Orest Kupyn^{1,3}, Volodymyr Budzan^{1,3}, Mykola Mykhailych¹, Dmytro Mishkin², Jiři Matas²

¹ Ukrainian Catholic University, Lviv, Ukraine {kupyn, budzan, mykhailych}@ucu.edu.ua

² Visual Recognition Group, Center for Machine Perception, FEE, CTU in Prague {mishkdmy, matas}@cmp.felk.cvut.cz

³ ELEKS Ltd.

摘要

我们提出 DeblurGAN, 一种用于运动去模糊的端到端学 习方法, 这种方法基于条件 GAN 和内容损失。 DeblurGAN 在结构相似性度量和视觉外观方面实现了最先进的性能。这 种去模糊模型的质量也以一种新颖的方式评估现实世界的问 题 - 对(去)模糊图像的物体检测。 该方法比最接近的竞争 对手 - Deep-Deblur [25]快 5 倍。 我们还介绍了一种新方 法, 用于生成锐利的合成运动模糊图像, 从而实现真实的数 据集增强。

有关模型,代码和数据集,请访问: https://github.com/KupynOrest/DeblurGAN

1. 介绍

这项工作是对单个照片的随机运动进行去模糊。最近 通过应用生成对抗网络(GAN)在图像超分辨率[20]和绘画 [45]的相关领域取得了重大进展[10]。 GAN 以能够保留图 像中的纹理细节,创建接近真实图像流形的解决方案并且在 感知上具有说服力而闻名。受近期图像超分辨率[20]和生成 对抗性网络图像到图像翻译[16]的启发,我们将去模糊视为 这种图像到图像翻译的特例。我们介绍 DeblurGAN - 一种 基于条件生成对抗网络[24]和多组件损失函数的方法。与以 前的工作不同,我们使用 Wasserstein GAN [2],梯度罚分 [11]和感知损失[17]。这推动了在感知上难以与真实清晰图 像区分的解决方案的产生,并且与使用传统 MSE 或 MAE 作为优化目标相比,这能够恢复更精细的纹理细节。



图 1: DeblurGAN 有助于对象检测。 YOLO [30]检测模糊图像(顶部), DeblurGAN 重新存储(中间)和来自 GoPro [25]数据集的锐利完全真实图像。

翻译



图 2: DeblurGAN 处理的 GoPro 图像[25]。 模糊 - 左, DeblurGAN - 中心, 锐利真实图像 - 右。

我们做了三个贡献。 首先,我们提出了一种损失和架 构,它可以获得最先进的动态去模糊结果,同时比最快的竞 争对手快 5 倍。 其次,我们提出了一种基于随机轨迹的方 法,用于从一组锐利的图像中以自动方式生成运动去模糊训 练的数据集。 我们表明,与仅对真实世界图像进行训练相 比,将其与用于运动去模糊学习的现有数据集相结合可以改 善结果。 最后,我们提出了一种新的数据集和方法,用于 评估去模糊算法,基于它们如何改善对象检测结果。

2. 相关工作

2.1. 图像去模糊

非均匀模糊模型的通用公式如下:

IB = k(M) *IS + N; (1) 其中 IB 是模糊图像, k (M) 是由运动场 M 确定的未 知模糊核.IS 是清晰的潜像, 表示卷积, N 是加性噪声。

去模糊问题分为两种类型:盲目和非盲目去模糊。早期的工作[37]主要集中在非盲目去模糊,假设这样做

模糊核 k (M) 是已知的。大多数依赖于经典的 Lucy-Richardson 算法, Wiener 或 Tikhonov 滤波器来执行反卷 积运算并获得 IS 估计。通常模糊函数是未知的,并且盲去 模糊算法估计潜在的清晰图像 IS 和模糊核 k (M) 。为每个 像素找到模糊函数是一个不适定的问题,并且大多数现有的 算法都依赖于启发式,图像统计和模糊源的假设。这些方法 系列通过考虑模糊在整个图像上的模糊来解决由相机抖动引 起的模糊。首先,根据诱导模糊核估计相机运动,然后通过 执行反卷积操作来反转效果。从 Fergus 等人的成功开始 [8], 过去十年来已经开发了许多方法[44] [42] [28] [3]。一 些方法基于迭代方法[8] [44],它通过使用参数先验模型改 进了每次迭代的运动内核和清晰图像的估计。然而,运行时 间以及停止标准对于那些类型的算法来说是一个重要问题。 其他人使用模糊函数的局部线性假设和简单的启发式来快速 估计未知内核。这些方法速度很快,但在一小部分图像上运 行良好。

最近, Whyte 等人 [40]开发了一种新颖的算法,



图 3: DeblurGAN 生成器架构。 DeblurGAN 包含两个跨步卷积块,其中有步幅 12,9 个残差块[13]和两个转置卷积块。 每个 ResBlock 由卷积 层,实例规范化层和 ReLU 激活组成。

对于基于曝光期间相机的旋转速度的模糊处理的参数化 几何模型的非均匀盲目去模糊。同时 Gupta 等人 [12]假设 模糊仅由 3D 相机移动引起。随着深度学习的成功,在过去 几年中,出现了一些基于卷积神经网络(CNN)的方法。 孙等人 [36]使用 CNN 估计模糊核,Chakrabarti [6]预测运 动核的复傅立叶系数在傅立叶空间中执行非盲去模糊,而 Gong [9]使用完全卷积网络移动运动流估计。所有这些方法 都使用 CNN 来估计未知模糊函数。最近,Noorozi [27]和 Nah [25]采用无内核的端到端方法,使用多尺度 CNN 直接 去除图像。Ramakrishnan 等人 [29]使用 pix2pix 框架[16] 和密集连接的卷积网络[15]的组合来执行盲目无核图像去模 糊。这些方法能够处理模糊的不同来源

2.2. 生成对抗网络

Goodfellow 等人介绍了生成对抗网络的概念[10],是在 两个竞争网络之间定义一个游戏:鉴别器和发生器。发生 器接收噪声作为输入并生成样本。鉴别器接收真实和生成 的样本,并试图区分它们。生成器的目标是通过生成无法 与真实样本区分开的令人信服的样本来欺骗鉴别器。生成 器 G 和鉴别器 D 之间的游戏是最小极大目标:

 $\min_{G} \max_{D} \mathop{\mathbb{E}}_{x \sim \mathbb{P}_{a}} [\log(D(x))] + \mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_{a}} [\log(1 - D(\tilde{x}))] \quad (2)$

其中 Pr 是数据分布, Pg 是模型分布, 由 x~ = G (z) 定义; z v P (z), 输入 z 是来自简单噪声分布的样本。 GAN 以其生成具有良好感知质量的样品的能力而闻名, 然 而, Vanilla 版的训练受到影响, 如[33]中所述,诸如模式崩溃,消失的梯度等许多问题。 最 小化 GAN 中的值函数等于最小化 x 上的数据和模型分布之 间的 Jensen-Shannon 差异。 Arjovsky 等[2]讨论了由 JS 散度近似引起的 GAN 训练中的困难,并提出使用 Earth-Mover (也称为 Wasserstein-1) 距离 W (q; p),WGAN 的价值函数是使用 Kantorovich-Rubinstein 二元性[39]构 造的:

$$\min_{G} \max_{D \in \mathcal{D}} \mathop{\mathbb{E}}_{x \sim \mathbb{P}_r} [D(x)] - \mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_q} [D(\tilde{x})]$$
(3)

其中 D 是 1-Lipschitz 函数的集合而 Pg 又是模型分布 这里的想法是批评值近似为 KW (Pr; Pθ), 其中 K 是 Lipschitz 常数, W (Pr; Pθ) 是 Wasserstein 距离。 在这 种设置中,鉴别器网络被称为批评者,它近似于样本之间的 距离。 在 WGAN Arjovsky 等人中强制执行 Lips-chitz 约 束。添加权重剪辑到[-c;c]。 Gulrajani 等[11]建议添加一 个梯度惩罚术语:

$$\lambda \mathop{\mathbb{E}}_{\tilde{x} \sim \mathbb{P}_{\tilde{x}}} \left[(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2 \right] \tag{4}$$

值函数作为强制 Lipschitz 约束的替代方法。 这种方法 对于发电机架构的选择是稳健的,并且几乎不需要超级调 整。 这对于图像去模糊是至关重要的,因为它可以使用新 颖的轻量级神经网络架构,而不是标准的 Deep ResNet 架 构,之前用于图像去模糊[25]。

2.3.条件对抗网络

生成性对抗网络已应用于不同的图像到图像转换问题,例如 超分辨率[20],样式转换[22],产品照片生成[5]等。 lsola[16]提 供了这些方法的详细概述,并提出了条件 GAN 架构,也称为 pix2pix。与 VanillaGAN 不同, cGAN 学习从观察图像 x 和随机噪声向量 z 到 y: G: x, z-→y 的映射。 lsola 等还对判别器设置了一个条件,并使用 U-net 架构[31]进行生成器和马尔可夫判别器,它允许在许 多任务上获得感知上优越的结果,包括从标签图合成照片, 从边缘图重建对象,以及着色年龄。

3. 本文算法

目标是在仅给出模糊图像 IB 作为输入的情况下恢复清晰 图像 IS,因此不提供关于模糊内核的信息。 剥离由经过训 练的 CNN Geg 完成,我们将其称为生成器。 对于每个 IB, 它估计相应的 IS 图像。 此外,在训练阶段,我们引入批评 网络 Dep 并以对抗方式训练两个网络。

3.1. 损失函数

我们将损失函数表示为内容损失和对抗损失之和:

$$\mathcal{L} = \underbrace{\mathcal{L}_{GAN}}_{adv \ loss} + \underbrace{\lambda \cdot \mathcal{L}_X}_{content \ loss}$$
(5)

λ在所有实验中等于 100。 与 Isola 等人 [16]不同, 我 们不对鉴别器进行条件限制, 因为我们不需要惩罚输入和输 出之间的不匹配。 对抗性损失大多数与条件性 GAN 相关的 论文使用 VanillaGAN 目标作为损失[20] [25]函数。 最近 [47]提供了一种使用最小方形 GAN [23]的替代方法, 它更 稳定并产生更高质量的结果。 我们使用 WGAN-GP [11]作 为判别函数, 它被证明对生成器架构的选择是稳健的[2]。 我们使用不同架构的预先实验证实了这些结果, 我们能够使 用比 ResNet152 轻得多的架构[25], 请参阅下一小节。 损 失计算如下:

$$\mathcal{L}_{GAN} = \sum_{n=1}^{N} -D_{\theta_D}(G_{\theta_G}(I^B)) \tag{6}$$

在没有 GAN 组件的情况下训练的 DeblurGAN 会聚,但会产生平滑且模糊的图像。

内容损失,内容损失函数的两个经典选择是 L1 或 MAE 丢失,原始像素上的 L2 或 MSE 损失。由于像素空间中可 能的解决方案的像素平均值,使用这些函数作为唯一的优化 目标导致生成的图像上的模糊伪像[20]。相反,我们采用了 最近提出的感知损失[17]。感知损失是一种简单的 L2 丢 失,但是基于生成的和目标图像 CNN 特征图的差异。它定 义如下:



图 4: DeblurGAN 培训。 生成器网络将模糊图像作为输入并产生 清晰图像的估计。 判别器网络采用恢复和清晰的图像并输出它们之 间的距离。 总损失包括判别器的 WGAN 损失和人均损失[17]。 感 知损失是锐利和恢复图像的 VGG-19 [34] conv3.3 特征图之间的差 异。 在测试时,只保留生成器。

$$\mathcal{L}_X = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^S)_{x,y} - \phi_{i,j}(G_{\theta_G}(I^B))_{x,y})^2$$

其中 i; j 是在 VGG19 网络中第 i 个最大化层之前的第 j 个转换(激活之后)获得的特征映射,在 ImageNet [7]上 预训练,Wi; j和 Hi; j 是维度功能图。在我们的工作中,我 们使用来自 VGG3; 3 卷积层的激活。更深层次的激活代表 了更高抽象的特征[46] [20]。感知损失侧重于恢复一般内 容[16] [20],



图 5: 生成的相机运动轨迹和模糊内核以及相应的模糊图像的示例。

而对抗性损失则侧重于恢复纹理细节。 在没有感知损失 或像素上使用简单 MSE 训练的 Deblur-GAN 不会收敛到有 意义的状态。

额外的正规化。我们还试图添加电视正规化和训练的模型,产生更差的性能 - 在 GoPro 数据集上 27.9 对 28.7w / o PSNR。

3.2. 网络架构

生成器 CNN 架构如图 3 所示。它类似于 Johnson 等人 提出的结构[17]用于风格转移任务。它包含两个带有步幅 1/2 的跨步卷积块,九个残余块[13] (ResBlocks)和两个转 置卷积块。每个 ResBlock 由卷积层,实例标准化层[38]和 ReLU [26]激活层组成。在每个 ResBlock 中的第一个卷积层 之后添加概率为 0.5 的丢失[35]正则化。另外,附含有被称 为 ResOut 的全局跳过连接。 CNN 学习模糊图像 IB 的残差 校正 IR,因此 IS = IB + IR。我们发现这样的公式使得训练 更快,并且使得模型更好地收敛。在训练阶段,我们定义了 一个判别器网络 D,它是 Wasserstein GAN [2],带有梯度 惩罚[11],我们称之为 WGAN-GP。判别网络的架构与 PatchGAN 相同[16,22]。除了最后一层之外的所有卷积层后 面都是 InstanceNorm 层, LeakyReLU [41]的参数α=0.2。

4. 运动模糊生成

没有简单的方法来获得用于训练的相应锐利和模糊图像 的图像对。获得用于训练的图像对的典型方法是使用高帧率 相机来使用来自视频的平均锐利帧来模拟模糊[27], 25]。 它允许创建逼真的模糊图像,但仅将图像空间限制为拍摄视 频中存在的场景,并使得缩放数据集变得复杂。 孙等人[36] 通过将清晰的自然图像与 73 种可能的线性运动核中的一种 进行卷积来创建合成模糊的图像,Xu 等[43]也使用线性运 动内核来创建合成模糊的图像。



图 6: 顶行:由 Fergus 等人估计的来自真实世界图像的模糊核[8]。 底行:通过我们的方法合成生成的内核。 我们的随机方法可以模拟 具有不同非线性水平的各种逼真模糊核。

Chakrabarti [6]通过对 6 个随机点进行采样并将样条拟 合到它们来创建模糊核。 我们更进一步提出了一种方法, 它可以模拟更真实和复杂的模糊内核。 我们遵循 Borac-chi 和 Foi [4]描述的随机轨迹生成的想法。 然后通过将子像素 插值应用于轨迹矢量来生成核。 每个轨迹矢量是复值矢 量,其对应于在连续域中的 2D 随机运动之后的对象的离散 位置。 轨迹生成是通过马尔可夫过程完成的,总结在算法 1 中。轨迹的下一个点的位置是基于先前的点速度和位置,高 斯扰动,脉冲扰动和确定性惯性分量随机生成的。

5. 训练细节

我们使用 PyTorch [1]深度学习框架实现了所有模型。 训练是在单个 Maxwell GTX Titan-X GPU 上使用三个进行的



图 7: GoPro 测试数据集的结果。从左到右:照片模糊, Nah 等 [25], DeblurGAN。



图 8: Kohler 数据集的结果。从左到右:照片模糊, Nah 等 [25], DeblurGAN。

不同的数据集。 我们称之为 DeblurGANWILD 的第一个模型是根据 1000 个 GoPro 训练数据集中的 256x256 大小的随机作物进行训练的

图像[25]缩小了两倍。 第二个 DeblurGANSynth 接受了来 自 MS COCO 数据集的 256x256 补丁的培训,该数据集通 过以前的方法呈现

Algorithm 1 Motion blur kernel generation.					
Parameters:					
M = 2000 - number of iterations,					
$L_{max} = 60 - \max$ length of the movement,					
$p_s = 0.001 - \text{probability of impulsive shake},$					
I – inertia term, uniform from (0,0.7),					
p_b – probability of big shake, uniform from (0,0.2),					
p_g – probability of gaussian shake, uniform from (0,0.7),					
ϕ – initial angle, uniform from (0,2 π),					
x - trajectory vector.					
1: procedure BLUR(Img, M, L_{max}, p_s)					
2: $v_0 \leftarrow \cos(\phi) + \sin(\phi) * i$					
3: $v \leftarrow v_o * L_{max}/(M-1)$					
4: $x = \operatorname{zeros}(M, 1)$					
5: for $t = 1$ to $M - 1$ do					
6: if randn $< p_b * p_s$ then					
7: $\operatorname{nextDir} \leftarrow 2 \cdot v \cdot e^{i * (\pi + (\operatorname{randn} - 0.5))})$					
8: else:					
9: $nextDir \leftarrow 0$					
10: $dv \leftarrow \text{nextDir} + p_s * (p_g * (\text{randn} + i * \text{randn}) *$					
$I * x[t] * (L_{max}/(M-1))$					
11: $v \leftarrow v + dv$					
12: $v \leftarrow (v/abs(v)) * L_{max}/(M-1)$					
13: $x[t+1] \leftarrow x[t] + v$					
14: Kernel \leftarrow sub pixel interpolation (x)					
15: Blurred image $\leftarrow conv(Kernel, Img)$					
16: return Blurred image					

部分。 我们还对合成模糊图像和野外拍摄的图像进行了 DeblurGANComb 的组合训练,其中合成生成的图像与高帧 率相机拍摄的图像的比例为 2:1。由于模型是完全卷积的 并且在图像块上训练,因此它们可以应用于任意大小的图 像。为了优化,我们遵循[2]的方法并在 D 上执行 5 个梯度 下降步骤,然后在 G 上执行一步,使用 Adam [18]作为求解 器。对于生成器和判别器来说,学习率最初设定为 10 4。 在前 150 个时期之后,我们在接下来的 150 个时期内将速 率线性衰减为零。 在推理时,我们遵循[16]的思想,并应用 Dropout 和 Instance Normalization。所有模型都经过批量 大小= 1 的训练,这表明在验证方面效果更好。 训练阶段用 于训练一个 DeblurGAN 网络需要 6 天。

6. 实验评估

6.1. GoPro 数据集

GoPro 数据集[25]包含 2103 对 720p 质量的模糊和清晰图像,取自各种场景。 我们将模型的结果与最先进的模型[36],[25]在标准指标上进行比较,并且还显示了

表 1:峰值信噪比和结构相似性度量,是 1111 张图像的 GoPro 测试数据集的平均值。所有模型都在线性图像子集上进行测试。 Nah 等人的最新结果 [25]在 gamma 子集上获得。

	Sun et al.	Nah et al.	Xu et al.	DeblurGAN		
Metric	[36]	[25]	[44]	WILD	Synth	Comb
PSNR	24.6	28.3/29.1*	25.1	27.2	23.6	28.7
SSIM	0.842	0.916	0.89	0.954	0.884	0.958
Time	20 min	4.33 s	13.41 s		0.85 s	

每个算法在单个 GPU 上的运行时间。 结果见表 1。 DeblurGAN 在结构化自相似性方面表现出优异的结果,接 近于峰值信噪比的最新技术水平,并通过目视检查提供更好 的结果。 与其他神经模型相比,我们的网络在像素空间中 不使用 L2 距离,因此不直接针对 PSNR 度量进行优化。 它 可以处理由相机抖动和物体移动引起的模糊,在内核估计方 法中不会受到常见伪影的影响,同时与多尺度 CNN 相比, 参数减少了 6 倍以上,从而大大加快了推理速度。 图 7 显 示了 GoPro 数据集上的测试图像。

6.2. Kohler 数据集

Kohler 数据集[19]由 4 个图像组成,每个图像用 12 个 不同的内核模糊。 这是用于评估盲目去模糊算法的标准基 准数据集。 通过记录和分析真实的相机运动来生成数据 集,该实际相机运动在机器人平台上回放,使得记录一系列 清晰图像以对 6D 相机运动轨迹进行采样。 结果在表 2 中, 类似于 GoPro 评估。

6.3. YOLO 上的对象检测基准

物体检测是计算机视觉中研究最充分的问题之一,应用于从 自动驾驶到安全的不同信号。 在过去几年中,基于深度卷积神 经网络的方法与传统方法相比显示出最先进的性能。 然而,这 些网络是在有限的数据集上进行训练的,在实际环境中,图像 会因不同的伪影而降级,包括运动模糊。与[21]和[32]相似,我 们研究了运动模糊对物体检测的影响并提出了建议。 一种在预 训练的 YOLO [30]网络上基于目标检测结果评估去模糊算法质 量的新方法。

为此,我们通过使用高帧率框架模拟相机抖动来构建一个锐利和模糊街景的数据集,

表 2:峰值信噪比和结构相似性度量,是 Kohler 数据集的平均值。徐等人 [44]和怀特等人 [40]是非 CNN 盲目去模糊方法,而 Sun 等人 [36] 和 Nah 等人 [25]使用 CNN。



(a) Blurred photo

(b) Nah et al. [25] (c) DeblurGAN 图 9: 去模糊前后的 YOLO 物体检测

(d) Sharp photo

在[14] [25] [27]之后,我们采用 240fps 摄像机拍摄的 5 到 25 帧之间的随机数,并计算中间帧的模糊版本作为这些帧 的平均值。所有帧都用= 2:2 进行 gamma 校正,然后采 用反函数来获得最终的模糊帧。总体而言,该数据集由 410 对模糊和清晰的图像组成,这些图像取自街道和停车位,具 有不同数量和类型的汽车。

模糊源包括相机抖动和汽车运动引起的模糊。数据集和 补充代码可在线获取。然后将清晰的图像输入 YOLO 网 络,并将视觉验证后的结果指定为基础事实。然后 YOLO 在模糊和恢复的图像版本上运行,并且计算获得的结果和地 面实况之间的平均召回和精确度。与标准 PSNR 度量相 比,该方法对应于实际问题的去模糊模型的质量,并与所生 成图像的视觉质量和清晰度相关联。通常,模糊图像的精度 较高,因为没有锐利的物体边界,并且没有检测到较小的物 体,如图 9 所示。

结果显示在表 3 中。在召回和 F1 得分方面 DeblurGAN 明显优于竞争对手。

7. 结论

我们描述了一种无内核盲目运动去模糊学习方法,并介 绍了 DeblurGAN,它是一种使用多分量损失函数优化的条 件性对抗网络。除此之外,我们 表 3: 使用 DeblurGAN 和 Nah 等人对模糊和恢复的照片进行 YOLO [30]物体检测的结果 [25]算法。相应清晰图像的结果被认为 是基本事实。 DeblurGAN 的召回率和 F1 得分高于竞争对手。

Method	prec.	recall	F1 score
no deblur	0.821	0.437	0.570
Nah <i>et al</i> . [25]	0.834	0.552	0.665
DeblurGAN WILD	0.764	0.631	0.691
DeblurGAN synth	0.801	0.517	0.628
DeblurGAN comb	0.671	0.742	0.704

实现了一种新方法,用于创建逼真的合成运动模糊,能够模 拟不同的模糊源。 我们根据物体检测结果引入了一个新的 基准和评估协议,并表明 DeblurGAN 显着有助于检测模糊 图像。

8. 致谢

作者得到了 ELEKS 有限公司, ARVI 实验室, 捷克科学基 金项目 GACR P103 / 12 / G084, 奥地利交通运输部, 创新和 技术部, 联邦科学, 研究和经济部以及 上奥地利州在 COMET 中心框架内, CTU 学生资助 SGS17 / 185 / OHK3 / 3T / 13。 我们感谢 Huaijin Chen 发现峰值信噪比评估中的错误。

参考文献

- [1] PyTorch. http://pytorch.org. 5
- [2] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein GAN. ArXiv e-prints. Jan. 2017, 1, 3, 4, 5, 7
- [3] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos. Bayesian blind deconvolution with general sparse image priors. In European Conference on Computer Vision (ECCV), Firenze, Italy, October 2012. Springer. 2
- [4] G. Boracchi and A. Foi. Modeling the performance of image restoration from motion blur. Image Processing, IEEE Transactions on, 21(8):3502 –3517, aug. 2012. 5
- [5] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Kr-ishnan. Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks. ArXiv e-prints, Dec. 2016. 3
- [6] A. Chakrabarti. A neural approach to blind motion deblurring. In Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016. 3, 5
- [7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009. 4
- [8] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. ACM Trans. Graph., 25(3):787–794, July 2006. 2, 5
- [9] D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. Van Den Hengel, and Q. Shi. From Motion Blur to Motion Flow: a Deep Learning Solution for Removing Heterogeneous Mo-tion Blur. 2016. 3
- [10] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Gen-erative Adversarial Networks. June 2014. 1, 3
- [11] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved Training of Wasserstein GANs. ArXiv e-prints, Mar. 2017. 1, 3, 4, 5
- [12] A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV'10, pages 171–184, Berlin, Heidel-berg, 2010. Springer-Verlag. 3
- [13] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learn-ing for image recognition. arXiv preprint arXiv:1512.03385, 2015. 3, 5
- [14] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf. Fast removal of non-uniform camera shake. In Proceedings of the 2011 International Conference on Computer Vision, ICCV '11, pages 463–470, Washington, DC, USA, 2011. IEEE Computer Society. 8
- [15] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recogni-tion (CVPR), pages 2261–2269, 2017. 3
- [16] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Imageto-image translation with conditional adversarial networks. arxiv, 2016. 1, 3, 4, 5, 7

- [17] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In European Conference on Computer Vision, 2016. 1, 4, 5
- [18] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. CoRR, abs/1412.6980, 2014. 7
- [19] R. Kohler, M. Hirsch, B. Mohler, B. Scholkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In Proceedings of the 12th European Conference on Computer Vision - Volume Part VII, ECCV'12, pages 27– 40, Berlin, Heidelberg, 2012. Springer-Verlag. 7
- [20] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunning-ham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and
 W. Shi. Photo-Realistic Single Image Super-Resolution Us-ing a Generative Adversarial Network. ArXiv e-prints, Sept. 2016. 1, 3, 4
- [21] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. An All-in-One Network for Dehazing and Beyond. ArXiv e-prints, July 2017. 7
- [22] C. Li and M. Wand. Precomputed Real-Time Texture Syn-thesis with Markovian Generative Adversarial Networks. ArXiv e-prints, Apr. 2016. 3, 5
- [23] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, and Z. Wang. Least squares generative adversarial networks, 2016. cite arxiv:1611.04076. 4
- [24] M. Mirza and S. Osindero. Conditional generative adversar-ial nets. CoRR, abs/1411.1784, 2014. 1
- [25] S. Nah, T. Hyun, K. Kyoung, and M. Lee. Deep Multiscale Convolutional Neural Network for Dynamic Scene Deblur-ring. 2016. 1, 2, 3, 4, 5, 6, 7, 8
- [26] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In International Conference on Machine Learning (ICML), pages 807–814, 2010. 5
- [27] M. Noroozi, P. Chandramouli, and P. Favaro. Motion De-blurring in the Wild. 2017. 3, 5, 8
- [28] D. Perrone and P. Favaro. Total variation blind deconvolution: The devil is in the details. In EEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014. 2
- [29] S. Ramakrishnan, S. Pachori, A. Gangopadhyay, and S. Ra-man. Deep generative filter for motion deblurring. 2017 IEEE International Conference on Computer Vision Work-shops (ICCVW), pages 2993–3000, 2017. 3
- [30] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You Only Look Once: Unified, Real-Time Object Detection. ArXiv e-prints, June 2015. 1, 7, 8
- [31] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolu-tional Networks for Biomedical Image Segmentation. ArXiv e-prints, May 2015. 4
- [32] M. S. M. Sajjadi, B. Scholkopf," and M. Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. 2017 IEEE International Conference on Computer Vision (ICCV), pages 4501–4510, 2017. 7
- [33] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Rad-ford, and X. Chen. Improved Techniques for Training GANs. ArXiv e-prints, June 2016. 3
- [34] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv e-prints, Sept. 2014. 4

- [35] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res., 15(1):1929–1958, Jan. 2014. 5
- [36] J. Sun, W. Cao, Z. Xu, and J. Ponce. Learning a Convo-lutional Neural Network for Non-uniform Motion Blur Re-moval. 2015. 3, 5, 7, 8
- [37] R. Szeliski. Computer Vision: Algorithms and Applications. Springer-Verlag New York, Inc., New York, NY, USA, 1st edition, 2010. 2
- [38] D. Ulyanov, A. Vedaldi, and V. S. Lempitsky. Instance normalization: The missing ingredient for fast stylization. CoRR, abs/1607.08022, 2016. 5
- [39] C. Villani. Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften. Springer Berlin Hei-delberg, 2008. 3
- [40] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Nonuniform deblurring for shaken images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010. 2, 8
- [41] B. Xu, N. Wang, T. Chen, and M. Li. Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853, 2015. 5
- [42] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2010. 2
- [43] L. Xu, J. S. J. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1, NIPS'14, pages 1790–1798, Cambridge, MA, USA, 2014. MIT Press. 5
- [44] L. Xu, S. Zheng, and J. Jia. Unnatural L0 Sparse Representation for Natural Image Deblurring. 2013. 2, 7, 8
- [45] R. A. Yeh, C. Chen, T. Lim, M. Hasegawa-Johnson, and M. N. Do. Semantic image inpainting with perceptual and contextual losses. CoRR, abs/1607.07539, 2016. 1
- [46] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. CoRR, abs/1311.2901, 2013. 4
- [47] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired imageto-image translation using cycle-consistent adversarial networks. arXiv preprint arXiv:1703.10593, 2017. 4