arXiv:2304.08269v1 [cs.MM] 17 Apr 2023

图像编解码器强幂等性评估

Qian Zhang⋆† *Tongda Xu*⋆1 *Yanghao Li*⋆ *Yan Wang*⋆2

⋆ 清华大学人工智能产业研究院（AIR），清华大学 † 西安交通大学

摘要

本文首先基于幂等编解码器提出强幂等编解码器的概念。编解码器的幂等性是指编解码器对重新压缩的稳定性。类似地，我们将编解码器的强幂等性定义为其对多质量重新压缩的稳定性，这是云端转码场景中编解码器的一项重要特性。我们给出了一个源分布已知的强幂等编解码器的详细示例。此外，我们制定了一套测试协议，用于评估图像编解码器的强幂等性。最后，我们评估了当前图像编解码器的强幂等性，包括传统编解码器和近期的神经编解码器。实验结果表明，当前图像编解码器在不同质量下进行多轮重新压缩时并不稳定，即使它们接近幂等。

索引词— 图像编解码器，幂等性

1. 引言

近年来，图像编解码器的幂等性屡次受到关注 [1, 2]。图像编解码器的幂等性是指编解码器在经过多轮重新压缩后保持重建质量的能力。更具体地说，我们将原始图像记为 $\boldsymbol{x}$，将编码器—解码器对记为 $f(.)$，将重建图像记为 $f(\boldsymbol{x})$。如果编解码器满足以下条件，我们就称其具有幂等性：

$$f(\boldsymbol{x}) = f(f(\boldsymbol{x})) \tag{1}$$

正如先前的研究所述， [1, 2], JPEG 等传统编解码标准 [3] 具有良好的幂等性，神经图像压缩的 [4, 5, 6]幂等性也可以得到改善。然而，在实际应用中，幂等编解码器仍不足以在多轮重新压缩过程中完全保留图像质量。例如，在社交网络场景中，用户可能使用不同的设备，并处于不同的网络条件下。不同用户可能获取图像后，使用自己的设备以不同的比特率对其进行重新压缩，然后再将图像转发给其他用户。因此，图像可能会以不同的比特率和质量因子被多次重新压缩。在这种情况下，幂等图像编解码器可能无法确保图像在不同质量下重新压缩时的图像质量。

为解决这一问题，我们提出强幂等性的概念。它是在考虑质量因子的基础上对幂等性的扩展。强幂等性对幂等性进行了推广，更适用于实际通信场景。我们正式定义了一种度量来评估强幂等性，并提供了测试协议。最后，我们使用该度量评估了当前图像编解码器的强幂等性。实验结果表明，即使是满足幂等性的传统编解码器，在不同质量因子下进行多轮重新压缩也颇具挑战。

2. 相关工作

对于传统编解码器，自 2000 [7]以来，人们便开始讨论其幂等性。当前的 JPEG、JPEG2000 和 MPEG 标准均不具备幂等性 [8], ，但可通过裁剪和调整变换矩阵来改善其幂等性 [9, 10, 11, 12]。另一方面，JPEG XS[13, 14] 在2017年明确要求具备多代鲁棒性，并将其定义为最多经过 10 次编解码循环后不会出现明显的质量下降。对于传统视频编解码器， [15]提出了一种通过裁剪补偿矩阵实现幂等性的 H.264 编码器。

对于较新的学习式图像压缩（LIC）， [1]首先建立了一个用于评估 LIC 方法幂等性的基准测试，并提出了一种用于提升鲁棒性的损失函数。在此基础上， [16]随后提出了压缩方法实现幂等性的两个条件。 [2] 采用直接量化、可逆性损失函数和通道松弛方法来减少代际损失。然而，以上工作都聚焦于相同质量下的多轮压缩，并未探讨多质量重新压缩时的鲁棒性。

3. 强幂等性

3.1. 强幂等性的定义与性质

我们希望用强幂等性来描述图像编解码器在不同质量因子下经受重新压缩时的稳定性。形式上，我们将原始图像记为 $\boldsymbol{x}$，将质量因子为 q 的编码器-解码器对记为 $f(。, q)$，并将重建图像记为

1. Qian Zhang 和 Tongda Xu 的贡献相同。2. 通信事宜请联系相应联系人。

$f(\boldsymbol{x},\ q)$。我们注意到，q 是质量因子，而非量化步长，因此较低的 q 意味着较低的质量。接着，如果编解码器满足以下条件，我们就称其具有强幂等性：

$$f(\boldsymbol{x}, q_{min}) = f(f(\boldsymbol{x}, q_1), ..., q_k),$$
$$\text{where } q_{min} = \min\{q_1, ..., q_k\}. \quad (2)$$

公式 2 表明，经过质量因子 q 不同的多轮重新压缩后，强幂等编解码器的重建结果与以最低质量因子进行单次压缩的重建结果相同。此外，强幂等编解码器具有以下性质：

1. **强幂等编解码器是幂等的**。显然，只要将 $q_1, ...q_k$ 设为相同值，强幂等性就意味着幂等性。

2. **强幂等编解码器的比特率隐式有界**。尽管只限制了重建质量，比特率也会受到隐式约束。考虑一个幂等编解码器 $f(.)$，那么压缩 $f(x)$ 的比特率在压缩过程中保持不变。由于强幂等编解码器也是幂等的，我们可以始终对 $f(\boldsymbol{x}, q_{min})$ 进行重新压缩，而不改变比特率。这意味着我们可以在公式 2 右侧增加质量为 q_{min} 的额外重新压缩，从而获得与压缩 $f(\boldsymbol{x}, q_{min})$ 相同的比特率。该性质保证，依次以 $q_1, ..., q_k$ 压缩 $\boldsymbol{x}$，所得比特率与压缩 $f(\boldsymbol{x}, q_{min})$ 相同。

3.2. 玩具示例

现在，我们通过一个强幂等编解码器的玩具示例，更好地理解其性质。假设我们有一个一维源信号 $x \sim \mathcal{U}(0,\ 1)$，它服从单位均匀分布。对于一个具有 4 个标量量化码字点、码率为 log 4 的编解码器，一种可能的选择是在区域 $[0, 1]$ 内均匀选取 4 个码字点 $x_1 = 1/8,\ x_2 = 3/8,\ x_3 = 5/8,\ x_4 = 7/8$。而 $[x_i - 1/8,\ x_i + 1/8]$ 内的任意值都会被量化为 x_i。我们假设该编解码器有 3 个质量点：log 4 是最高码率，对应质量 q_3；log 2 是中间码率，对应质量 q_2；log 1 是最低码率，对应质量 q_1。那么问题就变成了如何设计 q_2、q_1 的码字，使该编解码器达到强幂等。

考虑到我们将大小为 2 的码字点设计在大小为 4 的码字点之后，可以得到图1.A 所示的编解码器。该编解码器显然不是强幂等的。假设我们先使用 q_1 对 x 进行压缩，然后再使用 q_3 进行压缩，最终的重建结果将是 $3/8$ 或 $5/8$，而不是 $1/2$。然而，如果仅使用 q_1 对 x 进行压缩，重建结果将是 $1/2$。这是因为 q_1 的码字不是 q_3 的码字的子集。

考虑图1.B所示的另一种设计。对于 q_2 中码字数为 2 的情况，我们从 q_3的 4 个码字中取一个子集。类似地——

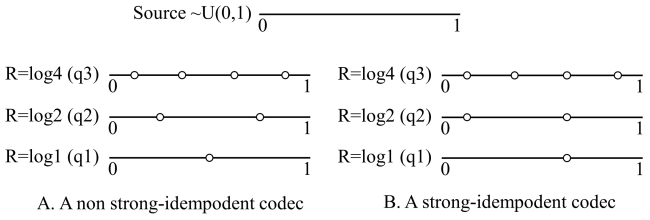


图1。A. 非强幂等编解码器的示例；B. 使用均匀源的强幂等编解码器。

对于 q_1 中码字数为 1 的情况，我们从 q_2的 2 个码字中取一个子集。而且，这种编解码器显然具有强幂等性。这是因为，对于任意两个质量点 q_i，q_j（其中$i < j$），q_i 的码字始终是 q_j的码字的子集。矢量量化也遵循类似的规则。

3.3. 设计度量

在设计度量之前，我们首先考察通用非幂等编解码器用于多轮重新压缩时会发生什么。对于通用编解码器，使用 $\{q_1, ..., q_k\}$ 进行多轮重新压缩所产生的失真不小于使用 q_{min} 进行单轮压缩所产生的失真。对于强幂等编解码器，这两种失真彼此相等，这也是我们能够达到的最佳结果。形式化地，我们有：

定理3.1。单轮压缩的失真不大于多轮压缩的失真。换言之，记 $d(.,.)$ 为一种失真度量，则有：

$$d(\boldsymbol{x}, f(\boldsymbol{x}, q_{min})) \leq d(\boldsymbol{x}, f(f(\boldsymbol{x}, q_1), ..., q_k)). \quad (3)$$

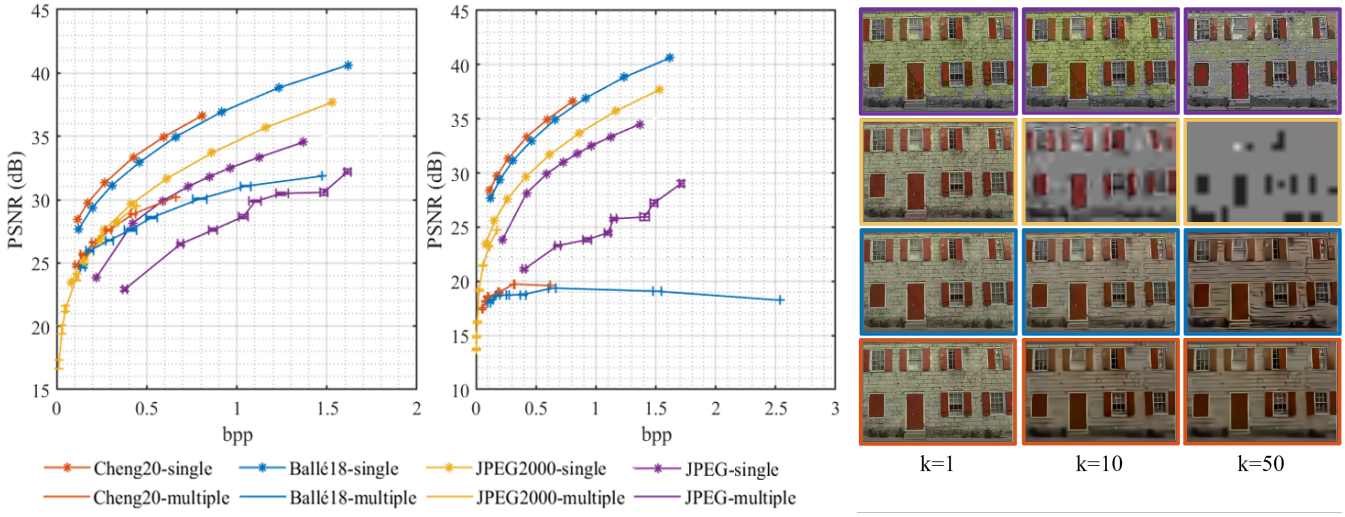
证明。我们采用反证法。首先，假设$d(\boldsymbol{x}, f(\boldsymbol{x}, q_{min})) > d(\boldsymbol{x}, f(f(\boldsymbol{x}, q_1)..., q_k))$成立。根据数据处理不等式，我们有：

$$I(\boldsymbol{x}, f(\boldsymbol{x}, q_{min})) \geq I(\boldsymbol{x}, f(f(\boldsymbol{x}, q_1), ..., q_k)), \quad (4)$$

这意味着存在一个可达成的率失真对，其码率为 $\leq I(\boldsymbol{x}, f(\boldsymbol{x}, q_{min}))$，失真为 $< d(\boldsymbol{x}, f(\boldsymbol{x}, q_{min}))$。这与率失真定理 [17] 的逆定理相矛盾。因此，假设不成立，证明完毕。

□

定理 3.1 告诉我们，多轮压缩的失真始终大于或等于单轮压缩的失真。根据强幂等性的定义，只有当编解码器满足强幂等性时，等式才成立。不过，需要注意的是，我们无法对实际比特率作出限定。尽管定理 3.1 的证明揭示了比特率下界之间的关系，但实际比特率之间的关系仍然未知。这意味着，对于非强幂等编解码器，



(a) (b)

图2。（a）在 Kodak 数据集 [18] 上，使用 Cheng20、Ball´e18、JPEG2000 和 JPEG 对图像进行多次压缩所得到的率失真曲线。左侧： $k = 10$。右侧： $k = 50$。（b）JPEG、JPEG2000、Ball´e18 和 Cheng20 模型在 $k = 1$，10，50 下的感知结果。 q_{min} 被选作可能的最低质量。

多轮压缩的比特率可能低于或高于单轮压缩的比特率。这也使我们无法使用广泛应用的 Bjontegaard 度量（BD指标） [19]。

由于描述强幂等性时需要使用失真，我们自然会考虑单轮压缩图像与多轮重压缩图像之间的失真，即 $d(f(\boldsymbol{x}, q_{min})\ f(f(\boldsymbol{x},\ q_1)\ .....,\ q_k))$

它有几个优点：

- **满足强幂等性时，该度量趋于 0**。根据公式 2 中强幂等性的定义，单轮压缩图像与多轮重压缩图像之间的失真为 0。如果使用原始图像与多轮重压缩图像之间的失真作为度量，就会失去这一性质。

- **该度量对原始图像与经过多轮重新压缩的图像之间的失真给出了上界**。给定 $d(.,)$ 满足三角不等式，我们有

$$d(\boldsymbol{x}, f(\boldsymbol{x}, q_{min})) + d(f(\boldsymbol{x}, q_{min}), f(f(\boldsymbol{x}, q_1), ..., q_k)) \geq d(\boldsymbol{x}, f(f(\boldsymbol{x}, q_1), ..., q_k)),$$

成立，这意味着我们得到了原始图像与经过多轮重新压缩的图像之间失真的上界。

另一个需要考虑的因素是，重压缩质量q_1、.....、 q_k 具有随机性。为了了解编解码器的平均表现，我们需要计算度量的期望。此外，重压缩轮次 k 也会影响表现。我们既关心编解码器的短期表现，也关心其长期表现。最后，还需要纳入最低质量q_{min} 。因为编解码器可能在高比特率下表现良好，但在低比特率下表现不佳。

综合以上因素，我们正式定义用于评估图像编解码器强幂等性的度量：

$$\rho(q_{min}, k) = \mathbb{E}[d(f(\boldsymbol{x}, q_{min}), f(f(\boldsymbol{x}, q_1), ..., q_k))], \quad (5)$$

其中， $\rho(q_{min}, k)$ 是我们的最终度量， q_{min} 是图像的最低质量， k 是重压缩轮次。此外， $q_1, ..., q_k$ 从 $\{q_{min}, ..., q_{max}\}$. 中均匀采样。正如我们将在下一节中展示的，不同的 q_{min} 和 k 会显著影响性能。实际应用中，我们通过简单的蒙特卡罗方法近似计算期望。

3.4. 测试协议

我们制定了一套测试协议，用于指导对图像编解码器强幂等性的评估。具体而言，我们给出以下流程：

- 首先，准备测试数据集。我们建议使用 Kodak 数据集 [18] 。
- 其次，选择 q_{min} 和 k 的列表，以评估每个ρ（q_{min}， k）。我们建议从编解码器的质量点中均匀抽取 8 个q_{min} 。对于 k，我们建议采用 $k = 10$ 和 $k = 50$ 来评估短期和长期性能。

	ρ（k，q_{min}），以均方误差衡量															
	$k=10$								$k=50$							
$q_{min}=$	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
传统																
JPEG	288.2	137.7	136.4	96.3	62.9	51.4	70.6	39.5	552.9	354.7	331.0	283.1	187.4	216.7	156.0	97.2
JPEG2000	1308	669.7	455.1	222.5	172.5	120.5	90.2	69.7	3098	3044	2396	1811	823.2	471.6	315.1	226.3
基于学习的																
Ball´e18 [5]	108.7	81.2	84.8	74.5	61.9	46.4	37.8	48.2	877.4	852.7	798.5	808.4	816.8	719.0	848.9	1342
Cheng20 [20]	123.6	109.3	95.9	83.4	62.9	41.2	-	-	1030	900.5	835.7	789.7	676.6	682.9	-	-

表1。使用 k 等于 10 和 50，测得在 Kodak 数据集上 Cheng20、Ball´e18、JPEG2000 和 JPEG 的 $\rho(q_{min},\ k)$ 均方误差（MSE）值 [18]。

•第三，从范围$\{q_{min}, q_{max}\}$内均匀抽取 k 个质量 q_i 。使用 q_{min}对图像进行一次压缩。然后，根据 q_i重复压缩图像，并评估一次压缩与ˆ次重复压缩之间的失真 ρ（q_{min}，k）。重复此过程b 次，并取失真的平均值，得到最终的 ρ（q_{min}，k）。

• 第四，对所有选定的 q_{min}，k 重复第三步，完成评估。

4. 实验

4.1. 实验设置

按照上述测试协议，我们评估两种传统编解码器和两种学习式图像压缩（LIC）模型的强幂等性。两种传统编解码器分别为 JPEG 和 JPEG2000。我们使用 libjpeg 评估 JPEG，使用 openjpeg1.5.1. 评估 JPEG2000。LIC 模型为采用均值尺度超先验的 Ball´e18 [5]，以及采用[21]预训练模型的 Cheng20 [20]。这两个模型被学习式图像压缩社区广泛用作基线。

对于 JPEG 和 JPEG2000，最低质量在每像素比特数（bpp）范围为 0 至 3.5 的区间内均匀采样。对于 JPEG，我们测试 8 $q_{min}=\{5, 15, 25, 35...75\}$。对于 JPEG2000，我们测试 8 $q_{min}=\{24, 26, 28, 30...38\}$。对于 Ball´e18 [5] 和 Cheng20 [20],，最低质量级别 q_{min} 遍历所有模型质量级别，它们分别为 $q_{min}=\{1, ..., 6\}$和 $\{1, ..., 8\}$。我们将参数 k（重新压缩次数）设为 10 和 50。对于每个 $\rho(q_{min},\ k)$ 的蒙特卡罗估计，我们对 JPEG2000、Ball´e18 [5] 和 Cheng20 [20]采样 $b=10$ 次。但由于 JPEG 的方差较大，我们对其采样 50 次。（见图2(a)）

4.2. 实验结果

评估结果如图2(a)、图2(b)和表1所示。图2(a)比较了直接压缩 q_{min} 与多轮压缩的率失真曲线，其中最低质量 q_{min}。

JPEG2000 的重新压缩率失真曲线与单次压缩非常接近。这意味着多轮压缩对 JPEG2000 的整体率失真性能影响甚微。然而，值得注意的是，JPEG2000 的绝对画质下降也非常快。在最高的 q_{min} 和 $k=10$ 下，JPEG2000 的峰值信噪比下降超过 5 dB。另一方面，JPEG [3] 的画质也显著下降，但其比特率保持不变。与传统编解码器相比，学习式图像压缩方法对多轮压缩更加敏感。当 $k=10$ 时，Ball´e18[5] 和 Cheng 20 [20] 的画质都会显著下降。而对于 $k=50$，Ball´e18 和 Cheng20 的率失真曲线不再单调递增，这意味着比特率越高，画质反而越低。

表1展示了以均方误差测得的 $\rho(q_{min},\ k)$值。与 $\rho(q_{min},\ k=10)$相比，$\rho(q_{min},\ k=50)$明显更低。最后几个点也有轻微波动。图2(b)展示了在测试区间内，最低质量级别q_{min}下进行多轮压缩的感知结果，其中 $k=1$, 10，50。随着 k 增加，传统编解码器的感知质量下降速度比学习式图像压缩更快，这与主观率失真曲线不符。这种现象在 JPEG2000 中尤为明显。可以看到，JPEG2000 图像经过$k=50$重新压缩后几乎无法辨认。

5. 结论

本文提出强幂等编解码器的概念，并研究其性质。我们提出一种度量，用于评估图像编解码器的强幂等性，并制定了测试协议。最后，我们评估了当前图像编解码器的强幂等性，包括两种传统编解码器和两种较新的神经编解码器。实验结果表明，当前的编解码器对不同质量下的多轮重新压缩非常敏感。我们也鼓励未来的研究致力于改善这一状况。

6. 参考文献

[1] Jun-Hyuk Kim、Soobeom Jang、Jun-Ho Choi 和 Jong- Seok Lee，《连续深度图像压缩的不稳定性》，载于第*28*届 *ACM* 国际多媒体会议论文集，2020年，第247–255页。[2] Litian Li、Zheng Yang 和 Ronggang Wang，《提升学习式图像压缩的多代鲁棒性》，*arXiv* 预印本 *arXiv:2210.17039*，2022年。[3] Gregory K Wallace，《JPEG 静态图像压缩标准》，*IEEE* 消费电子学汇刊，第38卷，第1期，第xviii–xxxiv页，1992年。[4]Jooyoung Lee、Seunghyun Cho 和 Seung-Kwon Beack，《用于端到端优化图像压缩的上下文自适应熵模型》，*arXiv* 预印本 *arXiv:1809.10452*，2018年。[5] Johannes Ball´e、David Minnen、Saurabh Singh、Sung Jin Hwang 和 Nick Johnston，《采用尺度超先验的变分图像压缩》，*arXiv* 预印本 *arXiv:1802.01436*，2018年。[6] David Minnen、Johannes Ball´e 和 George D Toderici，《用于学习式图像压缩的自回归先验与分层先验联合模型》，神经信息处理系统进展，第31卷，2018年。[7]Bhaskaran Vasudev、T Russell Hsing、Andrew G Tescher 和 Robert L Stevenson，《图像与视频通信和处理 2000》，图像与视频通信和处理 *2000*，第3974卷，2000年。[8] Rajan L Joshi、Majid Rabbani 和 Margaret A Lep- ley，《JPEG 和 JPEG 2000 多次压缩循环性能比较》，载于数字图像处理应用第二十三届。SPIE，2000年，第4115卷，第492–501页。[9]A Tanju Erdem 和 M Ibrahim Sezan，《MPEG 视频压缩标准的多代特性》，载于第一届国际图像处理会议论文集。IEEE，1994年，第2卷，第933–937页。

[10]Caspar Horne、Thumpudi Naveen、Ali Tabatabai、Robert O Eifrig 和 Ajay Luthra，《MPEG-2 4:2:2 配置文件特性研究——MPEG-2 在演播室环境中的应用》，*IEEE*视频技术电路与系统汇刊，第 6 卷，第 3 期，第 251–272 页，1996 年。

[11] Hani Z Sorial 和 William E Lynch，“变换编码图像的多代编码”，载于*Visual Communicationsand ImageProcessing'97*。SPIE，1997年，第3024卷，第1394–1405页。

[12] Y Yi、J Sun 和 S-y Yu，“关于图像编解码在质量无损情况下多代编码的研究”，《上海交通大学学报（中文版）》，第38卷，第4期，第0616–0619页，2004年。[13] Antonin Descampe、Joachim Keinert、Thomas Richter、Siegfried F¨oßel 和 Ga¨el Rouvroy，“JPEG XS：一种视觉无损、低延迟的轻量级图像压缩新标准”，载于*Applications of DigitalImageProcessing XL*。SPIE，2017年，第10396卷，第68–79页。[14] Thomas Richter、Joachim Keinert、Antonin Descampe、Gael Rouvroy 和 Alexandre Willeme，“适用于多代的稳健 JPEG XS 编码”，载于*2017*年*IEEE*国际多媒体研讨会。IEEE，2017年，第6–13页。[15] Ziyuan Zhu 和 Tao Lin，“幂等 H.264 帧内压缩”，*MultimediaToolsandApplications*，第46卷，第25–45页，2010年。[16] Jun-Hyuk Kim、Soobeom Jang、Jun-Ho Choi 和 Jong-Seok Lee，“逐次学习图像压缩：不稳定性的全面分析”，*Neurocomput-ing*，第506卷，第12–24页，2022年。[17] Thomas M Cover，《信息论基础》，John Wiley & Sons，1999年。[18] 伊士曼柯达，“柯达无损真彩色图像套件（PhotoCD PCD0992）”，网址 *http://r0k. us/graphics/kodak*，第6卷，1993年。[19] Gisle Bjontegaard，“计算 RD 曲线间的平均峰值信噪比差异”，*VCEG-M33*，2001年。[20] Zhengxue Cheng、Heming Sun、Masaru Takeuchi 和 Jiro Katto，“采用离散高斯混合似然和注意力模块的学习式图像压缩”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2020年，第7939–7948页。[21]Jean B´egaint、Fabien Racap´e、Simon Feltman 和 Akshay Pushparaja，“CompressAI：用于端到端压缩研究的 PyTorch 库和评估平台”，*arXiv*预印本 *arXiv:2011.03029*，2020年。