MoRIC: A Modular Region-based Implicit Codec for Image Compression

Gen Li*, Haotian Wu*†, Deniz Gündüz

* Equal contribution; † Project leader.

Project page: `https://eedavidwu.github.io/MoRIC/`

10.30.2025

Outline

1. Preliminary
2. C* coding
3. MoRIC scheme
4. Experimental Results
5. Conclusion and Future Work

1. Preliminary

From data representation to function representation

Parameterize a discrete signal as a **continuous** function.

Use neural networks to approximate the mapping from coordinates to signal intensities.

Shift the paradigm of data representation from feature-based to **function-based** representation.

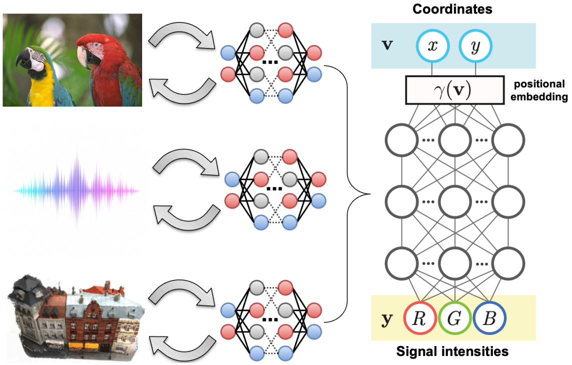


Fig. 1: Examples of INRs for various modalities.

1. Preliminary
From AE-based neural codec to overfitted codec

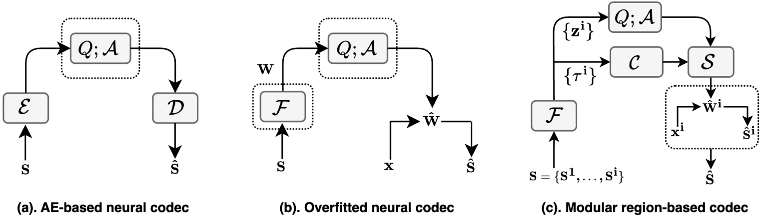


Fig. 2: Operational diagram of different compression models.

AE-based codecs leverage **advanced** architectures and **large-scale** datasets to achieve strong rate-distortion (RD) performance.

- An encoder-decoder pair maps the source to a quantized latent, which is entropy-coded to form the bitstream and decoded to reconstruct the signal.
- Developing **low-complexity**, **robust** codecs with **strong** RD performance remains an open challenge.

1. Preliminary

From AE-based neural codec to overfitted codec

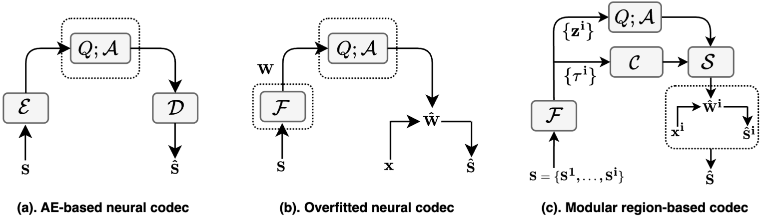


Fig. 2: Operational diagram of different compression models.

Overfitted codecs instead parameterize **each** data sample using **lightweight** neural functions, aiming for a **good, cheap, and fast** compression scheme.

- Each source sample is represented by a neural function and an alternative latent vector.
- Decoding complexity is extremely low, as no data-generalization is required.
- Outperforming some existing codecs, such as BPG, HEVC, and BMS.

1. Preliminary
From overfitted codec to MoRIC

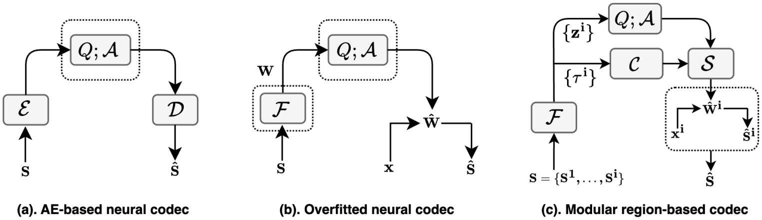


Fig. 2: Operational diagram of different compression models.

Motivations:

?

- Can an image be efficiently compressed through region-wise specialization rather than global overfitting? $\rightarrow$**local adaptability.**
- Can flexible modular INRs better capture diverse regional statistics and complexity? $\rightarrow$ **fine-grained modeling.**

Towards a **flexible**, **region-wised** codec with **fine-grained** RD control $\rightarrow$ **MoRIC**.

2. C* Chain Coding

A New Adaptive lossy compression strategy for contour

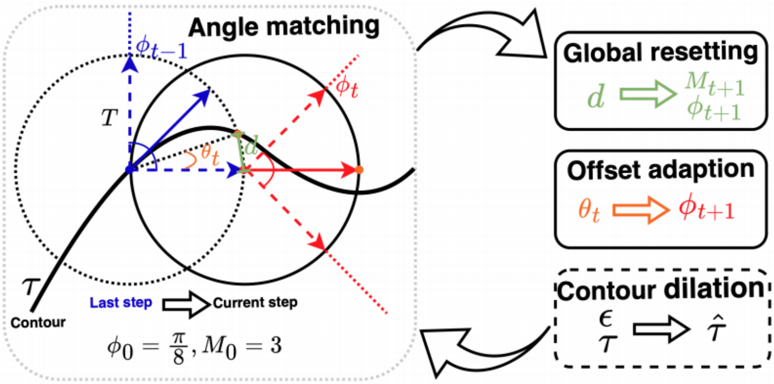


Fig. 3: Illustration of C* coding

2. C* Chain Coding

An adaptive chain coding method (called C^* coding) for efficient contour compression in region-based overfitted codecs.

C^* coding

C* coding approximates object contours using polygonal chains. Starting from an initial point, each contour segment is encoded by a fixed step length and a quantized offset angle relative to the previous direction.

It ensures (a) semantic coverage of target objects against artifacts, (b) robustness to high-curvature via adaptive dilation, and (c) a balanced RD trade-off. As illustrated in Fig. 3, C^* coding comprises four key modules: angle matching, global resetting, offset adaptation, and contour dilation.

3. MoRIC: A Modular Region-based Implicit Codec for Image Compression

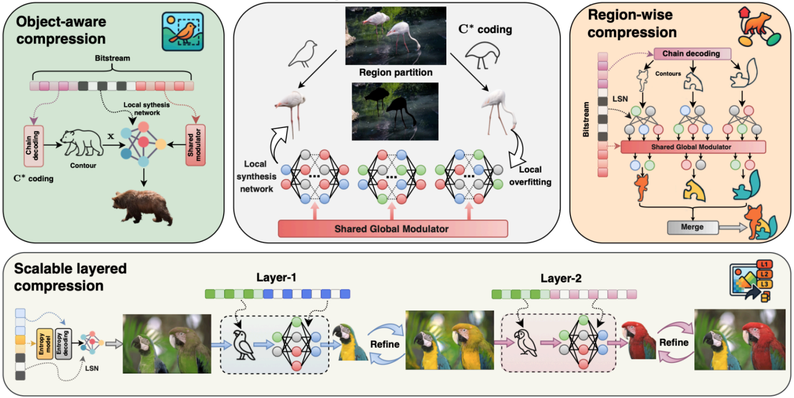


Fig.4: Illustration of MoRIC.

Long-term vision: Make overfitted compression “more than” single-image compression.

- From local to global: “Divide and conquer” $\rightarrow$ Coding an image like “playing LEGO”.
- INR has hallucinations/imaginations.

3. MoRIC Scheme

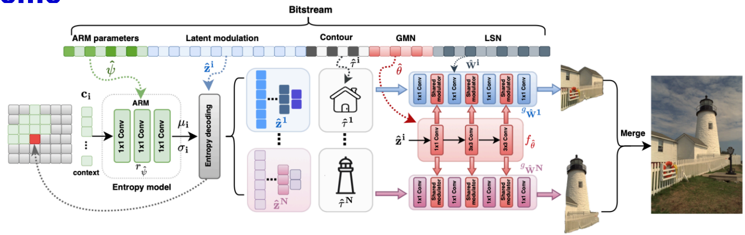


Fig. 5: Image decoding process in MoRIC.

Rate and distortion expressions:

$$\mathsf{R} = \mathbb{E}_{\mathbf{S}\sim p_s}\left[\sum_{i=0}^{N}\left[-\log_2 p_{\hat{\psi}}(\hat{z}^i) - \log_2 p(\hat{W}^i) - \log_2 p(\tau^i)\right] + R_{\hat{\psi}}\right], \mathsf{D} = \mathbb{E}_{\mathbf{S}\sim p_s}\left[d(\mathbf{S}, \bigcup_{i=1}^{N} g_{W^i}(\hat{z}^i, x^i))\right]$$

RD cost optimization: $\mathcal{L} = \mathsf{D} + \lambda \mathsf{R}(\hat{z}).$

3. MoRIC scheme

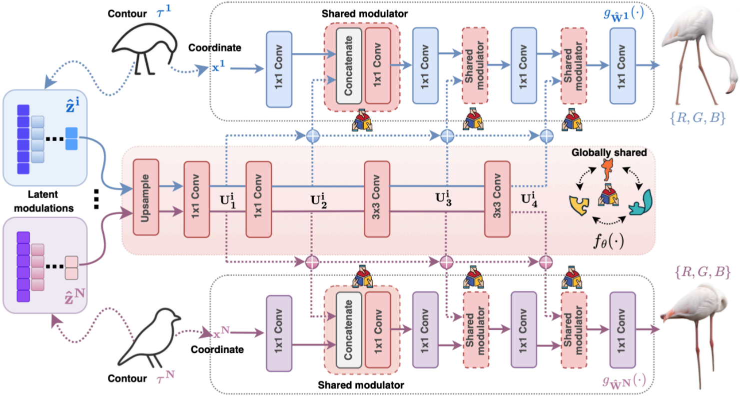
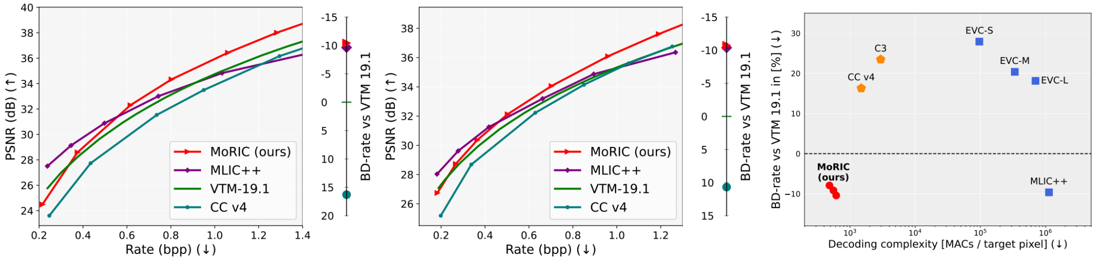


Fig. 6: Illustration of the Progressive Concatenated Modulation mechanism in MoRIC.

- Each Local Synthesis Network (LSN) is modulated by multi-scale latent features generated through the **Global Modulation Network (GMN)**.
- Progressive Concatenated Modulation mechanism enable shared-weight modulation $\mathcal{M}(\cdot)$ to inject global context while preserving local detail fidelity.

4.1 Single object compression



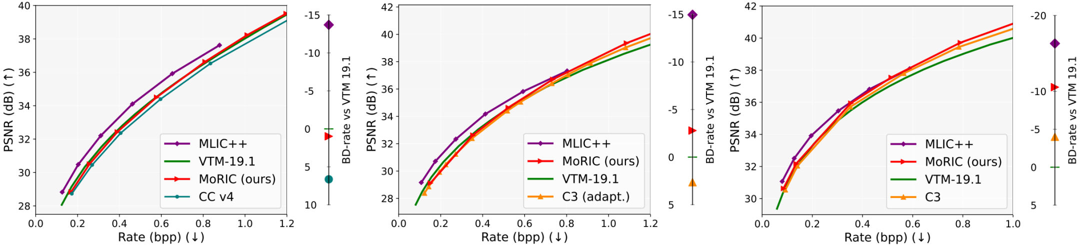
(a) Davis dataset (b) Kodak dataset (c) Davis dataset

Fig.7: RD performance of MoRIC for single-object compression.

Experimental Observations

- MoRIC achieves state-of-the-art RD performance for single-object compression on both DAVIS and Kodak.
- It achieves superior decoding efficiency with fewer than 800 MACs/pixel, outperforming all AE-based codecs while maintaining high-fidelity reconstructions through its shared-layer design.

4.2 Standard image compression



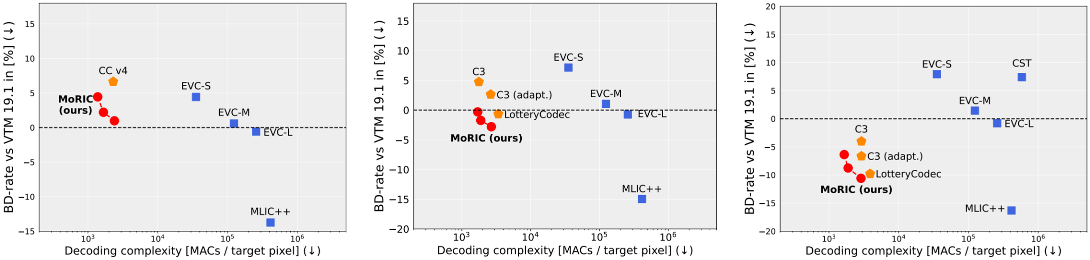
(a) Davis dataset (b) Kodak dataset (c) CLIC2020 dataset

Fig.8: RD performance of MoRIC for standard image compression.

Experimental Observations

- MoRIC achieves state-of-the-art RD performance among overfitted codecs for full-image compression, surpassing C3 and COOL-CHIC (v4).
- It further outperforms VTM-19.1 on Kodak and CLIC2020 datasets with BD-rate gains of –2.79% and –10.58%, respectively, while achieving over 100× lower decoding complexity than MLIC++.

4.2 Standard image compression



(a) Davis dataset (b) Kodak dataset (c) CLIC2020 dataset

Fig.9: BD-rate vs. decoding complexity.

Experimental Observations

- MoRIC delivers superior RD–complexity trade-off, achieving competitive BD-rate with fewer than 2000 MACs/pixel, the lowest among all overfitted codecs.
- It is the first implicit neural codec to surpass VTM-19.1 on Kodak in RD performance while maintaining extremely low decoding complexity.

4.3 Layered compression

Fig. 10: Visualization of layered compression.

- The background LSN $g_{\mathbf{W}^0}$ is first transmitted as layer 0, followed by $g_{\mathbf{W}^1}$ for the blue region (layer 1), and finally $g_{\mathbf{W}^2}$ for the red region (layer 2). We can see that colors and textures (e.g., red hat and golden ball) are refined with the additional LSN parameters.

4.4 Coding efficiency

Table: Average coding time for **standard full image compression** on Kodak dataset using an NVIDIA RTX 3090 GPU and Intel Core i9-10980XE CPU @ 3.00GHz. Orange indicates GPU computation, blue indicates CPU computation, and bold numbers highlight the best results.

Models	**Encoding time**	**Decoding time**
VTM-19.1	87.13 s	293.05 ms
EVC (S/M/L)	**22.43/36.19/47.87** ms	**19.89/24.12/31.91** ms
MLIC^{++}	271.91 ms	364.06 ms
C3 (fixed / adapt)	17.49/21.33 sec / 1k steps	**283.59** / **310.54** ms
MoRIC (single region)	24.70 sec / 1k steps	343.55 ms
MoRIC (two region)	26.41 sec / 1k steps	408.97 ms
MoRIC (three region)	32.06 sec / 1k steps	486.48 ms

5. Conclusion and Future Work

- We presented **MoRIC**, a modular region-based implicit codec that compresses images via **region-wise** overfitting, enabling fine-grained **rate–distortion control**.
- MoRIC achieves **state-of-the-art** RD performance in both single-object and full-image compression with **low**, **adaptive** decoding complexity.
- It can be extended as an alternative for **video coding**, offering better RD performance and control over complexity and rate.
- Project page: `https://eedavidwu.github.io/MoRIC/`

Scan for code and resources