

# 深度学习图像处理技术/高清放大技术

---

采用深度学习图像处理技术的  
佳能神经网络图像处理工具(Neural network Image Processing Tool)  
和神经网络升级工具(Neural network Upscaling Tool)  
之技术解读

# 目录

1. 采用深度学习的佳能图像处理技术.....	3
1.1 细致刻画“真实生活景象” .....	3
1.2 摄影原理 .....	3
1.3 摄影画质面临的固有挑战 .....	4
1.4 神经网络图像处理工具(Neural network Image Processing Tool) .....	5
1.5 深度学习的训练和推导 .....	11
1.6 训练数据 .....	12
1.7 摄影过程模拟 .....	13
2. 高清放大技术 .....	14
2.1 传统高清放大技术的概述和课题 .....	14
2.2 神经网络升级工具(Neural network Upscaling Tool)的概述 .....	14
2.3 神经网络升级工具(Neural network Upscaling Tool)的效果和用例 .....	14
3. 佳能高清放大技术的性能和优势 .....	17
3.1 目标画质 .....	17
3.2 实现目标画质 .....	18
3.3 与传统高清放大技术的对比 .....	19
4. 追求“真实生活景象”的佳能画质 .....	20

# 1. 采用深度学习的佳能图像处理技术

## 1.1 细致刻画“真实生活景象”

生活中的场景和情况瞬息万变，往日不会在我们眼前重现，可是相机却能将这些点点滴滴记录下来。无论是惊鸿一瞥的美景，还是回顾往昔的感动时刻，相机都能记录这些瞬间并化作照片留存。本白皮书将介绍由佳能开发的深度学习\*图像处理技术，来实现这些“真实生活景象”。

\*深度学习

一种模拟人脑神经网络的机器学习方法。通过使用大量数据训练计算机，计算机学习这些数据，从中总结特征，做出符合需要的推测和判断。

## 1.2 摄影原理

首先，我们来谈谈数码相机的摄影原理。下图1-1说明的是，数码相机在拍摄被摄体并处理为图像的过程中，影响图像画质的主要因素。

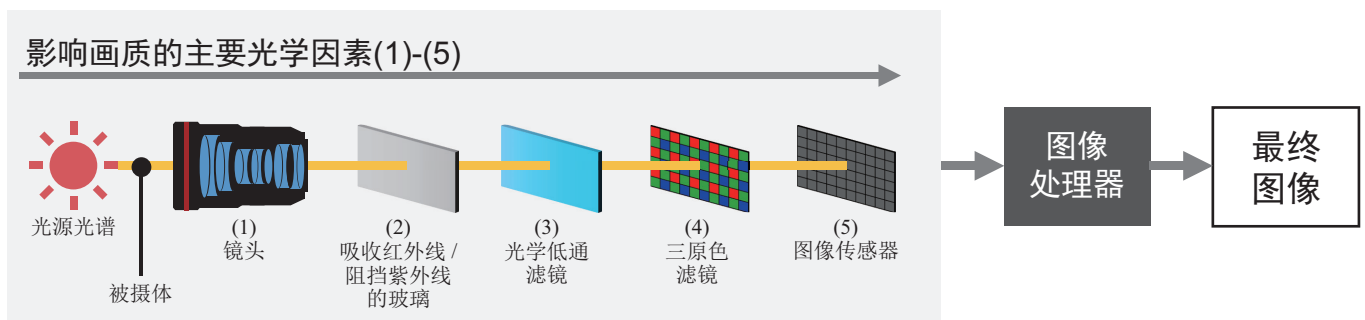
阳光、荧光灯、LED光或其他光源照射被摄体后，从被摄体发出的光，经由镜头(1)聚焦形成被摄体和场景的像。图像传感器(5)将光能转化为电信号，最终成像。

以下是整个过程的详细说明：

在图1-1所示的过程中，被摄体发出的光穿过镜头(1)中的多枚光学镜片，照射到相机机身(2)的玻璃滤镜上。这枚玻璃滤镜吸收红外线(IR)，阻挡紫外线光谱，从而消除可能会降低画质的不可见辐射波长。被摄体发出的光继续穿过光学低通滤镜(3)，这枚滤镜可以抑制高频摩尔纹图案、减少伪色现象，然后，光线穿过原色滤镜组(4)，此时针对每个单独的像素再次过滤光线，过滤后被摄体和场景的光线会分为红、绿、蓝(RGB)三原色信息。光线最终到达图像传感器(5)中光电二极管的光接收基底，光线中每种颜色(红、绿、蓝)的光强信息在此分到图像传感器的每个像素。

在成像过程的最后，被摄体/场景的光线的数字信息以电信号形式传输到图像处理器，由图像处理器执行多种图像处理功能。在此过程之后，当相机设定为保存RAW图像格式时，摄影师便可按需处理图像了，他们可以通过机内处理或软件应用程序调整多种RAW图像处理参数。

图1-1



备注：

## 1.3 摄影画质面临的固有挑战

要获得最佳的画质，摄影本身就面临许多挑战。例如，导致照片具有颗粒感的图像噪点、摩尔纹(干扰图案)，以及镜头固有的光学像差造成的图像模糊、伪色等(见图1-3至图1-9)。这些因素引入了原始场景中没有的视觉信息，对画质产生了影响。

由于图像传感器将光能转化为电信号的过程中产生误差，图像的每个像素都会产生有别于原始被摄体亮度和色彩的结果，我们称之为噪点。为了减少图像噪点，迄今为止已开发了很多降噪方法，但由于大多数方法在某些方面对画质产生了负面影响，比如在清除目标噪点时也去除了被摄体的某些细节，因此专业摄影师们也在不断寻找更好的解决方案。

拍摄的被摄体上有条纹或方格图案时，由于图像传感器的像素是规则排列的，再与被摄体的规律图案叠加，就不可避免地会产生波纹图案(摩尔纹)，尽管波纹图案并不存在于实际被摄体和场景中。

其他固有问题来自于色彩插值，图像传感器的像素利用色彩插值生成RGB数据。图像传感器的每个像素只检测光线三原色(红、绿、蓝)中的一种，对其余两种颜色的检测需要参考相邻像素的信息进行推断，这就是插值处理。采用这种处理，造成伪色(彩色摩尔纹)和锯齿状斜线(锯齿线)的风险则不可避免，尽管这些图案不存在于实际被摄体上。

虽然已经开发了很多方法来缓解画质下降问题，但这些方法也对画面分辨率\*1产生了负面影响，或者降低了色彩还原。

此外，相机捕获光线所用的镜头也会带来其他一些不可避免的问题，比如由镜头光学特性引起的多种像差\*2和衍射效应\*3。尽管高性能镜头使用凹透镜、凸透镜的不同组合和非球面镜片，以及特殊光学材料和镀膜来最小化这些像差，但理论上仍然不可能完全消除所有的像差及其影响。

这方面的一个例子是，使用广角镜头，选择较小f值(接近最大光圈)时，此时光学性能较低，拍摄的图像周围会有模糊趋势。

\*1：在本文档中，“可视分辨率”指图像的锐度或图像清晰度，而“分辨率”指图像的像素。

\*2：像差

像差是与光线在镜头中的折射相关的一种现象。像差的类型包括：球面像差(由镜片曲率造成多个焦点)、色差(由于不同波长光的折射率的差异造成)和彗形像差(会产生类似有尾巴的彗星的外围点)。其他的光学问题还有像散和失真，以及成像位置错位，这会导致模糊、失真和色偏。

\*3：衍射效应

衍射是镜片及材质影响光路并妨碍光传播时产生的一种现象。当光波经过光路中某个物体的边缘，间接绕到该物体的阴影区域时，衍射发生。衍射会降低画质。使用较大f值(光圈较小)进行拍摄可能会导致光在光圈边缘弯折，从而降低图像对比度和锐度。

备注：

## 1.4 神经网络图像处理工具(Neural network Image Processing Tool)

随着人工智能技术的不断进步，如今深度学习已得到广泛应用。佳能作为专长于相机和镜头的企业，一直致力于开发佳能自己的深度学习图像处理技术，直面摄影画质的固有挑战，实现“真实生活景象”。2022年11月，佳能发布了神经网络图像处理工具(Neural network Image Processing Tool)，这款图像处理软件将深度学习应用到降噪、色彩插值、像差和衍射校正(镜头模糊校正)三个领域。

图1-2

### 深度学习图像处理技术的三个领域



备注：

## 降噪

在影像行业中，至今对降噪也没有彻底的解决方案，佳能通过应用了深度学习的图像处理技术，致力于降噪和提高画质。实现这项技术需要使用大量画质好、难以处理的图像数据进行训练，以便生成清晰、高画质的图像。

然而，深度学习方法并不完美，根据拍摄情况的不同，可能会“错误校正”，在某些情况下这会比传统图像处理方法更差。

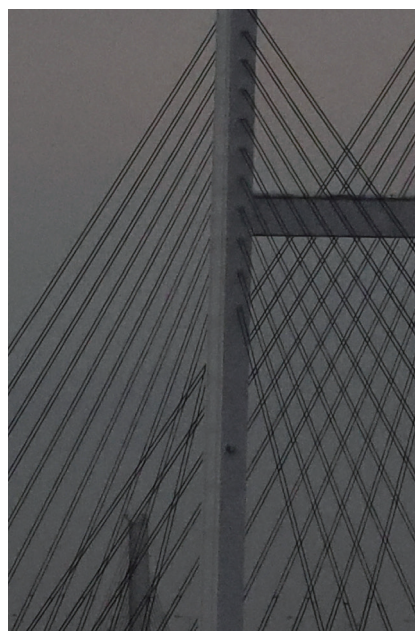
为了解决这一艰巨的挑战，佳能不仅修改了神经网络构造，还将佳能的专业知识应用到相机生成的图像噪点上，由此细化训练过程、训练数据和其他参数，开发出了可以生成清晰、高画质图像的神经网络降噪功能。

此功能可以消除在高ISO感光度下拍摄时随着图像亮度提高而增加的噪点，能够获得平滑的皮肤质感(此前会受到噪点影响)，再现细节并锐化有噪点图像中不易辨别的边缘。

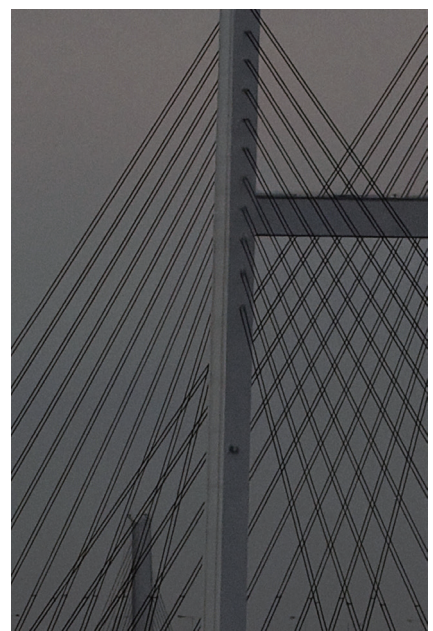
图1-3

### 降噪

再现图像细节



拍摄图像



神经网络降噪

\* 即使在高ISO感光度下也可以改善精细线条等细节的再现。

备注：

### 色彩插值

为了提高色彩插值的准确性，佳能使用了丰富的图像数据库，打造出神经网络去马赛克功能，通过此功能进行深度学习插值图像处理。在构建训练数据集时，开发人员甚至考虑了人类视觉的特征，即对亮度差异高度敏感，但对色彩变化反应不明显。

因此，通过对色彩插值中难以推断的被摄体进行针对性训练，抑制了假插值。与此同时，准确的插值现在可以用于具有挑战性的被摄体，例如容易产生伪色的条纹图案衬衫，容易产生锯齿线的斜线条，容易产生摩尔纹或伪色的动物皮毛/羽毛，提高了可视分辨率和色彩还原。

图1-4

色彩插值  
色彩还原性提高



拍摄的图像

神经网络去马赛克

\* 色彩插值得到改善，也增强了图像细节的色彩还原性以及可视分辨率。

图1-5

色彩插值  
抑制伪色(摩尔纹)



拍摄的图像

神经网络去马赛克

\* 很大程度上抑制了条纹衬衫照片中的伪色，提高了可视分辨率。

备注：

## 像差和衍射校正

在像差和衍射校正方面，佳能的神经网络镜头优化功能可以校正受光学问题影响的图像，并能大幅提高可视分辨率。在开发镜头时，佳能深知每支镜头设计中固有的像差和衍射效应。

因此，在上述知识的基础上，每支镜头的设计值都用到了下文所述的摄影模拟技术(见1.7节)中，生成大量深度学习数据并进行学习，从而校正了在图像周边等区域出现的多种模糊现象。

此外，现有数码镜头优化(DLO)\*4所执行的校正功能，在校正模糊后会增加噪点，神经网络镜头优化改进了DLO校正，可以校正模糊现象，却不增加噪点。

### \* 数码镜头优化

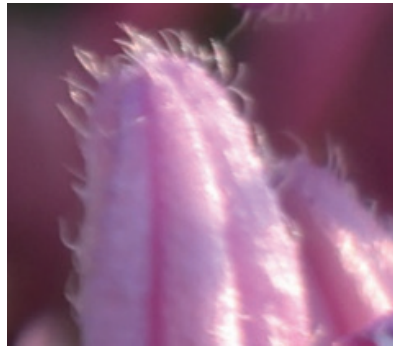
这是一种通过应用镜头特定设计值消除可视分辨率损失，从而提高画面分辨率的功能，残余像差和衍射的物理现象是与镜头成像性能有关的两个因素。可通过Digital Photo Professional软件使用数码镜头优化功能，并在RAW图像显像为JPEG和TIFF格式\*时应用校正。许多相机内置有数码镜头优化功能，可在记录JPEG图像时应用数码镜头优化效果。

\* EOS R系列相机和RF镜头，还有某些EOS数码单反相机和EF镜头上都有数码镜头优化功能。

图1-6

### 像差和衍射校正

提高可视分辨率



拍摄的图像



神经网络镜头优化

\* 校正光学像差导致的模糊可以更清晰地再现被摄体细节。

备注：



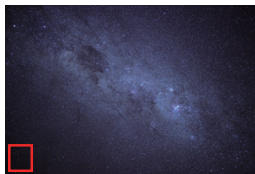
像差和衍射校正还可以有效校正高光细节丢失的被摄体区域中常见的模糊现象，由于图像信息丢失，所以校正往往比较困难。

然而，仅准备数据来训练如何校正高光细节丢失区域的模糊现象可能会意外导致错误校正。不仅如此，几乎没有与高光细节丢失区域的模糊校正相关的研究发表。因此，为了避免错误校正的风险，佳能的研究要求确定与校正相关的关键问题，并彻底阐明先前未知的发生原理。

训练数据通过计算机图形和其他方法得到改进，并对神经网络架构和图像后期处理进行了调整。经过反复试错，即使高光细节丢失区域中通常比较显眼的模糊现象，也能得到高精度校正。

图1-7

像差和衍射校正  
高光细节丢失图像的  
模糊校正



拍摄的图像

神经网络镜头优化

\* 校正高光细节丢失区域的模糊度可以很大程度上抑制天文摄影时容易发生的周边模糊现象。

图1-8

像差和衍射校正  
抑制色边



拍摄的图像

神经网络镜头优化

\* 高光细节丢失区域的色边也能通过高分辨率图像处理得到校正。

备注：

相较于单独应用深度学习图像处理技术实现的三个处理过程，将三者相结合则更能提高校正水平，细节刻画更真实，照片立体感更强、更清晰，实现了画质的整体提高。

例如，如果只单独应用深度学习图像处理的一个功能——像差和衍射校正(见1.4节)，那么只会校正由镜头像差和衍射现象引起的模糊，然而，在更清晰的图像中噪点仍然存在，尤其是在高ISO感光度图像中。

因此，如果能结合神经网络图像处理工具(Neural network Image Processing tool)的其他功能，比如像差和衍射校正功能以及降噪功能，在编辑图像时则只需使用一个软件，就可节省摄影师的时间和精力，最大限度地改善原始图像的整体画质。

图1-9



备注：

## 1.5 深度学习的训练和推导

深度学习图像处理技术同时应用于神经网络图像处理工具(Neural network Image Processing Tool)和神经网络升级工具(Neural network Upscaling Tool，稍后将在第2.1节进行说明)。

深度学习图像处理模型的编程和训练过程包括“训练”阶段和“推导”阶段，下文将具体介绍。在本文档中，如图1-10显示，对应佳能精心准备的由相机拍摄的真实照片数据集，希望被执行校正处理的图像称为“学生”图像。反之，佳能精心准备的旨在作为最终结果和深度学习校正处理目标的理想图像数据集的被称为“教师”图像(详见1.6和1.7)\*<sup>1</sup>。

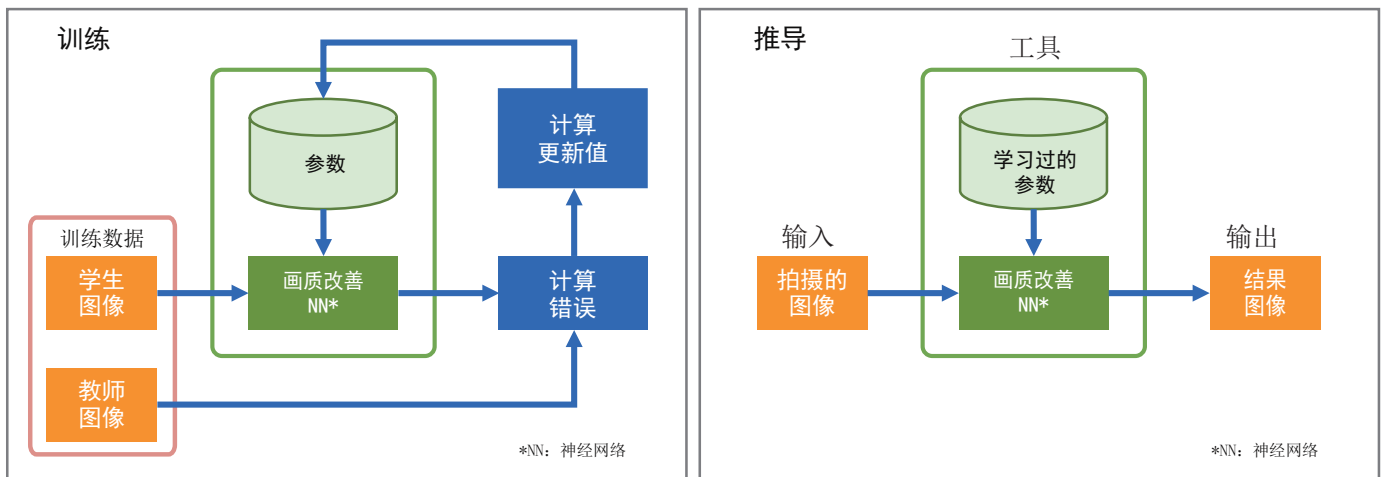
在“训练”阶段，将“学生”图像数据集输入到神经网络(NN)，与“教师”图像数据集进行比较，进而对图像校正进行评估。通过这种比较，神经网络(NN)处理参数会根据图像数据的差异进行更新，从而训练模型减少误差、提高画质。这个“训练”阶段过程会采用批量数据集进行，该数据集包含由佳能精心准备的在多种情况和环境中拍摄的图像，在误差变得十分小之前，训练过程会不断重复，即神经网络评估的校正图像，将十分接近“教师”图像的画质。请注意，“训练”阶段的工作是由佳能在产品开发过程中执行的，而不是在用户计算机上进行的。

在“推导”阶段，该模型会应用“已学习”参数来处理由相机拍摄的实际图像。与“训练”阶段相反，“推导”阶段的处理使用Digital Photo Professional软件应用程序，在安装了神经网络图像处理工具(Neural network Image Processing Tool)\*<sup>2</sup>或神经网络升级工具(Neural network Upscaling Tool)的用户计算机上执行。在显像RAW图像时(见图2-4)，会通过神经网络图像处理工具(Neural network Image Processing Tool)对RAW图像进行推导处理。而神经网络升级工具(Neural network Upscaling Tool)支持对JPEG和TIFF图像进行推导处理(见图2-1)。

\*1: 本文中使用的“教师图像”和“学生图像”等术语与训练数据进行了比较。在知识蒸馏中，即一种从大型训练模型中学习轻量级模型的方法，使用了术语“教师模型”和“学生模型”，但这里的“教师图像”和“学生图像”与它们不同。

\*2: 有关适用的型号，请参考本文档末尾的产品网站。

图1-10



备注:

## 1.6 训练数据

提高“训练”模型结果准确性的关键是准备足够数量的已配对的“学生”和“教师”图像，作为佳能精心准备的数据集，这可以引导模型学习所需的图像处理目标。如果提供的“训练”数据集不足，系统可能没有学习到参数来产生准确度高的结果。这种情况的发生是由于用于“训练”的“学生”图像与在推导阶段输入的实际使用的图像不相符。因此在实际使用中，如果条件不那么理想，模型会高度依赖于与有限的“训练”数据集更相符的实际图像。这一整体条件使得深度学习系统的性能在很大程度上依赖于大量高质量的“训练”数据来实现最佳的结果。

在准备大量高质量的“训练”数据时，佳能利用了作为影像行业领军企业的两个主要优势。

第一个优势是在长年累月相机和镜头开发过程中所积累的巨大图像数据库。海量的“训练”数据来自高分辨率RAW数据，这些RAW数据由相机拍摄所得，涵盖多种被摄体，而且RAW数据中的信息比JPEG等其他格式更多。

佳能的另一个优势是可以从相机拍摄的RAW图像中精准生成适合训练的成对“教师”和“学生”图像。“训练”数据的生成过程用到了拍摄过程模拟技术(见第1.7节)，这项技术建立在几十年相机和镜头开发的经验和专业知识的基础上。这项模拟技术根据之前阐述的摄影原理，使用计算机模拟，详细地再现了拍摄过程，通过相机拍摄的RAW数据来生成摄影图像。

佳能整合了这两大优势，创建了用于深度学习图像处理的海量“训练”数据，推动了神经网络图像处理工具(Neural network Image Processing Tool)和神经网络升级工具(Neural network Upscaling Tool)的开发(见1-7节)。

---

备注：

## 1.7 摄影过程模拟

神经网络图像处理工具(Neural network Image Processing Tool)有多种功能，其中之一是神经网络镜头优化，用于校正受像差和衍射影响的图像画质。以下将详细介绍摄影过程模拟，为深度学习创建理想数据集。

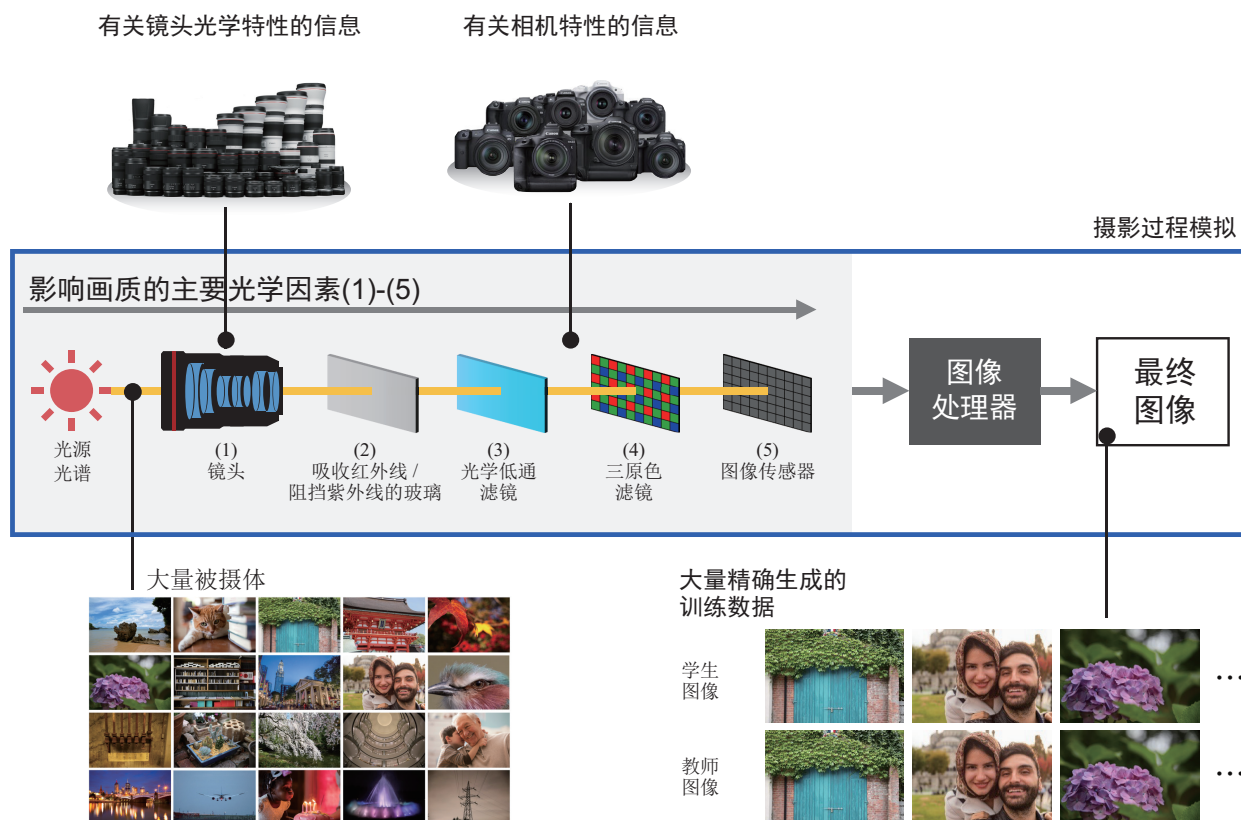
摄影过程模拟用于生成图像对，作为“训练”数据，具体来说，1)没有像差或衍射导致图像模糊的“教师”图像，和2)有像差和衍射现象的“学生”图像，类似于现实中使用数码相机拍摄的真实图像。

摄影过程模拟所用的基础图像，会使用前文所述的相机拍摄的RAW数据。这些RAW数据经过佳能自有技术进行处理，消除了像差和衍射引起的模糊，然后用作“教师”图像。之后，对“教师”图像进行仔细修改，再用作“学生”图像，修改包括多种镜头的像差和衍射效应等，以及相机中光学滤镜的特性，生成近似于实际拍摄的RAW数据。

而且，不仅可以模拟多种型号的相机和可更换镜头，还包括多种拍摄设置，所以能够得到大量理想的“训练”数据。由于佳能是相机和镜头的制造商，全程监管设计和生产，并且掌握这些设计值和特性，所以摄影过程模拟才成为可能。

如果没有佳能的自有技术，就很难准备出大量、充足的理想“教师”和“学生”图像。例如，将在现实环境中拍摄的图像用作“教师”图像，即使镜头像差和衍射效应最小，仍不可避免因拍摄过程产生画质下降，相较于使用佳能自有技术过程创建的图像，仍不完美。

图1-11



备注:

## 2. 高清放大技术

### 2.1 传统高清放大技术的概述和课题

在高清放大时，图像被转换为更高的分辨率。

至今为止最常见的高清放大方法是双三次插值，这种锐化和放大技术可通过图像中已知数据的亮度和色彩信息来估计必要的未知数据。双三次插值会预测必要的像素来增加像素数，然后通过从邻近的像素和周边的像素采样色彩和亮度信息来放大图像，被摄体的锐度不容易损失。

使用传统双三次插值进行高清放大的问题是，随着图像分辨率的提高，被摄体轮廓等物体边界区域，会因出现黑、白边缘，造成轮廓比原始图像\*5更粗。通过平衡轮廓粗细和可视分辨率之间的关系来校正这种现象，反而会降低可视分辨率，与原始图像相比可视分辨率出现损失。很多摄影师在野生动物摄影、体育摄影或者旅行街拍后，想要通过裁切图像来获得更有冲击力的构图，或者增加剩余像素用于放大打印，此时，高清放大后的照片对他们来说可能还是有些模糊。

\*5：详情请见“3.3 与传统高清放大技术的对比技术”。

### 2.2 神经网络升级工具(Neural network Upscaling Tool)的概述

2023年4月，佳能发布了神经网络升级工具(Neural network Upscaling Tool)，一款应用了深度学习技术的图像处理软件。这款软件可以保持原始JPEG或TIFF图像的可视分辨率，同时还将垂直和水平像素数翻倍，总像素为原来的四倍。

当图像被放大或裁切(见图2-1、2-4)时，使用传统高清放大有时会损失可视分辨率。相比之下，神经网络升级工具(Neural network Upscaling Tool)应用深度学习图像处理技术，在不改变色彩、亮度或噪点的情况下，生成接近原始被摄体可视分辨率的图像。

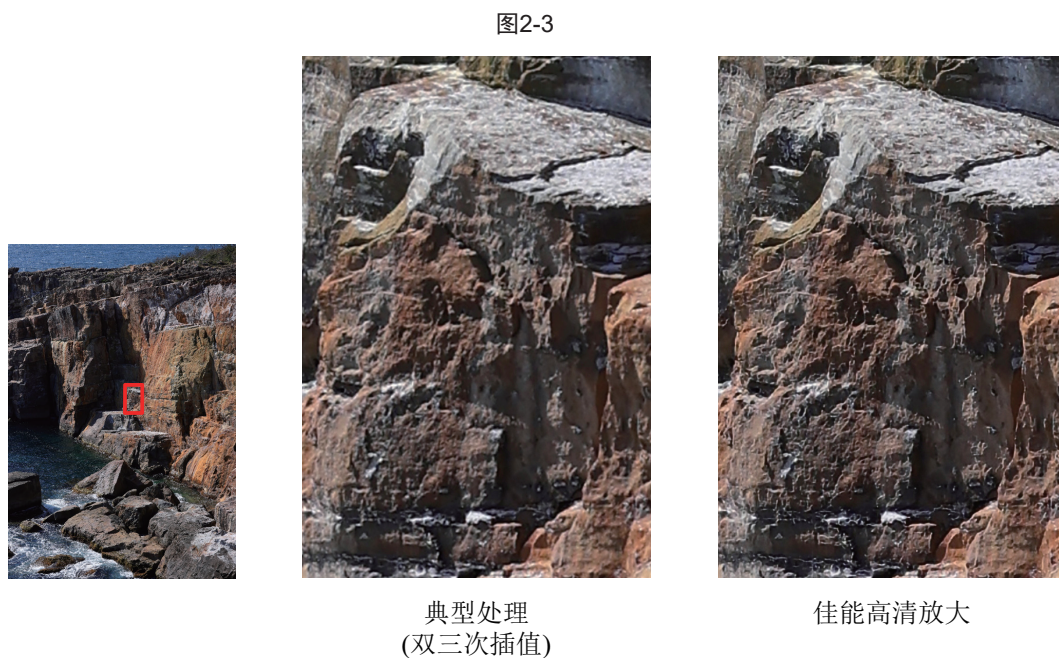
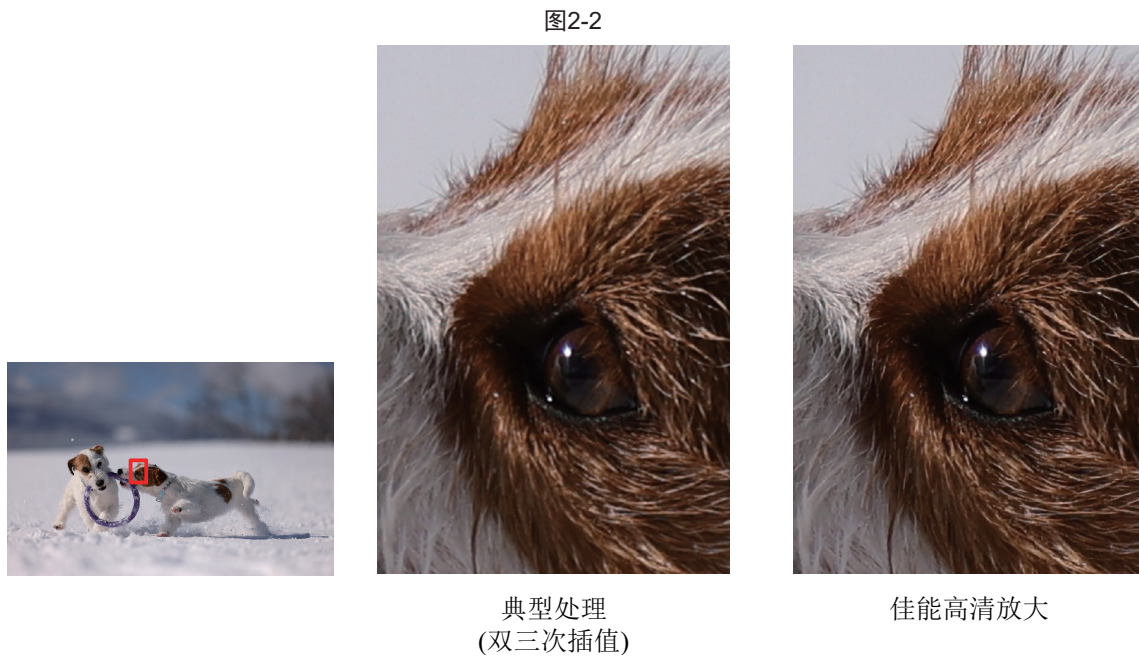
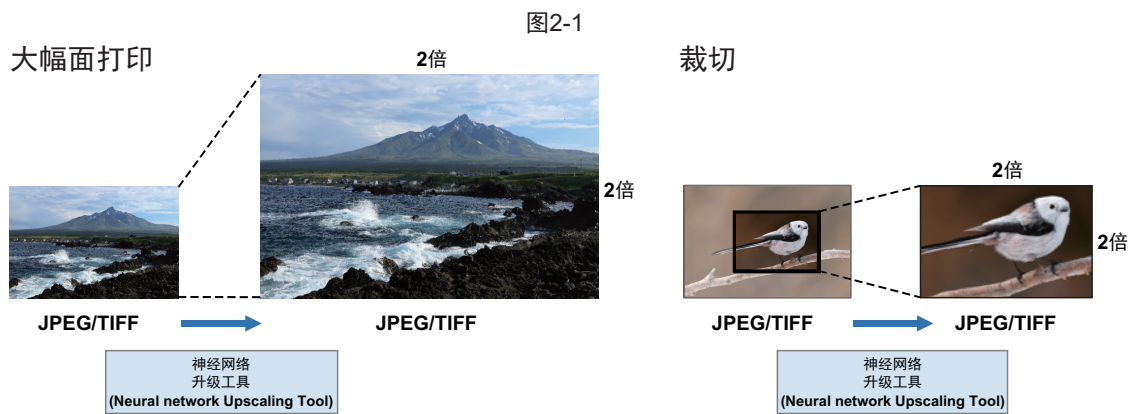
### 2.3 神经网络升级工具(Neural network Upscaling Tool)的效果和用例

神经网络升级工具(Neural network Upscaling Tool)的一个用途是为大幅面打印准备较高分辨率的图像。另一个用途是针对裁切后分辨率较低的图像，用户可能想要通过增加像素来保持锐度。

这些高清放大的图像不仅实现了被摄体边界模糊度最小化，还保持了等同于原始图像尺寸中展现的自然锐度，看起来就如同原始图像一般。在保持高可视分辨率的同时，放大效果在精细的动物皮毛、轮廓清晰的建筑物或文字、辽阔的广角风景照片中尤为显著。

\* 注意：该工具可以用于佳能相机或其他品牌相机所拍摄的照片。

备注：



备注:

使用神经网络升级工具(Neural network Upscaling Tool)，对经过Digital Photo Professional中的神经网络图像处理工具(Neural network Image Processing Tool)处理并保存的JPEG或TIFF图像进行高清放大时，可以完全发挥深度学习图像处理的效果。

图2-4

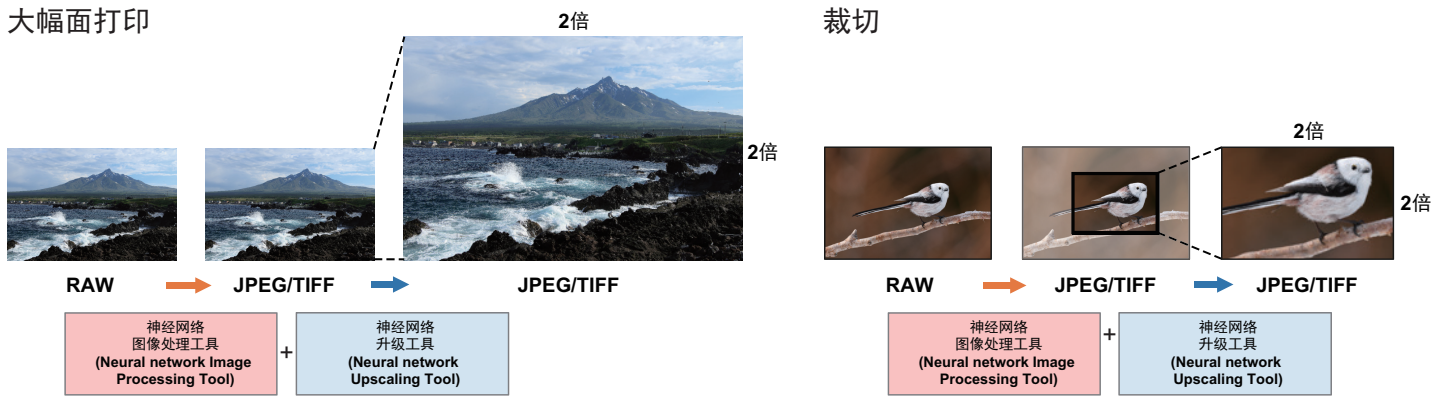
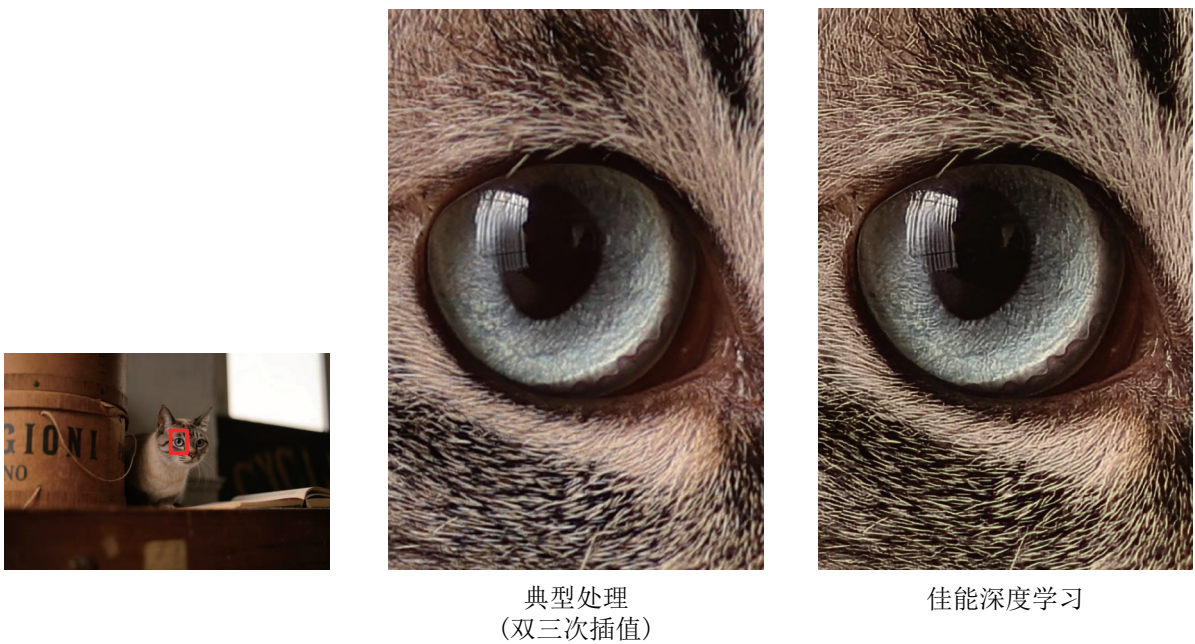


图2-5



备注：



## 3. 佳能高清放大技术的性能和优势

### 3.1 目标画质

佳能不懈追求的是，这种高清放大技术在将照片放大的同时，仍保持原来的可视分辨率。因此，通过深度学习图像处理技术将图像放大，处理后的图像接近原始图像的高分辨率。虽然深度学习在技术上足以将分辨率提高至超过原始图像分辨率的程度，但其结果就会偏离原始照片呈现的印象。神经网络升级工具(Neural network Upscaling Tool)在设计上密切关注色彩、亮度和噪点等细节，即使经过放大，也能保持原始照片呈现的印象。

---

备注：

### 3.2 实现目标画质

“训练”和“推导”以确保目标画质的技术说明如下。

在“训练”阶段，神经网络升级工具(Neural network Upscaling Tool)会使用“训练”数据，这些数据是通过与神经网络图像处理工具(Neural network Image Processing Tool)相同的摄影过程模拟技术精确生成的。上述摄影过程模拟技术用于生成图像对，用作“训练”数据，即对应放大后图像的高分辨率“教师”图像，和对应用户在现实世界实际拍摄的低分辨率“学生”图像。

之前介绍过的RAW数据会用作模拟过程的源图像。“教师”图像是通过编辑RAW数据获得的。之后“教师”图像将用来创建低分辨率的图像，对其应用多种相机光学滤镜的特性，生成类似于实际拍摄效果的“学生”图像。

这样，训练采用“教师”和“学生”图像(这两种图像忠实地再现了拍摄图像和放大图像之间的关系)，所以经修改的图像在放大后几乎保留了原始图像分辨率，而非不切实际地生成高分辨率图像。

在“推导”阶段，用户图像被提供给经过训练的神经网络，以生成放大的版本。虽然，我们以忠于现实的方式训练神经网络，但仔细检查画质可能还会发现某些场景有轻微的过度校正。

佳能在成像质量方面一直秉持高标准，追求生成的图像在色彩、亮度和噪点等细节都达到目标画质。为实现上述目标，还增加了另一个图像处理步骤来调整画质，从而使深度学习图像处理的运用更为可靠。

图3-1

#### 海量精确生成的训练数据用于高清放大

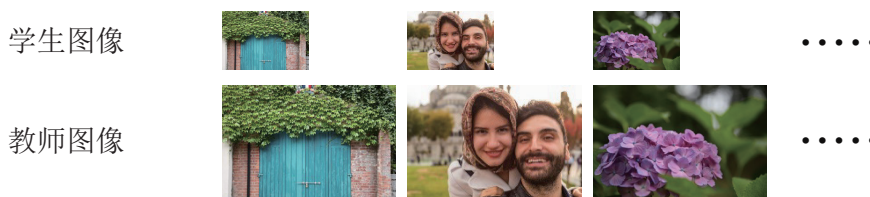
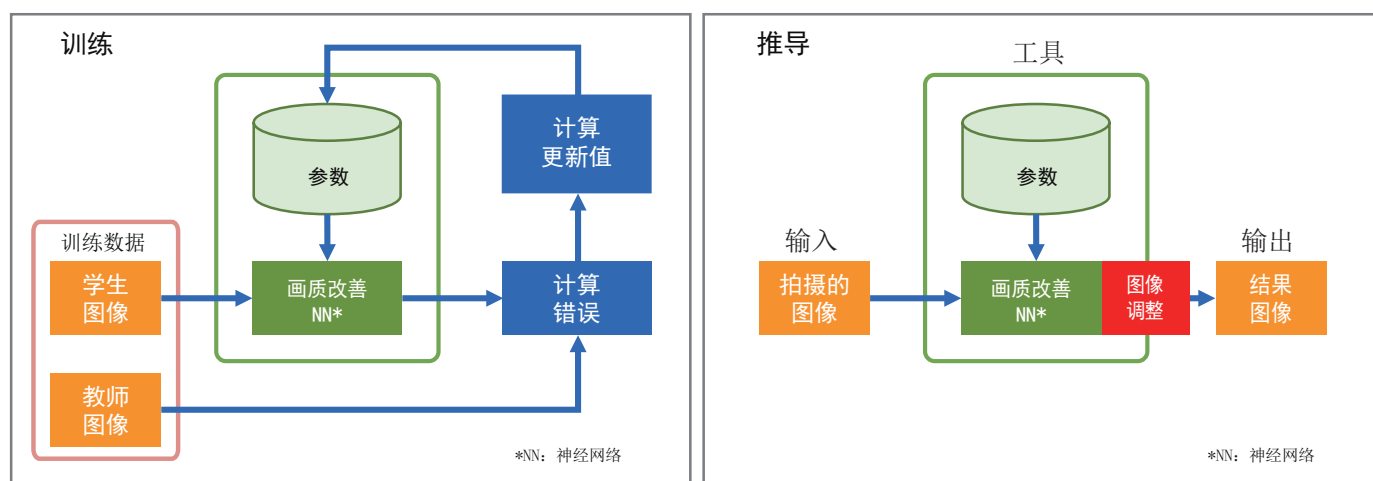


图3-2



备注：

### 3.3 与传统高清放大技术的对比

在此，我们利用两倍放大情况下的原理图来介绍采用深度学习图像处理的高清放大效果。

下图中右侧的线图(a)-(d)表示下图中左侧黑白边缘图像(A)-(D)中蓝线部分(“—”表示放大长度)在截面图上的每个像素的亮度(线剖面)。在线图(a)-(d)中，在水平方向上，用点表示像素位置(坐标)；在垂直方向上，向上表示亮度增加，向下表示亮度减小。

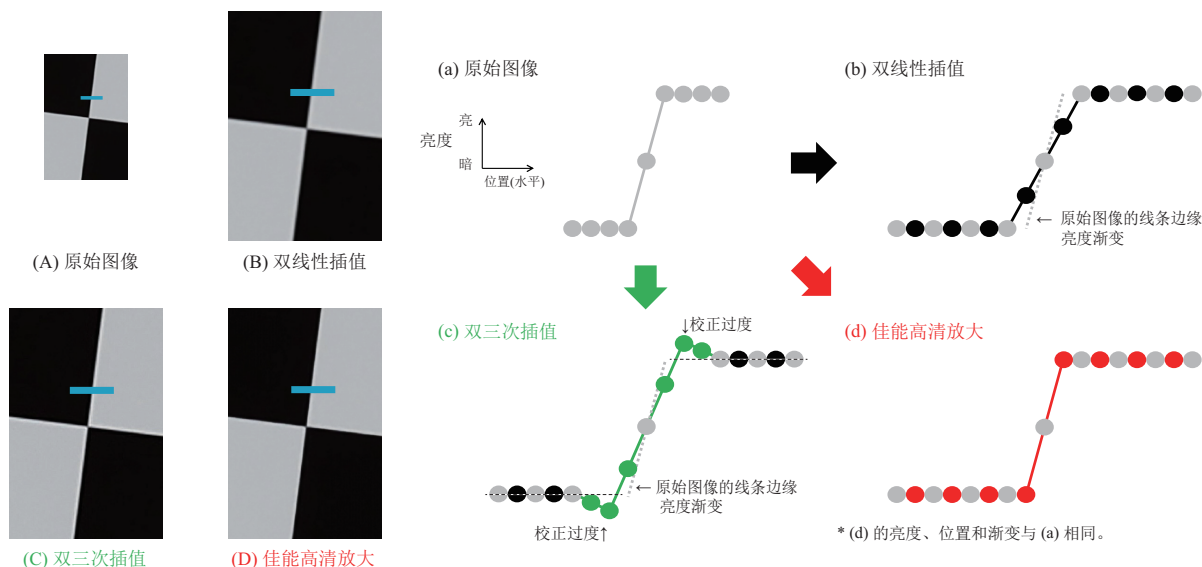
线图(a)是原始图像放大前的线剖面，可以看出蓝线的左侧较暗，右侧较亮。如果采用传统的双线性插值的放大方法进行放大，如线图(b)所示，原始图像在水平方向上拉伸两倍，并根据相邻像素(●)的亮度，在中间插入所需的像素值，生成新的像素(●)。这样一来，像素数虽然翻倍，但相较于原始图像，它改变了线条亮度渐变的对比度比值，导致边缘模糊。

另一种传统的放大方法是双三次插值，这种方法包括通过锐化来保持可视分辨率。如线图(c)所示，虽然与(b)双线性插值相比，双三次插值的线条亮度渐变得到了改善，图像更接近原始图像，但是在黑白像素相交(●)的边缘两侧，像素会被过度校正。

相比之下，如线图(d)所示，佳能高清放大处理技术采用深度学习来估算像素(●)，使线条亮度渐变更接近原始图像，同时保持边缘的亮度不变，由此生成清晰的放大图像且不会损失原始画质。因此，与传统的放大相比，被摄体边缘会更锐利，发丝会更精细，复杂的结构也会更清晰，质感会更强。

\* 注意：本文档说明的放大和锐化方法只是示例，还有其他几种方法可用。

图3-3

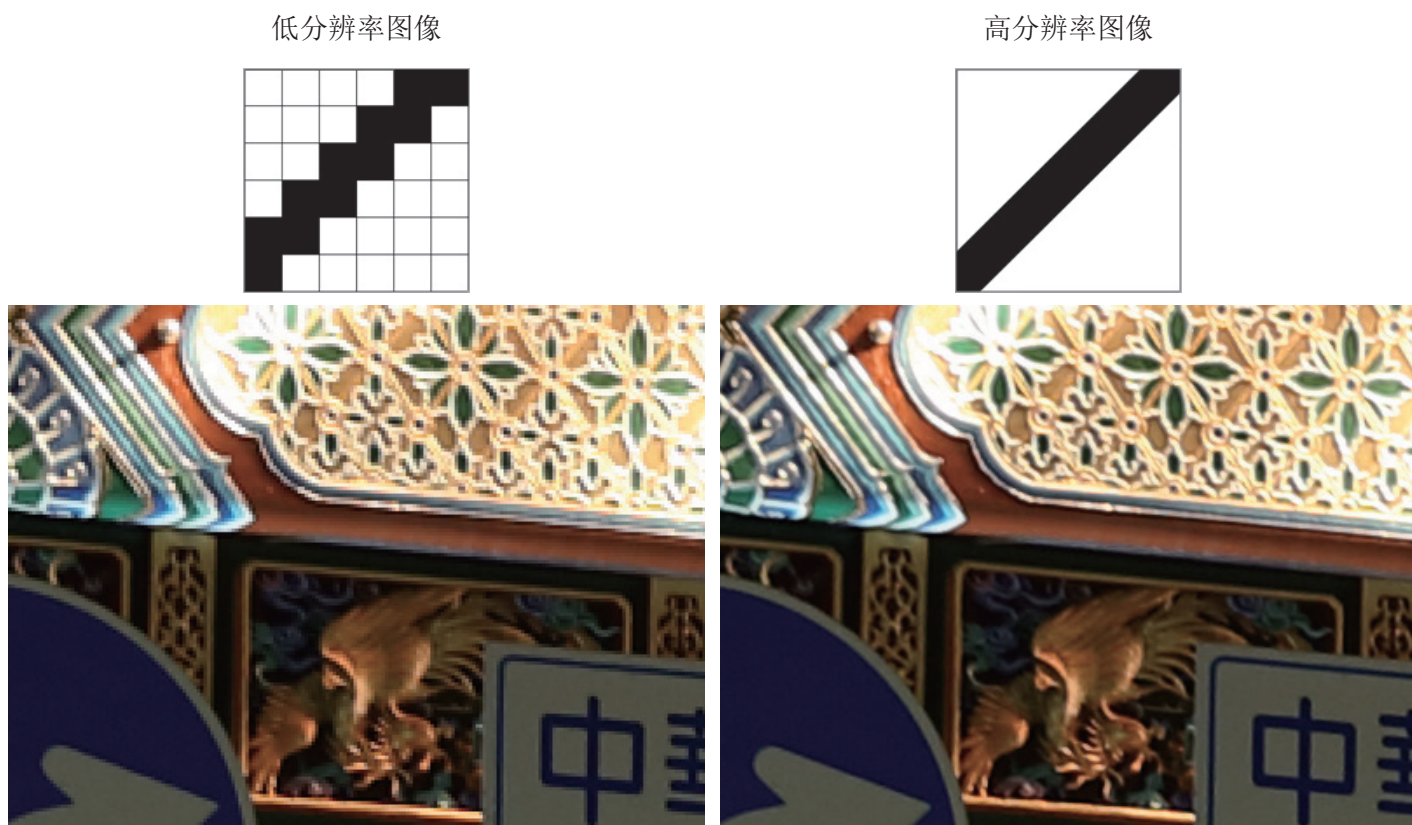


备注：

## 4. 追求“真实生活景象”的佳能画质

开发神经网络升级工具(Neural network Upscaling Tool)是为了在不改变色彩、亮度或噪点的前提下放大图像，细致刻画“真实生活景象”。在“训练”阶段会生成大量的成对“教师”和“学生”图像，前者通过之前介绍的忠于实际拍摄的摄影过程模拟技术，将垂直和水平像素翻倍。为了学习低分辨率“学生”图像对与高分辨率“教师”图像(图4-1)之间的关系，在“训练”阶段需要大量的数据集图像。具体来说，低分辨率的“学生”图像中包括有效像素无法呈现的被摄体细微构造的痕迹，它们以“锯齿信号”的形式存在。这为发现低分辨率图像和高分辨率图像之间关系(共性特征)提供了线索。从原理上说，锯齿是一种信号采样过程中发生的现象。比传感器像素小的被摄体结构看起来比像素不平滑。例如，在下图左侧的低分辨率图像中，锐化边缘呈现出“锯齿状”，在照片中形成了实际不存在的图案。当用户拍摄的图像通过深度学习图像处理技术进行高清放大时，神经网络会检测在源图像细节处发现的锯齿痕迹，并估算将其放大成高分辨率图像所需的像素。总之，从“训练”阶段开始的学习是为了避免渲染/生成错误的被摄体结构(此处的被摄体实际上是平滑的或空白的)。

图4-1



备注:

此外，近年来使用扩散模型或类似模型的人工智能技术已经问世，并日渐普及，这些技术可以根据文字描述的所需图像特征，生成全新的图像。例如，一张原本非常模糊、面部特征难以识别的低分辨率源图像，人工智能技术可以在此基础上生成具有睫毛、皱纹等细节的肖像。

在这种情况下，我们假设一张面部图像，分辨率很低，也无法识别。类似的高分辨率的面部图像有无数个版本，这些图像的分辨率一旦降低，就会变得和上述低分辨率图像相同。即使一张低分辨率的面部没有睫毛或皱纹等细微的面部特征痕迹，人工智能也能识别出这是一张面部图像，并能在其训练范围内生成类似的版本。因此，不能保证创建的高分辨率面部会与原始低分辨率面部的真实细节一致。应用人工智能技术，不仅可以提高人脸的分辨率，即使是被摄体结构因模糊导致无法识别的图像，也能变得清晰，例如下图4-2中的对比。

整个场景(A)的图像是对焦于最右侧的酒瓶拍摄的，随着画面向左侧的酒瓶移动，画面逐渐脱焦，不在景深(DOF)范围内，所以变得越来越模糊。下方的图像(B到E)对应整张照片(A)中的红框区域。因此，通过其他处理方法高清放大的图像(D)，虽然图像分辨率更高但没有必要，失去了虚化效果，导致产生的图像与实际瓶子标签状态不同，有悖于原始的摄影意图。相比之下，佳能深度学习图像处理技术(C)与这些方法截然不同。如前所述，图像处理技术反映了摄影过程模拟，并基于不同分辨率图像元素之间的关系，在不失去虚化表现力的前提下产生自然的结果。

图4-2



A: 整张图像  
\* 红框区域未合焦，所以模糊。



B: 典型处理  
(双三次插值)



C: 佳能深度学习



D: 其他方法的示例



E: 参考：合焦的图像

备注：

佳能是影像行业的领军企业，在开发产品时始终以摄影画质作为第一要务，致力于将“真实生活景象”还原到细微之处。佳能会继续追求技术进步，为佳能相机和镜头用户提供满意的摄影体验。

## [参考]

### 深度学习图像处理技术

<https://global.canon/en/technology/dl-ip/technology-2023.html>

### 高清放大技术

<https://global.canon/en/technology/dl-upscaling-2023.html>

### 神经网络图像处理工具(Neural network Image Processing Tool)

<https://sas.image.canon/st/nnip.html>

### 神经网络升级工具(Neural network Upscaling Tool)

<https://sas.image.canon/st/nnups.html>

- \* 本文档中描述的“深度学习图像处理技术/高清放大技术”概述了佳能公司截至2023年11月的技术。因此，因技术改进，本文档的内容在今后可能有所变更，敬请留意。
- \* 根据被摄体和拍摄条件的不同，效果和结果可能与本文档中的描述有所差异，敬请谅解。

---

备注：