得に手間である。既存マンガデータセット利用の場合、長所は容易に高品質データの入手可能なことで、短所はマンガの数が限定されていることである。

(1)の「紙マンガのスキャン」も試みた所、紙の傾きと劣化による色調のばらつきのため用いないこととした。(2)のスクリーンショットは、正規購入した電子書籍から手動でスクリーンショットを取得した。取得したマンガのタイトルおよびページ数を表1に示す。

表1 手動入手したマンガのページ画像

タイトル	略称	ページ数
チェンソーマン	CM	54
鋼の錬金術師	FA	58
アイシールド 21	ES	60

(3)の既存マンガデータセットとして Manga109[2][3] を用いた。Manga109 は、日本漫画のメディア処理の学術研究使用のために、東京大学 大学院情報理工学系研究科 電子情報学専攻 相澤・山崎・松井研究室によりとりまとめられたデータセットである。日本のプロ漫画家により描かれて出版された 109 作品（タイトル）について、1冊ずつ選ばれた計 109 冊のマンガ画像を保持する。109 冊のマンガは 1970 年代～2010 年代に出版された漫画で、幅広い対象読者層やジャンルを網羅している。

画像に加え 1 冊毎に作品のメタデータがあり、かつ全ページ画像に注釈 (annotation) が付与されている。メタデータには作品タイトル・作者名・出版社・出版年に加え、および登場する全キャラクターのリスト、各ページ（見開き）のリストも含む。各ページの注釈は、登場キャラクターの顔領域、キャラクターの全身領域、印字文字のテキスト、手書き文字のテキスト、コマ枠の領域情報を含む。メタデータと注釈は XML 形式で記述されている。アノテーションデータから欲しい対象物の座標データを取得し、画像データから該当部を取り出すことで任意のプロジェクトの切り抜き、保存が可能である。

3. マンガ画像分類器の構成

海賊版サイトでは、1 話分の連続したマンガページ画像が提供されることが多い。本研究で考えるマンガ画像分類器は、1 ページの画像、1 話のページ画像リスト、1 巻の画像リストを入力し、それらの画像（または画像リスト）から特徴量を抽出することで、分類や推定の処理する。

マンガページ画像が 1 枚与えられた場合の分類を考える。最も粒度の細かい分類としては、そのページ画像が、どのマンガタイトル（シリーズ）の何巻の何話の何ページかを特定するものである。しかしながら、マンガページ画像には微妙なズレ、欄外広告の有無などの影響があるため、タイトル・巻数・話数・ページまでの特定は困難かもしれない。最も粒度の粗い分類は、マンガのタイトル推定である。人間ならばマンガを見て「ドラゴンボール、ワンピース、ナルト」とタイトルを判別することは容易である。本研究の分類器でもタイトル推定は可能に思える。中間的な分類は、タイトルおよび巻数までの推定や、タイトル・何巻・何話までの推定になる。

次に、入力として 1 話の連続するページ画像からの分類を考える。分類器が推定するのは、タイトルのみ、タイトルと何巻か、タイトル・何巻・何話か、の 3 通りが考えられる。連続するページ数が多いと、条件付き確率で推定精度が向上すると予想される。

以上のを表 2 にまとめる。また想定するマンガ画像分類器の構成を図 1 に示す。

表 2 マンガ分類器の入出力

入力	1 ページ、1 話のページ画像、 1 巻のページ画像、1 タイトルの全ページ
分類器の出力	タイトル、巻、話、ページ

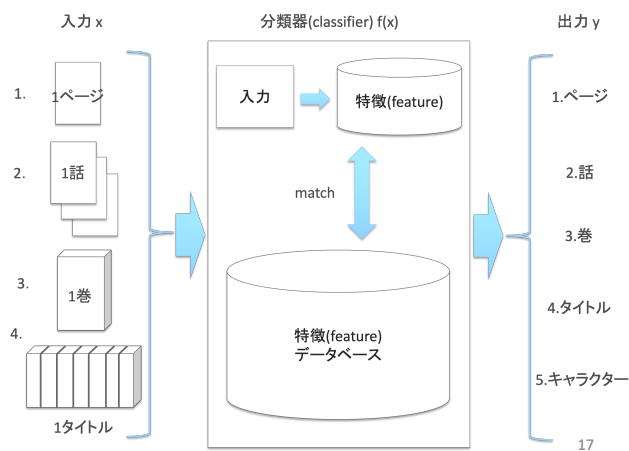


図 1 マンガ画像分類器の構成

4. マンガの特徴

画像分類に用いるマンガの特徴を分析する。

4.1 ページ, コマ, 吹き出し, 色

多くのマンガは, ページ (page), 話 (episode), 巻 (volume), 題名 (title) またはシリーズ (series) の構造を持つ。ページから話が, 話から巻, 巻からシリーズが構成される。本研究の対象は海賊版マンガでは1ページが1画像になっており, 話を単位として提供されている。本研究で対象とするマンガは, 1ページを最小単位とする。

ページの中に印刷された要素は, コマ, キャラクター, 吹き出しとテキストがある。コマの中には背景も描かれているし, コマの欄外には各種の情報 (ページ番号や広告など) が記載されている場合もある。

また, 漫画は基本的に白黒のみで構成されている。全てのページもしくは一部のページがカラーであったり, 灰色の濃淡が存在する場合もある。しかし, 基本的にそれらを白黒の2値化させてもマンガとしての情報が失われることはない。

4.2 キャラクター

マンガの読者に大きな影響・印象を与えるものは, 登場キャラクターである。我々がマンガを楽しむ場合, どのようなキャラクターが, いかなる状況で, どのような行動を取るのか, どのように感じているのか, を着目している。人間がマンガ画像の一部を見た場合, その画像が「どの作者の, 何というタイトルの, 誰というキャラクターなのか」を瞬時に判定している。画像から機械的にキャラクターを特定できれば, マンガ画像分類に有用である。

4.3 テキスト

吹き出し部分に, キャラクターのセリフが記述されている。キャラクターの心情を記述したモノログもある。擬態語や擬音語などもテキストで書かれている。擬態語や擬音語は, 作者の手書きで書かれることが多い。一方, 吹き出し内のセリフと, モノログは, 印刷書体で記載されることが多い。これらのテキスト情報も分類に有用である。

5. 白黒画素比によるページ推定

1つのマンガページ画像が, どのマンガのページかを当てる推定手法について, 各ページの画素の白黒画素数比 (= 白画素数/黒画素数) を用いる方法を考える。1ページのマンガ画像を $k \times k$ に分割し, 分割区域の白黒画素数を用いて数値化する。ここで k は, 3 から 49 までの奇数とした。各区域の数値化には以下の2つを用いる。

- (1) 区域の白黒画素数比
- (2) 前区域との現区域の白黒画素数比の差

白黒画素比率の並べ方を図2に示す。

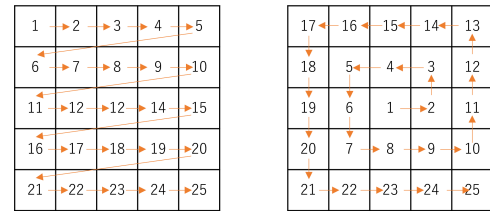


図2 区域の並べ方: 正順(左)と螺旋順(右)

表1に示したマンガについて, 電子書籍の手動スクリーンショット画像と, 海賊版サイトの画像を集めた。ページ数は表1に示した通り $54 + 58 + 60 = 172$ で, 正規版と海賊版が有るため2倍の344ページ有る。

各ページ画像を分割して正順で並べ, 各区域の白黒画素比率を算出してページを数値列化し, 次にページ画像 i, j の数値列のコサイン類似度 $\cos(i, j)$ を算出した。同一ページでは $\cos(i, i) = 1$ になり, $i \neq j$ では数値列が大きく異なる場合に $\cos(i, j)$ は0に近くなる。これらの全ペアのコサイン類似度を 344×344 の表に並べ, 値をヒートマップとして出力した。

出力されるヒートマップの概要を図3に示す。領域1,2,3,4に分かれる。1は海賊版画像の類似度, 4は正規版画像の類似度である。2,3は同じ図になり, 海賊版と正規版の類似度になる。領域2,3の対角線が白(高類似度)で, 上三角と下三角が黒(低類似度)であれば, 海賊版マンガの推定に良いことになる。

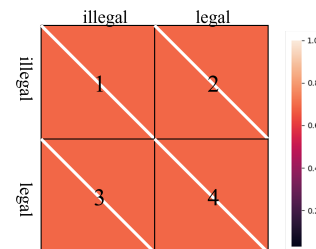


図3 ヒートマップ概要

5.1 正順で区域の白黒画素比

$k = 3, 5, 25, 49$ の場合のヒートマップを図4に示す。領域2の対角線が白く見えておらず, 海賊版の画像推定は出来ていない。

5.2 螺旋順で区域の白黒画素比の差

次に, 数値を前区画と現区画の白黒画素比の差にした。数値列の長さは $k^2 - 1$ になる。また順番は螺旋順とした。 $k = 3, 5, 25, 49$ の場合のコサイン類似度のヒートマップを図5に示す。領域2で対角線が白く, かつそれ以外が暗色となるのは $k = 25$ の場合である。

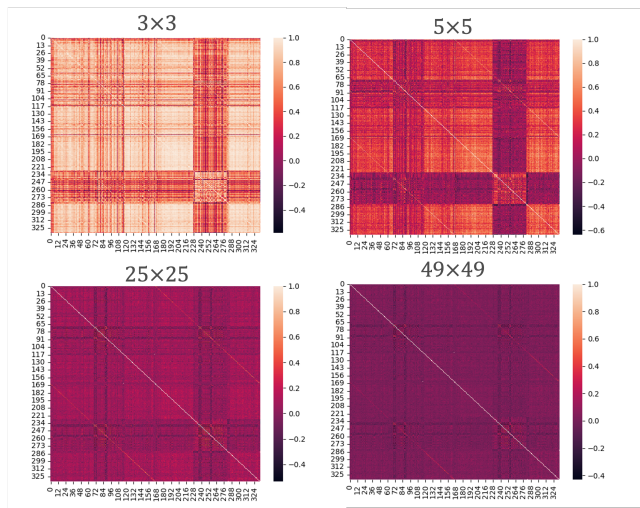


図 4 正順・白黒画素比

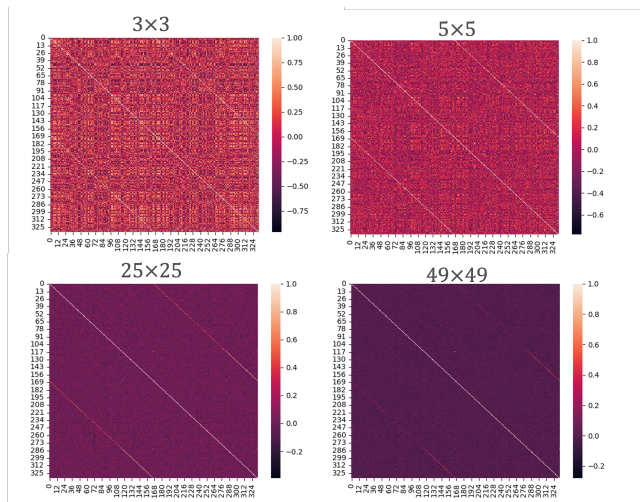


図 5 螺旋・白黒画素比の差

6. 機械学習によるキャラクター推定

4.2 節で述べた様に、マンガのタイトルを当てる場合、登場キャラクターは重要である。畳み込みニューラルネットワーク (CNN) による画像分類により、キャラクターの顔画像からキャラクター名を推定できるか試みた。

Manga109 にはキャラクター名 (ID) と、そのキャラクターの顔領域の情報がある。これを用いてキャラクター ID に対応する顔画像を抽出した。ただし解像度が低い (領域が小さい) 画像は学習や分類に適さないため、顔領域が 10000px 以上で、かつ画像数が 50 以上のものだけに限定する。その結果、学習および対象キャラクター数は 64、各作品の平均学習キャラクター数は 3.56、総学習画像枚数は 5742 枚、キャラクター毎の平均画像数は 89.72 となった。

これらの画像を VGG16[4] を用いた転移学習で学習した。VGG16 は 16 層 (プーリング層と Flatten 層はカウントしない) の畳み込みニューラルネットワークで、大規模な画像データセット ImageNet の学習を進めることで、1000 カテ

ゴリのマルチクラス分類を行うものである [5]。

10 交差検定による分類器精度 (accuracy) は 50% 以下で精度が悪い。顔画像の着目点 (アテンション) を図にしたものを 6 に示す。まず、マンガのキャラクターはデフォルトメが省略が多く、着目点が絞られていない。精度を高くするための改善が必要である。

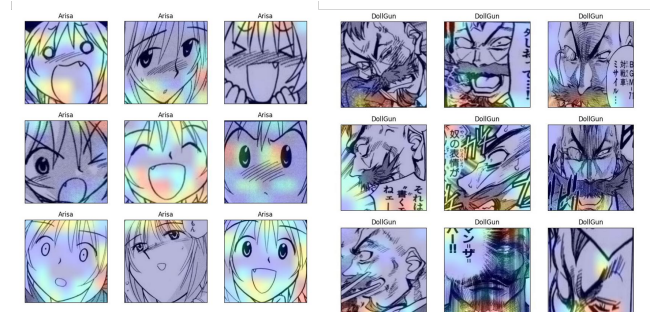


図 6 VGG16: キャラクターのアテンション部分

7. おわりに

本論文では海賊版マンガ対策に向け、パターン認識と機械学習によるマンガ画像分類手法について検討した。白黒画素比によるページ推定手法を提案し、マンガのページ画像に適用した。単純な手法であるにもかかわらず、小規模実験では高い精度で分類できた。次に VGG16 を用いて、顔部分のキャラクター名推定を試した。キャラクター顔画像の分類精度は高くなかった。

今後はデータを増やした場合での実験を行いたい。そのために出版社や著作者の協力を得たい。類似度や距離に別手法も適用した場合の比較や、画像のノイズ耐性も検討していきたい。

参考文献

- [1] 出版広報センター: 深刻な海賊版の被害 | 出版広報センター, <https://shuppankoho.jp/damage/>. (Accessed on 12/21/2022).
- [2] Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T. and Aizawa, K.: Sketch-based Manga Retrieval using Manga109 Dataset, *Multimedia Tools and Applications*, Vol. 76, No. 20, pp. 21811–21838 (online), DOI: 10.1007/s11042-016-4020-z (2017).
- [3] Aizawa, K., Fujimoto, A., Otsubo, A., Ogawa, T., Matsui, Y., Tsubota, K. and Ikuta, H.: Building a Manga Dataset “Manga109” with Annotations for Multimedia Applications, *IEEE MultiMedia*, Vol. 27, No. 2, pp. 8–18 (online), DOI: 10.1109/mmul.2020.2987895 (2020).
- [4] Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014).
- [5] チームカルポ: 物体画像認識と時系列データ処理入門 [TensorFlow2/PyTorch 対応第 2 版] NumPy/TensorFlow2(Keras)/PyTorch による実装ディープラーニング (単行本), 秀和システム (2021).