Towards Lossless Implicit Neural Representation via Bit Plane Decomposition

Woo Kyoung Han[1] Byeonghun Lee[1] Hyunmin Cho[1] Sunghoon Im[2*] Kyong Hwan Jin[1*]
[1]Korea University [2]DGIST
{wookyoung0727, byeonghun_lee, hyun_cho, kyong_jin}@korea.ac.kr, sunghoonim@dgist.ac.kr

Abstract

We quantify the upper bound on the size of the implicit neural representation (INR) model from a digital perspective. The upper bound of the model size increases exponentially as the required bit-precision increases. To this end, we present a bit-plane decomposition method that makes INR predict bit-planes, producing the same effect as reducing the upper bound of the model size. We validate our hypothesis that reducing the upper bound leads to faster convergence with constant model size. Our method achieves lossless representation in 2D image and audio fitting, even for high bit-depth signals, such as 16-bit, which was previously unachievable. We pioneered the presence of bit bias, which INR prioritizes as the most significant bit (MSB). We expand the application of the INR task to bit depth expansion, lossless image compression, and extreme network quantization. Our source code is available at https://github.com/WooKyoungHan/LosslessINR.

1. Introduction

Implicit neural representations (INRs), parameterizing the continuous signals with an artificial neural network (ANN), have been in the spotlight in various areas for recent years. From a signed distance function by Park et al. [28] to the best-known research by Mildenhall et al. [25] for radiance fields, INR shows promising performance in many fields [5, 12, 13, 17, 19, 24, 36, 39]. The fundamental principle of INR, which aims to train real-world signals with parameters operating in a range and domain of a continuous set, inspired various applications such as super-resolution [5, 19, 20] and a novel view synthesis [25].

However, research exploring precision close to the continuous range, i.e., analog, has not been actively pursued. Computers operate with digital signals, not analog ones, with values constrained by quantization, such as 8-bit or 16-bit for images and 24-bit for audio. Therefore, signal representation necessitates the concept of quantization, where bit precision-

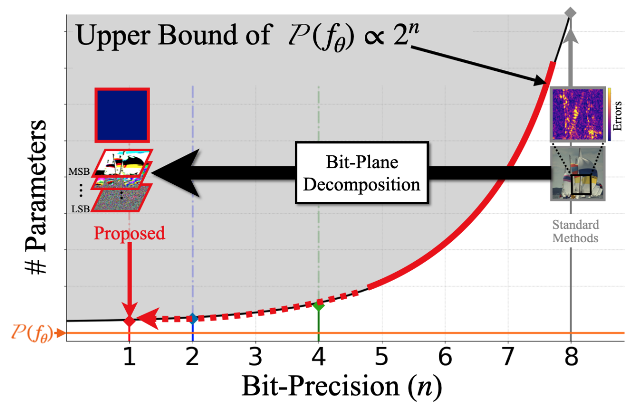

Figure 1. **Overview** of the proposed method and error maps at 1,500 iterations. The upper bound on the number of parameters ($\mathcal{P}(f_\theta) \propto 2^n$) of INR ($f_\theta$) grow proportionally to a bit-precision (n). We propose a bit-plane decomposition method, reducing the upper bound, enabling faster convergence, and ultimately achieving a lossless representation. The closer $\mathcal{P}(f_\theta)$ is to the upper bound, the faster it converges, enabling lossless representation

the number of bits required to represent the signal-serves a pivotal function. Furthermore, lossless representation is defined as satisfying the given bit precision across all input values. Although existing methods produce high-quality images, achieving complete lossless representation remains challenging, especially for high dynamic range images like 16 bits.

In this paper, based on "The implicit ANN approximations with described error tolerance and explicit parameter bounds" by Jentzen et al. [16], we quantify the upper bound of the size of an INR with given bit-precision. Fig. 1 show that the theoretical upper bound of model size increases as an exponential function proportional to required bit-precision. We suggest a method to reduce the quantified upper bound by bit-plane decomposition. We decompose the signal into bit-planes and represent them, leading to lossless representation. Our method is based on the hypothesis that INR reaches the target error—the maximum allowable error to ensure a lossless representation—as the upper bound approaches the model size. We validate our hypothesis through experiments in Fig. 8. As in Fig. 2, our method makes INR represent a lossless signal in a bit-for-bit manner, which was previously

*Corresponding author.

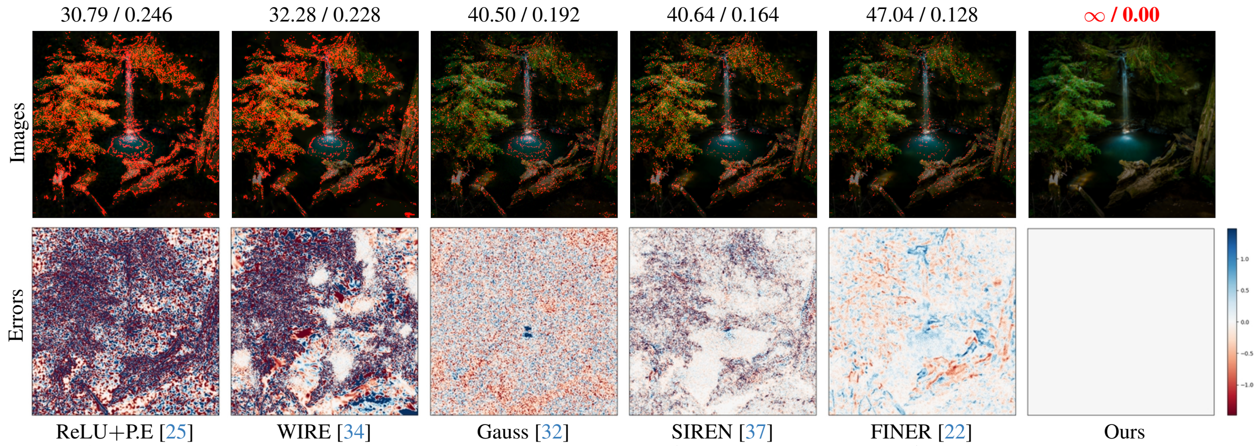

Figure 2. **Visual Demonstration** of representing 2D Image (PSNR(dB) ↑/ Bit-Error-Rate (BER) ↓ at top of images). ReLU with position encoding (P.E) [25], WIRE [34], gaussian activation [32], SIREN [37], FINER[22] and ours. We highlight the occurrence of significant errors with red dots.

unachievable.

We discovered that INRs learn the most significant bits (MSBs) faster than the least significant bits (LSBs) regardless of activations. We named the observed phenomenon '*Bit Bias*.' Additionally, we empirically show that the frequency of the bit axis also has a bias in learning. We demonstrate three applications utilizing our method: lossless compression through lossless representation, bit-depth expansion through a bit axis, and ternary INR through robustness on weight quantization.

In summary, our main contributions are as follows:

- We quantify the upper bound on the number of parameters of the implicit neural network based on the given bit precision.
- We propose a bit-plane decomposition for lossless implicit neural representation and validate our hypothesis that reducing bit precision lowers the upper bound on the number of parameters, leading to faster convergence to lossless representation compared to other networks.
- We discovered the existence of bit bias, where most significant bits converge faster than least significant bits, as in spectral bias.
- Our approach extends the application of INR to lossless compression, bit-depth expansion, and model quantization.

2. Related Work

Implicit Neural Representation INRs present signals with an ANN that takes spatial coordinates as input. Research to improve the performance of INRs has been conducted to enhance the low representing power of multi-layer perceptrons (MLP). Sitzmann et al. [37], addressed this challenge by employing a sine activation function and inspired researchers to apply various activations to INRs [22, 32, 34]. Research about enhancing the capacity of INR [6, 21, 26, 46, 47] have been conducted. Müller et al. [26], and Xie et al. [46] utilize hash-tables in INRs to enable faster training and accommodate larger signals compared to other methods. However, previous studies were not interested in making concrete lossless representations. The application of INR is also gaining attention. Specifically, several approaches [8, 9, 11, 38] extend the application of INR to lossy compression. Recently, the methods [11, 14] demonstrate remarkable performance by learning a Bayesian INR and encoding a sample.

Even by increasing the parameters significantly, INRs have difficulty achieving representations aimed at high accuracy as in Fig. 3. We focus on reducing the required model size based on bit precision and propose a bit-plane decomposition method to achieve lossless representation with a sufficiently sized model. Our approach offers new applications that were not proposed in existing INRs. We devise a lossless compression approach by combining lossless representation with existing methods. Using bit depth as an axis enables a bit-depth expansion through extrapolation. We propose a ternary INR that utilizes the robustness of our approach to weight quantization.

Spectral Bias Rahaman et al. [31] have shown the presence of a *spectral bias* which makes it challenging for INR to learn high-frequency components. To address this challenge, approaches [19, 25, 37, 41] that map coordinates into sinusoidal functions have been proposed. The prior works [25, 41] suggested fixed frequencies to solve spectral bias known as position encoding, while Lee and Jin [19] proposed a learnable position encoding which enables INR to learn continuous Fourier spectra. We reinterpret the spectral bias in the bit-plane aspect, which is our proposed method's core concept. Bit-plane is mainly used for image dequantization [12, 30] or vision model quantization tasks [48]. We found the existence of a similar phenomenon like spectral bias, which we call *bit bias*. Our method achieves lossless representations efficiently by mitigating this phenomenon.

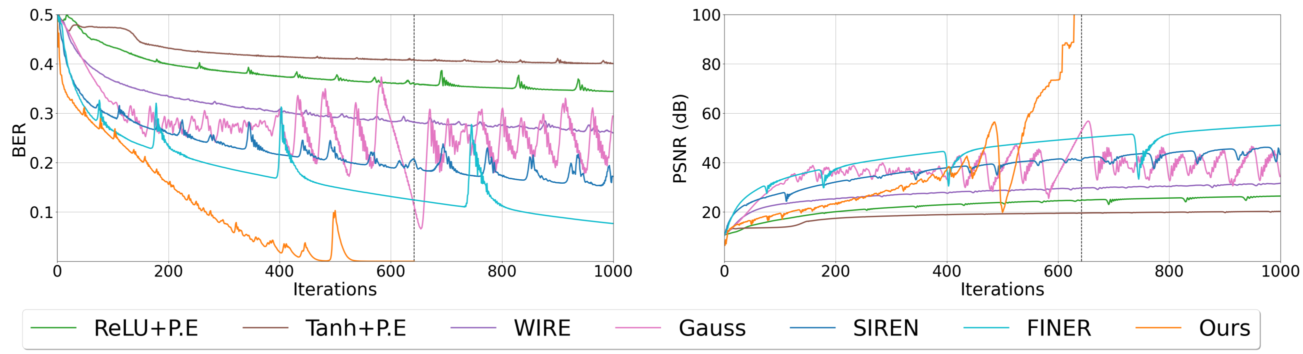

Figure 3. **Training curve** on a single image of DIV2K[1] dataset. Bit-Error-Rate(BER) (left) and PSNR (right). Vertical lines indicate the iteration when the model achieves lossless representation.

3. Method

In this section, we quantify the upper bound of the INR based on the given bit-precision, grounded in theory [16]. Our proposed bit-plane decomposition method reduces the quantified upper bound and accelerates the attainment of the target error bound, leading to lossless representation.

3.1. Preliminary

Quantization A quantization is an inevitable function for all signals to convert analog to digital, which is defined below:

$$\mathcal{Q}_n(\hat{x}) := \arg\min_x ||x - \hat{x}||_1 \quad (x \in Q_n), \tag{1}$$

where $n \in \mathbb{N}$ denotes bit precision, and $Q_n \subset \mathbb{Q}$ is a finite set. We assume the elements of Q_n have normalized and uniformly distributed (e.g., $Q_8 = \{0, \frac{1}{2^8-1}, \frac{2}{2^8-1}...1\}$ for 8-bit images). We set the range dimension of a function as 1 without loss of generality. Let $h : \mathbb{R}^d \to \mathbb{R}$ be a continuous and analog function in d-dimensional space and let h_n be a digital function with n-bits precision:

$$h_n : \mathbb{R}^d \xrightarrow{h} \mathbb{R} \xrightarrow{\mathcal{Q}_n(\cdot)} Q_n. \tag{2}$$

With our assumption, a ceiling of error $\epsilon(n)$ between h and h_n is a function of precision defined as below:

$$\epsilon(n) := \frac{1}{2(2^n - 1)}. \tag{3}$$

Explicit Bounds Jentzen et al. [16] have demonstrated the explicit upper bounds of the number of parameters of ANNs. This provides a specific number of parameters regarding the particular error tolerance proposed in the universal approximation theory (UAT) [7].

Let $L \in \mathbb{R}$ is a Lipschitz constant that satisfy $\forall \mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^d$ that $||h(\mathbf{x}_i) - h(\mathbf{x}_j)||_1 \leq L||\mathbf{x}_i - \mathbf{x}_j||_1$. Then, there exists MLP ($:= h_\theta$) that holds:

1. The number of parameters: $\mathcal{P}(h_\theta) \leq \mathfrak{C}\epsilon^{-2d} (:= \mathcal{U}_d)$,
2. The ceiling of the error: $\sup ||h_\theta(\mathbf{x}) - h(\mathbf{x})||_1 \leq \epsilon$,

where $\mathcal{P}(\cdot)$ is the number of parameters and $\mathfrak{C}$ is a constant determined by the condition of the domain. We provide details of the theorem and $\mathfrak{C}$ in the supplement material.

3.2. Problem Formulation

Lossless Representation A lossless representation requires having n-bit precision, where n remains identical to that of the ground truth digital signal at every point. The INR, coordinate-based MLP (h_θ), aims to parameterize a function h with trainable parameters, θ. Since our target is to represent h_n, the output of h_θ should map to Q_n for digital representation as Eq. (2). The parameterized function $h_\theta : \mathbb{R}^d \to \mathbb{R}$ achieves n-bit precision with respect to an analog function $h : \mathbb{R}^d \to \mathbb{R}$ at $(\mathbf{x}, h_n(\mathbf{x}))$, if and only if $\mathbf{x} \in \mathbb{R}^d$, the predicted output $\hat{h}_\theta(\mathbf{x})$ satisfies:

$$\mathcal{Q}_n(\hat{h}_\theta(\mathbf{x})) = h_n(\mathbf{x}) \leftrightarrow \hat{h}_\theta(\mathbf{x}) \in [h_n(\mathbf{x}) - \epsilon(n), h_n(\mathbf{x}) + \epsilon(n)]. \tag{4}$$

where $\hat{h}_\theta$ indicate predicted values. If a parameterized function h_θ satisfies Eq. (4) at $\forall \mathbf{x} \in \mathcal{X}$, it is defined to have n-bit precision with respect to an analog function $h : \mathbb{R}^d \to \mathbb{R}$. i.e.:

$$\sup_{\mathbf{x} \in \mathcal{X}} ||h_n(\mathbf{x}) - \hat{h}_\theta(\mathbf{x})||_1 \leq \epsilon(n). \tag{5}$$

The lossless representation is identical to make h_θ satisfy Eq. (5). According to UAT [7], it is known that if there is a sufficient number of parameters Eq. (5) be satisfied. However, efficient methods to achieve Eq. (5), especially for large n, have not been well studied.

Upper Bound Based on the Sec. 3.1, we derive the upper bound ($\mathcal{U}_d$). The 'upper bound' represents the threshold where more parameters don't improve the representation. $\mathcal{U}_d$ takes bit precision (n), the number of bits required to represent a signal, and signal's dimension (d) as a dependent variable:

$$\mathcal{U}_d(n) := \mathfrak{C}\epsilon(n)^{-2d} = \mathfrak{C}(2^{n+1} - 2)^{2d}. \tag{6}$$

In conclusion, the upper bound $\mathcal{U}_d(n)$ of the h_θ increases exponentially as a function of the given bit n. We hypothesize that h_θ achieves Eq. (5) faster more efficiently as $\mathcal{P}(h_\theta)$ approaches $\mathcal{U}_d$. We validate the hypothesis in Sec. 4.2 by adjusting n of target signals. We decrease the upper bound $\mathcal{U}_d(n)$ by reducing the required precision (n).

				Tanh+P.E [25]		ReLU+P.E [25]		WIRE [34]		Gauss [32]		SIREN [37]		FINER [22]		**Ours**	
16-Bit	Iterations (↓)			5000												**3450**(±877)	
	TESTIMAGES [2]	PSNR	SSIM (↑)	25.63	0.5447	35.91	0.8229	45.36	0.9341	69.40	0.9928	78.52	0.9969	80.17	0.9995	∞	1.0000
		RMSE	BER (↓)	3428.1	0.4224	1049.5	0.3794	359.17	0.3232	22.217	0.2167	7.7672	0.1544	6.4222	0.1498	0.0000	0.0000
	MIT-5k [3]	PSNR	SSIM (↑)	26.95	0.4644	37.43	0.8064	45.61	0.9241	66.90	0.9795	78.39	0.9970	86.48	0.9987	∞	1.0000
		RMSE	BER (↓)	2943.5	0.4195	880.50	0.3734	343.46	0.3202	29.611	0.2306	7.8885	0.1560	3.1079	0.1138	0.0000	0.0000
8-Bit	Iterations (↓)			1000												**790**(±109)	
	DIV2K [1]	PSNR	SSIM (↑)	21.08	0.4956	27.68	0.8124	35.69	0.9572	54.70	0.9940	47.19	0.9955	55.03	0.9989	∞	1.0000
		RMSE	BER (↓)	22.519	0.3921	10.533	0.3318	4.1864	0.2365	0.4694	0.0953	1.1144	0.1343	0.4519	0.0528	0.0000	0.0000
	Kodak [10]	PSNR	SSIM (↑)	23.94	0.5894	30.94	0.8473	37.86	0.9532	48.28	0.9864	47.28	0.9919	55.93	0.9985	∞	1.0000
		RMSE	BER (↓)	16.201	0.3709	7.2378	0.3036	3.2634	0.2234	0.9830	0.1456	1.1029	0.1382	0.4074	0.0447	0.0000	0.0000

Table 1. **Quantitative comparison** on 16-bit (top) and 8-bit (bottom) image fitting with existing INR methods. The iteration number of our methods indicates '$mean(\pm std)$' for the total dataset. The text color red/blue indicates the best and second-best, respectively.

3.3. Methodology

Bit-Plane Decomposition Let n-bit images be $\mathbf{I}_n : \mathbb{R}^2 \rightarrow Q_n^3$. By employing Eqs. (5) and (6), $\mathcal{U}_d(n)$ that required to ensure representing lossless $\mathbf{I}_8$ and $\mathbf{I}_{16}$ are significant numbers ($> 10^{10} \cdot \mathfrak{C}, 10^{20} \cdot \mathfrak{C}$, respectively). To this end, we suggest a *bit-plane decomposition* method for implicit neural representation. Instead of n-bit images, we decompose images into bit-planes and represent them. Bit-planes are binary images[1] $\mathbf{B}^{(i)} \in \{0,1\}^{H\times W\times 3}$ that satisfy:

$$\mathbf{I}_n = \frac{1}{2^n - 1}\sum_{i=0}^{n-1} 2^i \mathbf{B}^{(i)}, \quad (7)$$

where, i denote i-th least significant bit-plane. Our method reduces a bit precision n to 1, thereby reducing $\mathcal{U}_d(1) = 16\mathfrak{C}$. As a result, our approach brings the number of parameters closer to Eq. (6) and makes it easier to achieve Eq. (5).

A straightforward method for representing an n-bit signal is employing parallel sequence of INRs i.e. $[f_\theta^{(i)}]$, each representing bit-plane $[\mathbf{B}^{(i)}]$. Inspired by the recent dequantization approach proposed by Han et al. [12], we propose a method that employs an additional coordinate (i) to represent an n-bit image, as shown below:

$$\mathbf{B}^{(i)}(\mathbf{x}) \simeq f_\theta(\mathbf{x}, i), \quad (8)$$

where, $\mathbf{x} \in \mathbb{R}^2$ indicates spatial coordinate. Eq. (8) is motivated by the fact that each element of $[\mathbf{B}^{(i)}]$ is not independent but highly correlated. In other words, our proposed method considers an image as a 3-dimensional function with the bit coordinate.

Loss Function We optimize our parameters with the equation below:

$$\hat{\theta} = \arg\min_\theta \mathcal{L}(\mathbf{B}^{(i)}(\mathbf{x}), \hat{f}_\theta(\mathbf{x}, i)), \quad (9)$$

where, $\mathcal{L}$ indicates a loss function. In our approach, candidate loss function $\mathcal{L}$ to optimize parameters θ include general regression losses such as $||\cdot||_p$ with $p = 1, 2$. Furthermore, the optimization problem in our approach with $k = 1$ can be considered as a binary classification problem. We observed that the binary cross-entropy (BCE) loss effectively optimizes θ and conducted the ablation study in Sec. 5.

[1]Note that $\{0, 1\} = Q_1$

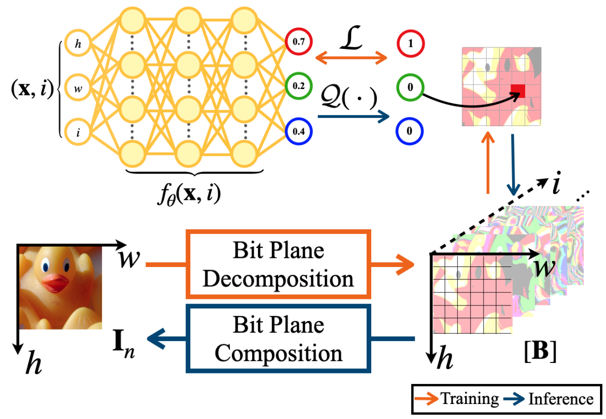

Figure 4. Overall process of our proposed method. We improve the performance of INR by lowering the upper bound of the number of parameters ($\mathcal{U}$) and achieve **lossless neural representation.**

Fig. 4 shows the overall process of our method. After training, we reassemble quantized images to n-bit precision representation using Eqs. (1), (2) and (7) i.e.:

$$\mathbf{I}_n(\mathbf{x};\theta) : \underbrace{\overbrace{(\mathbf{x}, i) \overset{f_\theta}{\mapsto} [\hat{\mathbf{B}}]}^{\text{Training}} \overset{\mathcal{Q}(\cdot)}{\longrightarrow} [\mathbf{B}] \overset{Eq.\ (7)}{\longrightarrow} \mathbf{I}_n}_{\text{Inference}}. \quad (10)$$

Note that $\hat{\mathbf{B}} \in \mathbb{Q}^{H\times W\times 3}$ which satisfy Eq. (4) for all coordinates. In summary, our method represent $\mathbf{I}_n$ as below:

$$\mathbf{I}_n(\mathbf{x};\theta) = \frac{1}{2^n - 1}\sum_{i=0}^{n-1} 2^i \mathcal{Q}(\hat{f}_\theta(\mathbf{x}, i)). \quad (11)$$

The extended description of the method is in the supplement material for n-ary representations of Fig. 8.

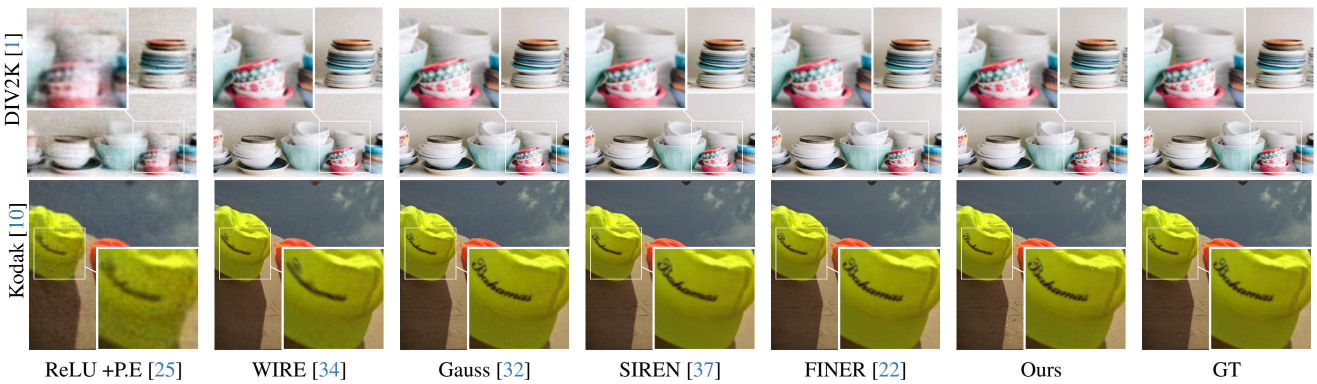

Figure 5. Qualitative comparison of under-fitted images (# of Iterations : 400) with existing methods.

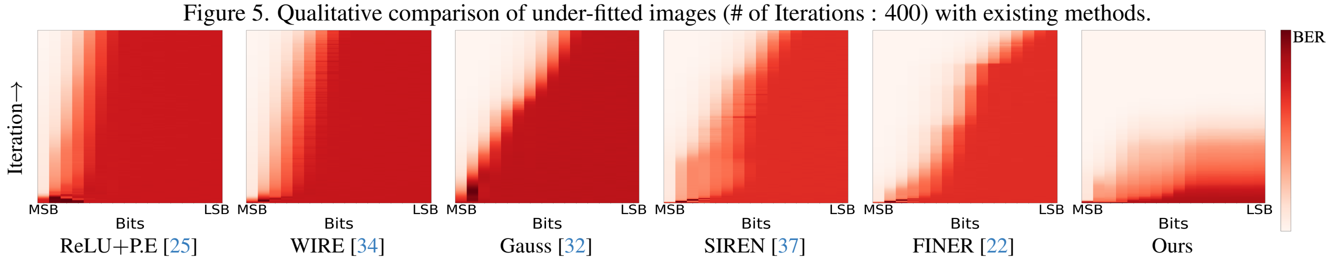

Figure 6. **Bit-Error-Rate** of each bit-plane on a TESTIMAGE [2]. The X-axis is for bit depth (MSB to LSB), and the Y-axis is for iteration.

	Kodak [10]		TESTIMAGES [2]	
	#Iter.(↓)	PSNR (↑)	#Iter.(↓)	PSNR (↑)
Instant-NGP [26]	2000	52.82	5000	54.92
Instant-NGP + **Ours**	1130	∞	4668	∞
DINER [46]	5000	39.59	5000	38.30
DINER + **Ours**	3347	∞	3915	∞
Gauss [32]	15000	100.48	50000	74.88
Gauss + **Ours**	7931	∞	29546	∞
FINER [22]	500	48.52	2000	56.58
FINER + **Ours**	428	∞	1464	∞

Table 2. Quantitative comparison results combining existing methods with ours. Coordinate encoding method (top) and activation modification method (bottom).

4. Experiments

4.1. Implementation Details

To validate our proposed method, we conduct experiments on MIT-fiveK [3] and TESTIMAGES1200 [2] dataset that require high dynamic range (0-65,535). The TESTIMAGES dataset includes 40 natural images. We select the last 1,000 images (with indices from 4,001 to 5,000) labeled by expert E in the MIT-fiveK dataset. We also conducted the representation experiments on general 8-bit imaged datasets: validation set of DIV2K [1], which includes 100 images, and Kodak [10], containing 24 images. All images are center-cropped and downsampled to a size of 256. Coordinates are normalized to $[-1, 1]$ as per prior works. Our method is compared with existing methods, including Tanh and ReLU activations with position encoding (+P.E) [25], wavelet [34], Gaussian [32], sine activation [37], and variable periodic activation [22]. All reported values, including baselines, are evaluated after the quantization (Eq. (1)). For a fair comparison, we take the average and standard deviation of the number of iterations and train baselines for a larger number than our average. We adopt the sine activation function for generality and use the BCE loss function unless otherwise stated. In Sec. 5, we conduct ablation studies on activation functions and loss functions. All networks have an identical number of parameters: 5 hidden layers, each with 512 dimensions, ensuring a fair comparison. We use NVIDIA RTX 3090 24GB for training and optimized all networks by Adam [18], with a 1e-4 learning rate.

4.2. Image Representation

Validation We quantify the theoretical upper bound $\mathcal{U}$ of INRs with a given bit precision. We provide experimental evidence supporting our hypothesis: if $\mathcal{P}(f_\theta)$ is close to the upper bound $\mathcal{U}_d$, then it is more efficient to achieve Eq. (5). We set all networks with the same number of parameters and set bit-precision (n) as a variable. The detailed figure is in the supplementary. Fig. 8 shows the experiment results for our hypothesis. In Tab. 3, the proposed method performs best against others regarding fast convergence. There are two reasons for fast convergence while the upper bound ($\mathcal{U}$) is higher than the second column of Tab. 3. First, $\mathcal{L}_{\text{BCE}}$ converges faster than $\mathcal{L}_{\text{MSE}}$ as in Fig. 11. Second, the experimental group uses the same number of layers and hidden parameters for quantized images while the bit bias exists in the image, which is inefficient.

Quantitative Results In Tab. 1, we report peak signal-to-noise ratio (PSNR(dB)), structural similarity index measure (SSIM), root mean squared error (RMSE) and bit-error-rate (BER) for evaluation on 16-bit and 8-bit image datasets. We pick the best values of metrics for each image during

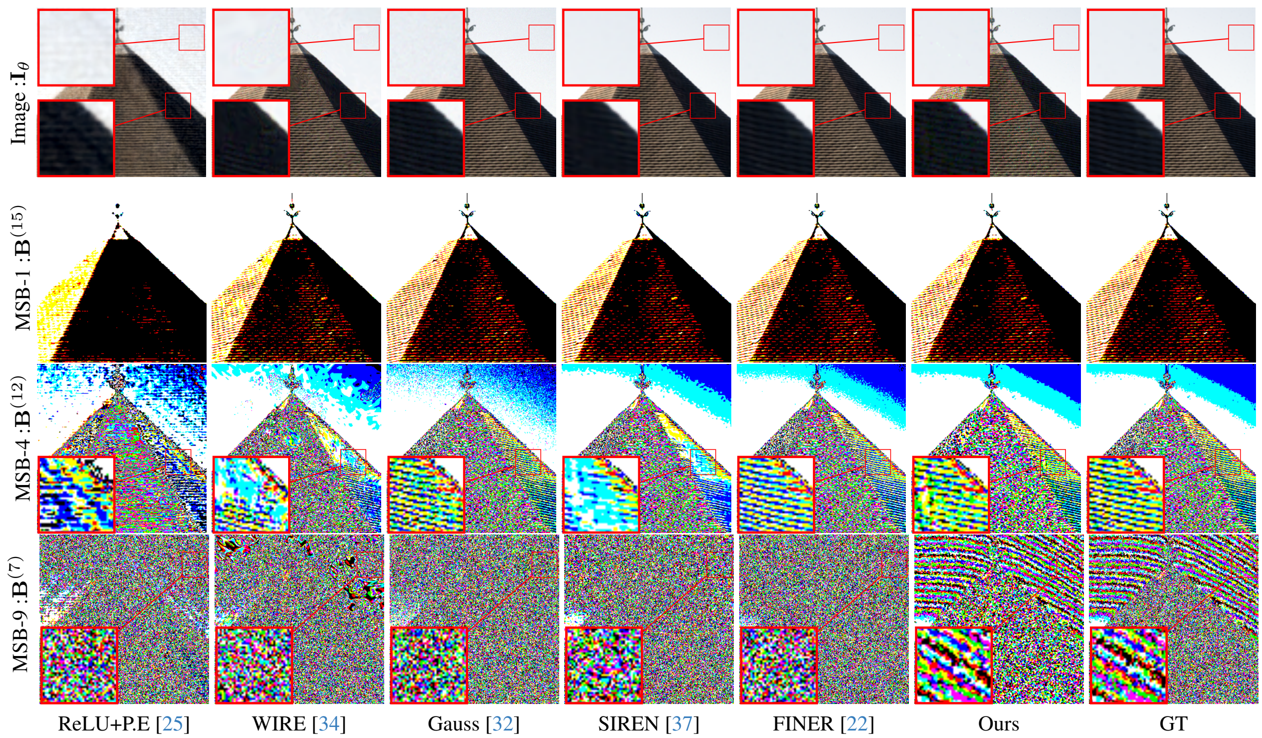

Figure 7. Qualitative comparison of an under-fitted image (# of Iterations : 400) and its bit-plane. The experiment was conducted on the 16-bit image of TESTIMAGE [2]. MSB-n indicates nth bit-plane from the MSB.

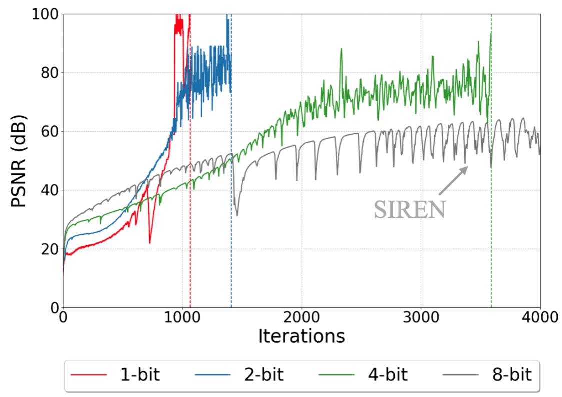

Figure 8. Comparison of convergence curve for an 8-bit image based on bit precision. Vertical dashed lines indicate the iteration when the model achieves lossless. We show that when $\mathcal{P}(f_\theta)$ is constant and close to $\mathcal{U}_d(n)$, convergence occurs effectively, enabling lossless representation.

Method	Bit Precision (n)	$\mathcal{U}_d(n)$ $(\cdot \times \mathfrak{C})$	Loss Function	PSNR (dB)	#Iter. (Mean ± std)
Experiment Group	1 2 4	16 $1.30K$ $0.81M$	$\mathcal{L}_{\text{MSE}}$	∞ (Lossless)	1233 ± 241 1288 ± 235 3852 ± 710
SIREN	8	$67.7G$	$\mathcal{L}_{\text{MSE}}$	102.8	5000
Proposed	1	64	$\mathcal{L}_{\text{BCE}}$	∞ (Lossless)	$\mathbf{778 \pm 83}$

Table 3. Quantitative result of our hypothesis test experiment on Kodak [10]. The number of iterations is proportional to $\mathcal{U}_d(n)$.

the training. RMSE and SSIM are calculated using integer value. Our method accomplishes lossless representation on all 16-bit and 8-bit images in experiments. Our method converges faster than all other baselines. Fig. 3 represents the results on a single 8-bit image. Note that minimizing BER is highly correlated to increasing PSNR but not equivalent. As a result, our proposed method has consistently low BER throughout the learning process, while the PSNR does not. Tab. 2 demonstrate that our proposed method can be applied to other existing INR approaches. We integrate our method with conventional approaches that use hash inputs [26, 46], as well as with efficient activation functions [22, 32]. For hash-based methods [26, 46], we follow the settings specified in their respective papers. We measure the average number of iterations for each result, ensuring that each baseline was trained sufficiently for a fair comparison. Our method is compatible with existing methods while achieving lossless representation.

Qualitative Results We report a visual comparison of under-fitted images in Fig. 5. Different artifacts occur during training, such as blurry artifacts for WIRE [34] and SIREN [37], or noise-like artifacts for Gauss [32]. Our method learns both high-frequency and low-frequency components faster than the others; however, salt-pepper impulse noise artifacts are present in the training stage. In Fig. 2, we present converged images with all baselines. Since the converged images are not easily discernible to human eyes, we highlight the occurrence of dominant errors ($\geq$ MSB-4). The error map indicates residual between the ground truth (GT). Our method

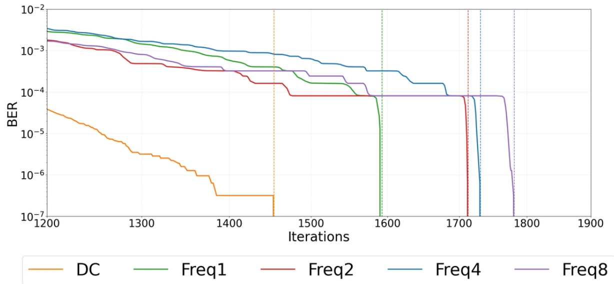

Figure 9. Quantitative comparison based on frequency to the bit axis. Vertical dashed lines indicate the iteration when the model achieves lossless.

	Weight	Model Size (Byte↓)	PSNR (dB↑)	#Iter.(↓)	#BitOps. (/Pixel)
Ours$^-$	$\{-1,0,1\}$	656.18K	∞	80K	50.66M
SIREN	FP32	1.27M	119.7	200K	405.2M

Table 4. Model size and weight & performance comparison in the aspect of model quantization.

makes INR represent no error in image representation.

4.3. Bit & Bit-Spectral Bias

Bit Bias In this section, we examine two different biases that we discovered: 1) '*bit bias*' and 2)'*bit-spectral bias*,' validating them through experiments. In Fig. 6, conducted on a single 16-bit image, the experiment quantitatively demonstrates the presence of bit bias across all tested baselines. Whether weights are assigned or not to the MSBs [2], there is a common challenge in representing the LSBs. Fig. 7 shows that fitting artifacts mentioned in Sec. 4.2 are related to bit bias. According to Fig. 7, MSBs are nearly indiscernible between INRs and GT. However, significant differences were observed in LSBs $\mathbf{B}^{(i)}(i \leq 13)$. In conclusion, the proposed method effectively reduces bit bias resulting in representing the signal's LSBs.

Bit-spectral Bias In the following, we study our method's bit axis and its bias. Spectral bias [31] exists on the bit axis. Therefore, specific pixel values are difficult to represent in our structure. The experiment is conducted on the 16-bit synthetic CYMK-RGB image, where we set frequency along the bit axis as a variable. Fig. 9 quantitatively shows the existence of bit-spectral bias. In conclusion, our method parameterizes specific values like 65,535 or 0 (DC) faster than high-frequency values. Implementation details and qualitative results are shown in the supplementary material.

4.4. Applications

We propose new applications by using our method. We introduce ternary INR with extreme weight quantization, bit-depth expansion using a bit axis, and lossless compression utilizing lossless representation. The implementation details are in the supplement material.

Ternary Implicit Neural Representation The intuitive question is whether 32-bit floating precision (FP32) parameters are necessary when parameterizing outputs with 1-bit precision. To address our concern, we design a ternary-weighted (1.58-bit) implicit neural representation for an image fitting that employs the novel method proposed by Wang et al. [45] and Ma et al. [23]. Each fully connected layer has ternary weights and calculates its output as follows:

$$y = \beta\gamma\tilde{\mathcal{W}}\tilde{x}, \quad (\tilde{\mathcal{W}} \in \{-1,0,1\}^{d_{\text{out}}\times d_{\text{in}}}), \qquad (12)$$

where $\tilde{x}$ is layer normalized and quantized values of input x and $\beta := \frac{1}{d_{\text{in}}d_{\text{out}}}||\mathcal{W}||_1$, $\gamma := ||x||_\infty$ as suggested in [45].

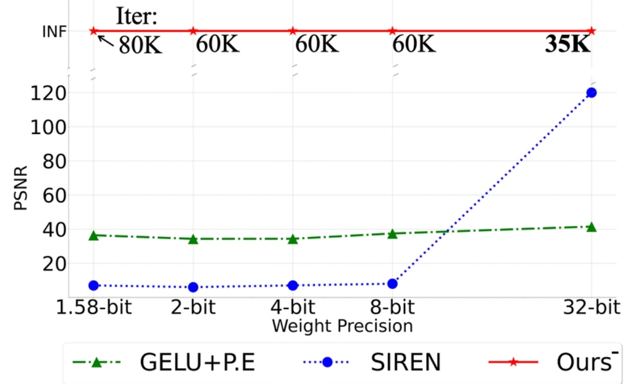

Figure 10. Quantitative comparison of performance (Y-axis) according to the parameter precision (X-axis)

PSNR(dB↑)|SSIM(↑))		8-bit → 16-bit	8-bit → 12-bit
ZP	Rule-based	52.92|0.9990	53.31|0.9990
MIG		55.91|0.9991	55.93|0.9991
BR [44]		52.98|0.9991	53.32|0.9991
BECNN [40]	Supervised	53.14|0.9986	N/A
BitNet [4]		53.60|0.9970	N/A
ABCD [12]		59.39|0.9997	59.37|0.9995
Ours	Self-supervised	55.92|0.9993	55.94|0.9997

Table 5. Quantitative comparison in the bit-depth expansion on TESTIMAGES[2]. Red and blue indicate the best and the second-best performance, respectively. 'N/A' indicates not applicable.

In Tab. 4 and Fig. 10, we show the performance comparison of weight-quantized INRs. As shown in Fig. 10, our proposed method accomplishes lossless image representation until ternary weights. The activation for our ternary INR should be the GELU [15] function, and we denote it as 'Ours$^-$'. Periodic activations must follow the strict weight initialization [22, 37]; breaking such initialization by the quantization makes the network collapse. The numbers under 'Ours$^-$' in Fig. 10 indicate the minimum iteration number for each model. In Tab. 4, We report model size, the number of parameters of f_θ (i.e. $\mathcal{P}(f_\theta)$) in bytes and the number of bit operations (BitOps). The proposed method requires less storage and BitOps than SIREN.

Bit Depth Expansion Our method conducts bit depth expansion by extrapolating the bit-axis. In Tab. 5, we conduct a quantitative comparison with existing methods. We train on the 8 MSBs of a 16-bit image and predict the lower 8 bits without using the 16-bit ground truth. To the best of our knowledge, our method is the first attempt at a self-supervised learning approach for bit depth expan-

[2]Baselines are equivalent to assigning weights to bits.

Bits Per Pixel (bpp)(↓)	MNIST	Fashion MNIST
PNG [33]	3.52(+36%)	5.78(-5%)
JPEG2000 [42]	6.75(+162%)	7.74(+27%)
WebP [35]	2.11(-18%)	6.60(+8%)
TIFF [29]	3.93(+52%)	6.76(+11%)
RECOMBINER[14]+**Ours**	2.58	6.11

Table 6. Quantitative comparison for lossless compression. Red and blue indicate the best and second-best performance, respectively.

	MNIST				Fashion MNIST			
	bpp	PSNR	SSIM	RMSE	bpp	PSNR	SSIM	RMSE
RECOMBINER [14]	4.20	48.60	0.994	0.945	9.06	56.64	0.996	0.375
RECOMBINER+**Ours**	**2.58**	∞	1.000	0.000	**6.11**	∞	1.000	0.000

Table 7. Quantitative comparison between RECOMBINER [14] and RECOMBINER with our method.

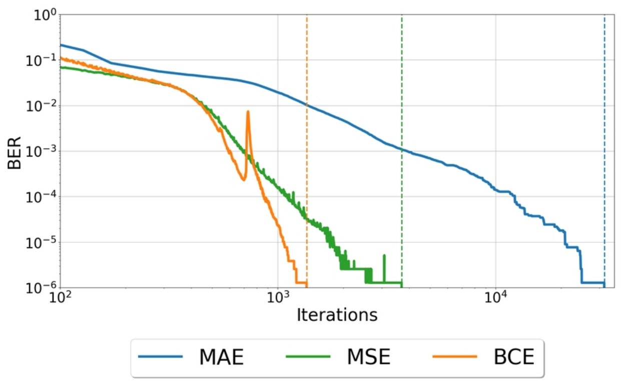

Figure 11. Quantitative ablation study on loss function of our method on TESTIMAGE[2]. Vertical dashed lines indicate the iteration when the model achieves lossless representations.

sion. Our method performs superior than existing rule-based algorithms or learning-based methods (BitNet [4] and BECNN[40]).

Lossless Compression We conduct lossless compression experiments by applying our method to the state-of-the-art INR compression method [14]. In Tab. 6, a simple combination of [14] and ours shows superior results compared to existing lossless image codecs such as PNG [33], JPEG2000 [42], WebP [35], and TIFF [35]. In Tab. 7, [14] cannot achieve lossless representation, even if the bpp is significantly increased.

5. Discussion

Ablation Study We conduct ablation studies for the loss function of the proposed method. We utilize a 16-bit sample image in the TESTIMAGE dataset [2] and reduce the model size to observe convergence speed. In Fig. 11, the performance of MSE is close to that of BCE, while MAE exhibits a slow convergence speed. Although achieving lossless representation through MSE or MAE is possible, BCE shows the fastest convergence speed.

We conduct an ablation study on the input dimension d. In Tab. 8, we extend a 3-dimensional coordinate to a 4-dimensional one by incorporating color as a coordinate. As Eq. (6), increasing a dimension increases the upper bound and makes INR converge slower than the proposed method.

Method	Proposed	Tested
Coord.	$\mathbf{x} = (h, w, i)$	$\mathbf{x} = (h, w, i, c)$
#Iter.(↓)	**790**	1438

Table 8. Quantitative ablation study of our method on Kodak [10] as the input dimension d increases ($3 \rightarrow 4$).

Text	Jack would become Eva's happy husband	
Method	PSNR(dB)(↑)	Prediction from [43]
GT Audio	-	Jack would become even happy ashon
SIREN [37]	68.54	Jark will become evil's haring ho
DINER [46]	85.19	Jar would become even hary ashon
Ours	∞	Jack would become even happy ashon

Table 9. Qualitative comparison on speech to text (STT) results of the represented audio using a pre-trained STT network [43].

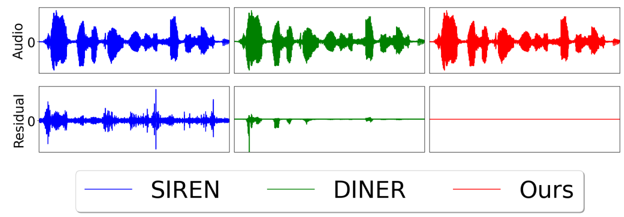

Figure 12. Qualitative comparison on representing the Librispeech [27] data with a floating point precision (FP32) and its residual.

Floating Point Representation Our approach has been discussed in the context of fixed precisions. We also verified whether our method can be applied to floating-point data, such as audio. The detailed formulation is in the supplement. In Fig. 12, we show the representation result of Librispeech data [27] and demonstrate that the FP32 data format is fitted by our method losslessly. In Tab. 9, we report the speech-to-text (STT) results predicted by pre-trained model [43] and compare our method with other methods [37, 46].

Limitation As mentioned by Jentzen et al. [16], when d exceeds 5, such as in radiance fields, the suggested upper bounds $\mathcal{U}_d$ increase extremely high ($\mathcal{U}_5(8) \simeq 1.23\mathfrak{C} \times 10^{27}$). Although our proposed approach performs better in representing low-dimensional data, the main drawback lies in predicting high-dimensional data. Further research is needed to explore parameter-efficient learning; thus, we demonstrate the use of recent techniques in the supplement material.

6. Conclusion

We quantify the upper bound of the size of INRs based on the given bit precision. Through bit-plane decomposition, we achieve lossless representation, which was previously unachievable. With experiments, we validate our hypothesis that "lowering the upper bound accelerates the achievement of lossless representation in INR." Furthermore, we reinterpret the concept of spectral bias from a digital computing perspective and explain new notions of 'bit bias.' Our method mitigates the bit bias and makes INR represent true LSBs, resulting in lossless representation. We demonstrate that our method enables true lossless representation in followed applications: ternary networks, lossless compression, and bit-depth expansion.

Acknowledgments

This work was partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00335741) and (RS-2024-00413303).

References

[1] Eirikur Agustsson and Radu Timofte. NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017.

[2] Nicola Asuni and Andrea Giachetti. Testimages: A large data archive for display and algorithm testing. *Journal of Graphics Tools*, 17(4):113–125, 2013.

[3] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input / output image pairs. In *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[4] Junyoung Byun, Kyujin Shim, and Changick Kim. BitNet: Learning-based bit-depth expansion. In *Asian Conference on Computer Vision*, pages 67–82. Springer, 2018.

[5] Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning Continuous Image Representation With Local Implicit Image Function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8628–8638, 2021.

[6] Zhang Chen, Zhong Li, Liangchen Song, Lele Chen, Jingyi Yu, Junsong Yuan, and Yi Xu. Neurbf: A neural fields representation with adaptive radial basis functions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4182–4194, 2023.

[7] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.

[8] Emilien Dupont, Adam Goliński, Milad Alizadeh, Yee Whye Teh, and Arnaud Doucet. Coin: Compression with implicit neural representations. *arXiv preprint arXiv:2103.03123*, 2021.

[9] Emilien Dupont, Hrushikesh Loya, Milad Alizadeh, Adam Goliński, Yee Whye Teh, and Arnaud Doucet. Coin++: Neural compression across modalities. *arXiv preprint arXiv:2201.12904*, 2022.

[10] Rich Franzen. Kodak lossless true color image suite. *source: http://r0k. us/graphics/kodak*, 4(2), 1999.

[11] Zongyu Guo, Gergely Flamich, Jiajun He, Zhibo Chen, and José Miguel Hernández-Lobato. Compression with bayesian implicit neural representations. *Advances in Neural Information Processing Systems*, 36:1938–1956, 2023.

[12] Woo Kyoung Han, Byeonghun Lee, Sang Hyun Park, and Kyong Hwan Jin. ABCD: Arbitrary bitwise coefficient for dequantization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5876–5885, 2023.

[13] Woo Kyoung Han, S. Im, J. Kim, and Kyong Hwan Jin. JDEC: Jpeg decoding via enhanced continuous cosine coefficients. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[14] Jiajun He, Gergely Flamich, Zongyu Guo, and José Miguel Hernández-Lobato. Recombiner: Robust and enhanced compression with bayesian implicit neural representations. *arXiv preprint arXiv:2309.17182*, 2023.

[15] Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016.

[16] Arnulf Jentzen, Benno Kuckuck, and Philippe von Wurstemberger. Mathematical introduction to deep learning: Methods, implementations, and theory. *arXiv preprint arXiv:2310.20360*, 2023.

[17] Chiyu "Max" Jiang, Avneesh Sud, Ameesh Makadia, Jingwei Huang, Matthias Niessner, and Thomas Funkhouser. Local Implicit Grid Representations for 3D Scenes. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[18] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[19] Jaewon Lee and Kyong Hwan Jin. Local texture estimator for implicit representation function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1929–1938, 2022.

[20] Jaewon Lee, Kwang Pyo Choi, and Kyong Hwan Jin. Learning local implicit fourier representation for image warping. In *European Conference on Computer Vision (ECCV)*, pages 182–200. Springer, 2022.

[21] David B Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. Bacon: Band-limited coordinate networks for multiscale scene representation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16252–16262, 2022.

[22] Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. Finer: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. *arXiv preprint arXiv:2312.02434*, 2023.

[23] Shuming Ma, Hongyu Wang, Lingxiao Ma, Lei Wang, Wenhui Wang, Shaohan Huang, Li Dong, Ruiping Wang, Jilong Xue, and Furu Wei. The era of 1-bit llms: All large language models are in 1.58 bits. *arXiv preprint arXiv:2402.17764*, 2024.

[24] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy Networks: Learning 3D Reconstruction in Function Space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[25] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020.

[26] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, 41(4):1–15, 2022.

[27] Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. Librispeech: an asr corpus based on public domain audio books. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 5206–5210. IEEE, 2015.

[28] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[29] Charles A Poynton. Overview of tiff 5.0. In *Image Processing and Interchange: Implementation and Systems*, pages 152–158. SPIE, 1992.

[30] Abhijith Punnappurath and Michael S Brown. A little bit more: Bitplane-wise bit-depth recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[31] Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the Spectral Bias of Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning*, pages 5301–5310. PMLR, 2019.

[32] Sameera Ramasinghe and Simon Lucey. Beyond periodicity: Towards a unifying framework for activations in coordinate-mlps. In *European Conference on Computer Vision*, pages 142–158. Springer, 2022.

[33] Greg Roelofs. *PNG: the definitive guide*. O'Reilly & Associates, Inc., 1999.

[34] Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. Wire: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18507–18516, 2023.

[35] Zhanjun Si and Ke Shen. Research on the webp image format. In *Advanced graphic communications, packaging technology and materials*, pages 271–277. Springer, 2016.

[36] Vincent Sitzmann, Michael Zollhoefer, and Gordon Wetzstein. Scene Representation Networks: Continuous 3D-Structure-Aware Neural Scene Representations. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019.

[37] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit Neural Representations with Periodic Activation Functions. In *Advances in Neural Information Processing Systems*, pages 7462–7473. Curran Associates, Inc., 2020.

[38] Yannick Strümpler, Janis Postels, Ren Yang, Luc Van Gool, and Federico Tombari. Implicit neural representations for image compression. In *European Conference on Computer Vision*, pages 74–91. Springer, 2022.

[39] Kun Su, Mingfei Chen, and Eli Shlizerman. Inras: Implicit neural representation for audio scenes. *Advances in Neural Information Processing Systems*, 35:8144–8158, 2022.

[40] Yuting Su, Wanning Sun, Jing Liu, Guangtao Zhai, and Peiguang Jing. Photo-realistic image bit-depth enhancement via residual transposed convolutional neural network. *Neurocomputing*, 347:200–211, 2019.

[41] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. In *Advances in Neural Information Processing Systems*, pages 7537–7547. Curran Associates, Inc., 2020.

[42] D Taubman. Jpeg 2000: Image compression fundamentals, standards and practice, 2002.

[43] Silero Team. Silero models: pre-trained enterprise-grade stt / tts models and benchmarks, 2021.

[44] Robert A Ulichney and Shiufun Cheung. Pixel bit-depth increase by bit replication. In *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts III*, pages 232–241. SPIE, 1998.

[45] Hongyu Wang, Shuming Ma, Li Dong, Shaohan Huang, Huaijie Wang, Lingxiao Ma, Fan Yang, Ruiping Wang, Yi Wu, and Furu Wei. Bitnet: Scaling 1-bit transformers for large language models. *arXiv preprint arXiv:2310.11453*, 2023.

[46] Shaowen Xie, Hao Zhu, Zhen Liu, Qi Zhang, You Zhou, Xun Cao, and Zhan Ma. DINER: Disorder-invariant implicit neural representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6143–6152, 2023.

[47] Guandao Yang, Sagie Benaim, Varun Jampani, Kyle Genova, Jonathan Barron, Thomas Funkhouser, Bharath Hariharan, and Serge Belongie. Polynomial neural fields for subband decomposition and manipulation. *Advances in Neural Information Processing Systems*, 35:4401–4415, 2022.

[48] Shien Zhu, Luan HK Duong, and Weichen Liu. Xor-net: an efficient computation pipeline for binary neural network inference on edge devices. In *2020 IEEE 26th international conference on parallel and distributed systems (ICPADS)*, pages 124–131. IEEE, 2020.