

## CNN を用いた近赤外線画像のセグメンテーションとカラリゼーション Semantic Segmentation and Colorization of Near Infrared Image using CNN

岩峰 晴也<sup>†1</sup> 三浦 大輝<sup>†1</sup> 黒木 修隆<sup>†1</sup> 廣瀬 哲也<sup>†2</sup> 沼 昌宏<sup>†1</sup>  
Haruya Iwamine Taiki Miura Nobutaka Kuroki Tetsuya Hirose Masahiro Numa

### 1. はじめに

近年、近赤外線カメラが夜間の監視目的で頻りに利用されている。しかし、そこから得られる画像は目に見えない近赤外線を疑似的に彩色したものであるため、人が知覚する通常のカラー画像と比べて違和感が大きいという問題点がある。

そこで、本稿では近赤外線画像の自然な着色を目的として、物体認識用のセグメンテーション技術と CNN を組み合わせた画像のカラリゼーション手法を提案する。

### 2. 先行研究

#### 2.1 概要

先行研究の 1 つに 3 層構造の CNN を用いた近赤外線画像のカラリゼーションに関する手法 [1] が提案されている。本来は近赤外線画像に色空間の概念はないが、カメラの出力は疑似的に着色された RGB 信号または YCbCr 信号である。このため、各成分を疑似 Y 成分、疑似 Cb 成分、疑似 Cr 成分と仮定する。近赤外線画像から得られる疑似 YCbCr 成分を 3 層 CNN に入力することで、着色された画像の YCbCr 成分が出力される。

#### 2.2 問題点

図 1 に従来手法を用いて近赤外線画像を着色した結果の一例を示す。従来手法では、服の色に対する肌の混入や影に青みを帯びているものが散見された。これは近赤外線画像のみの入力では CNN が適切に顔や服、およびそれらの境界線を認識できていないことが考えられる。



(a) カラー画像 (b) 近赤外線画像 (c) 処理画像

図 1 従来手法を用いて着色した画像の一例

<sup>†1</sup> 神戸大学大学院工学研究科,  
Graduate School of Engineering, Kobe University

<sup>†2</sup> 大阪大学大学院工学研究科,  
Graduate School of Engineering, Osaka University

表 1 3 層 CNN のパラメータ

	第 1 層	第 2 層	第 3 層
フィルタサイズ	5 × 5	5 × 5	3 × 3
特徴マップ数	256	128	3
学習係数	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>

### 3. 提案手法

図 2 に提案手法の処理の流れを示す。まず、voc-fcn8s [2] に疑似 RGB 成分を入力し、画像内の物体がピクセル単位でクラス分類されたセグメンテーション画像を得る。得られるセグメンテーション画像はラベル画像であるが、グレースケール画像に変換し 1ch で取り扱う。近赤外線画像の疑似 YCbCr 成分とセグメンテーション画像を合わせて 4ch とし、従来手法で用いた 3 層 CNN に入力する。最終的に、CNN から出力された YCbCr 成分を合成することで目的のカラー画像を得る。

### 4. 評価実験

#### 4.1 実験内容

本研究では、従来手法と提案手法を用いて近赤外線画像のカラリゼーションを行った。表 1 に 3 層 CNN のパラメータを示す。活性化関数は ReLU を使用した。

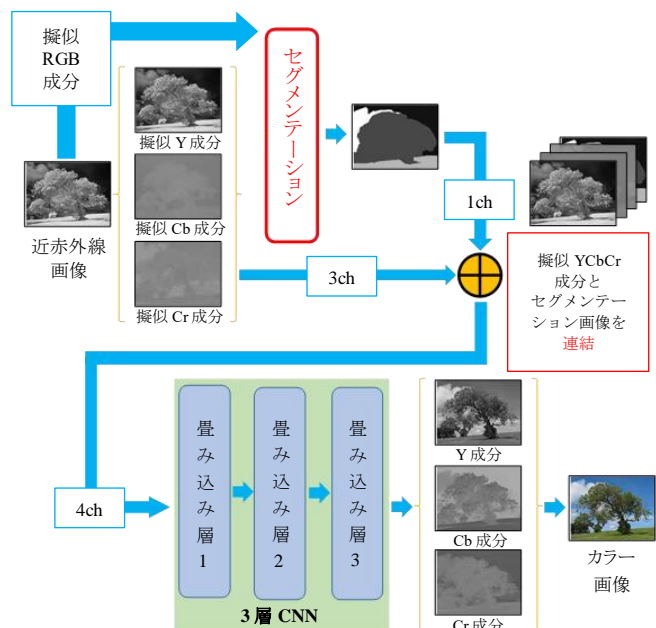


図 2 提案手法の処理の流れ

表 2 各手法で着色された評価用画像の PSNR

評価用画像	PSNR [dB]	
	従来手法	提案手法
(i)	23.04	<b>23.12</b>
(ii)	23.47	23.36
(iii)	23.46	<b>23.81</b>
(iv)	24.12	<b>24.90</b>
平均値	23.52	<b>23.80</b>

学習用画像は Sony の赤外線カメラ HDR-HC1 を使って上半身を赤外線モードとカラーモードで交互に撮影することで作成したデータセット 20 枚を用いる。また、評価用画像には学習用画像とは異なる 4 枚の赤外線画像を用いた。学習用画像及び評価用画像のサイズはいずれも  $640 \times 480$  pixel である。評価項目は、主観的評価と Peak Signal-to-Noise Ratio (PSNR) とする。PSNR とは元画像との類似度を示す客観的な指標であり、以下の式で求められる。

$$\text{PSNR} = 20 \log_{10} \frac{255}{\sqrt{\text{MSE}}} \quad (1)$$

ここで、幅  $m$ 、高さ  $n$  の出力画像を  $I$ 、理想画像を  $R$  とすると、MSE (平均二乗誤差) は、

$$\text{MSE} = \frac{1}{3} \times \frac{1}{mn} \sum_{l \in \{R, G, B\}} \sum_{i=1}^m \sum_{j=1}^n [I(i, j, l) - R(i, j, l)]^2 \quad (2)$$

と表される。

## 4.2 実験結果と考察

図 3 に実験結果の一例を示す。従来手法、提案手法における処理結果とカラー画像を比較すると、従来手法よりも服の色に対する肌色の混入量、影に対する青みを軽減できていることが分かる。これは、人間の輪郭を入力に加えることで、人間に特化した推論を行ったことが要因であると考えられる。

次に、表 2 に着色した評価用画像の PSNR を示す。提案手法では従来手法と比べて、平均で 0.28dB の向上が見られた。従来手法では画像全体で大きな面積を占めるカーテンを重視した学習を行ったことが原因で PSNR が低下したと考えられる。それに対して、提案手法では人間と背景を分けてより適切な学習ができたと考えられる。

## 5. まとめ

本稿では、赤外線画像に対して、より自然な着色を目的として、セグメンテーションを導入した CNN による赤外線画像のカラー化手法を提案した。評価実験の結果、従来手法と比較して平均で 0.28 dB、最高で 0.78 dB の PSNR 値の向上が見られ、服の色に対する肌色の混入や影の青みの軽減が確認できた。今後は屋外のより複雑な背景の元で提案手法の実用性を検証する必要がある。

### 参考文献

- [1] 三浦 大輝, 黒木修隆, 廣瀬哲也, 沼昌宏, “畳み込みニューラルネットワークを用いた赤外線画像のカラー化”, 第 16 回情報科学技術フォーラム (FIT), 2017.
- [2] L. Xu, J.S. Ren, C. Liu, J. Jia, “Deep Convolutional Neural Network for Image Deconvolution”, *Neural Information and Processing Systems (NIPS)*, 2014.



図 3 処理結果