

问题

### 技术背景

超先验编解码网络框架<sup>[1]</sup>

- 当前，手机拍摄的图像分辨率普遍提升至4K及以上，这也意味着更大量的图像信息需要存储和传输。为了提高图像信息的存储和传输效率，经过业界三十余年的努力，已发展出多种图像压缩技术，成为现实世界诸多应用不可缺少的基本组件。
- 传统图像压缩基于预测、变换和熵编码三种主要技术（如基于小波变换的JPEG2000、基于分块DCT变换和预测的H.266/VVC）。近年来，基于变分自编码器（VAE）结构的AI图像压缩方法<sup>[1]</sup>也获得学术界关注，发展迅速。
- 然而，与传统图像压缩方法相比，AI图像压缩方法的一个现实问题是它的计算复杂度要高出传统方法几十甚至上百倍。因此，目前这类方法并没有取得明显更具竞争力的压缩效率-计算开销折中，在手机等功耗敏感、算力受限的设备上的部署遇到许多困难和挑战，限制了它的应用范围。

### 技术挑战

**基于变分自编码器的AI图像编解码网络运算效率低：**传统图像压缩方法基于明确的数学原理，借助一系列简单的计算公式就能完成编码和解码；尤其对于解码而言，计算复杂度低（平均每个像素的解码使用几十到几百次运算）。AI图像压缩方法，基于卷积层的堆叠，算力较高。例如，JPEG-AI标准的VMuC 1<sup>st</sup> Stage模型<sup>[2]</sup>虽然在压缩效率上能超过H.266/VVC约8.5%，但每完成1个像素的解码平均需要使用大约6.9万次乘法运算，比传统解码方法高出两个数量级。因此，对于图像压缩任务，是否存在更高效的网络结构，能够明显地降低算力需求（例如每像素解码使用的运算次数下降到2千次左右）？

### 技术诉求

**目标：**基于公开的JPEG-AI VMuC算法原型，设计高效且轻量化的图像编解码网络，不限于网络结构、基础算子、训练方法等方面的创新和改进；相比于JPEG-AI VMuC 1<sup>st</sup> Stage模型，在不损失压缩效率的情况下，编解码算力降低35倍以上（算力开销接近传统方案）。

**参考文献：**

[1] J. Balle et al, "Variational image compression with a scale hyperprior", in Proceedings of the International Conference on Learning Representations (ICLR2018)

[2] ISO/IEC JTC 1/SC29/WG1 N100279, "Description of the JPEG AI Verification Model under Consideration and associated software integration procedure", July 2022.

background

技术背景

- 手机拍摄的图像分辨率普遍提升至4K及以上，这意味着更大量的图像信息需要存储和传输，因此为了提高图像信息的存储和传输效率，发展出多种图像压缩技术。
- **图像压缩是一种将图像文件的大小减小的技术**，主要用于减少图像在存储和传输过程中所需要的空间。常用的图像压缩技术包括无损压缩和有损压缩两种。无损压缩能够在不损失图像质量的前提下将图像文件的大小减小，常用的无损压缩技术有JPEG 2000，BPG等。有损压缩会在一定程度上损失图像质量，但能够更大幅度地减小图像文件的大小，常用的有损压缩技术有JPEG，WebP等。
  - 无损压缩中的JPEG 2000和BPG,完整地保留了图像的原始信息，使得压缩后的图像与压缩前的图像完全相同，并且不会因压缩而导致图像质量的下降<sup>[1]</sup>。
  - 有失真压缩技术中的JPEG和WebP，在压缩过程中会有一定的信息损失，所以压缩后的图像质量会相对较差，但是通常比无损压缩的文件更小。
  - 在实际应用中，图像压缩技术根据图像的特点进行选择使用。对于需要保证图像质量的场景，如医学图像和卫星遥感图像等，一般使用无损压缩技术来保证图像质量。对于需要尽可能减小文件大小的场景，如网络传输和手机拍照等，一般使用有损压缩技术来减小文件大小。还有一种技术叫做“无失真压缩”，就是在保证压缩文件尽量小的同时，压缩效率也比较高，但是在压缩过程中会有一定的质量损失。

## 图像压缩技术

- **传统图像压缩技术主要基于预测、变换和熵编码这三种技术。**这三种技术结合起来，能够更有效地减少图像的冗余信息，提高压缩效率。

- 预测技术是通过使用周围像素的信息来预测当前像素的值，从而减少冗余信息。
- 变换技术是通过将图像数据变换到频域来减少冗余信息。常用的变换技术有小波变换和离散余弦变换（DCT）。
- 熵编码技术是通过对变换后的图像数据进行熵编码来减少冗余信息。常用的熵编码技术有 Huffman 编码和 Arithmetic 编码。

- **一些例子：**

- JPEG 标准的发展始于1987年，它是一种基于 DCT 变换和熵编码的图像压缩算法，JPEG2000 基于小波变换进行压缩，能够在保证图像质量的前提下有效地压缩图像。
- H.264/AVC 和 H.265/HEVC 则基于分块 DCT 变换和预测技术进行压缩，能够高效地压缩视频。
- VVC (Versatile Video Coding) 基于基于分块DCT变换和预测技术进行压缩，能够比 H.265/HEVC更高效地压缩视频。
- 传统图像压缩算法有很多种，每种方法的发展历程和相关文献都是不同的。比如，JPEG 标准的发展始于1987年，它是一种基于 DCT 变换和熵编码的图像压缩算法。JPEG2000 标准则是在2000年发布，它是一种基于小波变换和熵编码的图像压缩算法。H.264/AVC 标准则是在2003年发布，它是一种基于分块 DCT 变换和预测编码的图像压缩算法。H.265/HEVC 标准则是在2013年发布，它是一种基于大块 DCT 变换和预测编码的图像压缩算法。

- 图像压缩技术的通用模型<sup>[2]</sup>

- 图像压缩的通用模型由编码器和解码器两个重要部分组成。编码器的工作原理是从输入图像  $f(x, y)$  经过通道传输后产生的压缩图像。压缩后的图像进入解码器系统，如下图<sup>[3]</sup>所示，在该系统中，压缩后的图像将被重建，生成一个输出图像  $f'(x, y)$ 。

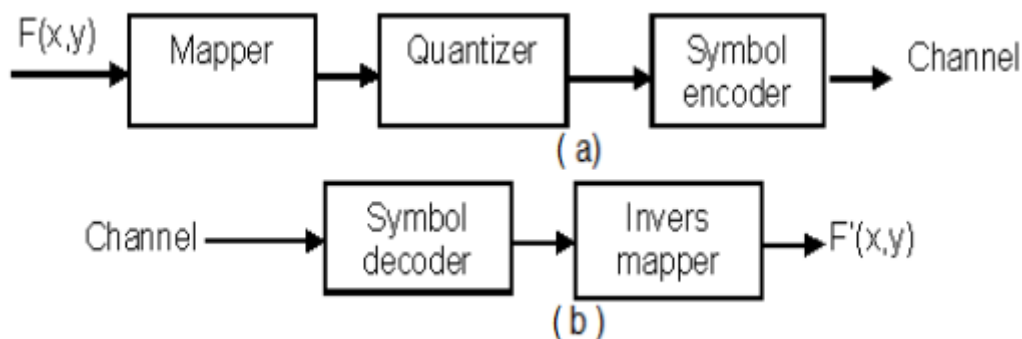


Figure 1 Image compression model: a) Source encoder b) Source decoder

- Mapper的工作是将输入图像转化为指定的格式，以消除像素间冗余。这个操作是可逆的，这意味着重建图像中的信息与原始图像相同。
- 量化器的工作是将输入数据转换为适合压缩的形式，以减少输入图像上存在的空间冗余。该操作将量化过程转化为前一阶段的结果。它的作用是不可逆的，这意味着重建图像中的信息与原始图像不同<sup>[4]</sup>。
- 符号编码器的工作是创建表示量化器输出的代码，以及它们映射成代码，以及将它们映射成代码。由此产生的结果可以摆脱不相关的数据，并且是可逆的。

## • AI图像压缩方法

- 在AI图像压缩算法的发展过程中，研究方向主要就两点降低复杂度、提升压缩率。主要有些下列这些研究贡献：
  - AI图像压缩算的发展始于深度学习技术的兴起。早期的工作主要集中在使用卷积神经网络(CNN)、反卷积网络 and 自编码器对图像进行压缩。
  - Ranzatp等<sup>[5]</sup> 最早对自编码器进行改进应用于图像压缩，其提出了一种基于能量函数的自编码器，并使用这种类型的网络来学习稀疏表示。作者通过在自编码器中引入能量函数来解决使用传统自编码器学习稀疏表示的问题，并将其应用于图像压缩。他们还提出了一种新的学习算法来最小化网络的能量，从而最大化稀疏性。
    - 使用自编码器进行图像压缩的方法，主要是利用自编码器的稀疏性质，将图像的高维像素空间映射到低维稀疏空间。在这种方法中，自编码器的编码器部分学习了图像的稀疏表示，并将其用于压缩图像。解码器部分则通过重构图像来还原压缩后的图像。
    - 然而，与传统图像压缩方法相比，使用自编码器进行图像压缩的方法存在一些缺陷，例如压缩率较低和计算复杂度较高。
    - 此外，还有一些其他的文献也提出了使用自编码器进行图像压缩的方法。这些文献通常展示了使用自编码器学习稀疏图像表示的优秀性能。
    - Theis等<sup>[6]</sup> 提出了一种基于压缩自编码器的损耗式图像压缩方法。文中的压缩自编码器包含一个编码器和一个解码器，编码器将图像压缩为一个低维表示，解码器将其还原为原始图像。这种方法通过在编码器中使用压缩感知重建损失来实现压缩。文章还提出了一种新的训练方法，以改进自编码器对图像的压缩能力。并进行了实验和评估，证明了这种方法在压缩效率和压缩质量方面优于现有方法。
  - Venetianer 等<sup>[7]</sup> 最早使用 CNN 方法对图像进行压缩和解码。Zeiler等<sup>[8]</sup> 则最早使用反卷积网络结构从高层特征中解码图像。他们在实验中证明了反卷积网络在解码图像时具有较高的性能。随后，反卷积网络在图像压缩领域得到广泛关注，并被用于许多不同的图像压缩模型中。在2016年左右，有许多研究者开始使用CNN来研究图像压缩，并在当时发表了许多关于这一领域的相关论文。如Gong等<sup>[9]</sup> 提出了一种使用向量量化来压缩深度卷积神经网络的方法。这种方法首先将网络的权重进行量化，然后

使用量化后的权重来重新训练网络。使用这种方法可以在不明显降低网络性能的前提下，将网络的尺寸和计算复杂度降低大约32倍。

- 随着深度学习技术的发展，研究者们开始使用更加复杂的网络结构，如变分自编码器 VAE 和生成对抗网络 GAN
  - Kingma等<sup>[10]</sup> 最早提出变分自编码器，在该文中，作者提出了使用变分自编码器（VAE）进行图像压缩的方法，并通过实验证明了这种方法的有效性。使用VAE进行图像压缩的方法主要是通过学习图像的潜在变量来压缩图像。这种方法在压缩率和压缩质量方面都有较高的性能。随后这种方法也被广泛的研究和使用，在降低复杂度，提高压缩率，改善压缩质量等方面有了很多的改进。
    - **基于变分自编码器（VAE）的 AI 图像压缩方法**在学术界受到了关注，发展迅速。
    - 变分自编码器（VAE）是一种生成模型，它通过学习图像的潜在空间表示来进行图像压缩<sup>[11]</sup>。它通过学习图像的潜在表示来进行图像压缩，并通过重建来保证图像质量。
    - VAE 结合了自编码器和生成对抗网络（GAN）的优点，能够在保证图像质量的前提下有效地压缩图像。VAE 通过学习图像的潜在表示来生成图像，并且可以通过调整压缩率来控制图像的质量和压缩率。
    - VAE 在图像压缩领域具有很大的潜力，因为它可以在保证图像质量的前提下有效地压缩图像。此外，VAE 具有很强的泛化能力，可以用于处理各种类型的图像。
    - 然而，VAE 目前还是一个新兴技术，在研究和应用上还有很多挑战需要克服，例如压缩速度，压缩率和质量之间的平衡等。
    - VAE 在图像压缩领域发展迅速，研究人员不断地提出新的方法和改进现有方法来提高 VAE 的性能<sup>[12]</sup>。例如，一些研究使用了高维注意力机制来提高 VAE 的压缩效率<sup>[13]</sup>。
    - 另一些研究使用了多尺度结构来提高 VAE 的性能<sup>[14]</sup><sup>[15]</sup><sup>[16]</sup>，这样可以更好地处理图像中的细节。还有一些研究使用了非局部连接<sup>[17]</sup><sup>[18]</sup><sup>[19]</sup> [chen\_end-to-end\_2021?] 和跳过连接<sup>[20]</sup><sup>[21]</sup> 来提高 VAE 的性能，这样可以更好地处理图像中的大片背景。
    - 总的来说，VAE 在图像压缩领域是一个很有潜力的技术，在未来可能会有更多的研究和应用。
    - Balle等<sup>[22]</sup> 提出了一种基于变分自编码器（VAE）的图像压缩方法，并引入了尺度超先验来控制压缩比率和图像质量之间的平衡。该文证明了这种方法在压缩效率和图像质量上优于其他现有的基于VAE的图像压缩方法。这篇文章是首先使用超先验来提高AI图像压缩算法效率的文章，这种方法在后来的研究中也广泛使用。该方法使用了一个尺度超先验来提高图像压缩的性能。这种方法首先使用 VAE 模型学习图像的潜在表示，然后在 VAE 的基础上，引入了一个尺度超先验



来控制压缩率。这个超先验可以通过调整控制编码器和解码器之间的平衡来控制压缩率。

- 在这种方法中,尺度超先验可以更好地控制压缩率和图像质量之间的平衡。这种方法能够在保证图像质量的前提下, 有效地压缩图像。它也具有较好的泛化能力, 可以用于处理各种类型的图像。
- 文章中还给出了一些实验结果, 显示了这种方法在压缩率和图像质量之间的平衡方面优于其他现有的方法。
- 总的来说, 这种方法是一种有前途的图像压缩方法, 它能够在保证图像质量的前提下有效地压缩图像, 并且具有很好的泛化能力。
- 尺度超先验是指在模型中加入关于尺度的额外信息的过程, 以帮助模型更好地学习和推理。在图像压缩领域, 尺度超先验可以用来控制压缩率和图像质量之间的平衡。通过加入尺度超先验, 可以使模型更好地控制压缩率, 使得图像质量更高。在图像压缩领域中, “尺度”通常指的是压缩率。尺度超先验中的尺度通常是指在模型中加入的一种控制压缩率的参数, 通过调整这个参数来控制压缩率, 从而控制图像的质量。
- GAN在图像压缩领域的应用较为广泛, 许多的研究都是沿用了GAN的基本思想去改进图像压缩模型。在Google Scholar中找到的较早使用GAN用于图像压缩的研究是 Santurkar等<sup>[23]</sup> 用生成对抗网络 (GAN) 来进行图像压缩的方法。具体来说, 该文提出了一种新的压缩方法, 称为“生成压缩”, 它使用GAN来生成压缩后的图像。文章中给出了这种方法的详细描述, 并对其进行了实验评估。结果表明, 这种方法在保证图像质量的同时, 能够实现高效的压缩率。Li等<sup>[24]</sup> 这篇文章提出了一种GAN压缩, 一种基于条件GAN的架构, 可以在保持高压缩比的同时降低网络的计算复杂度。文章还提出了一种新的训练方法来提高生成器的效率, 并使用实验结果证明了该方法的有效性。
- 研究人员也在寻求结合传统图像压缩的方法, 如小波变换<sup>[25]</sup> 和DCT变换来提高压缩效率并降低算力需求。此类方法也称之为混合压缩方法, 混合压缩技术是一种将每组方法的优秀属性组合在一起的技术。在高效的图像压缩算法中, 传统的算法和深度学习算法结合可以更好的利用各自的优势, 实现更高的压缩效率。Setyaningsih等<sup>[3]</sup> 综述了传统方法和AI图像压缩方法的混合应用的原理和研究现状。
  - 一种常见的方法是, 先使用传统图像压缩方法 (如DCT) 对图像进行预处理, 然后使用AI图像压缩算法来压缩已经与处理过的图像<sup>[26]</sup>。这样可以利用传统图像压缩方法的优势 (如高效性和稳定性), 并且使用AI算法来提高压缩效率。
  - 另一种方法是在AI图像压缩算法处理完图像之后, 使用传统图像压缩方法来进一步压缩图像。这样可以利用AI算法的优势 (如学习能力和模型复杂度), 并且使用传统图像压缩方法来提高压缩率。

## • AI图像压缩算法的优点

- AI图像压缩方法相对于传统图像压缩方法来说，其压缩效果更好，能够在较低的压缩率下得到更高的图像质量。这是因为AI图像压缩方法可以利用深度学习网络来学习图像的复杂结构和细节信息，并且能够在保证压缩率的前提下得到更高的图像质量。
- AI图像压缩方法还有很多其他优点，例如对于不同类型的图像具有很好的适应性，能够更好地处理复杂的图像结构和细节信息等，使得它在图像压缩领域中具有很大的潜力。
- 然而，目前AI图像压缩方法的应用范围受到计算复杂度和部署难度的限制，它主要应用在高性能计算机或云端环境中，在移动端或功耗敏感设备上的应用相对较少。
- 因此，目前在研究和开发中，研究者和工程师正在不断提高AI图像压缩方法的计算效率和部署可行性，使其能够在更广泛的场景中得到应用。

## • AI图像压缩方法面临的问题与可行的解决方案

- 目前AI图像压缩方法相对于传统图像压缩方法来说，其计算复杂度较高是一个现实问题。传统图像压缩方法通常采用预测、变换和熵编码等基于明确数学原理的技术，这些技术的计算复杂度相对较低。而AI图像压缩方法通常采用基于深度学习的方法，其计算复杂度要高出传统方法几十甚至上百倍，在算力有限的设备上的部署困难。
- 尽管如此，由于AI图像压缩方法的优秀性能，在研究和开发上也在不断探索更高效率的解决方案，例如采用移动端友好的网络结构，使用硬件加速器等。还有一些研究也在尝试结合传统图像压缩方法和深度学习方法，来提高压缩效率和降低计算复杂度。
- 另外，有一种方法是在线性的压缩率之间进行权衡，例如在高压压缩率的情况下使用传统图像压缩方法，在较低压缩率的情况下使用AI图像压缩方法，这样可以在保证压缩效率的同时降低计算复杂度。
- 总之，AI图像压缩方法虽然目前存在计算复杂度较高的问题，但研究者和工程师正在不断努力研究和改进，使得AI图像压缩更加具有实用性和可部署性。
- 目前存在一些研究是在尝试提出更高效的网络结构来降低AI图像压缩方法的算力需求。
- 例如，研究者们可以尝试使用轻量级的网络结构，例如MobileNet<sup>[27]</sup> [28]，ShuffleNet<sup>[29]</sup> 等进行编码和解码。这些网络结构具有较低的计算复杂度，可以在移动端和嵌入式设备上得到应用。
- 另外，研究者们也在尝试研究量化、剪枝和离线训练等技术来降低网络结构的计算复杂度，并提高解码效率。
- 具体来说，使用量化技术可以降低网络中权重和系数的精度，减少网络中运算次数，从而降低算力需求。剪枝技术可以删除网络中不必要的节点和连接，减少网络的参数量和计算复杂度。离线训练技术可以将训练过程放在高性能的计算机上进行，从而降低部署时的计算复杂度。
- 尽管这些技术能够降低网络结构的计算复杂度，但是它们也可能影响到网络结构的性能。因此，研究者需要权衡各种技术对网络性能和计算复杂度的影响，找到最优的技术组合。

- 另外一些研究者也在尝试使用新的编码和解码方法来降低算力需求。例如，使用非局部编码器或解码器可以减少网络的计算复杂度。使用局部连接结构或多尺度编码器/解码器可以提高网络的局部特征建模能力并降低算力需求。
- 还有一些研究是在尝试使用结合了传统图像压缩方法和AI图像压缩方法的网络结构来提高压缩效率并降低算力需求。例如,使用基于DCT变换的传统压缩方法编码低频部分，使用基于卷积网络的AI压缩方法编码高频部分。这种方法可以利用传统方法的高效性来减少编码低频部分的计算复杂度，同时利用AI方法的高压缩效率来增强高频部分的压缩效果。
- 还可以尝试使用类似于蒸馏学习<sup>[30]</sup>的方法来训练网络，即将一个大型网络（如JPEG-AI标准的VMuC 1st Stage模型）的知识转移到一个较小的网络上，从而降低算力需求。
- 另外, 研究者可以考虑结合传统图像压缩方法和 AI 图像压缩方法，使用传统方法编码低频部分，使用 AI 方法编码高频部分。这种方法可以利用传统方法的高效性来减少编码低频部分的计算复杂度，同时利用 AI 方法的高压缩效率来增强高频部分的压缩效果。
- 此外, 可以尝试使用一些类似于线性回归的算法来进行预测, 从而减少网络的参数量和计算量, 达到降低算力需求的目的。
- 设计高效且轻量的图像编解码网络是一个复杂的问题，需要在网络结构、基础算子、训练方法等多个方面进行创新和改进, 需要综合考虑多种方法，并在实际应用中不断进行优化。
- 研究人员正在不断尝试各种方法来提高AI图像压缩的算力效率，还有很多可能会更高效的方法正在研究中。

- **JPEG-AI VMuC(verification model under consideration)**

- JPEG-AI VMuC（正在考虑的JPEG-AI验证模型）的体系结构如下图所示<sup>[31]</sup>。

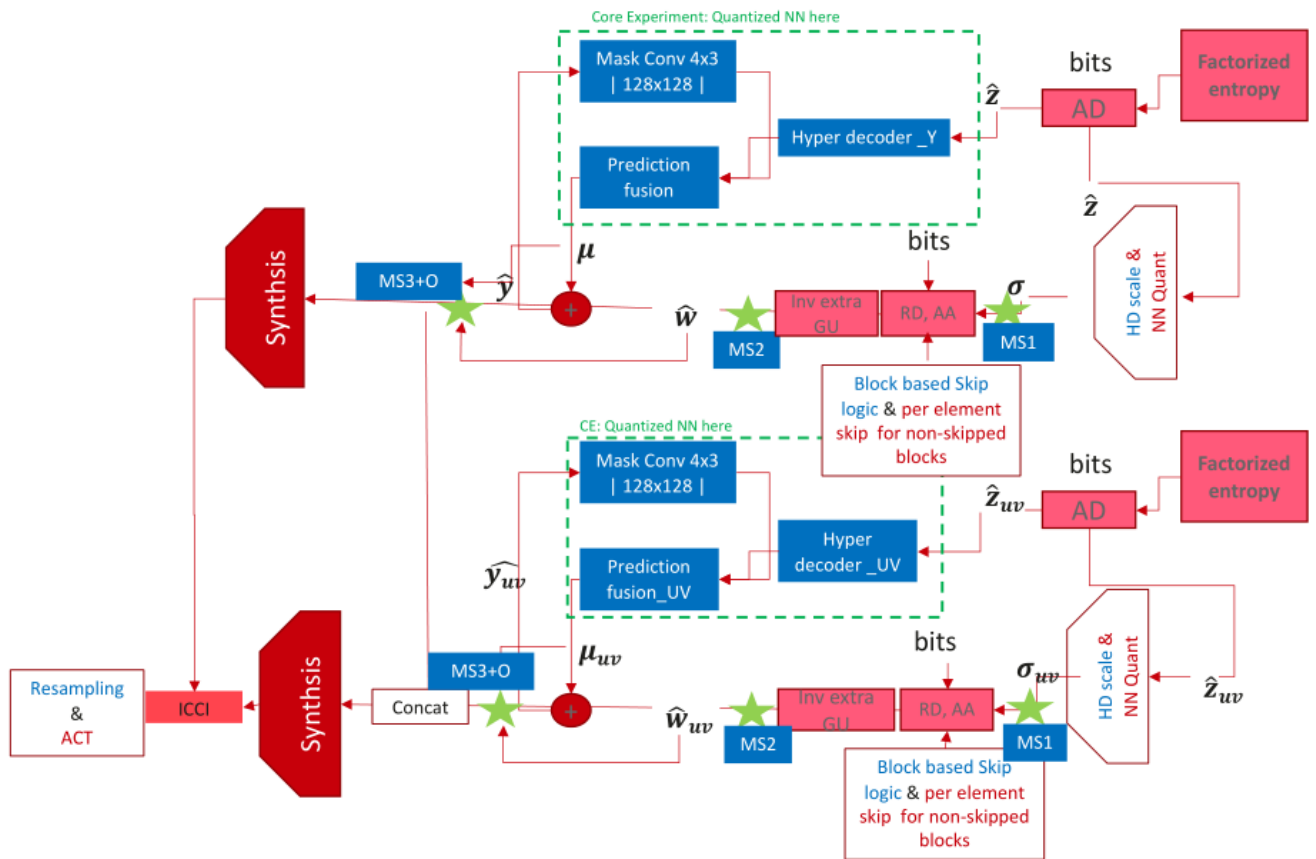


Figure 2 - JPEG AI VMuC high-level architecture

- 分解式熵（Factorized entropy）是一种图像压缩算法，它将图像中的熵（即随机性）从一个总体中的分解成多个单独的因子。这些因子通过压缩技术独立地对数据进行压缩，从而最大限度地减少图像的大小。在分解式熵图像压缩中，图像中的熵被划分为独立的因子，这些因子可以被单独处理。因此，分解式熵算法可以有效地减小图像的大小，同时保持图像的质量。
  - 图像中的熵（Entropy）是指图像信息的随机性和不确定性的度量。它通常用来衡量图像的复杂度，用于图像的压缩和评估图像的质量。在图像压缩中，熵较高的图像通常需要更多的信息来编码，因此需要更多的存储空间或更多的带宽来进行传输。因此，图像的熵通常被用作图像压缩算法的度量，并用来比较不同的图像压缩算法。
- 模数转换（AD），模数转换在分解式熵方法中被用来在熵编码的基础上进行改进，从而提高编码效率。模数转换是将原始数据转换为一个新的数据表示，其中可以利用更有效的数学模型进行熵编码。这一步的目的是提高编码效率，使得更多的数据可以在同样的数据长度下进行压缩。模数转换是图像压缩的重要组成部分，因为它可以减少数据量，降低数据存储和传输的成本。模数转换可以通过多种不同的算法实现，其中常见的算法包括树形编码，游程编码，差分编码等。
- HD scale（High Dynamic Scale）是图像压缩中的一个术语，其定义为在压缩过程中图像中灰度级别变化的程度。HD scale被用于评估图像压缩算法的效果。数值越大，表明灰度级别变化越大，因此图像中可见的细节更多，图像质量更高。相反，数值越小，图像中的细节就越少，图像质量就越低。
- NN Quant（Neural Network Quantization）是图像压缩中的一种算法。它是一种神经网络量化算法，旨在通过使用神经网络来优化图像压缩的效率。通常，在图像压缩过程中，图像数据会被编码



为更小的数据，以便节省存储空间。NN Quant算法通过使用神经网络模型，来识别图像中重要的特征，并为这些特征保留更多的精度，从而实现更高的图像压缩效率。

- Quantized NN 阶段中的“超解码器（Hyper decoder）”是指对前一阶段生成的量化信息进行解码的算法中的特定组件或过程。超解码器的目的是反转量化过程，并尽可能将图像数据恢复到其原始形式。**超解码器的具体设计和实现可以对图像压缩算法的整体性能和质量产生重大影响。** Hyper decoder是一个深度学习模型，它可以从  $\hat{z}$ （即模数转换之后得到的数据）中提取高级特征，并将其转换为特征图。
- Quantized NN 阶段中的Mask Conv是一种卷积操作，其中使用了一个 Mask 来对卷积的计算结果进行控制。卷积运算可以用来提取图像的特征，而 Mask Conv 操作可以在不影响卷积结果的基础上，控制卷积的影响范围，从而改善图像压缩算法的效果。
- Quantized NN 阶段中的预测融合（Prediction fusion）是指将图像数据的不同预测或估计进行组合以产生最终估计或预测的过程。图像压缩算法中“预测融合”的具体含义和目的将取决于所使用算法的具体细节。Prediction fusion是一种对多个预测进行融合的技术，其中多个预测可以是来自Hyper decoder的多个输出。最后， $\mu$ 可以是Prediction fusion的输出，即最终的预测结果。
- Hyper decoder\_Y 和 Hyper decoder\_UV 是图像压缩算法中“Quantized NN”阶段的两个部分。它们之间的区别可能在于它们所处理的图像数据的颜色空间不同。Hyper decoder\_Y 通常用于处理 Y 分量，这是图像的灰度分量，而 Hyper decoder\_UV 通常用于处理 UV 分量，这是图像的色彩分量。这两个部分可能有不同的设计和参数，以适应不同的图像数据。
  - Y分量描述黑白图像，UV分量描述的是整块的色彩。这里依据的是“大面积着色原理”。人眼视觉特性的研究表明，人眼对黑白图像的细节有较高的分辨力，而对彩色图像的细节分辨力较低，这即所谓的“彩色细节失明”。在彩色图像传送过程中，只有大面积部分需要在传送其亮度信息的同时还必须传送其色度成分。颜色的细节部分(对应于信号的高频部分)，可以用亮度信号来取代。
  - 灰度图像（gray image）是每个像素只有一个采样颜色的图像，这类图像通常显示为从最暗黑色到最亮的白色的灰度，尽管理论上这个采样可以任何颜色的不同深浅，甚至可以是不同亮度上的不同颜色。灰度图像经常是在单个电磁波频谱如可见光内测量每个像素的亮度得到的，用于显示的灰度图像通常用每个采样像素8位的非线性尺度来保存，这样可以有256级灰度（如果用16位，则有65536级）。
- MS1、MS2、MS3+O（缩放单元），缩放单元是图像压缩算法中一个重要的组成部分。它的作用是对图像进行尺寸缩放，以减小图像的数据量。通常，缩放单元使用高斯滤波器或其他图像处理技术对图像进行降采样，然后对降采样后的图像进行量化。缩放单元通常在图像压缩算法的前期进行，并且可以有效地减小图像的数据量，提高图像压缩效率。
- RD（Rate-Distortion）表示码率与失真度之间的折衷关系，常被用做评估图像压缩算法的效果。具体地，码率表示压缩后图像的大小，失真度表示压缩前和压缩后的图像的差异。码率与失真度的折衷关系可以帮助我们找到一种压缩方法，使得图像尽可能的小，并且与原图像的差异尽可能的小。
- AA（Average Absolute Difference）是一种评估图像失真度的方法。它通过计算压缩前后图像的每个像素之间的绝对差值的平均值来表示失真度。具体地，如果图像中每个像素的差值都很小，那么

失真度就较小；反之，如果图像中每个像素的差值都很大，那么失真度就较大。

- “Block based Skip logic”是图像压缩算法中的一种方法，其作用是在压缩过程中跳过那些不需要进行处理的块。在图像压缩中，通常对图像进行分块，在进行压缩时对每个块进行处理。在使用“Block based Skip logic”时，如果一个块中的数据变化不大，则可以跳过该块，不对其进行处理，从而降低压缩的复杂度。这样的方法在图像压缩的处理效率方面有着良好的提升。
- “Per element skip for non-skipped blocks”是图像压缩算法中的一种技术，它指在图像块不被跳过（即不通过“Block based skip logic”算法）的情况下，对图像元素进行跳过的处理。这样的处理能够更加细致地跳过不重要的图像信息，从而提高压缩效率。
- Synthesis 指的是合成操作，其目的是将压缩过的图像数据解码为一个近似原图的图像，通常是在经过一系列复杂的步骤，如预测、量化、变换等。在这个过程中，一些丢失的图像信息将被恢复或重建，从而得到一个比原图更加简化但是可读性很高的图像。
- ICCI (Inter-Channel Correlation Index) , ICCI 是一种图像压缩算法中用于评估图像通道之间相关性的指标。具体地，ICCI 表示不同通道之间的相关性，因此在图像压缩算法中可以用于评估某种压缩方案的效果。
- Concat 是图像压缩算法中用于拼接两个或多个不同的信号或数据的技术。它可以将多个信号或数据连接在一起，以形成一个新的数据序列。这个技术通常在图像压缩算法中用于合并多个不同的编码数据块，以形成单个的编码图像数据。
- 重采样 (Resampling) 是指降低分辨率或改变图像大小的过程。这通常是为了减少表示图像所需的数据量，并使图像更易于处理或传输。
- ACT (非对称卷积变换, Asymmetric Convolutional Transform) 是一种使用卷积神经网络来减小图像数据大小的图像压缩技术。它被称为“非对称”，因为编码和解码过程使用网络的不同部分。编码过程减小数据大小，而解码过程将数据扩展回原始大小。ACT可用于实现高压压缩率，同时保持良好的图像质量。

## Reference

- [1] MARCELLIN M W. GORMISH M J. BILGIN A. et al. An overview of JPEG-2000[C]//Proceedings DCC 2000. Data compression conference. IEEE, 2000: 523-541.
- [2] GONZALEZ R C. Digital image processing[M]. Pearson education india, 2009.
- [3] SETYANINGSIH E. HARJOKO A. Survey of hybrid image compression techniques[J]. International Journal of Electrical and Computer Engineering, 2017, 7(4): 2206.
- [4] ZAINELDIN H. ELHOSSEINI M A. ALI H A. Image compression algorithms in wireless multimedia sensor networks: A survey[J]. Ain Shams Engineering Journal, 2015, 6(2): 481-490.

- [5]  
RANZATO M. POULTNEY C. CHOPRA S. et al. Efficient learning of sparse representations with an energy-based model[J]. Advances in Neural Information Processing Systems, 2006, 19.
- [6]  
THEIS L. SHI W. CUNNINGHAM A. et al. Lossy image compression with compressive autoencoders[J]. arXiv Preprint arXiv:1703.00395, 2017.
- [7]  
VENETIANER P L. WERBLIN F. ROSKA T. et al. Analogic CNN algorithms for some image compression and restoration tasks[J]. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 1995, 42(5): 278-284.
- [8]  
ZEILER M D. KRISHNAN D. TAYLOR G W. et al. Deconvolutional networks[C]//2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2010: 2528-2535.
- [9]  
GONG Y. LIU L. YANG M. et al. Compressing deep convolutional networks using vector quantization[J]. arXiv Preprint arXiv:1412.6115, 2014.
- [10]  
KINGMA D P. WELLING M. Auto-encoding variational bayes[J]. arXiv Preprint arXiv:1312.6114, 2013.
- [11]  
DUAN X. LIU J. ZHANG E. Efficient image encryption and compression based on a VAE generative model[J/OL]. Journal of Real-Time Image Processing, 2019, 16(3): 765-773.  
<https://doi.org/10.1007/s11554-018-0826-4>.
- [12]  
GUO T. WANG J. CUI Z. et al. Variable rate image compression with content adaptive optimization[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR) workshops. 2020.
- [13]  
ZHANG L. LIU Y. Remote sensing image generation based on attention mechanism and VAE-MSGAN for ROI extraction[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.  
<https://doi.org/10.1109/LGRS.2021.3068271>.
- [14]  
CHOU J. HATHI G. Generated loss, augmented training, and multiscale VAE[J]. arXiv Preprint arXiv:1904.10446, 2019.
- [15]  
COUR T. BENEZIT F. SHI J. Spectral segmentation with multiscale graph decomposition[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05): Vol. 2. IEEE, 2005: 1124-1131.
- [16]

STARCK J L. MURTAGH F D. BIJAOU A. Image processing and data analysis: The multiscale approach[M]. Cambridge University Press, 1998.

[17]

CHEN T. LIU H. MA Z. et al. Neural image compression via non-local attention optimization and improved context modeling[EB/OL]. arXiv, 2019. <https://arxiv.org/abs/1910.06244>.

[18]

LIU H. CHEN T. SHEN Q. et al. Practical stacked non-local attention modules for image compression. [C]//CVPR workshops. 2019: 0.

[19]

LIU H. CHEN T. GUO P. et al. Non-local attention optimized deep image compression[J]. arXiv Preprint arXiv:1904.09757, 2019.

[20]

SHIN C. LEE H. SON H. et al. Expanded adaptive scaling normalization for end to end image compression[C]//Computer vision—ECCV 2022: 17th european conference, tel aviv, israel, october 23–27, 2022, proceedings, part XVII. Springer, 2022: 390-405.

[21]

WANG T. QIAO M. LIN Z. et al. Generative neural networks for anomaly detection in crowded scenes[J]. IEEE Transactions on Information Forensics and Security, 2018, 14(5): 1390-1399.

[22]

BALLÉ J. MINNEN D. SINGH S. et al. Variational image compression with a scale hyperprior[J]. arXiv Preprint arXiv:1802.01436, 2018.

[23]

SANTURKAR S. BUDDEN D. SHAVIT N. Generative compression[C]//2018 picture coding symposium (PCS). IEEE, 2018: 258-262.

[24]

LI M. LIN J. DING Y. et al. Gan compression: Efficient architectures for interactive conditional gans[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 5284-5294.

[25]

李万臣. 王炼. 基于小波变换和混合神经网络的图像压缩算法[J]. 应用科技, 2006, 33(1).

[26]

VARSHNEY T. SAHOO S. BHATTACHARYYA C. DCT-VAE: Capturing low level and high level features for image generation[J]. 2021.

[27]

YU S. MAZAHARI A. JANNESARI A. Auto graph encoder-decoder for neural network pruning[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 6362-6372.

[28]



YU S. MAZAHARI A. JANNESARI A. Auto graph encoder-decoder for model compression and network acceleration[J]. arXiv Preprint arXiv:2011.12641, 2020.

[29]

MOON A S. LEE S. CHO S H. et al. An efficient neural network based on early compression of sparse CT slice images[C]//2021 international conference on platform technology and service (PlatCon). IEEE, 2021: 1-5.

[30]

PHUONG M. LAMPERT C. Towards understanding knowledge distillation[C]//International conference on machine learning. PMLR, 2019: 5142-5151.

[31]

JTC 1/SC29/WG1 N100279 I ISO/IEC. "Description of the JPEG AI verification model under consideration and associated software integration procedure", 96th meeting, online.[M]. 2022.