

GAN および Template Matching を用いた水滴除去手法の検討

有本麻梨香¹ 原潤一² 渡辺裕^{1,2}

概要：近年、画像の欠落した領域を自動修復する技術として、画像修復が注目を集めている。画像修復を施す対象として特にニーズが高まっているものが、窓やカメラレンズに付着して写真に写り込んだ水滴の除去である。撮影時の天候や水滴の背景等により除去対象となる水滴の形状や色は大きく異なる。そのため、従来の水滴除去手法では完全に水滴を検出・除去してクリアな視界を手に入れるまでの精度を現状では達成できていない。そこで、本研究では単一の静止画に対する水滴除去の精度向上を目指し、深層学習を用いた水滴除去手法である Attentive GAN に対し、水滴検出および水滴除去のそれぞれを改善した水滴除去手法を提案する。

キーワード：水滴除去, 画像修復

A Study on Raindrop Removal Method Using GAN and Template Matching

MARIKA ARIMOTO^{†1} JUNICHI HARA^{†2} HIROSHI WATANABE^{†1,2}

1. まえがき

近年、画像の欠落した領域を自動修復する技術として、画像修復が注目を集めている。画像修復を施す対象として特にニーズが高まっているものが、窓やカメラレンズに付着して写真に写り込んだ水滴の除去である。活用例としては、ドライブレコーダーや監視カメラに付着した雨滴の除去が期待されているが、画像処理のアプローチによる水滴除去は未だ実用化に至っていない。従来にも Attentive GAN [1]を始めとする様々な水滴除去手法が提案されてきた。しかし、撮影時の天候や水滴の背景等により除去対象となる水滴の形状や色は大きく異なり、完全に水滴を検出・除去してクリアな視界を手に入れるまでの精度を現状では達成できていない。そこで、本研究では単一の静止画に対する水滴除去の精度向上を目指し、深層学習を用いた水滴除去手法である Attentive GAN に対し、水滴検出および水滴除去のそれぞれを改善した水滴除去手法を提案する。

2. 関連技術

2.1 Attentive GAN

Attentive GAN とは、GAN(Generative Adversarial Network) をベースとした水滴除去モデルである。Attentive GAN の生成器は recurrent network により入力画像から水滴領域を検出し、Attention Map を生成する Attentive-Recurrent Network と、Attentive-Recurrent Network で検出された水滴領域を除去し、水滴のない画像を生成する Contextual Autoencoder の二つのサブネットワークから成る。識別器は、カーネルが (3,3) の七つの畳み込み層、1024 の全結合層、シグモイド活性化関数を持つ一つのニューロンから成る。図 1 に Attentive GAN の構造を示す。

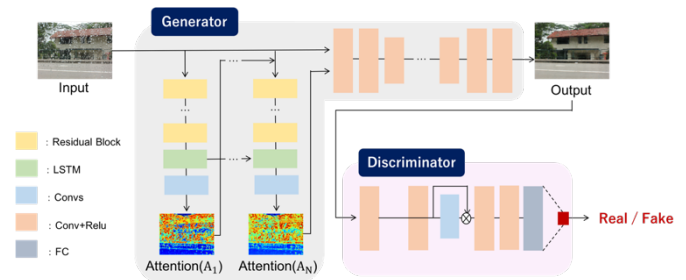


図 1 Attentive GAN の構造

2.2 DeepFill v2

DeepFill v2 [2]とは、風景や人物を対象とした GAN ベースの画像修復手法の一つである。DeepFill v2 の Generator は Gated Convolution と呼ばれる畳み込み層から構成される Encoder-Decoder ネットワークである。Gated Convolution では学習を通して動的に特徴量を選択し、マスクをアップデートしながら畳み込みを行う。これにより自由形状のマスクに対して、深層部でもマスク情報を失うことなく学習することが可能となる。Discriminator には、SN-PatchGAN が採用されている。SN-PatchGAN はカーネルサイズ 5、ストライド幅 2 の六つの畳み込み層から成り、各畳み込み層から出力された特徴マップに対して、入力された画像が本物か偽物かの判定を行う。また、各畳み込み層では、学習の安定化を目的とした spectral normalization を行なっている。図 2 に DeepFill v2 の構造を示す。

2.3 テンプレートマッチング

テンプレートマッチングとは、入力画像中からテンプレート画像と最も類似する領域を探索する処理である。類似度の計算方法の一つに ZNCC(零平均正規化相互相関)があ

¹ 早稲田大学大学院基幹理工学研究所

² 早稲田大学国際情報通信センター

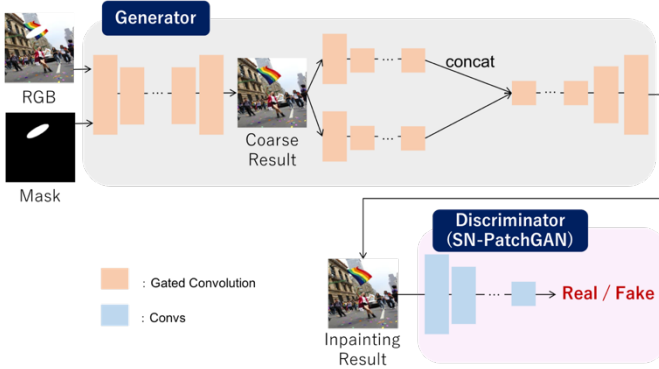


図 2 DeepFill v2 の構造

り, 最大値 1.0 に近いほど類似度が高いことを表す. ZNCC を式(1)に示す.

$$ZNCC(d_x, d_y) = \frac{\sum \sum [(I(d_x + x, d_y + y) - \mu_I)(T(x, y) - \mu_T)]}{\sqrt{\sum \sum [I(d_x + x, d_y + y) - \mu_I]^2} \sqrt{\sum \sum [T(x, y) - \mu_T]^2}} \quad (1)$$

ここで, $\sum \sum = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1}$ である. μ_I は入力画像の平均値, μ_T はテンプレート画像の平均値, w はテンプレート画像の幅, h はテンプレート画像の高さを表す. また, 入力画像の画素値が $I(x, y)$, テンプレート画像の画素値が $T(x, y)$, 走査位置が d_x, d_y である.

3. 提案手法

Attentive GAN には, 背景を反射して輝度変化のある水滴を検出しきれない問題がある. そこで, 背景の逆像を映した水滴を検出するテンプレートマッチングを Attentive GAN の Attentive-Recurrent Network に追加して水滴検出を行う手法を提案する.

また, 同一のマスク領域に対する除去精度は Contextual Autoencoder より DeepFill v2 が上回ることが予備実験で確認されている. そこで, Attentive GAN の Contextual Autoencoder を DeepFill v2 に変更して水滴除去を行う手法を提案する. 提案手法の構造を図 3 に示す.

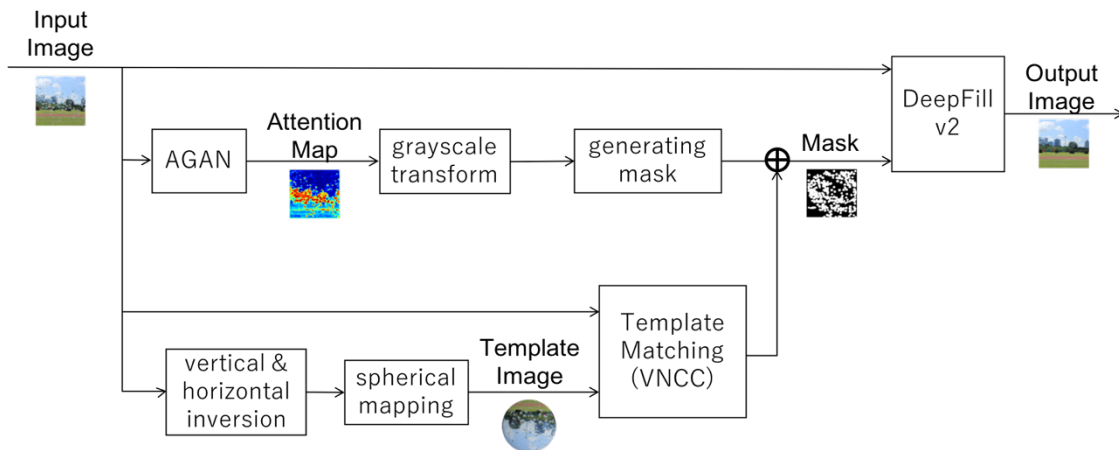


図 3 提案手法の構造

提案手法は検出した水滴領域のマスクを生成する水滴検出部と, 与えられたマスク領域を除去して周囲の画素値を参照しながら補完を行う水滴除去部から成り立つ.

3.1 水滴検出部

水滴検出部では, 水滴が付着した入力画像を Attentive GAN に入力し, Attention Map を生成する. Attention Map をグレースケール画像に変換し, 任意の閾値以上の RGB 値を抽出してマスク画像を生成する. 同時に, 入力画像を上下左右反転させ, 楕円面マッピングを行い, 背景の逆像を映した水滴を再現したテンプレートを作成する. 入力画像とテンプレート画像のテンプレートマッチングを行い, 検出された水滴領域がグレースケール変換された Attention Map において任意の閾値以上の RGB 値をとるピクセルを含む場合にのみ, 検出された領域が水滴領域であると判断する. テンプレートマッチングの検出領域が Attention Map から水滴領域であると判断した場合は Attention Map から生成されたマスク画像にマスク領域を追加し, そうでない場合はマスク領域を追加しない処理を行う. これにより, テンプレートマッチングで水滴の誤検出が起きても最終的な水滴検出領域には含まれないようにする. テンプレートマッチングで検出された領域ごとに以上の処理を繰り返すことにより, 最終的に水滴領域を示すマスク画像を生成する.

3.1.1 水滴テンプレート画像の作成

水滴テンプレートは入力画像を上下左右反転させ, 楕円面にマッピングして作成する. マッピングする楕円面の媒介変数表示は極座標系を用いて以下の式(2)で求まる.

$$\begin{aligned} x &= a \sin \theta \cos \varphi \\ y &= b \sin \theta \sin \varphi \\ z &= c \cos \theta \\ 0 &\leq \theta \leq \pi, 0 \leq \varphi \leq \pi \end{aligned} \quad (2)$$

一般的に, 水滴がガラスに付着した際の接触角は 20~30 度とされる[8]. そこで, 接触角 30 度の水滴と同様の形状とするため, マッピングする楕円面は $a = b = 1, c = 0.2679$

とする。z 軸方向から見た仰角を 0 度、x, y 軸方向の方位角を 90 度にして楕円面を表示した画像を水滴テンプレートとする。

3.1.2 テンプレートマッチングの類似度計算方法

入力画像とテンプレート画像間の類似度を計算する方法として ZNCC を用いる。これは、ZNCC が照明変化にロバストである特徴があり、水滴テンプレートと水滴付着画像間の類似度を計算するには適した方法だと考えられるためである。

3.2 水滴除去部

水滴付着画像と水滴検出部で生成されたマスク画像を DeepFill v2 に入力する。DeepFill v2 では、水滴付着画像におけるマスク領域を除去し、除去された領域の画素値を(R, G, B)=(255, 255, 255)に変換する。その後、マスク領域の周囲の画素値を参照しながら自動的にマスク領域の画素を補完する深層学習ネットワークを用いた処理を行う。これにより、水滴を除去した画像を生成する。

4. 実験

4.1 データセット

データセットには、水滴除去の評価に特化した Attentive GAN dataset のテスト用データセットである“test_a”を用いる。Attentive GAN dataset は、屋外の風景シーンが撮影された水滴なしの画像と同じ風景シーンに水滴が付着している画像のペアが含まれ、テスト用データセット“test_a”は 248 ペア用意されている。

4.2 提案手法による水滴除去実験

4.2.1 Attentive GAN による水滴検出およびマスク生成

Attentive GAN により水滴が付着した入力画像の水滴領域を検出した Attention Map を生成し、グレースケール画像に変換した。RGB の各値が任意の閾値以上であるピクセルは(R, G, B, A)=(255, 255, 255, 1)、閾値未満であるピクセルは(R, G, B, A)=(0, 0, 0, 0)の RGBA 画像に変換した。なお RGBA 画像において、透明度を表すアルファ値は 0~1 をとり、0 のとき完全に透明、1 のとき完全に不透明となる。実験では、閾値を 90 とした。

4.2.2 Attentive GAN による水滴検出およびマスク生成

背景の逆像を映した水滴を再現したテンプレートを作成した。3.1.1 節で述べたように、入力画像を上下左右反転させて楕円面にマッピングして得られた画像を水滴テンプレート画像とした。なお水滴テンプレートの画像サイズは 10×10 [pix]とし、水滴以外の領域のピクセルは(R, G, B, A)=(0, 0, 0, 0)の RGBA 画像とした。

4.2.3 テンプレートマッチングおよびマスク生成

入力画像と作成した水滴テンプレート画像のテンプレートマッチングを行った。実験では、ZNCC の閾値を 0.55 とした。また、テンプレートマッチングにより検出された領域が、4.2.2 節で作成されたマスク画像において(R, G, B,

A)=(255, 255, 255, 1)のピクセルを含む場合にのみ、検出された領域が水滴領域であると判断する。テンプレートマッチングの検出領域が Attention Map から水滴領域であると判断した場合は Attention Map から生成されたマスク画像にマスク領域を追加し、そうでない場合はマスク領域を追加しない処理を行う。図 4 に入力画像、テンプレートマッチングによる水滴検出結果、Attention Map とテンプレートマッチングの水滴検出結果から生成されたマスクを示す。

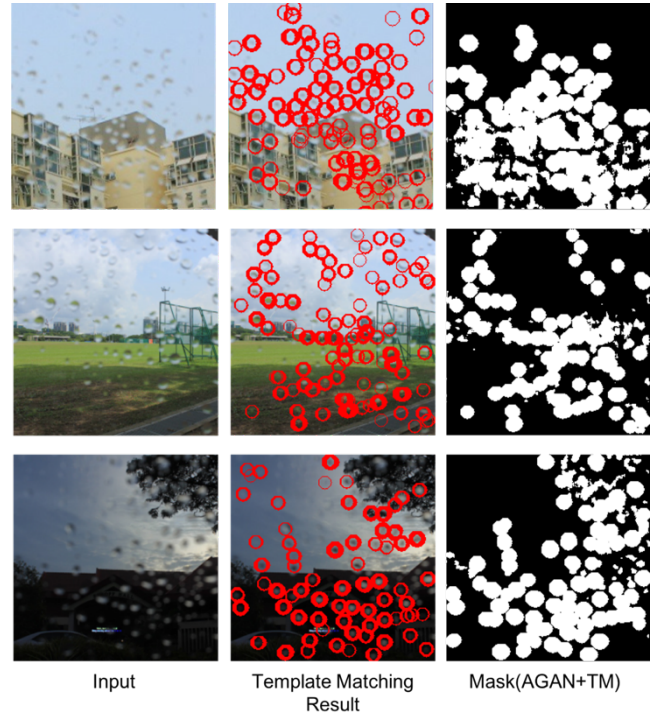


図 4 (左から) 入力画像、テンプレートマッチングによる水滴検出結果、Attention Map とテンプレートマッチングにより生成されたマスク画像

4.2.4 DeepFill v2 による水滴除去

水滴が付着した入力画像と 4.2.3 節で生成したマスクを DeepFill v2 に入力し、与えられたマスク領域に対して画像修復を行い、水滴除去結果を出力した。また、Attentive GAN から生成された Attention Map にテンプレートマッチングの水滴検出結果を追加することの有効性を確かめるために、Attention Map から生成したマスク画像単体を入力画像とともに DeepFill v2 に入力し、水滴除去結果を出力して比較した。図 5 に入力画像、Attentive GAN による水滴除去結果、Attentive GAN のみで水滴検出して DeepFill v2 で水滴除去した結果、Attentive GAN およびテンプレートマッチングにより水滴検出して DeepFill v2 で水滴除去した結果 (提案手法)、Ground Truth を示す。

4.3 評価実験

各手法の水滴除去結果を評価するため、水滴除去結果と水滴が付着していない Ground Truth 画像間の PSNR, SSIM の平均値をそれぞれ算出した。評価には、“test_a”からランダムに 30 枚を選択して得られた各手法の水滴除去結果を

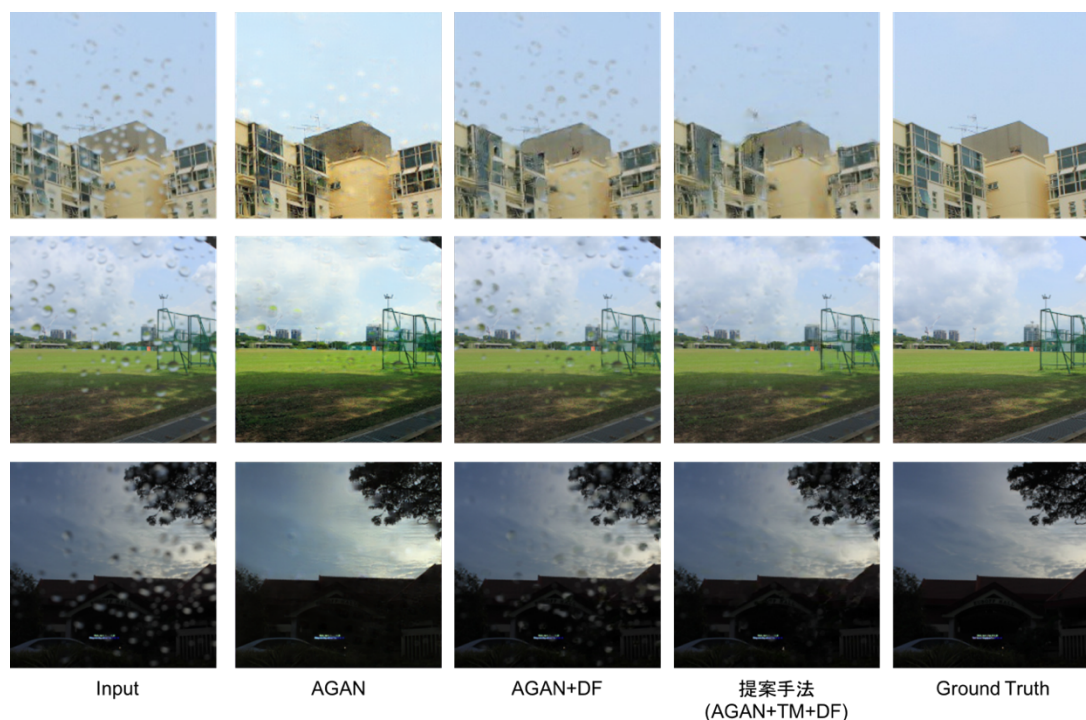


図5 (左から) 入力画像, Attentive GAN による水滴除去結果, Attentive GAN および DeepFill v2 による水滴除去結果, 提案手法による水滴除去結果, Ground Truth

用いた. 表1に Attentive GANによる水滴除去結果, Attentive GANのみで水滴検出して DeepFill v2で水滴除去した結果, 提案手法により水滴除去した結果と Ground Truth 間の PSNR, SSIM の平均値を表1に示す.

表1 AGAN dataset “test_a”の評価結果(30枚)

	PSNR (↑)	SSIM (↑)
AGAN	22.3011	0.8408
AGAN + DF	21.4744	0.8162
提案手法 (AGAN + TM + DF)	23.1711	0.9056

表1から, 提案手法は従来手法の Attentive GAN より PSNR, SSIM が向上し, 高い精度で水滴除去できることが確認できた.

また, Attentive GAN から生成された Attention Map のみをマスク化して DeepFill で水滴除去した結果は, Attentive GAN より精度が低くなった. Attentive GAN は Attention Map を元に入力画像全域のピクセルに対して RGB 値を再構築して水滴除去結果を出力する. そのため, 水滴の一部が検出できなかった領域も必ず再構築されることとなり, 水滴検出の精度の低さが結果に影響しづらいつと考えられる.

それに対して Attentive GAN と DeepFill v2 を組み合わせた手法の場合, Deep Fill v2 は二値のマスクで示されたマスク領域に対してのみ再構築を行うため, 水滴が検出されなかった領域に対しては再構築がされず, 一部検出できた領域に対して再構築を行ったとしても, 隣り合うピクセルの

影響を受けて水滴の RGB 値を再構築してしまうこととなる. Attentive GAN から出力された Attention Map から閾値以上の輝度を抽出して DeepFill v2 に入力する方法では水滴領域を完全にマスク化することができず, Attentive GAN より精度が低くなったと考えられる.

一方で, 提案手法では Attention Map の結果とテンプレートマッチングの結果を組み合わせることで最終的な水滴検出結果とすることにより, 水滴全域を検出できるようになったため, Attentive GAN よりも高い精度の水滴除去が可能になったと考えられる.

5. むすび

本研究では, 単一の画像に対する水滴除去手法である Attentive GAN の水滴検出にテンプレートマッチングを追加し, 水滴除去を Contextual Autoencoder から DeepFill v2 に変更した水滴除去手法を提案した. 実験の結果, 提案手法は Attentive GAN dataset “test_a”の30枚に対して PSNR, SSIM が Attentive GAN より向上し, 高い精度で水滴除去できることを確認した.

参考文献

- [1] R. Qian, R. T. Tan, W. Yang, J. Su, and J. Liu, “Attentive Generative Adversarial Network for Raindrop Removal from a Single Image,” Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp.2482-2491, May 2018.
- [2] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, “Free-Form Image Inpainting With Gated Convolution,” 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp.4470-4479, Oct. 2019.