色彩特徴を入力に用いた畳み込みニューラルネットワークによる 漫画の自動彩色

廣芝 和之*† 大垣 慶介† 小田桐 優理† 古澤 知英*†

†株式会社ドワンゴ 〒 104-0061 東京都中央区銀座 4-12-15 歌舞伎座タワー E-mail: †{chie furusawa,kazuyuki hiroshiba,keisuke ogaki,yuri odagiri}@dwango.co.jp

あらまし 白黒漫画の着色には多大な労力が必要なため、着色工程の自動化が求められている、本稿では多層畳み込 みニューラルネットワークを用いた自動着色手法を提案する。白黒写真のような自然画像については、ニューラルネッ トワークを用いて色を復元する研究がなされてきたが、自然画像と白黒漫画への着色では異なる問題がある。自然画 像の着色と異なり、創作物である漫画中ではキャラクターの髪や眼などの色が多様であるため、形状からの色の推定 が難しい。この課題を解決するため、本手法ではヒストグラムやパレットといった色彩特徴を用いて創作者の意図し た色を反映させる。これにより、白黒漫画を鮮やかに着色できた。

キーワード 自動彩色, Colorization, 漫画, ディープラーニング, Deep Learning

1. はじめに

日本の漫画は白地の紙に黒のペンやインクで描かれているも のが多いが、最近は特に海外でカラー漫画が人気なため、白黒 漫画のカラー化が出版社や著者らから着目されている。そこで、 白黒漫画のカラー版を出版することにより、新たな読者から着 目されたり、より豊かな表現で読者を楽しませられることが期 待できる。

白黒漫画に着色する作業は多大な労力が必要なため、着色工 程の自動化が求められている。漫画の着色の自動化には、二段 階の工程を踏む研究がなされてきた。まず一段階目に、白黒漫 画の線を元に人や物の領域を推定し、続いて各領域に対して色 を推定する. Sato らは、線で区切られた領域を分割して隣接す る領域同士のグラフ構造を得たあと、グラフ構造を元にして参 照カラー画像の色を白黒画像に転写する手法を提案した[1]. し かしこの手法では、参照画像と白黒画像で隣接する領域が異な るとグラフ構造が破綻して転写ができない点が問題となる.

一方、自然画像の自動着色については、多層畳み込みニュー ラルネットワークによる着色手法[2]が提案された。この手法 は、領域分割などの前処理を行わずに End-to-end で色の推定 を学習させるため、領域分割の誤りが影響せず、 グラデーショ ンなどの領域内の塗り分けも可能となる。しかしこの手法を白 黒漫画に適用すると、漫画中のキャラクターの各パーツ (髪や 眼など) や背景の色が鮮やかではなく, 使用者の意図しない色 に着色される(図1). これは、創作物である漫画は自然画像 と異なり、キャラクターの髪や眼などの色の種類が実際の人間 に比べて豊富であるため、髪や眼の形が同じでも色が一意に定 めることが難しいためである.

本論文では, 色彩特徴を与えることで, キャラクターの各 パーツの色を定めにくい問題を解決する。使用者の意図の反映





(b) 目標画像







(ヒストグラム入力)

図 1 入力画像および目標画像と, 既存手法 [2] と提案手法の着色結果 の比較. 実際のカラー漫画画像 (b) および同一の入力画像 (a) に対する着色結果 (c),(d). 既存手法によるもの (c), 提案手 法によるもの (d). (c)浅月 舞

と作業量の多さのトレードオフを考慮し、2 つの色彩特徴を提 案する. パレットは, 各色の有無を示す二値化された色彩特徴 で、色の選択についてのみ調整ができる。 ヒストグラムは、各 色の面積の情報を含んだ色彩特徴で、色に加え、各色の度数を 調整できる。先程のニューラルネットワークを用いた学習手法 を白黒漫画に適用し、ネットワークの入力にいずれかの特徴を 追加することで、色彩特徴の反映を実現する

^{*}筆頭著者として同等に貢献した

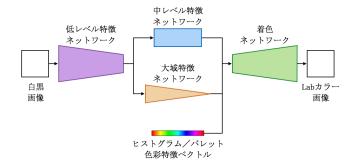


図 2 提案手法のニューラルネットワークの全体構造. 白黒画像を入力とし、ネットワークは、低レベル特徴ネットワーク、中レベル特徴ネットワーク、大域特徴ネットワーク、着色ネットワークの4つのネットワークにより構成される. ヒストグラムやパレットの色彩特徴ベクトルは、中レベル特徴ネットワークおよび大域特徴ネットワークと共に結合層で統合される.

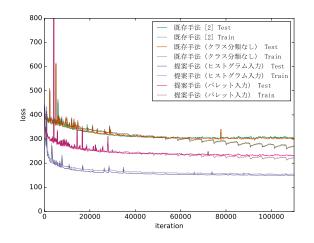


図 3 既存手法 [2] と提案手法の予測誤差を比較. Train は訓練誤差を, Test はテスト誤差を示す.

2. 手 法

本稿で提案するニューラルネットワークの全体構造を図2に示す.本章では、まずニューラルネットワークの構造について述べ、続いて本稿の提案手法の要であるヒストグラムおよびパレットの入力手法について述べる.

2.1 多層畳み込みニューラルネットワークを用いた自動着色

本論文では、自然画像の自動着色において性能の良い手法[2]のモデル構造を元にネットワークを構築する.[2]のモデルでは、画像の大域特徴と局所特徴を考慮した多層畳み込みネットワークを用いて白黒画像をカラー画像に自動着色する。このモデルは4つのネットワーク(低レベル特徴ネットワーク、中レベル特徴ネットワーク,大域特徴ネットワーク,着色ネットワーク)から構成されており、画像全体から抽出した大域特徴ベクトルと、より小さな画像領域から抽出した局所特徴ベクトルを結合層で一つのベクトルに統合し、着色ネットワークに入力する。また、大域特徴ベクトルを入力として、画像に対応するラベルを推測するクラス分類タスクを行う。

2.2 色彩特徴ベクトル入力のネットワーク構造

ヒストグラムやパレットの色彩特徴ベクトルは,他の特徴ベクトルと同様に結合層で統合したあと,着色ネットワークに入力する(図 2).既存手法の結合層では,28 × 28 × 256 次元の局所特徴ベクトルと 256 次元の大域特徴ベクトルが入力され,大域特徴ベクトルが 28 × 28 にブロードキャストされる形で局所特徴ベクトルに組み込まれ,各ピクセルの色推定に利用する.色彩特徴ベクトルも大域特徴ベクトルと同様に色推定に利用するため,28 × 28 にブロードキャストして局所特徴ベクトルに組み込む.これにより,色彩特徴ベクトルは,白黒画像の特徴ベクトルとは独立した状態でネットワークに入力することが可能で,なおかつ着色ネットワークに入力されるため着色画像生成の性能向上が期待できる.

2.3 ヒストグラム

カラー画像を RGB それぞれ 6 段階にした 6×6×6次元の3 変量ヒストグラムを 216次元ベクトルに並び替えて、ヒストグラム特徴ベクトルとして用いる。カラー画像はもともと RGB それぞれ 256 段階だが、入力の次元を小さくするために6 段階に減色する。また、学習結果が画像サイズに依存しないようにするために、ヒストグラムは全度数の合計値で除算して正規化する。ヒストグラムは面積の情報を含むため、度数を調整することで着色結果を調整できる。

2.4 パレット

ヒストグラムの他に、ヒストグラムを二値化したパレット特 徴ベクトルを用いた。3変量ヒストグラムの度数が0と等しい かで二値化したベクトルを、パレット特徴ベクトルとして用い た。パレットは着色に用いる色を選択するだけで良いため、ヒ ストグラムの度数を調整するよりも簡単に着色結果を調整で きる。

2.5 出力画像の色空間

既存手法のモデルは目標画像の輝度画像(Lab 色空間の L)を入力に想定しており、彩度画像(Lab 色空間の a と b)を入力白黒画像と結合した画像を出力するが、提案モデルは二値化された白黒画像を入力に想定しているため、Lab 色空間の L と a と b の 3 次元画像を出力する.

3. 実 験

3.1 データセット

対応関係のある白黒漫画とカラー漫画のデータセットが少ないため、学習データはニコニコ静画サービス (注1) に投稿されたイラスト画像を用いた。これらのイラスト画像に、LBP 特徴を用いたカスケード分類器による物体検出 (注2) を利用した顔検出を適用し、顔検出できた約 160000 枚の画像を顔の位置でトリミングして学習に用いた。これは着色推定を学習しやすくするためである。これにより、白黒漫画のキャラクターの着色に不要な背景や服のみの画像を省き、画像のスケールやキャラクターの大きさを揃えることができた。

(注1): ${\tt http://seiga.nicovideo.jp/}$

(注2): $https://github.com/nagadomi/lbpcascade_animeface$

着色結果





(a)入力画像



(b) 既存手法[2]

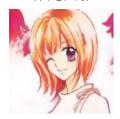




(c)参照画像

(d) 提案手法 ((c)を入力)





(e)参照画像

(f) 提案手法 ((e)を入力)





(g)参照画像

(h) 提案手法 ((g)を入力)

図 4 参照画像をヒストグラム入力に用いた提案手法による着色結果 と,目標画像や既存手法の比較.同一の入力画像に対する着色結 果. 既存手法によるもの(b). 各参照画像(c),(e),(g)を用 いて提案手法により着色したもの(d),(f),(h). ©浅月 舞 (a-d,f,h) , ⓒダイナミック 太郎 (e) , ⓒ南澤 久佳 (g)

3.1.1 入力白黒画像と入力ヒストグラム

学習モデルに入力する白黒画像は, カラー画像に画像処理を 施して作成した。まずカラー画像から Lab 色空間における輝度 を算出してグレースケール画像を得た. このグレースケール画 像を学習済みモデルを適用したい白黒漫画に寄せるために,大 津二値化[3]をグレースケール画像に適用し、二値化した白黒 画像を作成して学習の入力にした.

学習時に入力するヒストグラムは、目標とするカラー画像か ら算出した.

3.1.2 クラス分類のラベル

クラス分類に用いる画像に対応したラベルは、それぞれの画 像に描かれたキャラクターの名前などの 428 種類のタグ情報を





(a) 目標画像

(b) パレット入力による 着色結果((a)を参照)





(c)ヒストグラム 度数の変更なし

(d) ヒストグラム 度数を全て同値に変更





(e)ヒストグラム 度数をランダムに変更

(f)ヒストグラム 度数を調整

図 5 色彩特徴の計算方法ごとの着色結果の比較、目標画像をパレット 入力した着色結果 (b). 目標画像から得られたヒストグラム度 数を変更せずに入力し着色した結果 (c). 目標画像から得られ たヒストグラム度数を全て等しい値に変更し着色した結果 (d). 目標画像から得られたヒストグラム度数をランダムに変更し着 色した結果 (e). 目標画像から得られたヒストグラム度数を調 整し着色した結果 (f) ©伊藤 伸平 (a), ©浅月 舞 (b-f)

用い、これらのラベルを推測するマルチラベル分類問題を学習 タスクに与えた.

3.2 予測誤差の遷移

既存手法[2]で学習した場合と、ヒストグラムおよびパレッ トを追加で入力して学習した場合の予測誤差の遷移を比較した (図3). 学習データ・テストデータ共に、ヒストグラムおよび パレットを入力する提案手法の予測誤差が既存手法を下回った. また、パレットを入力するよりヒストグラムを入力した方が誤 差が小さいことがわかった。以降の結果は、100000 イテレー ションの学習済みモデルを利用して作成した.

3.3 着色結果

白黒画像の着色を既存手法で学習させたモデルと、ヒストグ ラムを入力する提案手法で学習させたモデルを利用してテスト 画像を着色した(図1). Manga109[4]のカラー表紙画像から 白黒画像を作成し、テスト画像として用いた。既存手法を用い て白黒漫画を着色する学習を行ったモデルを利用すると, 色鮮 やかさが欠けるような色合いになる点で目標の画像とは異なる 着色結果となった。一方、目標となる画像のヒストグラムを補 助入力として与えることで、肌の色、髪の色、背景の色がそれ ぞれ目標画像と同じように着色され、目標画像に近い画像が生



図 6 既存手法および目標画像を色彩特徴として用いて提案手法で漫画の着色結果の比較、既存手法によるもの (c). 目標画像のヒストグラムを用いて提案手法により着色したもの (d). 目標画像のパレットを用いて提案手法により着色したもの (e). ⑥長谷川 裕一 (1), ⑥ 桜野 みねね (2,4), ⑥新沢 基栄 (3), ⑥八神 健 (5), ⑥浅月 舞 (6)

成された. このことから、学習モデルにヒストグラムを入力する提案手法は、目標とする着色画像の生成に貢献することが分かった.

モデルに入力するヒストグラムやパレットを変更することで、同一の白黒画像を用いて異なる着色画像が生成されることが期待できる。目標画像とは異なる参照画像のヒストグラムやパレットを学習済みモデルに入力すると、目標画像と大きく異なり、更に参照画像の色合いと似た着色画像が生成された(図 4(f)).また、色合いが似ているだけでなく、参照画像の背景や髪の色を転写したような着色画像も生成された.一方で、色合いは似ているが、背景や髪の色が参照画像と異なる画像が生成される場合もあった(図 4(h)).

3.4 ヒストグラムの微調整

色を推定する学習が、各色の面積情報に強く依存していると 仮定すると、白黒画像に合わせて各色のヒストグラム度数を調 整することで、モデル使用者が意図するように着色結果を修正 できると考えた。そこで、パレットを入力する場合と度数を調 整したヒストグラムを入力する場合の着色画像を比較した(図 5).

ヒストグラムの調整を行わない場合,参照画像のヒストグラムをそのまま入力して生成した着色画像に比べて、パレットを入力して生成した着色画像の方が参照画像に近い着色結果が得られた。しかし、入力するヒストグラムを調整した場合は、ヒストグラムをそのまま入力した場合やパレットを入力した場合の着色画像よりも、目標画像に近い着色画像が得られることがわかった。

既存手法[2]と提案手法を用いて、Manga109のカラー表紙画像から作成した白黒画像を着色した結果を図6に、提案手法を用いて Manga109の白黒漫画を着色した結果を図7に示す.

3.5 学習の詳細

データセットの画像サイズは 256 × 256 ピクセルに統一した. 学習モデルの汎化性能を向上させるため, ランダムに 224 × 224 ピクセルに切り出した画像を学習に用いた. また, 50%の確率で画像を左右に反転させた. バッチサイズは 30 に固定した. 最適化手法は Adam [5] を用いた.

4. 考 察

4.1 ヒストグラムの頑健性に関して

ヒストグラムには各色の面積比の情報が含まれているため、 色に関する特徴量が多く、その結果、目標により近い着色画像 を生成することができた。しかし、実際の白黒漫画中のキャラ クターの髪や背景の面積はキャラクターの向きや状況によって 変化するため、ヒストグラムを用いて良い着色画像を得るには その都度度数を微調整する手間がかかってしまう。学習時に入 力するヒストグラムの度数をランダムに増減すれば、ヒストグ ラムの度数に対して頑健性を獲得し、学習済みモデルの使用時 の微調整の手間が緩和される可能性がある。また、データ増強 の役割を果たして精度が向上する可能性もある。

4.2 背景や髪の色が逆転して着色されることに関して

目標とした着色画像と近い着色結果になることがあったが、髪や背景などの領域の着色結果が意図したものと逆転した結果もいくつか見られた(図 4). この理由として、次の 3 つの可能性として考えた. (1) 学習に用いたデータセットに存在しないような領域と色の組み合わせの着色画像は生成されないような学習が行われることで、補助入力に関わらずデータセットに存在するような組み合わせの着色画像が生成されたため. (2) 入力白黒画像の白/黒に輝度の小さい/大きい色が塗られにくい学習が行われることで、例えば白黒画像の白背景を暗い色に着色した画像や、白黒画像の黒い髪に明るい色を着色した画像が生成されにくいため. (3) ヒストグラムを入力して学習すると、各領域の面積とヒストグラムの度数が対応するような学習が行われることで、各領域の面積やヒストグラムの度数の順序が目標と異なったため.

- (1) や(2) は、学習する画像データに関する課題であり、データ数を増やしたり、入力白黒画像の白黒を反転した画像も学習の入力に加えることによって解決できる可能性がある。しかし、白黒を反転すると黒地に白で線を描いたような画像となり、白地に黒で線を描く実際の白黒漫画画像と大きく異なるため、白黒反転した画像を用いることによって白黒漫画の着色結果が悪くなってしまう可能性も考えられる。
- (3) は、各領域に対して目標の色を明示的に指定する学習を行うことで、解決できる可能性がある。これは、一部が目標のカラー画像に着色された白黒画像を入力すれば良い。また、学習済みモデルを使用する際に画像の一部を着色することで、目標の色を指定するようなインターフェイスを構築できる可能性がある。更に、この手法を適用すれば、(1)や(2)も同時に解決できる可能性がある。

4.3 着色結果の色が滲んで見えることに関して

生成された画像を確認すると、背景や服の色が滲んでいるものがあった(図 7 (3-d)や (3-e)). これは、広範囲な背景などを同じような色で大雑把に着色する学習が優先され、塗りムラなどの細かい学習ができていないことが原因として考えられる. 細かい箇所も学習するように、例えば領域の境界周りだけ重みを増やしたり、あるいは Generative Adversarial Networks (GAN) [6] [7] を導入することで解決する可能性がある.

4.4 クラス分類による画像の大域特徴抽出に関して

今回用いた手法は既存手法 [2] に従い、大域特徴を入力として、画像に描かれたキャラクターの名前などのタグ情報を推測するクラス分類問題も学習タスクに加えた。しかし、このクラス分類を行わなかった場合と比べても、予測誤差の遷移や着色画像の生成結果の差はあまり見られなかった(図 3). クラス分類を行うためには、分類のためのラベルを用意し、ハイパーパラメータを増やす必要があるが、学習結果にあまり関係しないため、学習に必須では無いと思われる.

5. ま と め

本稿では、ヒストグラムやパレットといった色彩特徴を用いることで、創作者の意図した白黒漫画の色を推定する手法を提案した。実験により、提案手法は既存手法に比べて着色結果の予測誤差が小さくなり、使用者の意図に合うような鮮やかな着色結果を得られることが確かめられた。

ヒストグラムの各色の割合を入力する白黒画像に合わせて調整することで、より鮮やかな着色結果を得ることができた。一方で、パレットは使用する色を選択して簡単に作成できるため、大量の白黒漫画を着色する場合に有用だと考えられる。

入力する画像によっては、髪や背景などの領域の着色結果が 意図したものと逆転し、目標と大きく異なる着色結果になるこ とがあった。この課題は、各領域に対して目標の色を明示的に 指定する学習を行うことで解決できると考えており、今後はこ のような実験を行う予定である。

文 献

- Sato, K., Matsui, Y., Yamasaki, T., and Aizawa, K. "Reference-based manga colorization by graph correspondence using quadratic programming." SIGGRAPH Asia 2014 Technical Briefs, 2014.
- [2] Iizuka, S., Simo-Serra, E., and Ishikawa, H. "Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification." ACM Transactions on Graphics 35.4 (2016).
- [3] 大津展之. "判別および最小 2 乗規準に基づく自動しきい値選定法.", 電子情報通信学会論文誌 D 63.4 (1980).
- [4] Matsui, Y., Ito, K., Aramaki, Y., Yamasaki, T., and Aizawa, K. "Sketch-based Manga Retrieval using Manga109 Dataset.", arXiv preprint arXiv:1510.04389, 2015.
- [5] Kingma, D., and Ba, J. "Adam: A Method for Stochastic Optimization." The International Conference on Learning Representations, 2015.
- [6] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. "Generative adversarial nets." Advances in Neural Information Processing Systems, 2014.
- [7] Radford, A., Metz, L., and Chintala, S. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434, 2015.

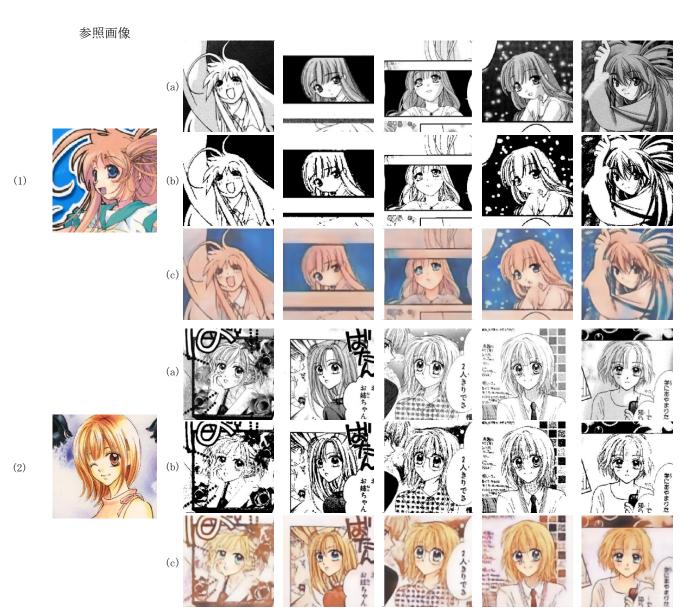


図 7 提案手法を用いて白黒漫画を着色した結果。グレースケールの画像 (a) を二値化した入力 画像 (b) と、参照画像 (1-3) のヒストグラムを入力して (b) を着色した結果 (c) . ⓒ 桜野 みねね (1) 、 ⓒ浅月 舞 (2)