

ディープラーニング画像処理技術/ アップスケーリング技術

ディープラーニング画像処理技術を使った
キヤノンの Neural network Image Processing Tool / Neural network Upscaling Tool の技術探求

目次

1. キヤノンのディープラーニング画像処理技術.....	3
1.1 「あるがままの情景」を細部まで再現.....	3
1.2 撮影の原理.....	3
1.3 写真の原理上避けられない課題.....	4
1.4 Neural network Image Processing Tool.....	5
1.5 ディープラーニング「学習」と「推論」.....	11
1.6 学習データ.....	12
1.7 撮影シミュレーション.....	13
2. アップスケーリング技術.....	14
2.1 従来のアップスケーリングの概要と課題.....	14
2.2 Neural network Upscaling Toolの概要.....	14
2.3 アップスケーリングの効果/ユースケース.....	14
3. キヤノンのアップスケーリングの特長/優位性.....	17
3.1 目標とした画質.....	17
3.2 目標画質の実現.....	18
3.3 従来处理との効果比較.....	19
4. 「あるがままの情景」を追求するキヤノンの写真画質.....	20

1. キヤノンのディープラーニング画像処理技術

1.1 「あるがままの情景」を細部まで再現

ある瞬間、ある場所の光景は、二度と訪れることはない。しかし、それはカメラで記録することができる。見たことのない絶景や、後で見返すと記憶がまざまざとよみがえる感動の瞬間など、カメラは素晴らしい瞬間を写真として残してくれる。本ホワイトペーパーでは、本来の「あるがままの情景」の写真を実現するためにキヤノンが開発した「ディープラーニング* 画像処理技術」について説明する。

* ディープラーニング

人間の脳の神経回路を模した数学モデルである「ニューラルネットワーク」を用いて、大量のデータをコンピューターに学ばせ、コンピューター自身が学習してデータから特徴を導き出し、望ましい推測・判断などの結果を得られるようにする手法

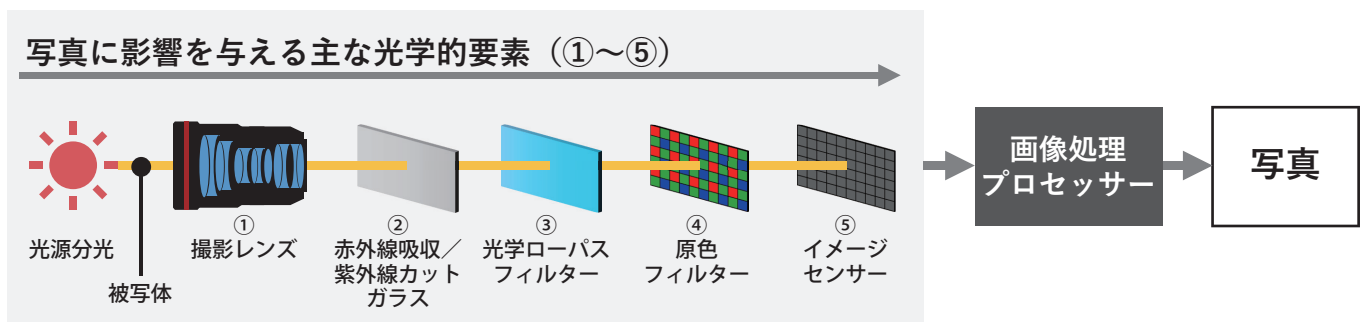
1.2 撮影の原理

はじめにカメラの撮影原理について説明する。下図は、被写体がカメラを通して写真になるまでに、画質に影響を与える主な要素を説明している。

太陽や蛍光灯、LED 光源などにより照明された被写体からの光を、「①撮影レンズ」で集光して像をつくり、「⑤イメージセンサー」で光を電気信号に変換する。その際、紫外線や赤外線といった、人間の眼には見えない波長の光をカットする「②赤外線吸収／紫外線カットガラス」や、モアレを抑制する「③光学ローパスフィルター」、カラー画像を得るための「④原色フィルター」を透過して、RGB（赤 / 緑 / 青）に分離する。

イメージセンサーの画素ごとに割り振られた色（RGB（赤 / 緑 / 青））ごとの光の強度情報が、電気信号として画像処理チップ（映像エンジン）に送信され、さまざまな画像処理で構成される現像処理を行うことで、写真に仕上げられる。また、カメラの設定で RAW データを保存している場合、カメラ内や現像ソフトを用いて、さまざまなパラメーターを調整しながら、好みの仕上がりに現像することもできる。

Fig. 1-1



1.3 写真の原理上避けられない課題

写真の画質には、避けることのできない幾つかの課題がある。例えば、写真がざらついた感じに見えるノイズや、本来はないはずのモアレ（干渉縞）、レンズの原理に起因する像のボケなど、写真に影響を与える光学的要素により、見ている光景にはない情報が写り込んでしまうことがある（Fig.1-3～Fig.1-9参照）。

イメージセンサーで光のエネルギーを電気信号に変換する過程で発生するエラーにより、画像の画素ごとに、ノイズと呼ばれる本来の被写体の明るさや色と違う結果が発生する。

ノイズを除去する機能は、これまでも多数開発されてきたが、ノイズの除去にともない、被写体のディテールまでが失われ、画像の品位が低下してしまうことがほとんどで、プロフォトグラファーからも改善の要望が多い領域であった。

なお、イメージセンサーは、画素の一つひとつが規則的に並んでいるため、縞やチェック柄などを撮影すると、規則的なパターンの重ね合わせとなり、本来は無いはずの、まだら模様（モアレ）が発生するという、原理的な課題がある。

また、イメージセンサーの1つの画素は、赤（R）／緑（G）／青（B）の、光3原色のうちの1色だけを検出し、残りの2色に関しては、周りの画素の情報を参照して推測することで、1つの画素のRGBデータが生成される（＝色補間処理）。この処理によって発生する、実際の被写体には存在しない「偽色（色モアレ）」や、斜めの線がギザギザに見えてしまう「ジャギー」についても、原理上避けることができない。

これまで、これらの画像劣化を抑える処理が数多く開発されてきたが、劣化を抑えることにより、解像感^{*1}や色再現性の低下が生じることは避けられなかった。

さらに、光をカメラに取り込むレンズには、光学の理論上、収差^{*2}や回折ボケ^{*3}といった逃れられない課題がある。高性能なレンズでは、凹凸レンズの組み合わせや、非球面レンズ、特殊な材料などを用いて、収差を最小限にとどめることに成功しているが、完全に除去することは原理上不可能である。

例えば、広角レンズを小さいF値（絞り開放側）で使用した場合、レンズの中心から外れた周辺部分の画質は、レンズの光学性能が低下して、ボケやすい傾向がある。

*1：本書では、「解像感」は画像のシャープさ、「解像度」は画像の画素数の意味で用いている

*2：収差

光学レンズが光の屈折を利用することから生じる現象。レンズが球面であるため焦点が一点に集まらない「球面収差」、光の波長ごとの屈折率の違いによる「色収差」、周辺部の点が彗星のように尾を引いた形状になる「コマ収差」、そのほかにも非点収差や歪曲収差などがあり、結像位置がズレるなどにより、像の「ボケ」や「ゆがみ」「色ずれ」の原因となる

*3：回折ボケ

光が障害物の端を通過する際に、障害物の影に光が回り込む現象。大きいF値（小絞り側）で撮影する場合、光が回り込み、像のシャープネスやコントラストが低下する

1.4 Neural network Image Processing Tool

AIの技術が進み、ディープラーニング技術がさまざまに応用されるようになった今、キヤノンは、カメラ、レンズを知り尽くしたリーディング企業として、本来の「あるがままの情景」の写真を実現するため、写真の原理上避けられない「カメラに残された課題」の解決に向け、自らディープラーニング画像処理技術の開発に正面から取り組んでいる。

2022年11月、キヤノンは、ディープラーニングを用いた画像処理ソフト「Neural network Image Processing Tool」をリリースし、ノイズリダクション、色補間、収差・回折補正（レンズのボケ補正）の3つの領域に対応した。

Fig. 1-2

ディープラーニング画像処理技術の3つの領域



ノイズリダクション

ノイズリダクションについては、これまで決定的な手法がない中、キヤノンは、劣化の少ない画像データ、処理の難しい画像データを、大量に用意できる強みを十二分に生かしながら、ディープラーニングを行うことで、クリアで高品位な画像を得られるノイズリダクション画像処理を突き詰めてきた。

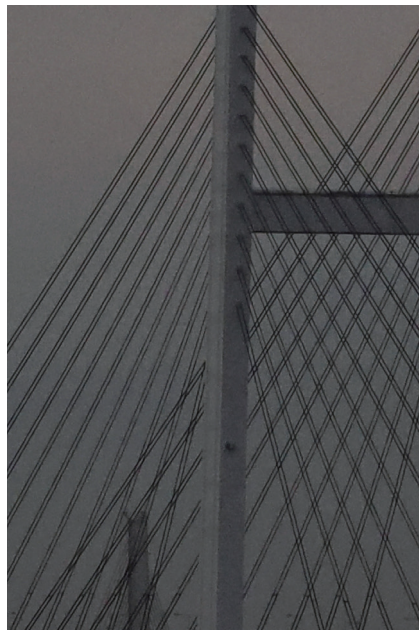
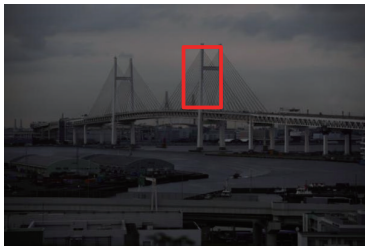
しかしながら、ディープラーニングは万能ではないため、撮影シーンによっては、部分的に従来の画像処理よりも悪くなる、「誤補正」となることがある。

高いハードルであるこの課題に対し、キヤノンは、ニューラルネットワークそのものの構造に変更を加える一方、学習プロセスを見直すとともに、学習データを工夫するなど、カメラで生じるノイズを熟知した、キヤノンならではのノウハウをつぎ込むことにより、クリアで高品位な画像が得られる「Neural network Noise Reduction」機能を確立した。

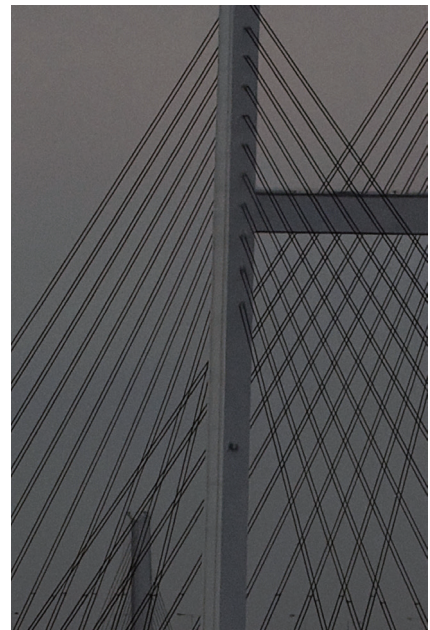
これにより、高感度撮影時において、光の情報を増幅させると同時に増幅されていたノイズの除去や、ノイズの発生によって阻害されていた、なめらかな肌の質感（スキントーン）の表現、ディテールの再現などが可能となった。

Fig. 1-3

ノイズリダクション ディテール再現



撮影画像



Neural network Noise Reduction

*高感度撮影でも細い線などのディテールの再現性が向上する

色補間

色補間については、キャノンの豊富な画像データベースを活用し、色補間のディープラーニング画像処理技術「Neural network Demosaic」を確立。人間の視覚が明るさの違いに敏感に反応し、色の変化にはあまり反応しないという視覚特性までを考慮して、学習データセットが構築されている。

その結果、色補間処理において、推測が難しい被写体を重点的に学習させることで、誤補間を抑制している。これにより、縞模様のシャツの偽色、斜め線で目立つことが多かったジャギー、ペットの写真などで目立ったモアレや偽色なども的確に補間し、解像感や色再現性の向上を実現している。

Fig. 1-4

色補間
色再現性向上



撮影画像

Neural network Demosaic

*色補間性能の向上により、細部の解像感や色再現性も向上する

Fig. 1-5

色補間
偽色(色モアレ)低減



撮影画像

Neural network Demosaic

*縞模様のシャツの撮影で発生していた偽色を強力に抑制し、より解像感の高い画像が得られる

収差・回折補正

収差・回折補正については、レンズによる画像の劣化を補正し、解像感を大幅に向上させる「Neural network Lens Optimizer」を開発。キヤノンは、レンズを開発する段階で、レンズごとに発生する収差や回折ボケを詳細に熟知している。

そのため、レンズごとの設計値をもとに、後述する撮影シミュレーション技術（1.7 参照）を駆使し、ディープラーニングの学習データを大量に生成して学習することで、画像周辺部などの細部に発生していた、さまざまなボケを補正することに成功。また、従来のデジタルレンズオプティマイザ* では、ボケを補正するとノイズが増幅されていたが、「Neural network Lens Optimizer」では、ノイズを増やすことなく、ボケのみを補正することができるようになっている。

* デジタルレンズオプティマイザ

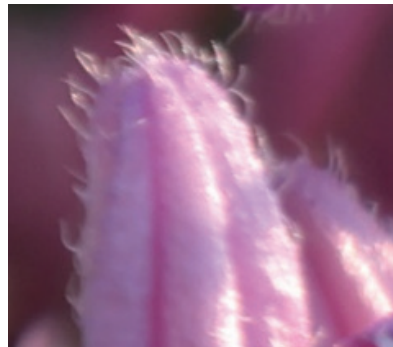
レンズの結像性能に関わる、残収差や物理現象である回折現象による解像劣化を、それぞれのレンズの設計値を用いて除去し、画像の解像感を上げる機能。Digital Photo Professional で利用することができ*、RAW から JPEG や TIFF 等の画像に現像する時に適用される。対象となる画像は、対象レンズで撮影された RAW 画像。なお、多くの EOS カメラにもデジタルレンズオプティマイザ機能が搭載されており、効果を適用した JPEG 画像を撮影することもできる

* デジタルレンズオプティマイザは、EOS R シリーズカメラと RF レンズ、および一部の EOS デジタル一眼レフカメラと EF レンズで利用可能)

Fig. 1-6

収差・回折補正

解像感向上



撮影画像



Neural network Lens Optimizer

*収差によるボケを補正することにより、被写体の細部が鮮明に描写される

さらに、この収差・回折補正は、光が明るすぎて画像情報が得られず補正自体が困難であった、被写体の白飛び部分で目立ちやすいボケについても、補正効果を発揮する。

しかしながら、白飛び部分のボケの補正を学習するデータを用意するだけでは、思わぬ誤補正が起きてしまうことがある。また、この白飛びに関連するボケの補正の研究発表はこれまでほとんどなかったため、乗り越えるべき課題を自ら新たに発見し、その未知であった原理まで解明する必要があった。

CGの活用などにより、学習データの改善に取り組む一方、ニューラルネットワークの構造や、実行後の画像処理にもさまざまな工夫を施すとともに、試行錯誤を繰り返し、白飛び部分で目立ちやすいボケに対しても、高精度なボケ補正を実現している。

Fig. 1-7

収差・回折補正

白飛びを含む被写体に対する補正



撮影画像



Neural network Lens Optimizer

*白飛び部分のボケ補正により、星空画像の周辺部で発生しやすいボケに対しても補正効果が大幅に向上する

Fig. 1-8

収差・回折補正

色にじみ低減



撮影画像



Neural network Lens Optimizer

*白飛び部分に発生する色にじみも高解像度に補正

ディープラーニング画像処理技術により実現した、3つの領域の画像補正は、単独で実行するよりも、組み合わせることで、より高いレベルの補正が可能となり、細部の描写や写真の立体感など、大幅な画質向上を実現する。

例えば、収差・回折補正は、ノイズを増やさないようにボケを補正できるが、ノイズが残っていると、画像の細部で収差・回折補正の効果が発揮されない。そのため、収差補正機能や回折補正機能に加えて、ノイズ低減機能など、ニューラルネットワークイメージプロセッサの付加機能を組み合わせることで、1つのソフトウェアを使用するだけで、撮影者の編集の手間を省き、原画全体の画質を向上させ、最大限に高めることができる。

Fig. 1-9



1.5 ディープラーニング「学習」と「推論」

次に「Neural network Image Processing Tool」や、2.1で後述する「Neural network Upscaling Tool」を実現するための、ディープラーニング画像処理技術について説明する。

ディープラーニング画像処理には、「学習」と「推論」の2つの段階があり、順を追って説明する。なお、本資料では、Fig.1-10で示すように、実際の撮影に相当する画像を「生徒画像」、ディープラーニングを通して最終的に得たい正解画像を「教師画像」と呼ぶこととする（詳細は1.6、1.7参照）^{*1}。

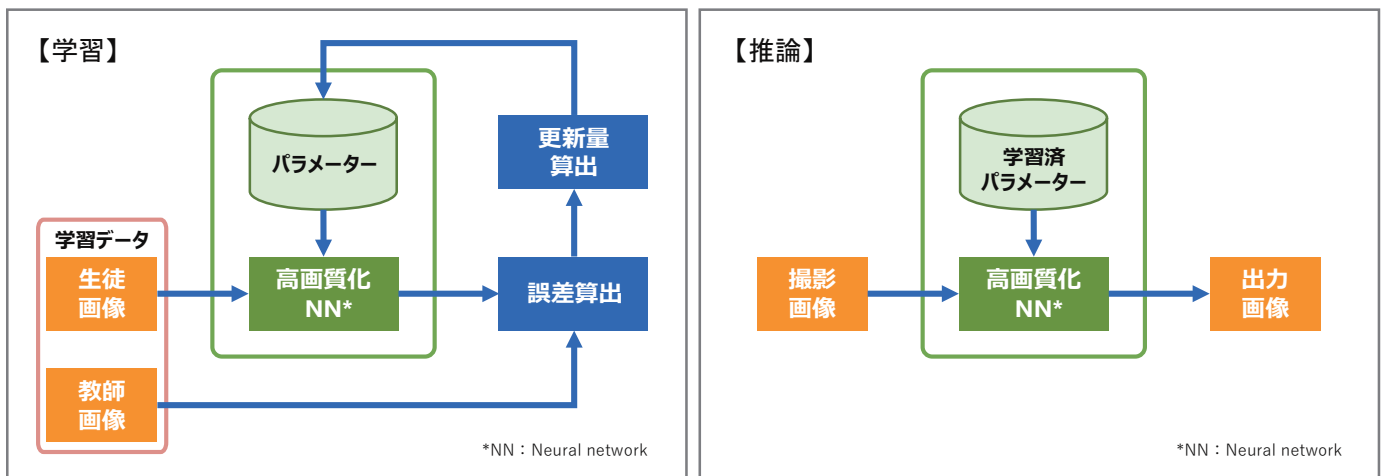
「学習」の段階では、「生徒画像」をニューラルネットワーク（NN）に入力して、ニューラルネットワークが推定した画像と「教師画像」を比較し、その誤差が小さくなるように、ニューラルネットワークのパラメータを更新する。これを大量のさまざまな画像に対して行い、この誤差が十分に小さくなる、すなわちニューラルネットワークが推定する画像が、十分に教師画像に近いものになるまで繰り返す。なお、この「学習」の段階は、キヤノンが製品開発する過程で実施するものであり、ユーザーのパソコン上で学習が行われることはない。

学習済みパラメータを用いて、実際に撮影した画像を処理する段階を「推論」と呼ぶ。これは Digital Photo Professional を介して、「Neural network Image Processing Tool」^{*2}や「Neural network Upscaling Tool」をインストールした、ユーザーのパソコン上で行われる処理である。「Neural network Image Processing Tool」では、RAW に対して現像時に推論処理が実行される（Fig.2-4 参照）。また、「Neural network Upscaling Tool」では、JPEG / TIFF に対して推論処理を実行することができる（Fig.2-1 参照）。

*1：ここでは学習データに対して対比的に「教師画像」「生徒画像」という表現を用いている。規模の大きい学習済みモデルから、軽量なモデルを学習する方法として知られる知識蒸留において、「教師モデル」「生徒モデル」という表現が用いられるが、それとは異なる

*2：対象機種は、巻末に記載の製品サイトを参照のこと

Fig. 1-10



1.6 学習データ

「学習」の精度を高めるには、目的の画像処理に適した「教師画像」と「生徒画像」のペア(= 学習データ)を、どれぐらい用意できるかが大きなカギを握っている。すなわち、「学習」に使われる生徒画像が、「推論」で入力される実際の撮影画像に対応していなければ、実際の使用で高精度な結果が得られるパラメータを学習したことにはならない。そのため、ディープラーニングの性能は、「学習データ」の質により、大きく左右される。

質の高い学習データを大量に用意するにあたり、キヤノンには大きな強みが二つある。

一つは、長年のカメラ/レンズの開発で蓄積してきた、キヤノンが保有する膨大な画像データベースである。多様な被写体がカバーされるとともに、JPEG などよりも詳細な情報を持ち合わせた、カメラで撮影された高精細な「RAW データ」を元に、大量の学習データを生成することができる。

もう一つの強みは、カメラで撮影された RAW データから、学習に適した「教師画像」と「生徒画像」のペアを、高精度に生成する技術である。その技術は、長年にわたるカメラ/レンズ開発の経験とノウハウによって確立された、撮影シミュレーション技術を用いて、学習データを生成する技術である。この撮影シミュレーション技術(1.7 章)は、前述した撮影の原理に沿った撮影プロセスを精密に再現し、カメラで撮影された RAW データから撮影画像を、コンピューターシミュレーションで生成することができる。

上述した2つの強みを組み合わせること(1.7 参照)により、ディープラーニング画像処理技術を開発するための、膨大な学習データを生成することが可能となり、「Neural network Image Processing Tool」、
「Neural network Upscaling Tool」の開発につながっている。

1.7 撮影シミュレーション

撮影プロセスのシミュレーションについて、「Neural network Image Processing Tool」の収差・回折補正機能である「Neural network Lens Optimizer」を例に説明する。

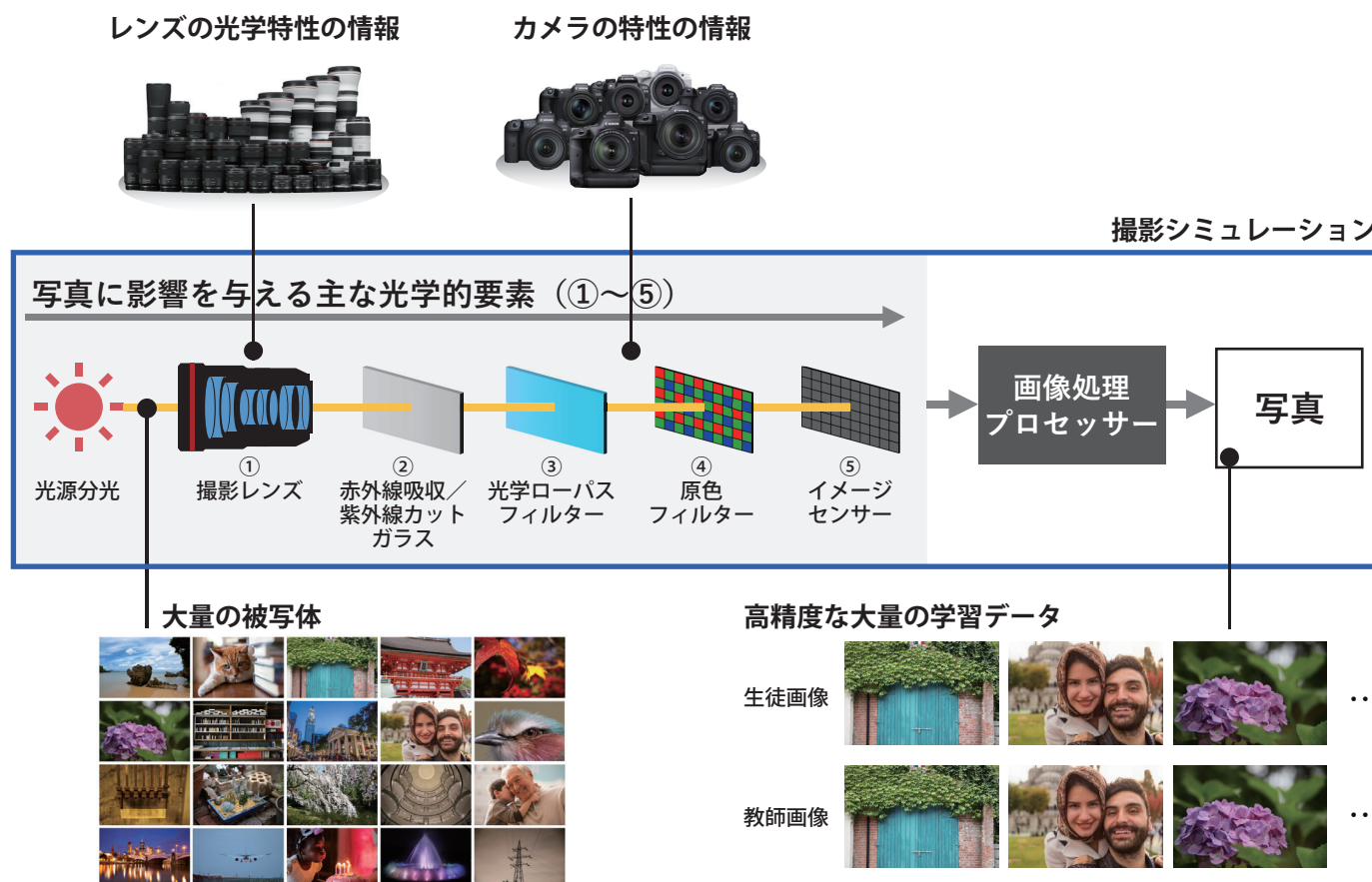
ここでは、学習データとして、収差や回折によるボケの無い「教師画像」と、実際の撮影画像に相当する、収差や回折によるボケのある「生徒画像」のペアを、シミュレーションにより生成する。

シミュレーションのベースに用いる被写体の画像は、前述したカメラで撮影されたRAWデータを利用する。ここで用いるRAWデータに対して工夫を加え、収差や回折によるボケが無い状態にしたものを「教師画像」とする。この教師画像に対して、さまざまなレンズの収差や回折といった特性と、カメラ本体による各種フィルターの特性を反映させることで、実際の撮影を模したRAWデータを生成し、「生徒画像」として使用する。

また、カメラや交換レンズの機種、さらに撮影の設定などにも対応でき、理想的な学習データを大量に揃えることができる。これは、設計から製造までを行う、キヤノンだからこそ保有している設計値や、特性データによって実現できる強みである。

この技術が無ければ、理想的な「教師画像」と「生徒画像」を大量に用意することは非常に困難である。例えばレンズの収差や回折の影響ができるだけ少ない撮影条件で実際に撮影した画像を「教師画像」としても、そのままでは少なからず撮影プロセスによる劣化が加わってしまう。

Fig. 1-11



2. アップスケーリング技術

2.1 従来のアップスケーリングの概要と課題

アップスケーリングとは、画像を元の解像度より、高い解像度へと変換することを指している。

アップスケーリングの手法で、これまで一般的に使われてきたのが「バイキュービック補間」と呼ばれる方法である。画像の拡大によって足りなくなる画素を、その隣だけでなく、周りの画素の色や輝度の情報を使って、新たに作り出すため、被写体の解像感が損なわれにくくなっている。

しかし、拡大時に解像感を維持しようとする、輪郭などの「境界部分」では、白や黒のふちができて輪郭が太くなってしまうため、輪郭の太さと解像感のバランスを適度に調整した結果、元の画像よりも解像感が損なわれてしまうという課題があった*。デジタルカメラで野生の動物を遠くから撮影したり、旅行先で街の景色を撮影したりして、撮った写真を後で大きくプリントした時や、写真の一部を切り取って（トリミングして）拡大した時に、元の写真よりも少し「ぼやっ」とした印象を受けたユーザーは、少なくないと思われる。

* 詳細は「3.3 従来处理との効果比較」参照

2.2 Neural network Upscaling Tool の概要

2023年4月、キヤノンは、ディープラーニングを用いた画像処理ソフト「Neural network Upscaling Tool」をリリースした。このソフトを利用することで、入力した画像（JPEG / TIFF）の解像感を保ちながら、縦横の画素数を2倍、全画素数を4倍に変換することができる。

従来のアップスケーリングにおいては、撮影画像を拡大 / トリミング（Fig.2-1、Fig.2-4 参照）したときに、解像感が失われる場合があったが、「Neural network Upscaling Tool」では、これまで培ったディープラーニング画像処理技術により、色、輝度、およびノイズ感を変えることなく、被写体本来が持つ解像感に迫る画像を、生成できるようになっている。

2.3 アップスケーリングの効果 / ユースケース

大判プリントを行うために、より高い解像度が必要な時や、トリミングにより画素数が減った画像に対して、高い解像感を維持したまま解像度を上げたい時に、「Neural network Upscaling Tool」を利用する。

被写体の境界部分のボケを抑え、画像拡大前のサイズで見えていた時の、自然な解像感を保ちながら、被写体本来が持つリアリティに満ちた、アップスケーリング画像を実現している。特に、動物の細かな毛並みや、輪郭のはっきりした構造物や文字、広角で撮影した広範囲の風景写真などでは、解像感の高いアップスケーリングの効果を実感することができる。

* これまでに発売されたキヤノンのカメラや、他社のカメラで撮影した画像にも対応している

Fig. 2-1

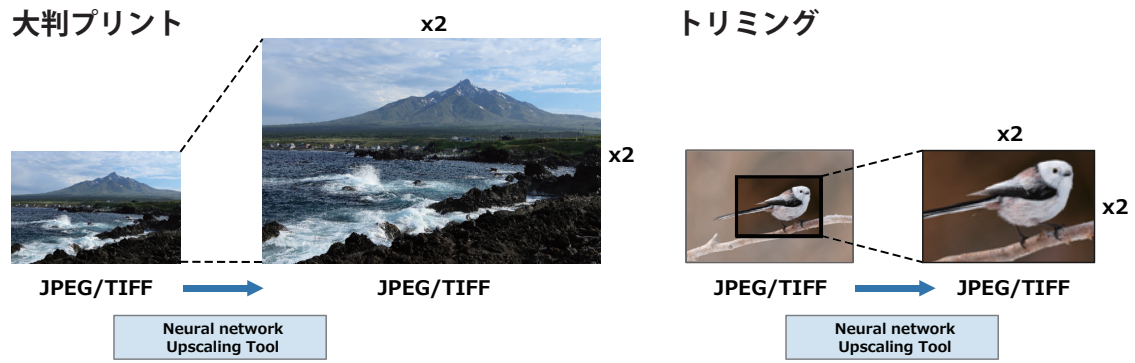


Fig. 2-2



一般的な処理(バイキュービック補間) キヤノンのアップスケーリング

Fig. 2-3



一般的な処理(バイキュービック補間) キヤノンのアップスケーリング

さらに、前述した「Neural network Image Processing Tool」を使用して、「Digital Photo Professional」で画像調整を行って保存した JPEG / TIFF に対し、「Neural network Upscaling Tool」を使ってアップスケーリングを行うと、ディープラーニング画像処理の効果が最大限に発揮される。

Fig. 2-4

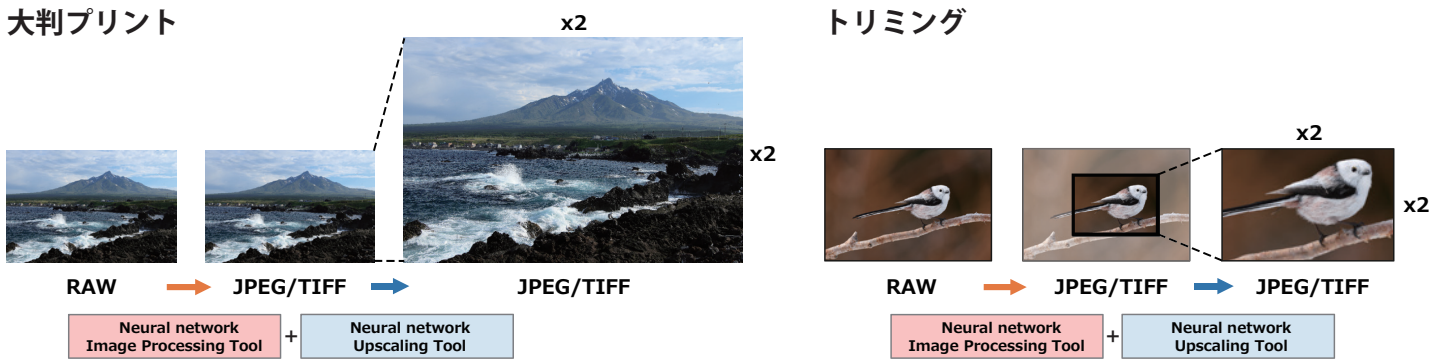
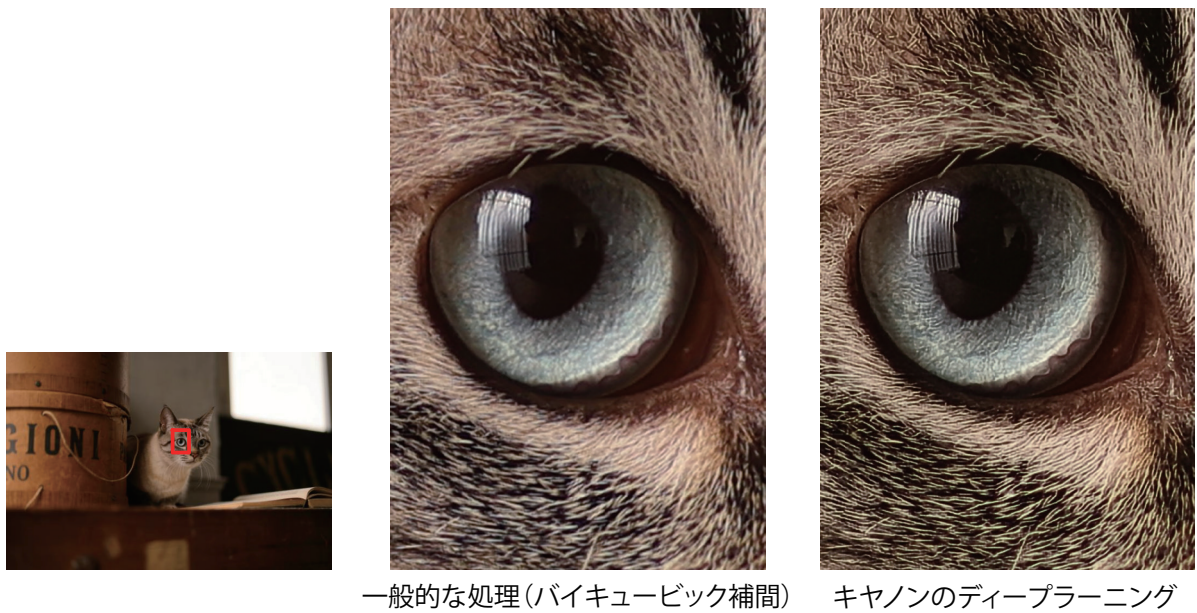


Fig. 2-5



3. キヤノンのアップスケーリングの特長 / 優位性

3.1 目標とした画質

キヤノンは、撮影した写真を拡大しても元の解像感を損なわない、アップスケーリングを目指してきた。言い換えると、撮影画像をディープラーニング画像処理で拡大した画像が、その画像本来の高い解像度に近づくように処理することである。画像本来の解像感を超えた、過度な解像感の向上は、ディープラーニングを使えば技術的には可能であるが、元の写真の印象からは逸脱したものとなる。「Neural network Upscaling Tool」では、拡大後も元の写真の印象を維持するため、色、輝度、およびノイズ感といった、細部にまで徹底的にこだわった設計が行われている。

3.2 目標画質の実現

ここでは、目標とする画質を確実に実現するための「学習」と「推論」に関する技術を説明する。

「学習」の段階では、「Neural network Upscaling Tool」においても、前述した「Neural network Image Processing Tool」と同様に、撮影シミュレーション技術を用いた、高精度な学習データを用いる。学習データとして、拡大後に相当する高解像な「教師画像」と、実際の撮影画像に相当する低解像な「生徒画像」のペアをシミュレーションにより生成する。

シミュレーションのベースとなる素材画像は、前述したRAWデータを利用する。このRAWデータに工夫を加えたものを「教師画像」とする。その「教師画像」を用いて、カメラ本体による各種フィルターの特性を付与した、低解像度の画像をつくることで、実際の撮影を模した「生徒画像」を生成することができる。

このように、撮影画像と拡大画像の関係を忠実に再現した、「教師画像」と「生徒画像」を用いて学習することで、架空の高解像画像を「生成」するのではなく、拡大後の解像度本来の画像に近づくように「補正」することができる。

「推論」の段階では、撮影画像を学習済みのニューラルネットワークに入力し、拡大画像を出力する。ニューラルネットワークは、前述の通り、現実にも忠実に学習されるようになっているが、厳しい目で画質を見ると、シーンにより、わずかながら過補正になる場合がある。

画質にこだわるキヤノンとしては、安心してディープラーニング画像処理を利用できるよう、画質を調整する画像処理の工程を加えることで、出力画像の色、輝度、およびノイズ感の細部にいたるまで、目標画質の実現を徹底している。

Fig. 3-1

アップスケーリング用の高精度な大量の学習データ

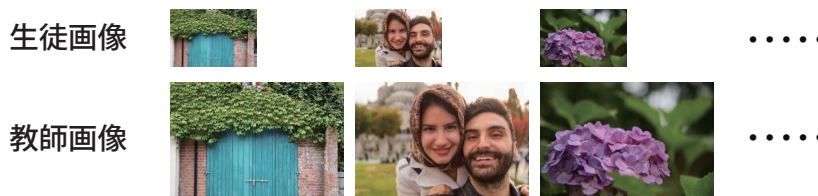
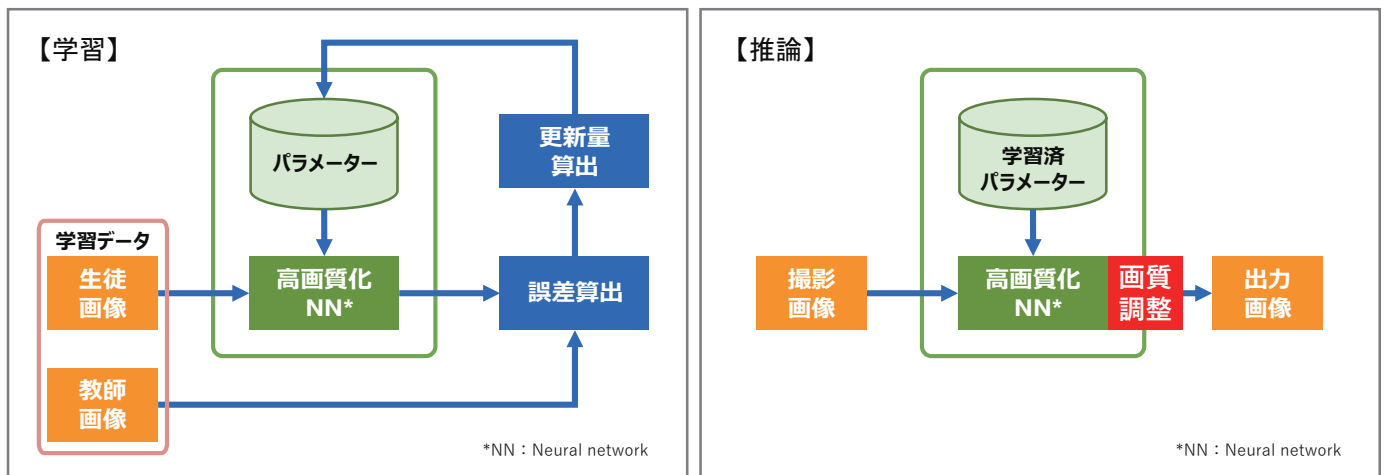


Fig. 3-2



3.3 従来処理との効果比較

ディープラーニング画像処理を用いたアップスケーリングの効果について、2倍の拡大を例に模式図を用いて説明する。

下図右側の (a) ~ (d) のグラフは、下図左側に示した画像 (A) ~ (D) の、白と黒のエッジ画像の青線部分を断面で見た ("—" 部分を拡大表示)、画素ごとの明るさ (ラインプロファイル) を示している。また、(a) ~ (d) の各グラフの横方向は、画素の位置 (座標) を丸印で表しており、縦方向は上が明るく下が暗いことを表している。

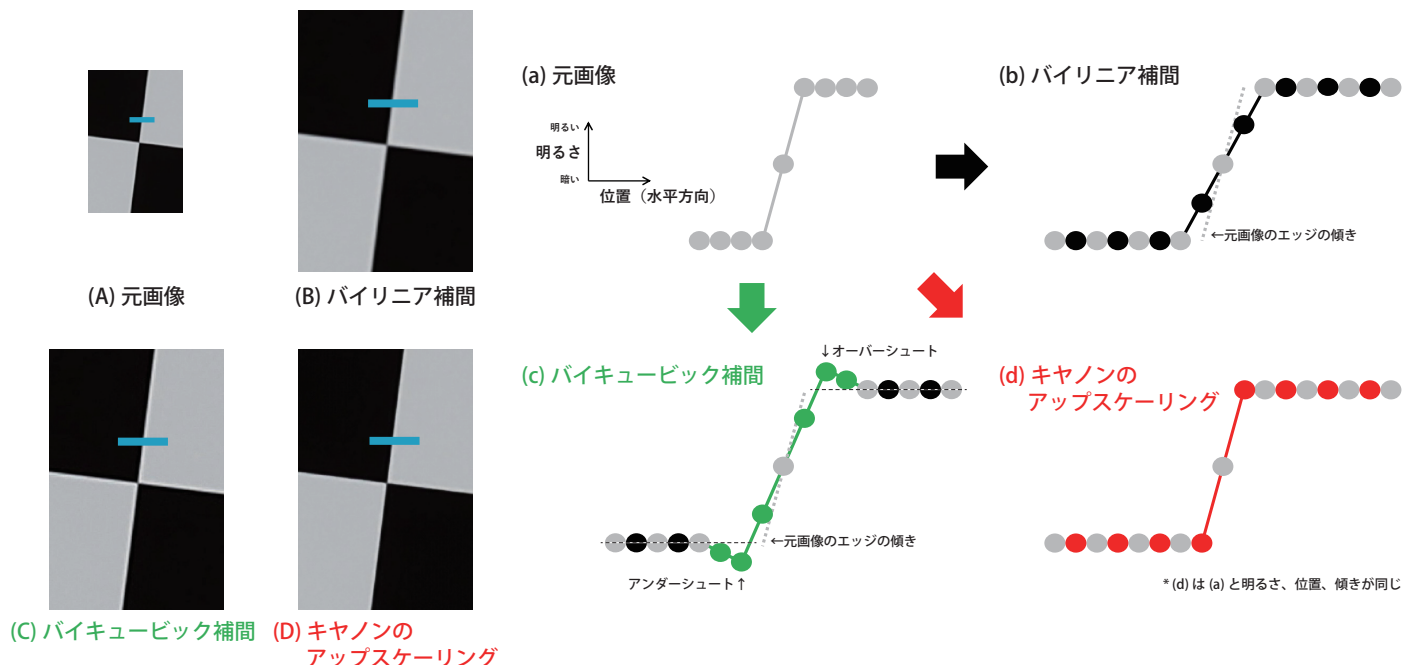
(a) のグラフが、拡大前の元画像のラインプロファイルで、エッジの左が暗く、右が明るいことがわかる。これを従来の拡大方法の一つである、「バイリニア補間」で拡大すると、(b) のグラフのように、元画像が横方向に2倍引き伸ばされ、その間の不足した画素を、隣接する画素●の明るさから補間して、新たな画素●が生成される。その結果、画素数は2倍になるが、元画像に比べエッジの傾きが緩くなっており、エッジがボケてしまう。

また、別の従来の拡大方法として、解像感を維持するために、シャープネスの効果を含めるようにつくられた、「バイキュービック補間」で拡大すると、(c) のグラフのように、エッジの傾きが改善して元画像に近づくが、エッジの両端の画素●が過剰に補正される、沈み (アンダーシュート) や浮き (オーバーシュート) が発生する。

これに対し、キヤノンのアップスケーリングでは、(d) に示すように、ディープラーニングで画素●を推定し、エッジの傾きを元画像に近づけるとともに、エッジ周辺の明るさも維持することにより、元画像の品位を損なうことなく、くっきりとしたアップスケール画像を得ることができる。その結果、従来のアップスケーリングよりも、被写体のエッジがシャープになる、髪の毛が細くなる、細かな構造がはっきりして質感が向上するなどの効果が得られる。

* 本説明における拡大方法やシャープネスの手法は一例であり、他にも複数の手法がある

Fig. 3-3



4. 「あるがままの情景」を追求するキヤノンの写真画質

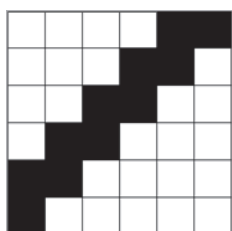
「Neural network Upscaling Tool」は、本来の「あるがままの情景」の写真を実現するために、画像の色、輝度、およびノイズ感を変えずに、画像を拡大するように開発が行われている。「学習」の段階では、前述した「実際の撮影に忠実な撮影シミュレーション」により、画素数が縦横とも2倍違う「教師画像」と、「生徒画像」のペアを大量に生成し、各ペア画像の細部の特徴の関係を大量に学習する (Fig.4-1)。

詳しくは、「生徒画像」である低解像度画像には、画素で捉えきれなかった被写体の細かな構造が、「折り返し信号」としてわずかに残っており、これが低解像度画像と高解像度画像の関係性（特徴）を見つける手がかりとなる。「折り返し信号」とは、信号のサンプリングで原理的に発生する現象であり、センサーの画素より細かい被写体の構造が、画素よりも荒い構造として見えてしまうものである。

例えば、下図左側の「低解像度の画像」のように、シャープなエッジがガタついて見える「ジャギー」など、実際には存在しないパターンとして写真に現れる。ディープラーニング画像処理で、実際に撮影画像をアップスケーリングする際は、入力された撮影画像の細部に残されたわずかな折り返し信号を、ニューラルネットワークが捉え、高解像度画像に適した画素を推定し、拡大画像が出力される。言い換えると、例えば、のつぺりとした何も無いところに、偽の被写体構造を描く（生成する）ことはしないように、学習が行われる。

Fig. 4-1

低解像度の画像



高解像度の画像



また近年、拡散モデルなどを用いて、例えば、欲しい画像の特徴をテキスト入力することで、新たに画像自体を生成する AI 技術や、人相も特定できないほどの低画素な粗い（または大きくボケた）顔画像から、例えばまつ毛や小じわまで生成する AI 技術などの製品化が進展し、一般的な利用が増えてきている。

ここに個人が特定できないほど、非常に低解像な顔画像が1つあるとする。高解像な顔画像に対して解像度を下げたときに、先ほどの低解像な顔画像と同一になる高解像な顔画像には、無数のバリエーションが存在する。低解像な顔画像に、まつ毛や小じわのような細かな顔の構造の痕跡が一切なくても、AI は顔の画像であることを認識し、学習した範囲から、そのバリエーションのうちのある1つを生成する。そのため、高解像化された顔画像が、本来の顔と合致する保証は一切ない。これは顔に限らず、例えば、被写体の構造がわからないほど大きくボケた画像を、先鋭化するような場合にも同じことが言える。なお、以下に比較のための作例を示す。

この全体画像 (A) の写真は、一番右のワインボトルにピントを合わせて撮影した一枚であり、左側のボトルに行くにしたがって、徐々に焦点深度から外れてやわらかくボケていく。全体画像 (A) に赤枠で示した、ボケたボトルの領域の画像を並べて示す。なお、他の手法でアップスケーリングした場合は、解像感が上がりすぎるなど、本来と異なる結果が生じる場合がある。そのため、他の手法でアップスケーリングされた (D) は、解像感が上がりすぎており、ボケの表現力が損なわれるとともに、そのボトルの実際のラベルとも異なった画像が生成されている。一方、キヤノンのディープラーニング画像処理 (C) は、これらの技術とは一線を画し、前述の通り、撮影シミュレーションを反映した、解像度の関係性に基づいた補正処理となっているため、ボケの表現力を損なうことなく自然な仕上がりとなっている。

Fig. 4-2



A: 全体画像

*赤枠の部分はピントが合っていないためボケている



B: 一般的な処理
(バイキュービック補間)



C: キヤノンのディープラーニング



D: 他の手法の例



E: ピントが合った画像 (参考)

キヤノンは映像のリーディングメーカーとして、「あるがままの情景」を細部まで再現する、キヤノンの写真画質を実現することを第一に開発を行ってきた。これからも技術を進化させ、キヤノンのカメラ/レンズを使ったユーザーが、「幸せを感じる」撮影体験を追求していく。

参考

- ディープラーニング画像処理技術
<https://global.canon/ja/technology/dl-iptechnology-2023.html>
- アップスケーリング技術
<https://global.canon/ja/technology/dl-upscaling-2023.html>
- Neural network Image Processing Tool 製品サイト
<https://sas.image.canon/st/nnip.html>
- Neural network Upscaling Tool 製品サイト
<https://sas.image.canon/st/nnups.html>

※ 本書に記載した「ディープラーニング画像処理技術／アップスケーリング技術」は、2023年11月時点における、当社の技術概要を説明したものである。そのため、本書に記載した内容を、将来に渡り保証するものではない。

※ 被写体や撮影条件などにより、本書に記載した内容と、効果や結果が異なる場合がある。