

CNNを用いた画像鮮明化システムの実装 Image enhancement system using CNN

大極 康平[†] 林 亮佑[†] 河野 浩之[†]
Kohei Daikoku Ryosuke Hayashi Hiroyuki Kawano

1. 画像鮮明化が必要とされる場面

現在、製品検査や監視カメラによる安全確保、ITS交通システムにおける道路管理などで画像が用いられる。例えば、表1の監視カメラの画像による被疑者特定の端緒では、2020年に画像が刑法犯数の約10%において手がかりとなっている。

しかし、画像は天候や照明条件により常に鮮明ではないので手がかりとして使用が困難な場合がある。そこで画像鮮明化システムを使用することで手がかりとして使用できる画像を増やすことが可能である。同様に、製品検査などでも画像が不鮮明なことによる問題を画像鮮明化システムにより解決できる。

表1 監視カメラの画像による被疑者特定の端緒[1]

| 年 | 刑法犯数 | 画像が手がかりになった件数 |
|------|---------|---------------|
| 2018 | 36,963件 | 1,401件 |
| 2019 | 33,779件 | 1,531件 |
| 2020 | 33,009件 | 3,493件 |

2. 画像鮮明化手法の先行研究

2.1で画像鮮明化の先行研究、2.2で先行研究の課題と低照度画像鮮明化手法の提案、2.3で低照度補正ネットワーク、2.4で色改善ネットワークについて述べる。

2.1 先行研究

表2 画像鮮明化手法の先行研究

| 著者 | 研究内容 |
|-----------|---|
| Shenら[2] | CNNを使用して低照度画像と通常照度画像の対応関係を学習するMSR-netを提案 |
| Huangら[3] | 反射率マップを補正画像として推定する、Attentionメカニズムを導入しているネットワークを提案 |
| Maeda[4] | GANによる学習にペア画像を必要としない超解像手法を提案 |

画像鮮明化の先行研究を表2に示す。画像鮮明化手法にはShenらやHuangらのような画像の照度を改善する研究の他にMaedaの研究のように画像の超解像を行う研究も行われている。本研究では、これらの画像鮮明化手法の中でも画像の照度改善に焦点を当てる。

2.2 先行研究の課題と低照度鮮明化手法の提案

ShenらのMSR-netでは晴れた空のような非常に滑らかな領域ではハロー効果の影響を受けるといった課題があり、

[†] 南山大学 Nanzan University

Huangらのネットワークでは出力画像に影が見られるなど、出力画像の全体的な明るさが最適な状態に達していない課題がある。つまり、先行研究の提案手法は、画像の色を適切に復元できない課題がある。本研究では、この問題を解決するために、暗い画像を明るくする低照度補正ネットワークと画像の色を適切に復元する色改善ネットワークを合わせた低照度改善ネットワークを提案する。提案手法では、色改善ネットワークを組み合わせることにより、先行研究の問題であった色の問題を解決することを狙う。

図1に、提案する低照度鮮明化ネットワークを示す。はじめに低照度画像を低照度補正ネットワークに入力し、低照度画像を明るくした画像である低照度補正画像を取得する。次に取得した低照度補正結果画像を色改善ネットワークに入力し、低照度補正結果画像を色改善した結果である色改善結果画像を取得し、これを最終出力とする。Pythonの深層学習フレームワークを使用してネットワークを構築し、入力画像に低照度画像、正解画像に通常照度画像を使用して学習を行う。

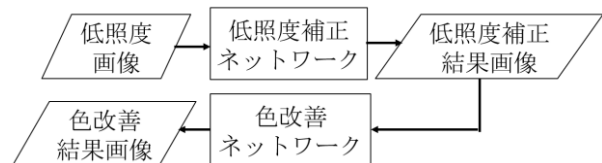


図1 提案する低照度鮮明化ネットワーク

2.3 低照度補正ネットワーク

低照度補正ネットワークを図2に示す。

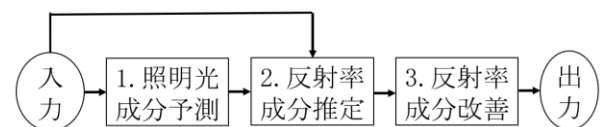


図2 低照度補正ネットワーク

低照度補正ネットワークは、1. 照明光成分予測、2. 反射率成分推定、3. 反射率成分改善の3ステップで低照度画像を明るくする。照明光成分予測では、畳み込みオートエンコーダによって、入力された低照度画像から照明光成分を予測する。反射率成分推定では、レティネックス理論[5]に基づいて、予測した照明光成分と入力された低照度画像から反射率成分を推定する。反射率改善成分では、複数の畳み込みにより、推定した反射率成分をより明るくする。

2.4 色改善ネットワーク

色改善ネットワークを図3に示す。色改善ネットワークは、畳み込みの回数を多くすることで深いネットワークを実現し、損失関数にコサイン類似度や色差などの画像間の色の類似に関する指標を使用し学習させることで色改善

を行う。しかし、ただ畳み込み層を増やしてネットワークを作成するだけでは、勾配消失問題が発生してしまう。よって、色改善ネットワークにおいて深いネットワークを実現させる際にはスキップ接続を使用することで勾配消失問題を回避する。

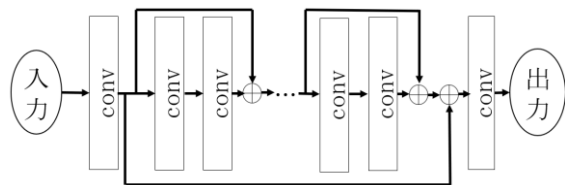


図 3 色改善ネットワーク

3. 実験環境

3.1でフレームワークの比較と選定, 3.2で使用するデータセットについて述べる。

3.1 フレームワークの比較と選定

本研究ではネットワークの構築にTensorflowとKerasを使用する。Tensorflow, Kerasの特徴を表3に示す。使用する理由として、Tensorflowは処理が高速であること、GPUでの処理を行うことができることが挙げられ、KerasはTensorflowと一緒に使用できることや高レベルのAPIが使用することでディープラーニングの実装が容易になることから以上の二つを使用する。

表 3 使用するフレームワーク

| フレームワーク | 特徴 |
|------------|--------------------------------------|
| Tensorflow | 高, 低レベルのAPIが使用可能 多くのプラットフォームで使用可能 |
| Keras | ディープラーニングを比較的短いソースコードで実装が可能 |

3.2 使用するデータセット

次に画像データセットの特徴を表4に示す。データセットにMIT-Adobe FiveKとLOL Dataset, SICEを使用する。これらのデータセットは、低照度画像と通常画像がペアになっている為、低照度画像の鮮明化を行うネットワークの学習に適している。ダウンロードした画像は、幅を150ピクセル、高さを100ピクセルにリサイズして使用する。

また、LOL DatasetとMIT-Adobe FiveK, SICEの画像だけでは6,074枚であり、ネットワーク学習に十分なデータ量に達しないためデータ拡張を行った。データ拡張の例を図4に示す。データ拡張では上下反転, 左右反転, 上下左右反転を通常画像と低照度画像に行い画像を生成することで、24,296枚にデータ量を増やし、24,000枚を訓練用データ、残りの296枚を検証用データとして使用する。



元の画像 左右反転 上下反転 上下左右反転
図 4 データ拡張の例

表 4 使用するデータセット

| データセット | 画像の枚数 |
|-----------------|-------------|
| MIT-Adobe FiveK | 5,000組のペア画像 |
| LOL Dataset | 485組のペア画像 |
| SICE | 589組のペア画像 |

4. 作成したネットワークの詳細

4.1で作成したネットワーク全体の説明, 4.2で畳み込みオートエンコーダを用いて低照度画像を明るくするネットワークの説明, 4.3で色改善ネットワークの説明, 4.4でネットワークの学習条件について述べる。

4.1 作成したネットワークの全体構造

作成したネットワークの全体図を図5に示す。作成したネットワークは、畳み込みオートエンコーダを用いた低照度画像を明るくするネットワークと残差ユニットを用いた色改善ネットワークで構成される。

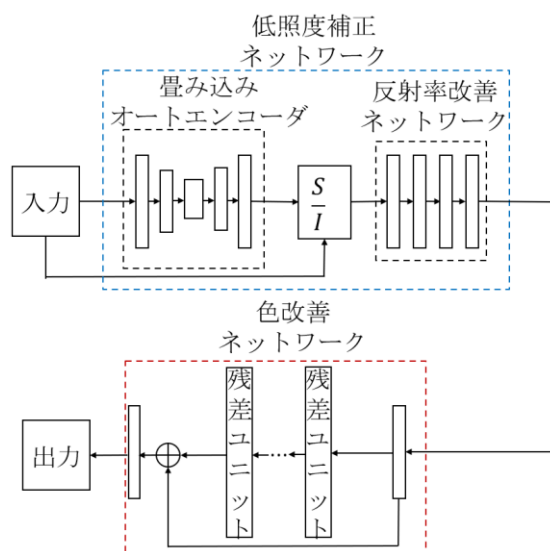


図 5 作成したネットワークの全体像

4.2 畳み込みオートエンコーダを用いた低照度画像を明るくするネットワーク

レディネックス理論によると、画像は照明光成分と反射率成分に分解することができる。照明光成分とは被写体に照射された照明光の成分であり、反射率成分とは照明に依存しない画像成分である。そして、レディネックス理論は次の式(1)のように定式化される。

$$S = I * R \quad (1)$$

S は画像を表し、 I は照明光成分、 R は反射率成分、 $*$ はピクセル単位の乗算を表す。

図6に、畳み込みオートエンコーダのネットワーク図を示す。この畳み込みオートエンコーダは、畳み込み層と最大プーリング層でエンコードをし、転置畳み込み層でデコードをするネットワークである。エンコードで画像のデータを圧縮し、デコードで画像を復元することでぼやけた画像を取得する。レディネックス理論では、照明光成分は低

周波特性を持つ性質があり、ぼやけた画像は画像全体にわたって低周波特性を持っているので照明光成分として使用する。そして畳み込みオートエンコーダのフィルタ数(A, B, C)を決定するために複数のパラメータで比較を行った。フィルタ数を(A, B, C)=(32, 64, 32)とした。

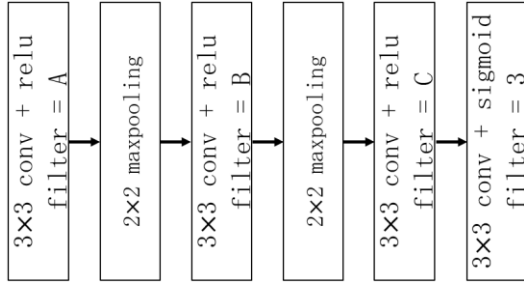


図 6 畳み込みオートエンコーダ

反射率成分を求める処理を行うソースコードの一部を、ソースコード1に示す。このコードは、入力画像とオートエンコーダで取得した照明光成分を使い、式(1)を計算することで反射率成分を取得している。入力画像と照明光成分の値の範囲は0~1であるが、除算を行ったときに0除算を行ってしまう可能性があるため、除算をして反射率成分を求める前に入力画像と照明光成分の値の範囲を1~256へと変換している。

ソースコード 1 反射率を求めるコード

```

1 def get_Reflectance(inputs):
2   #inputs[0]は入力画像, inputs[1]は照明光成分
3   img1 = 255 * inputs[0] + 1
4   img2 = 255 * inputs[1] + 1
5   res = img1 / img2
6   return res
    
```

図7に、反射率成分改善ネットワークのネットワーク図を示す。

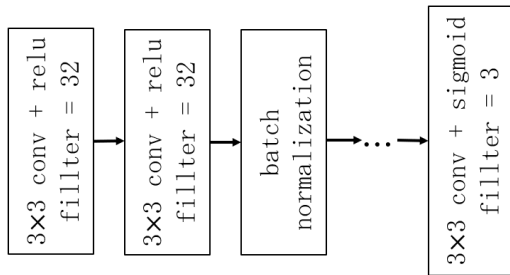


図 7 畳み込み層による反射率改善ネット

反射率成分改善ネットワークは、反射率成分に対し畳み込み層を用いることでより明るい画像を取得している。活性化関数はReLU関数を使用し、最後の層のみSigmoid関数を使用する。さらにこのネットワークでは2層の畳み込み層と1回のバッチ正規化が1セットとする。そのため、セット数を変更した時の性能を比較し、セット数を3に設定した。

4.3 色改善ネットワーク

図8に、色改善ネットワークのネットワーク図を示す。この色改善ネットワークでは残差ユニットと呼ばれるネッ

トワークを使用して深いネットワークを構築している。深いネットワークにして畳み込みの回数を増やすことで、色をより改善する。

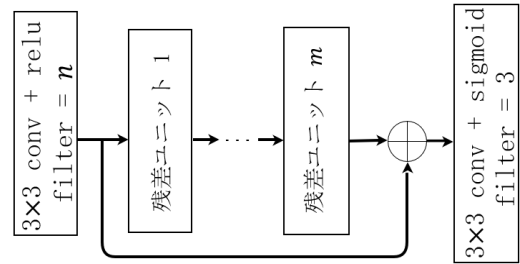


図 8 色改善ネットワーク

図9に、残差ユニットの構造を示す。この残差ユニットでは、ある層に与えられた信号を少し上位の層の出力に追加するスキップ接続を使用することで、層の最適な出力ではなく、層の入力と出力の差分を学習すればよくなるので、最適化がしやすくなる。これにより、深いネットワークを構築できるようになる。

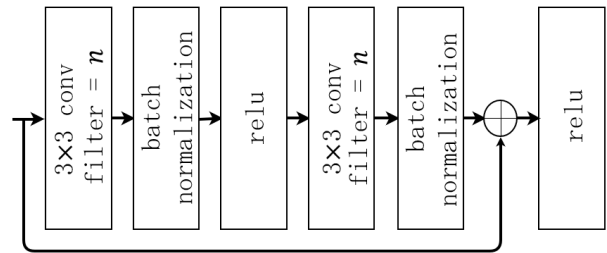


図 9 残差ユニット

加えて、色改善ネットワークの最適なパラメータを決定するために残差ユニットの個数、畳み込み層のフィルタ数を変更し比較をした。結果、フィルタ数nを8、残差ユニットの個数mを10とした。

4.4 学習条件

低照度補正ネットワークの学習条件として、最適化手法はAdam、学習率は0.001、エポック数は270、バッチサイズは32とした。そして、学習に使用する損失関数には、SSIMを損失関数にした式(2)のSSIM_{Loss}を使用する。

$$SSIM_{Loss}(\hat{x}, x) = 1 - \frac{(2\mu_{\hat{x}}\mu_x + C_1)(2\sigma_{\hat{x}x} + C_2)}{(\mu_{\hat{x}}^2 + \mu_x^2 + C_1)(\sigma_{\hat{x}}^2 + \sigma_x^2 + C_2)} \quad (2)$$

ここで、 \hat{x} は正解画像、 x はネットワークによる予測画像、 $\mu_{\hat{x}}$ 、 μ_x はそれぞれ \hat{x} 、 x の平均画素値、 $\sigma_{\hat{x}}$ 、 σ_x はそれぞれ \hat{x} 、 x の画素値の標準偏差、 $\sigma_{\hat{x}x}$ は共分散、 C_1 、 C_2 は定数を表す。SSIM_{Loss}では、正解画像と予測画像の構造が一致するほど値は0に近づき、一致しないほど値は1に近づく。

色改善ネットワークの学習条件として、最適化手法はAdam、学習率は0.001、エポック数は180、バッチサイズは32とした。そして、学習に使用する損失関数には、SSIM_{Loss}とCOS_{Loss}、ColorDiff_{Loss}、CIQI_{Loss}を合わせたColorImprove_{Loss}を使用した。ColorImprove_{Loss}では、画

像のSSIMを損なわず、色改善を行うことを狙う。 $ColorImproveLoss$ の式を式(3)、 COS_{Loss} の式を式(4)、 $ColorDiffLoss$ の式を式(5)、 $CIQILoss$ の式を式(6)に示す。

$$ColorImproveLoss(\hat{x}, x) = 4 * SSIM_{Loss}(\hat{x}, x) + 3 * COS_{Loss}(\hat{x}, x) + 2 * ColorDiffLoss(\hat{x}, x) + CIQILoss(x) \quad (3)$$

$$COS_{Loss}(\hat{x}, x) = -\frac{(\hat{x}, x)}{|\hat{x}||x|} \quad (4)$$

$$ColorDiffLoss(\hat{x}, x) = \frac{1}{H} \frac{1}{W} \sum_{i=1}^H \sum_{j=1}^W \{(\hat{R}_{ij} - R_{ij})^2 + (\hat{G}_{ij} - G_{ij})^2 + (\hat{B}_{ij} - B_{ij})^2\} \quad (5)$$

$$CIQILoss(x) = -\left(\sqrt{\sigma_\alpha^2 + \sigma_\beta^2} + 0.3 * \sqrt{\mu_\alpha^2 + \mu_\beta^2}\right) \quad (6)$$

ここで、 \hat{x} は正解画像、 x はネットワークによる予測画像を表す。 H は画像の高さ、 W は画像の幅、 \hat{R}_{ij} 、 \hat{B}_{ij} 、 \hat{G}_{ij} は正解画像 \hat{x} の*i*行*j*列番目のピクセルの各RGB値を表し、 R_{ij} 、 B_{ij} 、 G_{ij} はネットワークによる予測画像 x の*i*行*j*列番目のピクセルの各RGB値を表す。 σ_α 、 σ_β 、 μ_α 、 μ_β はRGB画像の各カラーチャンネル(R 、 G 、 B)を用いて $\alpha = R - G$ 、 $\beta = 0.5(R + G) - B$ としたとき、 σ_α 、 σ_β は α 、 β の各分散、 μ_α 、 μ_β は α 、 β の各平均を表す。

$CIQILoss$ は-1に近いほど正解画像の色と類似していることを示す。 $ColorDiffLoss$ は値が小さいほど正解画像の色と差が無いことを示す。 $CIQILoss$ は値が大きいほど画像が鮮やかであることを示す。

5. 色改善ネットワークの有無による性能比較

低照度補正ネットワーク(以下低照度補正netとする)と提案するネットワーク(以下色改善netありとする)を比較することで色改善ネットワークにより画像の色が改善されているのか確認する。比較では、両ネットワークに低照度画像を入力し、その出力画像と正解画像を用いてPSNR、SSIM、コサイン類似度、色差(LAB)、CIQIで計測を行うことで出力画像が正解画像とどの程度類似しているか確認する。PSNRは値が大きいほど画質の劣化が少ないことを示す。SSIMは値が1に近いほど画像の構造が類似していることを示す。コサイン類似度は値が1に近いほど画像の色が類似していることを示す。色差は値が小さいほど画像の色が類似していることを示す。CIQIは値が高いほど画像の色が鮮やかであることを示す。

ここで使用する正解画像はGoogle検索で入手した画像28枚と現実世界で撮影した通常照度画像2枚であり、低照度画像はGoogle検索で入手した画像に対してガンマ補正を行った画像28枚と現実世界で撮影した低照度画像2枚である。

評価値ごとの平均値を表5、出力画像の一部を図10に示す。表より色改善ネットワークは低照度補正ネットワークと比較してPSNRは1.04%、SSIMは0.31%、コサイン類似度は0.15%、色差は0.15%、CIQIは23.4%向上している。コサイン類似度、色差、CIQIは色の評価値であり、これらすべてが向上していることから色改善ネットにより画像の色を改善していることが確認できた。また、出力画像では低

照度が改善されており、さらにShenらの課題であったハロー効果などの歪みも見られなかった。

表5 低照度画像 net と色改善 net あり

| 評価指標 | 低照度補正net | 色改善netあり |
|---------|----------|----------|
| PSNR | 16.7055 | 16.8817 |
| SSIM | 0.79640 | 0.79884 |
| コサイン類似度 | 0.97012 | 0.97157 |
| 色差(LAB) | 17.8805 | 17.8529 |
| CIQI | 11.9402 | 15.6003 |



正解画像

低照度画像



低照度補正net

色改善netあり

図10 低照度補正 net と色改善 net ありの出力画像例

6. 結び

低照度補正ネットワークと色改善ネットワークを組み合わせた低照度画像鮮明化システムを提案した。これにより、低照度画像を明るくし、色改善を行うことができた。さらに、出力画像ではハロー効果のような歪みがない画像を取得できた。

謝辞

本研究はJP17K00432の助成を受けた。

参考文献

- [1] 警視庁, “警視庁の統計,” https://www.keishicho.metro.tokyo.lg.jp/about_mpd/jokyo_tokei/tokei/, 参照 Oct. 2021.
- [2] Liang Shen, Zihan Yue, Fan Feng, Quan Chen, Shihao Liu, Jie Ma, “MSR-net: Low-light Image Enhancement Using Deep Convolutional Network,” arXiv:1711.02488v1 [cs.CV], Nov. 2017.
- [3] Wei Huang, Yifeng Zhu, Rui Huang, “Low Light Image Enhancement Network With Attention Mechanism and Retinex Model,” IEEE Access, Vol.8, pp. 74306-74314, Apr. 2020.
- [4] Shunta Maeda, “Unpaired Image Super-Resolution Using Pseudo-Supervision,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 291-300, 2020.
- [5] Edwin Land, John McCann, “Lightness and Retinex Theory,” Journal of the Optical Society of America, Vol.61, No. 1, pp. 1-11, Jan. 1971.