Optimization of C3 Entropy Model Architecture via FNLIC Two-Phase Fitting Mechanism

黄程榆，林芷萱，吳柏翰，江瑞秋
Department of Electrical Engineering, National Chung Cheng University, Taiwan.

Abstract

Optimization-based single-image compression methods such as C3 achieve high reconstruction quality through per-image overfitting, but their entropy models are initialized from scratch and lack dataset-level statistical priors, leading to high uncertainty and inefficient bitrate usage. Inspired by FNLIC, which stabilizes lossless compression via probability decomposition, this work transfers the concept of prior decomposition to latent-domain lossy compression. We propose a hybrid entropy modeling architecture that integrates a pre-fitted latent autoregressive prior with C3's original entropy network, reinterpreted as an instance-specific residual model. A learnable scale map adaptively fuses the two branches, improving entropy estimation, coding efficiency, and convergence while preserving C3's single-image optimization framework.

1. Introduction

Recent advances in learned image compression have demonstrated that neural networks can significantly outperform traditional codecs by jointly optimizing rate and distortion through end-to-end training. Most existing approaches adopt an autoencoder-based framework, where an encoder transforms an input image into a compact latent representation, which is then entropy-coded and transmitted to the decoder for reconstruction. These methods have shown strong performance in lossy compression settings, especially when trained on large-scale datasets.

Among these approaches, C3 [1] represents a distinct paradigm that departs from dataset-level training. Instead of relying on a pretrained global model, C3 performs single-image overfitting, optimizing both the synthesis network and the entropy model directly on the target image. This optimization-based strategy allows C3 to adapt precisely to the characteristics of each individual image and achieve high reconstruction quality. However, this flexibility comes at a cost. Since the entropy model in C3 is randomly initialized for each image, it lacks any dataset-level statistical prior, forcing the model to relearn the latent distribution from scratch during encoding. As a result, entropy modeling becomes highly uncertain in the early stages of optimization, leading to inefficient bitrate usage and slow convergence.

In contrast, recent progress in lossless image compression has shown that incorporating explicit prior decomposition can significantly improve entropy modeling. In particular, FNLIC [2] proposes a probability decomposition framework that separates the image distribution into a general prior, learned offline from large datasets, and an instance-specific residual, optimized online for each image. By combining these two components through a lightweight adaptive mechanism, FNLIC achieves efficient lossless compression while preserving exact reconstruction.

Motivated by the success of FNLIC, this project explores whether its core principle—prior decomposition—can be transferred from pixel-domain lossless compression to latent-domain lossy compression. Specifically, we investigate how introducing a dataset-level latent prior into C3's entropy model can reduce uncertainty, improve coding efficiency, and retain the advantages of single-image overfitting.

In this work, our contributions can be summarized as follows:

- We transfer the principle of probability decomposition from pixel-domain lossless compression to latent-domain lossy compression, establishing a novel connection between FNLIC-style prior modeling and optimization-based neural compression.
- We propose a hybrid entropy modeling architecture for C3 that integrates a pre-fitted dataset-level latent prior with single-image overfitting, effectively reducing entropy uncertainty without sacrificing per-image adaptability.
- We introduce an adaptive prior fusion mechanism based on a learnable scale map, which dynamically balances shared latent statistics and instance-specific residuals during encoding.

These contributions improve entropy modeling efficiency and optimization stability while fully preserving C3's original synthesis and decoding framework.

2. Related Work

2.1. Learned Image Compression

Learned image compression has been extensively studied in recent years and has demonstrated remarkable performance improvements over traditional hand-crafted codecs. Most existing methods follow an autoencoder-based framework optimized using a rate–distortion objective, where an encoder maps the input image to a compact latent representation and a decoder reconstructs the image from entropy-coded latents. To improve entropy modeling, early works introduce hyperprior models to capture spatial dependencies and uncertainty in latent variables. Subsequent approaches further incorporate autoregressive models to exploit conditional dependencies among latent features, leading to improved compression efficiency.

Despite their strong rate–distortion performance, these methods are typically trained on large-scale datasets and aim to learn a single global model that generalizes across diverse image content. As a result, their entropy models primarily rely on dataset-level statistics and lack flexibility when adapting to individual images.

2.2. Single-Image Optimization-Based Compression

In contrast to dataset-trained compression models, single-image compression methods adopt an optimization-based paradigm that performs per-image overfitting. Representative approaches such as C3 [1] directly optimize both the synthesis network and the entropy model on the target image, enabling highly adaptive modeling tailored to individual image

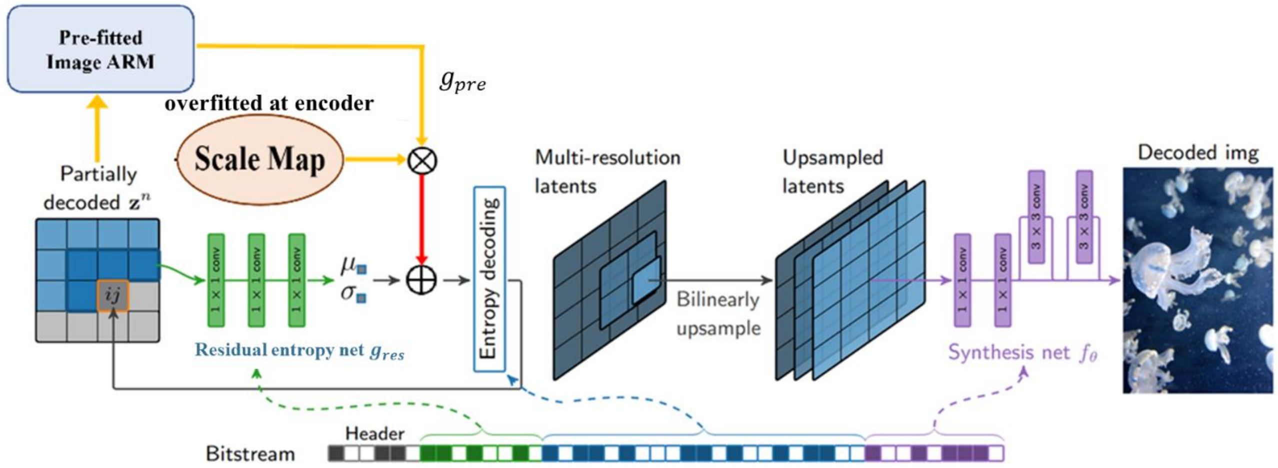

Figure 1. Overview of the proposed hybrid entropy modeling architecture. A pre-fitted latent autoregressive prior provides dataset-level statistics, while the original entropy network of C3 is reinterpreted as a residual model optimized via single-image overfitting. A learnable scale map adaptively fuses the general prior and instance-specific residual. The synthesis network and decoding pipeline of C3 remain unchanged.

characteristics and often achieving high reconstruction quality.

However, this paradigm also introduces challenges. Since the entropy model in C3 is randomly initialized for each image, it lacks access to dataset-level statistical priors. Consequently, the entropy model must learn latent distributions from scratch during encoding, leading to increased uncertainty, inefficient bitrate allocation, and slower convergence in the early stages of optimization.

2.3. Probability Decomposition in Lossless Image Compression

Recent advances in lossless image compression have revisited pixel-level probabilistic modeling using neural autoregressive architectures. FNLIC [2] introduces a two-phase fitting mechanism that explicitly decomposes the image probability distribution into a general prior learned offline from large-scale datasets and an instance-specific residual optimized online for each image. By combining these two components through a lightweight adaptive formulation, FNLIC demonstrates that probability decomposition can stabilize entropy modeling and reduce coding cost while strictly preserving exact reconstruction fidelity.

3. Method

3.1. Overview

To address the absence of dataset-level statistical priors in the entropy model of C3 [1], we propose a dual-branch hybrid entropy modeling architecture inspired by the probability decomposition framework introduced in FNLIC [2]. The key idea is to decompose latent entropy modeling into a shared dataset-level prior and an instance-specific residual, while fully preserving

C3's original single-image optimization paradigm.

An overview of the proposed architecture is illustrated in Figure 1. Importantly, the proposed design modifies **only the entropy modeling component** of C3. The synthesis network, quantization process, and decoding pipeline remain unchanged, ensuring that the core structure and decoding behavior of C3 are preserved.

3.2. Pre-fitted Latent Autoregressive Prior

To introduce dataset-level statistical priors into C3, we incorporate a pre-fitted latent autoregressive model (ARM) that operates directly in the latent domain. The objective of this module is to provide a shared and stable prior over latent variables before single-image overfitting is performed.

Specifically, the input to the pre-fitted ARM is the quantized latent representation $\hat{z}$ produced by the C3 encoder. Since entropy coding is defined over discrete symbols, conditioning the prior on quantized latents ensures strict consistency between encoding and decoding. The pre-fitted ARM predicts general distribution parameters Θ_{pre}, such as the mean and variance of a Gaussian distribution, for each latent position.

This model is trained offline using a large collection of latent representations extracted from a baseline C3 model applied to a training dataset. During the encoding of a target image, the parameters of the pre-fitted ARM are kept frozen, and no gradients are propagated through this branch. Consequently, the pre-fitted ARM functions as a dataset-level latent prior that captures statistical regularities shared across images, without being influenced by instance-specific optimization.

3.3. Residual Entropy Modeling via Single-Image Overfitting

While the introduction of a dataset-level prior improves entropy model initialization, preserving C3's ability to adapt to individual images remains essential. Therefore, we retain C3's original entropy network and reinterpret its role within the proposed architecture.

Instead of modeling the entire latent distribution from scratch, the original C3 entropy network is treated as a residual entropy model. This branch is optimized through single-image overfitting and is responsible for capturing instance-specific deviations from the general prior provided by the pre-fitted ARM. Its output, denoted as Θ_{overfit}, encodes image-dependent characteristics that cannot be explained by dataset-level statistics alone.

By delegating the modeling of shared latent statistics to the pre-fitted branch, the optimization burden during encoding is substantially reduced. This design enables C3 to retain its strong per-image adaptability while avoiding unnecessary relearning of common latent structures across images.

3.4. Adaptive Prior Fusion with a Learnable Scale Map

To combine the dataset-level prior and the instance-specific residual, we introduce an **adaptive prior fusion mechanism** based on a **learnable scale map** $\boldsymbol{\alpha}$.

At each latent location, the final distribution parameters used for entropy coding are computed according as below:

$$\Theta = \alpha \odot \Theta_{pre} + \Theta_{overfit}, \qquad (1)$$

where $\odot$ denotes element-wise multiplication.

The scale map α has the same spatial resolution as the latent feature map and can be defined either per-channel or per spatial location. It is optimized during the single-image encoding process and transmitted as side information in the bitstream. Conceptually, α acts as a *prior trust gate* that dynamically controls the relative contribution of the dataset-level prior and the instance-specific residual.

In regions where latent statistics closely follow common patterns, the entropy model relies more heavily on the pre-fitted prior. Conversely, for regions with unique or complex structures, the residual branch dominates. Importantly, this fusion strategy does not correspond to a mixture of probability distributions; instead, the residual branch refines the parameters of a single parametric distribution, ensuring probabilistic consistency with the original C3 entropy model.

Through this adaptive fusion mechanism, the proposed method effectively leverages stable shared priors while preserving the flexibility of single-image overfitting, resulting in a more efficient and robust entropy modeling process.

4. Experiments

4.1. Experimental Setup

Training Settings For the proposed hybrid entropy modeling framework, the pre-fitted latent prior is trained offline using latent representations extracted from a baseline C3 model. During training, images are processed in a single-image optimization manner following the C3 paradigm. All experiments are conducted using PyTorch with the Adam optimizer. For single-image optimization, the learning rate is set to 0.01, and each image is optimized for up to 5,000 iterations unless otherwise specified. Mean squared error (MSE) is used as the reconstruction loss, combined with a latent regularization term weighted by the rate–distortion parameter λ. The experiments are performed on a single NVIDIA **RTX 3090 GPU**, and each image is processed independently without batch training.

Test Datasets We evaluate the rate–distortion performance of the proposed method on the **Kodak dataset** [3], which consists of 24 natural images with a resolution of 768×512. Kodak is a widely used benchmark for evaluating learned image compression methods and allows direct comparison with prior work. For fair evaluation, all images are processed at their original resolution. When necessary, reflection padding is applied to ensure that image dimensions are divisible by the latent downsampling factor. Rate–distortion curves are obtained by varying the rate–distortion

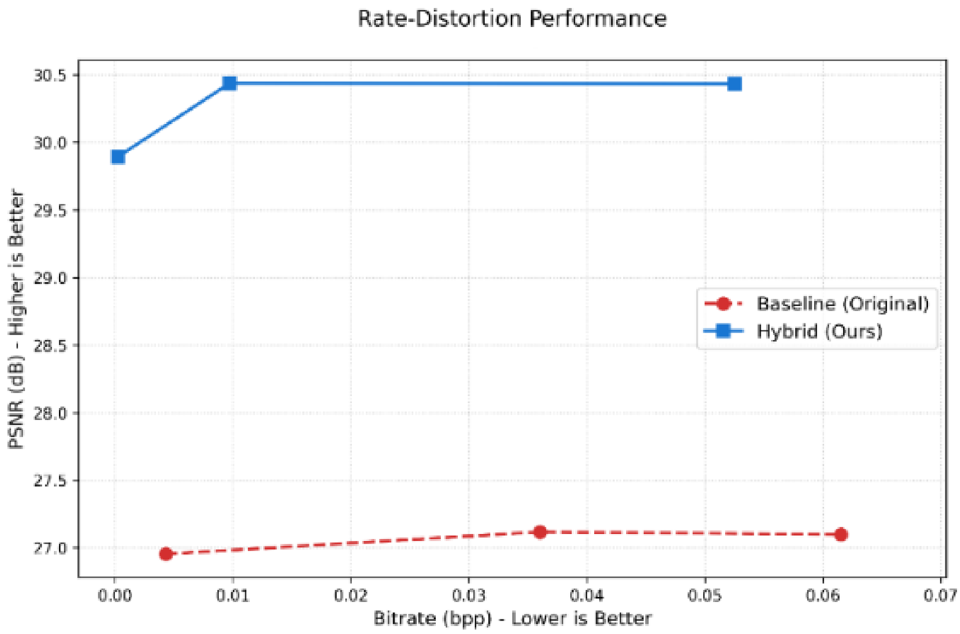

Figure 2. Rate–distortion performance comparison between the original C3 baseline and the proposed hybrid entropy modeling method. PSNR is plotted against bitrate (bpp), where higher PSNR and lower bitrate indicate better compression performance.

parameter λ, and performance is reported in terms of PSNR versus bitrate (bits per pixel).

4.2. Rate–Distortion Performance

Figure 2 shows the rate–distortion (RD) performance comparison between the original C3 baseline and the proposed hybrid entropy modeling method on the Kodak dataset. The RD curves are obtained by varying the rate–distortion parameter λ and reporting PSNR versus bitrate (bits per pixel), where higher PSNR and lower bitrate indicate better compression performance.

As illustrated in Figure 2, the proposed hybrid method consistently outperforms the baseline C3 across the evaluated bitrate range. In the low-bitrate regime, where entropy modeling uncertainty during the early stages of single-image optimization has a pronounced effect, the proposed method achieves substantially higher reconstruction quality. For example, at its lowest operating point, the hybrid model reaches approximately **29.9 dB**, whereas the baseline C3 remains below **27 dB** within the same low-bitrate region. This notable gap highlights the effectiveness of introducing a dataset-level latent prior to stabilize entropy estimation.

At medium bitrates (approximately 0.01–0.05 bpp), the proposed hybrid method maintains a clear and stable performance margin. Specifically, the hybrid approach achieves PSNR values around **30.4–30.5 dB**, while the baseline remains near **27.1 dB** at comparable bitrate levels. This consistent improvement indicates that the residual entropy model effectively complements the pre-fitted prior by capturing image-specific latent characteristics beyond shared statistical patterns.

Even at higher bitrates, where the baseline benefits from increased optimization freedom, the proposed method continues to demonstrate superior performance. This suggests that the benefit of FNLIC-inspired prior decomposition extends beyond early optimization stages and remains effective throughout the rate–distortion spectrum.

Overall, these results confirm that integrating a dataset-level latent prior into the entropy modeling component of C3 leads to consistent and robust rate–distortion improvements. Importantly, the observed gains are achieved without modifying the synthesis network or decoding pipeline, validating that combining a shared latent prior with instance-specific residual entropy modeling is an effective strategy for enhancing optimization-based neural image compression.

5. Conclusions

This work addresses a key limitation of optimization-based single-image compression methods, namely the absence of dataset-level statistical priors in the entropy model of C3. Inspired by the probability decomposition framework of FNLIC, we transfer this principle from pixel-domain lossless compression to latent-domain lossy compression and propose a hybrid entropy modeling architecture that integrates a pre-fitted latent autoregressive prior with C3's original entropy network.

The proposed method decomposes the latent probability model into a shared dataset-level prior and an instance-specific residual, which are adaptively fused using a learnable scale map while preserving C3's original synthesis and decoding pipeline. Experimental results on the Kodak dataset show that the proposed approach consistently improves rate–distortion performance over the C3 baseline, with particularly strong gains at low bitrates. These findings demonstrate that incorporating dataset-level latent priors is an effective strategy for enhancing optimization-based neural image compression.

References

[1] H. Kim, M. Bauer, L. Theis, J. R. Schwarz and E. Dupont, "C3: High-Performance and Low-Complexity Neural Compression from a Single Image or Video," 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2024, pp. 9347-9358, doi: 10.1109/CVPR52733.2024.00893.

[2] Z. Zhang, Z. Chen and S. Liu, "Fitted Neural Lossless Image Compression," 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2025, pp. 23249-23258, doi: 10.1109/CVPR52734.2025.02165.

[3] Eastman Kodak. Kodak lossless true color image suite (photocd pcd0992), 1993.