

# 自己注意機構を用いたグレースケール画像の自動着色

下川 ゆふ<sup>†</sup> 黒崎 正行<sup>††</sup> 尾知 博<sup>††</sup>

† 九州工業大学大学院情報工学府

†† 九州工業大学大学院情報工学研究院

## 1. はじめに

近年、ユーザの介入を必要としない、畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)を用いた白黒画像の自動着色技術が多く研究されている[1][2]. これらの手法では、ある程度の自然なカラー画像を手軽に得ることができるが、その画像の多くは着色領域に誤りがあったり、全体的に色が褪せていたりするなど、未だ多くの課題を残している. 本研究では、既存手法を基に、正しい着色領域でより鮮やかな着色を行うモデルの実現を目的とする.

## 2. 提案モデルの概要

本研究で提案する自動着色モデルを図1に示す. このモデルは、[1]のモデルをベースとし、6つのネットワークで構成されている. 低レベル特徴ネットワークでダウンサンプリングとパッチ単位ごとの self-attention[3]が繰り返された後、大域特徴ネットワークで特徴ベクトル、中レベル特徴ネットワークで特徴マップがそれぞれ抽出される. 結合ネットワークではこの特徴ベクトルと特徴マップ結合することで、大域特徴と局所特徴を考慮した新たな特徴マップを生成する. 色付けネットワークでは、入力画像のサイズの1/2になるまでアップサンプリングと畳み込みが繰り返され、self-attention[4]の処理を行った後、[2]で提案されている多クラス分類を用いた色推定の結果が出力される. また、クラス分類ネットでは、入力画像のクラスラベルが出力される.

提案モデルの中間層と出力層の活性化関数には、それぞれ ReLU 関数と Softmax 関数を用いる. また、損失関数には交差エントロピー関数を用いる. このとき、色のクラス分類の損失には色の希少性に基づいた重み付けを行うことで、色の出現頻度による学習の偏りを抑制する. したがって、モデル全体の誤差 $L$ は式(1)で表される.

$$L = - \sum_i y_i^{class} \log \hat{y}_i^{class} - \sum_{h,w} v(y_{h,w}^{color}) \sum_c y_{h,w,c}^{color} \log \hat{y}_{h,w,c}^{color} \quad (1)$$

$y_i^{class}, y_{h,w,c}^{color}$  は正解データ,  $\hat{y}_i^{class}, \hat{y}_{h,w,c}^{color}$  は推定出力,  $v(y_{h,w})$  は、色の各クラスに掛かる重みを示す.

## 3. 提案モデルの学習

提案モデルの学習を行う. データセットには food-101 データセットを用い、75,750 枚の学習データの内、60,600 枚を学習データ、15,150 枚を検証データとする.

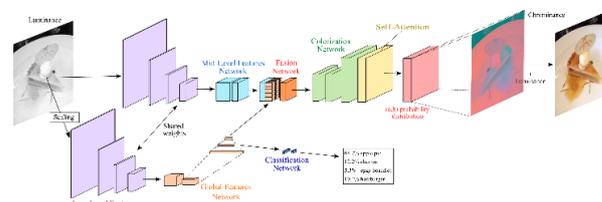


図1. 提案モデル



図2. シミュレーション結果

モデルの学習は、バッチサイズ 8 として、約  $2.0 \times 10^3$  イテレーション行う. 入力画像のサイズは  $224 \times 224$  とし、ランダムに回転や拡大・縮小の処理を画像に行う.

また、事前学習として、低レベル特徴ネットワーク、大域特徴ネットワーク、クラス分類ネットワークのみを用いたクラス分類タスクの学習をバッチサイズ 32 として、約  $2.0 \times 10^6$  イテレーション行う.

## 4. 結果

図2に学習後のモデルに画像を入力した際の結果の一部を示す. 図2を見ると、全体的に茶色に着色されており、色の学習が十分でないことが分かる. しかし、どの画像も食べ物を中心に着色されており、クラス分類の事前学習によりある程度領域の認識ができていているということが分かる.

## 5. まとめ

大域特徴と局所特徴を考慮した自動着色モデルの提案と実装を行った. シミュレーションより、提案モデルによって白黒画像からカラー画像が得られることと事前学習によって領域の学習が進んでいることが確認できた.

## 参考文献

- [1] S. Iizuka, E. S. Serra, and H. Ishikawa, "Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification," ACM Trans. Graph., vol.35, no.4, 2016
- [2] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," ECCV, pp.649-666, 2016
- [3] H. Zhao, J. Jia, and V. Koltun, "Exploring Self-attention for Image Recognition," CVPR, 2020
- [4] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-Attention Generative Adversarial Networks," PMLR, 2019