LoRaSeek: Boosting Denoising Ability in Neural-enhanced LoRa Decoder via Hierarchical Feature Extraction

Khang Nguyen[1], Yidong Ren[1], Jialuo Du[1], Jingkai Lin[1]
Maolin Gan[1], Shigang Chen[2], Mi Zhang[3], Chunyi Peng[4], Zhichao Cao[1]
[1]Michigan State University [2]University of Florida [3]The Ohio State University [4]Purdue University

Abstract

In this paper, we propose LoRaSeek, a lightweight and reliable LoRa denoising framework that enhances signal quality and robustness for neural-enhanced LoRa decoding. LoRaSeek integrates a hybrid architecture combining Convolutional Neural Networks (CNNs), Transformers, and a hierarchical U-Net to effectively capture multi-scale, multidimensional features of LoRa chirp signals. To maintain efficiency, we integrate a lightweight Transformer block that supports various LoRa configurations while keeping computational overhead low. Additionally, we incorporate dual attention-based skip connections to preserve chirp signal properties across different scales. Experiments across diverse LoRa configurations show that LoRaSeek achieves 2.04–3.86 dB signal-to-noise ratio (SNR) gains over standard decoding methods and up to 3.03 dB improvement over state-of-the-art neural-enhanced LoRa decoding methods while reducing model storage by up to 7.4× and inference time by up to 1.6×.

CCS Concepts

• **Networks** → **Network protocols**; **Network algorithms**; • **Computing methodologies** → **Hierarchical representations**.

Keywords

LoRa, AIoT, Deep Learning for Wireless System

ACM Reference Format:
Khang Nguyen[1], Yidong Ren[1], Jialuo Du[1], Jingkai Lin[1], Maolin Gan[1], Shigang Chen[2], Mi Zhang[3], Chunyi Peng[4], Zhichao Cao[1]. 2025. LoRaSeek: Boosting Denoising Ability in Neural-enhanced LoRa Decoder via Hierarchical Feature Extraction. In *The 31st Annual International Conference on Mobile Computing and Networking (ACM MOBICOM '25), November 4–8, 2025, Hong Kong, China.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3680207.3765241

ACM MOBICOM '25, Hong Kong, China

ACM ISBN 979-8-4007-1129-9/2025/11
https://doi.org/10.1145/3680207.3765241

1 Introduction

Low-Power Wide-Area Networks (LPWANs) introduce a new paradigm for the Internet of Things (IoT), enabling long-distance, low-power wireless communication. Among LPWAN technologies, LoRa is widely adopted, operating in unlicensed frequency bands to support applications in smart cities [40], precision agriculture [7, 37, 39, 41, 50]. As of May 2024, Semtech [46], a leading LoRa chip manufacturer, reports that LoRa accounts for over 30% of global LPWAN connections, with 6.9 million gateways and more than 350 million end nodes.

In a LoRa network, end nodes transmit data directly to any gateway within range, which then forwards the data to the cloud via backhaul links. Under line-of-sight conditions, LoRa end nodes can communicate with gateways over tens of miles. However, real-world obstacles such as buildings, trees, and hills introduce significant signal attenuation, compromising transmission reliability [7, 40, 41]. Addressing weak LoRa signals is therefore critical to sustaining its long-range, low-power communication capabilities [24].

The core principle of LoRa decoding is to coherently accumulate signal energy into a single peak in the frequency domain, surpassing random noise. To improve weak signal decoding, various studies exploit redundancy in symbols [58], packets [55], antennas [18], or gateways [10] to accumulate more signal energy. For instance, XCopy [55] coherently combines multiple retransmitted packet copies to strengthen the energy peak of each symbol. However, these approaches come at the cost of either reduced data rates or increased infrastructure complexity.

In contrast, NELoRa [12, 24] introduces neural-enhanced LoRa decoding, reducing the signal-to-noise ratio (SNR) required for symbol decoding without relying on redundancy. Its key insight is that while the overall accumulated signal energy peak may be buried in noise, short time windows within the signal can still exhibit peaks that surpass temporally random noise. However, there is no free lunch: NELoRa's model size and inference time present major deployment challenges

on edge gateways for real-time decoding. For instance, decoding a single LoRa symbol takes up to 8.16 seconds on a Raspberry Pi. GLoRiPHY [45] attempts to mitigate these issues by employing an encoder-decoder model to reduce model size and inference time. However, its maximum SNR gain remains comparable to NELoRa's (e.g., 0.52 dB difference), suggesting that the encoder-decoder structure does not retain more information than NELoRa's Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM). This raises a crucial question: What is the optimal neural network architecture to effectively capture the temporal-spatial signal peak patterns while minimizing storage and inference overhead?

In this paper, we propose LoRaSeek, a hierarchical neural network architecture that advances LoRa weak signal decoding beyond NELoRa while optimizing model size and inference time. LoRaSeek leverages a hierarchical structure to efficiently denoise weak LoRa signals, capturing both fine-grained local details and global contextual features from noisy inputs. However, ensuring optimal performance across diverse LoRa configurations presents three key challenges.

Challenge #1: The frequency of a LoRa chirp symbol increases linearly within a channel. Data modulation is achieved by shifting the chirp's initial frequency. LoRa employs spreading factors (SF) to balance data rate and communication reliability, adjusting the on-air time of a chirp symbol, which ranges from 1.025 ms to 32.8 ms in a typical 125 kHz channel. Due to the relatively long on-air time and noise variations across temporal scales, developing an effective denoising model is challenging. A hierarchical architecture capable of capturing multi-scale spectral and temporal features is essential for robust noise reduction.

Challenge #2: Each LoRa chirp symbol is unique, with critical data encoded through subtle variations such as small frequency shifts or cyclic offsets. For instance, when the spreading factor (SF) is set to 12, there are 4096 possible symbols to classify. While aggressive noise removal improves clarity, it risks erasing essential features or distorting the spectral structure. Conversely, conservative approaches preserve signal details but leave residual noise. Achieving effective noise suppression while maintaining these key chirp characteristics for accurate decoding is a significant challenge. Furthermore, at ultra-low SNR levels, distinguishing between noise and signal features in the LoRa domain becomes increasingly difficult.

Challenge #3: In the current LoRa network architecture, LoRa packets are demodulated at LoRa gateways before the decoded data is forwarded to cloud servers. While shifting the decoding process to the cloud [48] is possible, the high cost of establishing high-speed backhaul connections in rural areas makes it more practical to implement a neural-enhanced LoRa decoder at the edge LoRa gateways. However, unlike GPU-powered cloud servers, LoRa gateways are resource-constrained with limited computational power. This makes it highly challenging to develop a model that is both computationally efficient and reliable. The challenge is further compounded by the increasing signal length for higher SFs, requiring a solution that balances accuracy, efficiency, and low latency without relying on overly complex architectures unless optimized for deployment.

Solution #1: To address the first challenge, we propose a UNet-based denoising architecture inspired by the U-shaped design, integrating local and global feature extraction for multi-scale learning. Long LoRa chirp signals exhibit both localized noise perturbations and broader spectral distortions, making it crucial to leverage both local and global contexts for effective recognition. Our UNet incorporates CNNs and Transformers to enhance denoising performance: the CNN efficiently captures short-term fluctuations, while the Transformer models global patterns such as chirp linearity and cyclic offsets. These components are strategically placed at multiple stages of the encoder and decoder to ensure robust feature extraction and denoising across scales.

Solution #2: For the second challenge, we propose a selective denoising approach that focuses on informative features within specific short time windows rather than processing the entire chirp signal. This prevents over-denoising and ensures reliable decoding. To achieve this, we introduce a dual-attention skip connection mechanism that selectively filters and fuses relevant features for reconstruction. Our dual-attention mechanism consists of two key components: (1) Channel Attention, which emphasizes the most relevant signal chirps while suppressing noise-dominated channels. (2) Spatial Attention, which captures the temporal energy distribution to enhance feature representation. Integrating attention, our model effectively balances noise reduction and signal preservation, improving decoding robustness.

Solution #3: To address the third challenge, we take a two-step approach to optimize the computational efficiency of our UNet-based denoising architecture. 1) Progressive Downsampling and Upsampling: Leveraging the U-shaped hierarchical structure, we progressively reduce high-dimensional data to lower dimensions while simultaneously increasing the number of feature channels the network learns. This ensures efficient representation learning while preserving critical information. 2) Efficient Transformer Design: To mitigate the quadratic complexity of traditional Transformer architectures, we integrate an efficient Transformer block with linear complexity. Additionally, we enhance the Transformer with a local context module, which captures fine-grained dependencies while maintaining efficiency. This approach significantly reduces computational overhead without compromising performance, making the model well-suited for resource-constrained LoRa gateways.

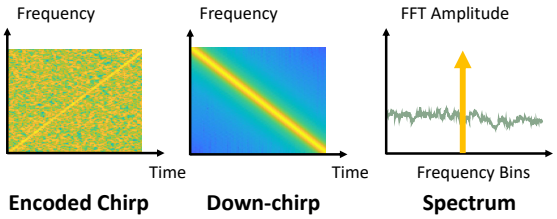


Figure 1: LoRa CSS modulation and demodulation.

We implement LoRaSeek with commercial-off-the-shelf (COTS) LoRa radio and USRP N210 [44] as the LoRa gateway and evaluate it over emulated datasets and real-world deployments. Experiments show that LoRaSeek improves SNR gain and inference time simultaneously compared to NELoRa [24] and GLoRiPHY [45].

We summarize our main contributions as follows:

- We propose a novel hierarchical feature extraction approach to improve weak signal denoising in neural-enhanced LoRa decoders, enhancing both decoding reliability and computational efficiency, making it lightweight enough to run on edge nodes.
- We propose LoRaSeek, a UNet-based hierarchical denoising model designed to capture both local and global features in the temporal-spatial domain. To further enhance decoding robustness and computational efficiency, we incorporate an adaptive attention mechanism and optimize the model modules.
- We implement LoRaSeek and evaluate its performance in various environments. The results show up to 3.03 dB improvement over state-of-the-art neural-enhanced LoRa decoding methods. Meanwhile, the storage is reduced up to 7.4× and inference time is reduced up to 1.6×.

2 Background and Motivation

2.1 Standard LoRa Physical Layer

In practical wireless communication, the received signal is inevitably contaminated by additive noise. Let $n(t)$ represent this noise, often modeled as a zero-mean Gaussian random process with variance σ^2. The actual received signal is $y(t) = y_e(t) + n(t)$. Even under ultra-low SNR conditions, LoRa can reliably recover data through its Chirp Spread Spectrum (CSS) modulation. In this scheme, a base up-chirp is defined over a symbol duration T with a frequency that linearly sweeps from $-\frac{BW}{2}$ to $\frac{BW}{2}$, where BW denotes the channel bandwidth. As shown in Figure 1, the modulated chirp signal is given by $y_e(t) = C(t) \cdot \exp\{j2\pi f_s t\}$. At the receiver, demodulation is performed via a process called "dechirp". The receiver multiplies the received signal by a locally generated down-chirp, the complex conjugate of the base chirp. The dechirped signal becomes

$$z(t) = y(t) \cdot C^*(t) = \exp\{j2\pi f_s t\} + n(t) \cdot C^*(t),$$

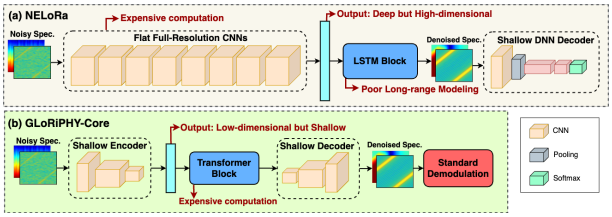


Figure 2: Architectures of existing neural-enhanced LoRa decoders. (a) NELoRa. (b) GLoRiPHY-Core.

where the first term is the desired signal and the second term represents the noise after dechirp. To extract the transmitted symbol, the receiver applies a Fast Fourier Transform (FFT) over the symbol duration T:

$$Z[n] = \int_0^T z(t) \exp\left\{-j\frac{2\pi n}{T}t\right\} dt.$$

In the frequency domain, the energy of the dechirped yellow signal is concentrated in the frequency bin corresponding to the offset f_s. The transmitted symbol is then identified by finding the bin with the maximum energy: $\hat{f}_s = \arg\max_n\{|Z[n]|\}$. However, if the noise level is excessively high, the signal's energy peak can be overwhelmed by noise. Suppose the amplitude of the noise term $n(t) \cdot C^*(t)$ in a given FFT bin is on the order of $\sigma\sqrt{T}$ (with σ^2 being the noise variance). When the condition $T \lesssim \sigma\sqrt{T}$ is met, or equivalently when the SNR is too low, the dechirp and subsequent FFT fail to yield a discernible peak, leading to erroneous decoding.

2.2 Neural-enhanced LoRa Decoding

Recent works [12, 13, 24, 45] enhance LoRa decoding by primarily using deep neural networks (DNN) to mitigate noise and interference. The chirp decoding problem is converted to an initial frequency classification task given the SF configuration. The rationale behind neural-enhanced decoding is to replace the single-dimensional feature space of the frequency-domain energy-based dechirp decoding by a multi-dimension, multi-resolution feature space. Firstly, a neural-enhanced decoder takes the multi-dimensional spectrograms of both amplitude and phase, reflecting the information from energy and channel as inputs. The complementary information provides more distinguishable patterns among different chirp symbols [24]. Moreover, by generating the spectrograms with the Short-Time Fourier Transform (STFT), it adds a temporal dimension by splitting a chirp symbol into a series of shorter chirp slices so that even if the aggregate noise over the whole symbol is overwhelming, some chirp slices may survive due to temporarily weak noises. They may exhibit sufficient features for decoding. The neural-enhanced decoder can decode ultra-low SNR chirp symbols by extracting the spatial and temporal features. Given LoRa chirps' large scale and long range, neural networks are more effective than white-box signal processing for chirp slicing.

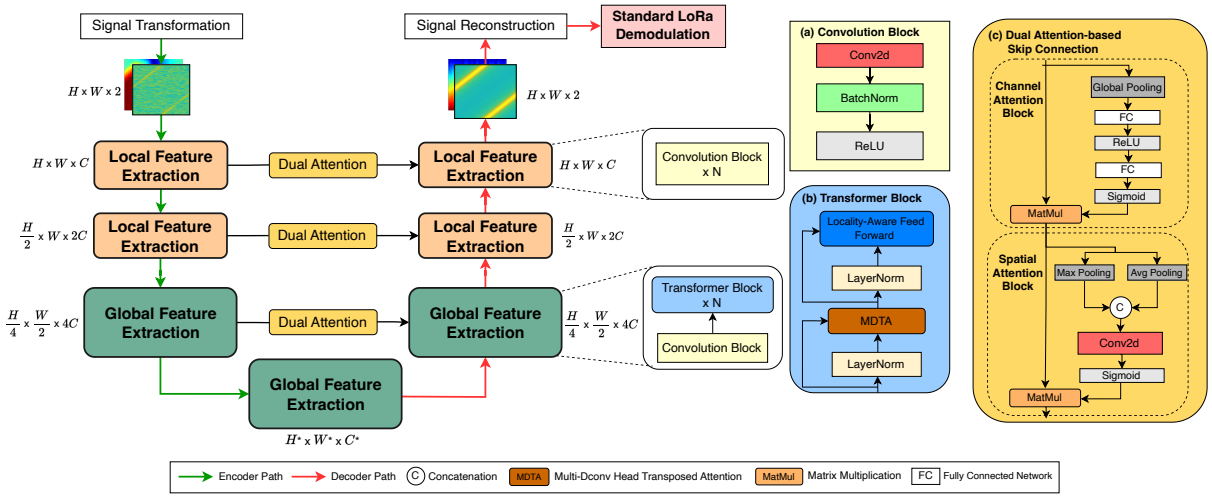


Figure 3: Overview of LoRaSeek Architecture. LoRaSeek adopts a U-shaped hierarchical denoising structure that extracts essential chirp signal patterns and reduces noise at multiple scales. (a) Convolution Block. (b) Transformer Block. (c) Dual Attention-based Skip Connection module.

NELoRa [12, 24], in Figure 2(a), stacks different CNN and LSTM blocks to capture intricate signal characteristics, significantly improving accuracy over traditional methods. However, the flat architecture causes exponential growth in computation at higher SF, restricting deployment in resource-constrained IoT devices. Additionally, CNNs and LSTMs struggle with modeling long-range temporal dependencies, capping achievable decoding performance. GLoRiPHY [45] addresses some limitations of NELoRa by utilizing a CNN encoder-decoder with a Transformer bottleneck. By effectively leveraging the LoRa frame preamble, GLoRiPHY reduces computational overhead and improves robustness at lower spreading factors. However, in Figure 2(b), GLoRiPHY-Core only encodes shallow representations of the LoRa chirp signal, causing the Transformer bottleneck to fail to model extremely long-range dependencies at higher SFs ($SF \geq 10$) while incurring high computational costs and limiting performance. We empirically evaluated NELoRa and GLoRiPHY-Core on SNR gains over dechirp and inference time per LoRa packet. Using high-SNR data at SF 7 and 10 with generated Gaussian noise, our results on a GPU server (Section 5.1) show NELoRa's SNR gain drops from 2.03 to 1.31 dB with inference time increases from 7.1 to 34.6 ms. GLoRiPHY shows a dramatic decline from 2.71 to 0.5 dB, increasing inference time from 6.7 to 9.1 ms.

2.3 Motivation

The limitations of existing neural-enhanced decoders motivate us to design novel architectures to effectively capture spatial chirp-slice features and model long-range temporal dependencies without exponential computational growth under different LoRa configurations. Instead of the shallow structure in NELoRa and GLoRiPHY, our core ideas come from a hierarchical DNN architecture with hybrid attentive CNNs and Transformers, boosting the denoising ability significantly while achieving high computation efficiency.

3 LoRaSeek Design

3.1 Overall Architecture

As illustrated in Figure 3, LoRaSeek follows a hierarchical U-shaped structure. We progressively denoise the chirp signal across multiple scales, transitioning from high to low resolutions during encoding, and reconstruct it via the corresponding upsampling in the decoding phase. From the noisy time-domain chirp signal, we convert it to a dual-spectrogram representation by the *Signal Transformation* (Section 3.2.1) module. The encoder then gradually downsamples the spectrogram through different stages by a sequence of *Local and Global Feature Extraction* (Section 3.2.2) modules to generate multi-scale feature maps. At each stage's transition, the feature map undergoes an aggressive downsampling, where either frequency or time dimension is reduced by half. In parallel, the decoder path mirrors the encoder's structure in reverse, upsampling the feature map to the original resolution. Local blocks capture short-term distortions, while global blocks capture long-range dependencies and noise through LoRa's inherently long chirp. To ensure a high-quality denoising process, LoRaSeek integrates *Dual Attention-based Skip Connections* (Section 3.4), which selectively filter and maintain core low-level features from the encoder at various resolutions, effectively fusing them with the decoder stage to enhance signal reconstruction. These blocks preserve chirp information at multiple scales and both local and global regions for effective denoising. Finally, the module *Signal Reconstruction* (Section 3.5) converts the denoised spectrogram into the representation of the signal in the time domain by the inverse short-time Fourier transform (iSTFT).

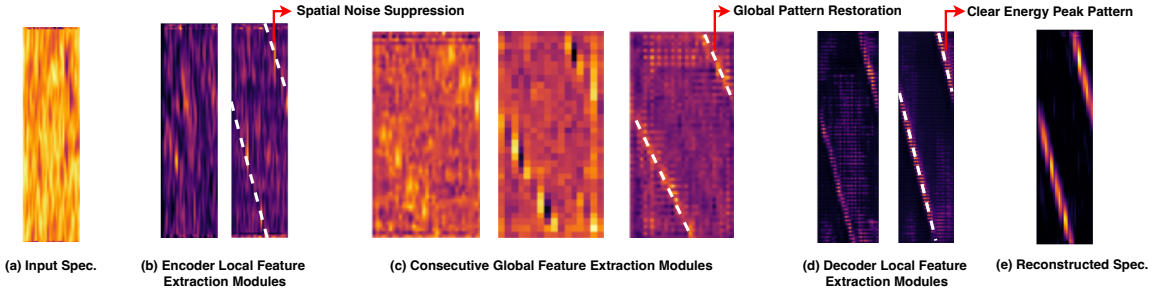


Figure 4: Visualizations of intermediate results across different feature extraction blocks at -20 dB SNR. (a) Input spectrogram; (b) Outputs of Encoder Local Feature Extraction modules; (c) Outputs of three Global Feature Extraction modules; (d) Outputs of Decoder Local Feature Extraction modules; (e) Reconstructed spectrogram.

As a result, we perform the standard LoRa demodulation to decode the transmitted information.

3.2 Local and Global Feature Extraction

3.2.1 Signal Transformation. To create a suitable input representation for LoRaSeek, we first create a dual-channel spectrogram input from the time-domain LoRa chirp $x(n)$. The process begins with computing STFT on $x(n)$ and concatenating the real and imaginary parts of the STFT result as:

$$S(m,k) = \sum_n x(n)W[n-mH]e^{-j2\pi kn/N}$$
$$\mathbf{X} = [\mathrm{Re}\{S(m,k)\}, \mathrm{Im}\{S(m,k)\}]$$

W is the *Hann* window with size m, $x(n) \in \mathbb{C}^n$ is transformed and reshaped into 3D feature input tensor $\mathbf{X} \in \mathbb{R}^{H\times W\times 2}$, where H and W are sampling points and frequency bins.

3.2.2 Local and Global Feature Extraction. Under ultra-low SNR conditions, LoRa chirp symbol is contaminated with noise spanning both time and frequency domains. For effective signal denoising, we argue that both local and global features are essential. Local features encode fine-grained information such as instantaneous change across time and frequency domains, maintaining chirp integrity and removing short-term, high-frequency distortions. Global features learn long-range sweep structures of the extended chirp duration and identify large-scale background noise over the signal. Therefore, an effective LoRa chirp denoising design must balance local and global feature learning for reliable decoding accuracy. Recent studies have shown that CNNs [23, 30, 32, 36] excel at local feature extraction by applying a sliding window to the input. On the other hand, Transformers [27, 31, 52, 61] are effective in modeling long-distance relationships using the self-attention mechanism. By combining the strengths of both architectures, we propose a denoising hierarchical structure where local and global feature extraction modules are harmoniously integrated throughout the network.

Overall Pipeline: LoRaSeek feeds the dual-channel input spectrogram $\mathbf{X} \in \mathbb{R}^{H\times W\times 2}$ through a series of local and global feature extraction modules. At each local-local or local-global stage transition, the feature input is continuously downsampled by halving either the time or frequency dimensions, facilitating a multi-scale representation. The shallow feature $\mathbf{F}_0 \in \mathbb{R}^{H\times W\times C}$, where C is often 16 or 32, from early local extraction is transformed into deeper levels with dimensions $\frac{H}{4} \times \frac{W}{2} \times 4C$. The deepest feature map $\mathbf{F}_d \in \mathbb{R}^{H^*\times W^*\times C^*}$ is obtained in the bottleneck, with further halving of all spatial dimensions and an increase in C (e.g., from $4C$ to $8C$). The specific downsampling strategy relies on different SF configurations, maintaining adaptive extraction modules for different LoRa settings. The decoder, following the reverse structure, takes the low-resolution latent feature $\mathbf{F}_d$ to aggressively recover the denoised output $\hat{\mathbf{X}}$ with original resolution.

Local Feature Extraction: We stack multiple Convolution blocks to expand the receptive field while preserving meaningful temporal and spectral patterns. The locality nature of the convolution block also prevents the network from losing small key features. As shown in Figure 3, each Convolution block consists of a standard convolution layer, followed by batch normalization and ReLU activation.

Global Feature Extraction: In Figure 3, our global feature extraction module includes an early Convolution block followed by a sequence of Transformer blocks. The Convolution block downsamples the feature map in the spatial dimension while increasing the channel dimension, reducing redundancy, and enhancing feature representation. The Transformer block takes an enriched feature map and models the long-range global dependencies across time and frequency domains. This process allows LoRaSeek to distinguish the core chirp patterns from the critical noise that spans the entire symbol.

We evaluate the contribution of each feature extraction block to the denoising process in Figure 4. In Figure 4(b), the encoder's Local Feature Extraction modules effectively utilize the spatial information for noise suppression. Due to CNN's localized nature, the distinct energy peak and chirp pattern are distorted. In Figure 4(c), the overall chirp structure and global energy pattern start aggregating across the time dimension after three Global Feature Extraction layers.

With further noise reduction in the decoder and the information fusion from the skip connection, Figure 4(d) illustrates the clear peak distribution of the core energy of a distinct chirp symbol, accurately preserving the key signal. Finally, as shown in Figure 4(e), we derive the denoised spectrogram with distinct characteristics of the chirp signal.

3.3 Transformer Block

Although the Transformer architecture can model global dependencies of the chirp symbol, it suffers from a major computational limitation: in Transformer's standard Multi-Head Self Attention (MHSA) mechanism [52], the computational complexity grows quadratically with the input resolution, i.e., $\mathcal{O}(W^2H^2)$ for an input of size $H \times W$. This quadratic scaling incurs excessive time and memory consumption when processing LoRa chirp symbols in large SF (e.g., 12). To address this challenge, we employ two core components in the LoRaSeek Transformer block: Multi-Dconv Head Transposed Attention and Locality-Aware Feed Forward Network. Before each component, Layer Normalization is applied to preprocess the input, and the output is concatenated with a skip connection after each component.

Multi-Dconv Head Transposed Attention: We adopt the Multi-Dconv Head Transposed Attention (MDTA) mechanism [61] to replace MHSA. As shown in Figure 5(a), MDTA applies self-attention (SA) across channels instead of spatial dimensions, further reducing the computational complexity to linear scaling, i.e., $\mathcal{O}(CC)$ for input size $H \times W \times C$. Moreover, this design incorporates more locality into the transformer, improving the denoising context information.

Locality-Aware Feed Forward Network: As Section 3.2.2 shows, both local and global information are necessary for chirp signal denoising. Therefore, we further inject more locality to LoRaSeek Transformer block in the feed-forward network component. Several studies [27, 53, 60, 61] have demonstrated superior performance gain of the Transformer architecture when locality is introduced into the feed-forward network. In Figure 5(b), we use 1×1 convolutions followed by a 3×3 depth-wise convolution to capture local spatial information for each channel. We use another 1×1 convolution to project the feature back to the original dimension.

3.4 Dual Attention-based Skip Connection

Skip connections have been widely used to reduce information loss and performance degradation as the neural network depth increases [6, 15, 22]. Under ultra-low SNR, LoRa chirp noise spreads across time and frequency. Directly concatenating skip connections with the decoder path can amplify the noise energy and limit decoding accuracy. Compared with the encoder path, deeper layers in the decoder extract complex signal representations through hierarchical CNN and Transformer blocks. This semantic difference can impede the neural network's ability to maximize denoising performance.

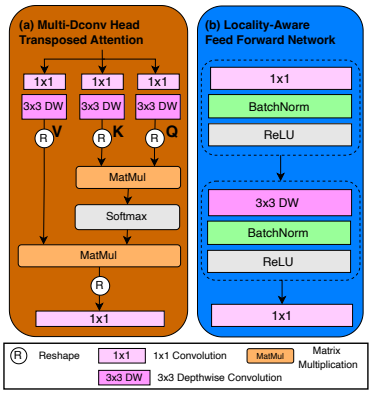


Figure 5: Core Components of the Transformer Block. (a) Multi-Dconv Head Transposed Attention. (b) Locality-Aware Feed Forward Network.

To address these technical challenges, we incorporate a Dual Attention-based Skip Connection. As illustrated in Figure 3(c), our skip connection contains a channel attention block followed by a spatial attention block, utilizing multi-scale concatenations in LoRaSeek's U-shaped architecture. Given the encoder feature map $\mathbf{F}_l \in \mathbb{R}^{H_l \times W_l \times C_l}$ at stage l, our intuition is to determine which features to retain and forget across channel and spatial domains for the denoised output.

Channel Attention Block: Our channel attention block is a Squeeze-and-Excitation (SE) block [19]. The goal is to highlight the important feature channels and suppress the less important channels in the denoising procedures. As shown in Figure 3(c), we apply Global Average Pooling (GAP) to the feature map $\mathbf{F}_l$ to generate a channel descriptor $\mathbf{z}_c \in \mathbb{R}^{C_l}$, followed by two fully connected (FC) layers with weights $\mathbf{W}_1 \in \mathbb{R}^{(C_l/r) \times C_l}$ and $\mathbf{W}_2 \in \mathbb{R}^{C_l \times (C_l/r)}$. ReLU and sigmoid activations are then used to compute attention map $\mathbf{A}_c \in \mathbb{R}^{1 \times 1 \times C_l}$. The output feature map $\mathbf{F}'_l \in \mathbb{R}^{H_l \times W_l \times C_l}$ is computed as:

$$\mathbf{F}'_l(i,j,c) = \mathbf{A}_c(c) \cdot \mathbf{F}_l(i,j,c), \quad \forall c \in \{1, 2, \dots, C_l\}$$

where i, j are the spatial positions, c is the channel index, and $A_c = \sigma(\mathbf{W}_2 \cdot \mathrm{ReLU}(\mathbf{W}_1 \cdot \mathbf{z}_c))$.

Spatial Attention Block: After the channel attention block, spatial attention is used to prioritize the relevant spatial information, such as concentrated chirp signal energy, and reduce intensely noisy regions. Utilizing Spatial Attention mechanism in Convolutional Block Attention Module (CBAM) [54], in Figure 3(c), we apply average pooling and max pooling for $\mathbf{F}'_l$ to compute spatial descriptors $\mathbf{F}'_{l,\max}, \mathbf{F}'_{l,\text{avg}} \in \mathbb{R}^{H_l \times W_l \times 1}$, concatenated and convolved them with a 7×7 convolution layer followed by a sigmoid activation to generate a spatial attention map $\mathbf{A}_s \in \mathbb{R}^{H_l \times W_l \times 1}$. The output feature map $\mathbf{F}''_l(i,j,c)$ is computed as:

$$\mathbf{F}''_l(i,j,c) = \mathbf{A}_s(i,j) \cdot \mathbf{F}'_l(i,j,c), \quad \forall c \in \{1, 2, \dots, C_l\}$$

where i, j are the spatial positions, c is channel index, and $\mathbf{A}_s = \sigma\left(Conv_{7\times7}\left(\text{Concat}\left(\mathbf{F}'_{l,\max}, \mathbf{F}'_{l,\text{avg}}\right)\right)\right)$.

3.5 Signal Reconstruction

To ensure compatibility with existing LoRa systems, we utilize standard LoRa decoding techniques. Therefore, we apply iSTFT to reconstruct the time-domain signal, preserving amplitude and phase for precise decoding. We convert the denoised spectrogram $\hat{\mathbf{X}}$ into complex-valued spectrogram as:

$$\hat{S}(m,k) = \hat{X}_{\text{Real}}(m,k) + j\hat{X}_{\text{Imag}}(m,k),$$

where m is the frame index, k is the frequency bin, and $\hat{X}_{\text{Real}}$ and $\hat{X}_{\text{Imag}}$ are the real and imaginary parts.

We then reconstruct the time-domain signal by applying iSTFT, which is defined as:

$$\hat{x}(n) = \sum_m \sum_k \hat{S}(m,k) e^{j2\pi kn/N} W[n - mH]$$

where $W[n]$, N, and H are the window function, FFT size, and hop size used in STFT to maintain consistency.

3.6 Weak Signal Packet Detection

We also need to improve the packet detection procedures for extremely low SNR. Traditional detection [24] performs FFT over the preambles and adds their power together to generate a power peak, which often fails in low-SNR scenarios.

Firstly, we observe the Sampling Frequency Offsets (SFO), making the length of a symbol different from its theoretical value. The small frequency shifts caused by SFO will slightly misalign the power peaks of each preamble chirp, and by taking into account the SFO, we can align them and generate a higher power peak. To do this, we multiply a singletone with the designated frequency to shift the frequency of the symbol. The initial phases of the preamble symbols follow a quadratic pattern [24]. After accounting for the frequency shifts, the phases will follow a linear pattern, and we can search for the gradient that maximizes the added power. Thus, we can mitigate the phase differences between preamble chirps and add the preamble power constructively, achieving a more noise-resistant packet detection.

Secondly, we use a two-stage detection to perform accurate Carrier Frequency Offsets (CFO) and Timing Offset (TO) detection. The first coarse-stage detection uses the traditional up-down detection method [57]. After that, we remove the CFO and TO from the symbol by multiplying the singletone and shifting the samples. Then, we perform a fine-grained search since the initial frequencies of the preamble, and start frame delimiter (SFD) may not fall onto discrete FFT bins. Using the CFO and TO results from the previous stage, each symbol in the preamble, and SFD is now windowed correctly. Each symbol's power will be added constructively into one frequency bin, centered closely to zero. Instead of using FFT, we perform a heuristic search on a small interval around that peak based on Brent's method [5] to achieve the highest accuracy while keeping the computational cost low. Thus, we can compute CFO and TO at sub-integer levels to remove them accurately.

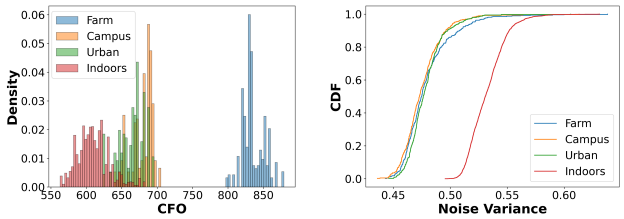
Figure 6: Hardware implementation.

Thirdly, after we perform accurate CFO and TO removal, we observe that if the symbol's frequency does not fall onto discrete FFT frequency bins, the power of FFT bins may be lower than the actual height of the power peak. The symbol may also not start at discrete samples, leading to a small time offset. The STO in the first step also accounts for the time offsets. Since time offset translates to frequency shifts after dechirp, we can mitigate the CFO, TO, and SFO by shifting each symbol by the frequencies. Through these procedures, we can detect extremely weak LoRa packets with high success rates and accurately remove the CFO, STO, and TO to extract high-quality LoRa symbols from the payload while maintaining a small computational overhead. The algorithm runs in parallel with demodulation, sharing the same GPU, requiring 147 ms to process one second of signal at SF10. It introduces only 12 ms detection delay after the end of the preamble, significantly shorter than the payload duration, allowing real-time packet detection.

4 Implementation

Hardware Implementation: We test the performance of LoRaSeek with real-life hardware. As shown in Figure 6, the transmitter is based on Semtech SX1276 [47], and the receiver is a USRP N210 SDR [44] with SBX 400-4400 RF daughterboards [43]. We use a ThinkPad E460 laptop and UHD+GNU-radio software to control the receiver. We use a sampling rate of 1 MS/s. Experiments are conducted at 470MHz central frequency, 3 bandwidths (e.g., 125 kHz, 250 kHz, 500 kHz), and 6 SFs (e.g., from 7 to 12). Each preamble contains 8 upchirps. Overall, we collected approximately 60,000 packets at high SNR (>30 dB).

Chirp Signal Dataset: We conduct experiments in four different environments: indoor, campus, urban, and farm. The collected data contains different levels of multipath, interference, and channel fading effects. Additionally, to generate signals with different SNR levels for evaluation, we add Additive White Gaussian Noise (AWGN) into the collected dataset



(a) Channel Multipath (b) Environmental Noise

Figure 7: Comparison of CFO and noise differences across environments.

[12, 24], rendering over 40 million chirp symbols from -40 dB to 15 dB.

Environmental Characterization: The four data collection environments exhibit different channel conditions. The indoor scenario, a furnished office, exhibits the strongest multipath and the highest interference, while the campus environment, a semi-urban setting with fewer buildings and more greenery, and the urban environment, characterized by dense high-rise buildings and busy streets, lie between these two extremes. We quantify multipath and interference intensity across the four environments.

Multipath effects are characterized using the CFO of received packets. Multipath introduces additional transmission delays. Because LoRa symbols are chirps with monotonically increasing instantaneous frequency, delayed multipath components exhibit lower frequency, thereby reducing the observed CFO. Hence, stronger multipath propagation corresponds to lower CFO. Using the same device across all environments to ensure a constant oscillator frequency, we measure CFO under each setting. In Figure 7a, the results show that indoor signals yield the lowest CFO, indicating the strongest multipath, followed by campus, while the farm exhibits the highest CFO, namely the weakest multipath.

Interference levels are assessed by analyzing noise characteristics. Interference produces frequency-selective power fluctuations, which manifest as uneven spectral spikes and increased variance in noise power. To quantify this effect, we first remove the signal component from each received packet, then compute the variance of the residual frequency-domain noise power. As shown in Figure 7b, the indoor environment exhibits the highest noise variance, attributable to coexisting wireless transmitters. In contrast, the other three environments show substantially lower variance, consistent with their relatively sparse wireless activity. These measurements confirm significant diversity in both channel and interference characteristics across the four environments, thereby capturing a wide spectrum of real-world LoRa deployment conditions.

Model Training and Testing: For each SF and BW, we split the dataset into training and test subsets, reserving 80% for training and the remaining 20% for testing. We use a batch size of 32 and a learning rate of 1×10^{-4}. The Adam optimizer is used with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. Lower SF, which contains fewer chirp symbols and smaller input dimension, shows faster convergence rates than higher SF. As we introduce the model to various extreme SNR levels (e.g., −40 dB), this can serve as a form of adversarial training.

Dual Loss Function: We propose a dual Mean Squared Error (MSE) reconstruction loss to optimize both spectrogram and time-domain representations. Let $\hat{\mathbf{X}}$ denote the denoised spectrogram, we collect the groundtruth high-SNR spectrogram $\mathbf{X}_{gt}$. Similiarly, we denote the reconstructed time-domain chirp signal as $\hat{x}(n)$ and the corresponding high SNR groundtruth $x_{gt}(n)$. The final loss function is:

$$\mathscr{L} = \alpha \cdot \text{MSE}(\hat{\mathbf{X}}, \mathbf{X}_{gt}) + \beta \cdot \text{MSE}(\hat{x}(n), x_{gt}(n)),$$

where α and β represent weights to maintain balance between the spectrogram and time-domain losses.

Curriculum Learning: We utilize curriculum learning [4], where we initially train LoRaSeek on high SNR level datasets before gradually presenting more challenging low SNR conditions. This systematic SNR decrease enhances LoRaSeek's robustness across varying noise levels while capturing essential signal characteristics in clean, optimal conditions.

Model Compression: We further compress LoRaSeek to enhance efficiency and capability within resource-constrained LoRa gateway environments. We add more convolution layers to further downsample the input dimension and implement structural pruning [16]. We preserve layers with the largest L1-norms. We incorporated mixed-precision training with half-precision (FP16) floating-point arithmetic and ONNX Runtime [9] to reduce model complexity, while fine-tuning to maintain performance.

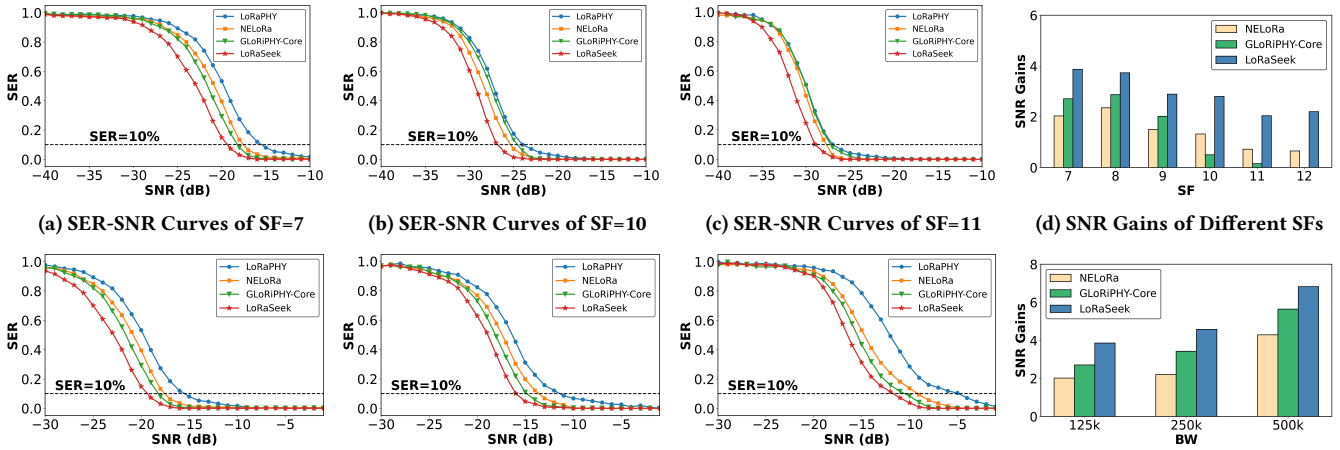
5 Evaluation

Baseline Methods: (1) **LoRaPHY [57]:** A non-neural method which is believed to be used in the default physical layer of commercial LoRa gateways. It applies dechirp and FFT to each symbol and then adds the power of the two signal power peaks. Their work provides two methods: "Abs+" and "Ps+", but since [11] pointed out that the two methods are equivalent, we refer to them as LoRaPHY.

(2) **NELoRa [24]:** A neural-based method which first applies a neural network to denoise the spectrogram of each symbol and then uses another network to classify symbols.

(3) **GLoRiPHY-Core [45]:** A neural-enhanced method that uses CNNs encoder-decoder with a transformer bottleneck to denoise the noisy LoRa signals. After denoising, it uses inverse STFT to reconstruct the signal from the denoised spectrogram and uses standard LoRa demodulation.

(4) **LoRaTrimmer [11]:** The state-of-the-art LoRa weak signal decoding method based on a probabilistic model, which is parallel with our neural-enhanced method. It replaces the



(a) SER-SNR Curves of SF=7 (b) SER-SNR Curves of SF=10 (c) SER-SNR Curves of SF=11 (d) SNR Gains of Different SFs

(e) SER-SNR Curves of BW=125K (f) SER-SNR Curves of BW=250K (g) SER-SNR Curves of BW=500K (h) SNR Gains of Different BWs

Figure 8: Overall performance of different methods under. (a)-(c) demonstrate SER-SNR curves under different SFs at BW=125 kHz. (d) shows SNR gains of different SFs at BW=125 kHz over LoRaPHY. (e)-(g) illustrate SER-SNR curve under different BWs at SF=7. (h) shows SNR gains of different BWs at SF=7 over LoRaPHY.

FFT in dechirp with a probabilistic phase-jump searching to optimally add the energy and remove out-band noise. LoRaTrimmer assumes the distribution of the chirp phase-jump is uniform. We compare LoRaSeek with LoRaTrimmer separately in Section 5.5 and 5.6.

Evaluation Metrics: (1) **Symbol Error Rate (SER):** After decoding each symbol, we calculate its SER by computing its accuracy with ground truth. (2) **SNR Gains:** We calculate SER at different SNRs. Thenwe set the SER threshold at 10%, a reasonable threshold for stable communication. We find the lowest SNR (in decibels) to achieve SER $\leq$ 10%, using interpolation when necessary.

5.1 Overall Performance

Setup: We evaluate LoRaSeek's performance with different LoRa configurations in the indoor environments, including 6 SFs (e.g., 7 to 12) and 3 BWs (e.g., 125K, 250K, 500K).

Results: Figure 8a-h shows the results. Given the BW is 125 kHz, Figure 8a-c shows the SER across different SNRs at small SF (e.g., 7) and large SF (e.g., 10, 11) scenarios. We measure the SNR gains of different demodulation techniques, shown in Figure 8d. LoRaSeek outperforms all the baseline methods with consistent SER improvements at all SNR levels. Specifically, in terms of SNR gains, our method achieves 2.04 to 3.86 dB gain over LoRaPHY, 1.32 to 1.83 dB gain over NELoRa, and 0.86 to 3.03 dB gain over GLoRiPHY-Core. We observe a clear performance that as the SF increases, NELoRa's performance declines, obtaining only approximately 0.7 dB gain over LoRaPHY when the SF is 11 or 12. For GLoRiPHY-Core, the decline is even more dramatic, with only 0.5 dB and 0.15 dB gain over LoRaPHY at SF set to 10 and 11 and no performance enhancement at SF=12. At SF=12, the lowest SNR required for GLoRiPHY-Core to achieve an SER of 10% is 0.83 dB lower than that of LoRaPHY, translating to a 3.03 dB behind LoRaSeek. This performance degradation results from the lack of robust network design to fully capture LoRa chirp signal multi-scale dependencies and high-dimensional signal characteristics at large SFs. In contrast, these results confirm LoRaSeek's efficiency in learning multi-dimensional signal features and maintaining robustness under ultra-low SNR. We further examine the performance of the BWs change when SF=7. In Figure 8e-h, LoRaSeek demonstrates consistent decoding improvements and the highest SNR gain over LoRaPHY in all BWs configurations. The largest SNR gain is BW=500 kHz, inducing 6.82 dB gain over LoRaPHY, 2.53 dB gain over NELoRa, and 1.18 dB gain over GLoRiPHY. For LoRa node battery saving, NELoRa achieves an extra 1.39 years at a 2 dB gain for SF7 BW125 kHz. LoRaSeek achieves a 3.86 dB gain, suggesting comparable or greater enhancement. At SF10, LoRaSeek can decode LoRa packets using lower SFs in ultra-low SNRs, potentially exceeding NELoRa's 272% battery life gain.

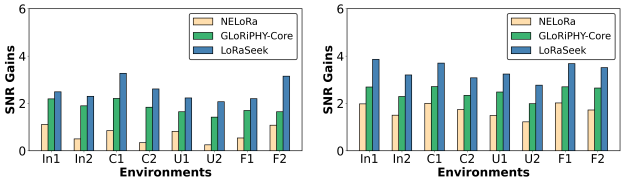
Remark: LoRaSeek maintain a consistent superior SNR gain over baseline methods: 2.04 to 3.86 dB gain over LoRaPHY, 1.32 to 1.83 dB gain over NELoRa, and 0.86 to 3.03 dB gain over GLoRiPHY-Core.

5.2 Robustness in Different Environments

Setup: We test LoRaSeek in indoor settings (e.g., offices) and outdoor settings (e.g., urban areas, campuses, and farms) with diverse noise patterns. Experiments were conducted with an SF=7 and a BW=125 kHz with different LoRa transmitters and gateways. We propose two evaluation scenarios: (1) directly using pretrained models from the indoor dataset (Section 5.1) to infer on a new dataset, and (2) fine-tuning all pretrained models on the new dataset for adaptation.

Table 1: Computational overhead of ML-based LoRa demodulation in terms of model parameters, storage overhead, and inference time for demodulating a packet with 16 chirp symbols on GPU (Full/Compressed model).

	Parameters (M)						Storage Overhead (MB)						Inference Time (ms)					
Method	SF7	SF8	SF9	SF10	SF11	SF12	SF7	SF8	SF9	SF10	SF11	SF12	SF7	SF8	SF9	SF10	SF11	SF12
NELoRa	8.0/5.1	18.2/12.4	52.2/40.0	174.4/49.3	636.0/63.1	2427.3/78.8	32.1/20.6	72.9/45.1	208.2/58.1	697.8/152.9	1051.2/241.1	3548.3/300.9	7.1/5.9	10.6/8.5	18.2/14.8	34.6/20.7	70.7/24.4	149.9/31.4
GLoRiPHY-Core	3.2/2.1	3.2/2.2	3.4/2.5	13.7/9.5	16.9/12.7	29.5/25.3	12.2/8.1	12.5/8.5	13.3/9.4	52.5/36.5	64.6/48.6	112.8/96.7	6.7/4.8	7.3/5.8	8.1/6.0	9.1/6.6	11.9/7.1	17.4/8.7
LoRaSeek	0.5/0.3	1.8/1.4	1.8/1.4	3.7/2.4	5.6/3.3	5.7/3.3	2.1/1.2	7.2/5.3	7.2/5.3	14.2/9.4	21.7/13.0	21.8/13.0	3.8/3.0	6.0/4.6	6.8/4.8	7.9/5.2	9.0/5.6	14.7/6.2



(a) Unseen Environments (b) Seen Environments

Figure 9: Comparison of SNR gains across different environments. (a) Performance of the pretrained models on new environments without fine-tuning. (b) Performance after fine-tuning on environment-specific data. Settings: In (Indoor), C (Campus), U (Urban), F (Farm).

Table 2: Time consumption to demodulate 16-chirp symbols on Raspberry Pi (ms, Full/Compressed model).

Method	SF7	SF8	SF9	SF10	SF11	SF12
NELoRa	13934/2023	24460/3579	46866/6694	130524/15455	NA/20977	NA/39452
GLoRiPHY-Core	834/492	1050/790	1455/1085	3740/2652	4930/4507	9432/8463
LoRaSeek	577/390	911/728	1351/980	3212/2078	4125/3248	8637/6590

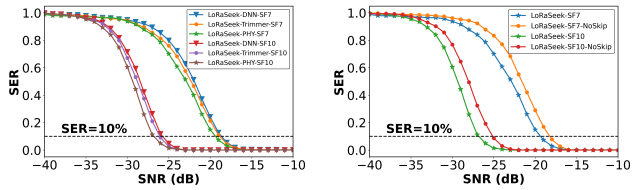
Results: Figure 9 displays LoRaSeek's SNR gains over LoRaPHY across different environments and locations. Without prior training or fine-tuning, LoRaSeek delivers the highest SNR gains over LoRaPHY, outperforming all baseline methods. As depicted in Figure 9a, when the pretrained model is directly applied to new environments without fine-tuning on a target dataset, LoRaSeek obtains SNR gains ranging from 1.67 to 3.15 dB over LoRaPHY, 1.38 to 2.11 dB over NELoRa, and 0.3 to 1 dB over GLoRiPHY-Core. As shown in Figure 9b, when fine-tuned on each dataset, LoRaSeek further enhances performance, achieving SNR gains of 3 to 3.86 dB over LoRaPHY, 1.5 to 1.88 dB over NELoRa, and 0.67 to 1.17 dB over GLoRiPHY-Core. Overall, LoRaSeek consistently improves SNR gains and decoding accuracy in diverse noise conditions, demonstrating its robustness in adapting to different hardware configurations and environments.

5.3 Computational Overhead

Setup: We evaluate the computational cost of LoRaSeek and other ML-based demodulation models based on model size, storage overhead, and inference time for demodulating a 16-chirps packet. Inference time is averaged over 100 times on a PC with one NVIDIA RTX 3090 GPU. We also test on Raspberry Pi to emulate the limited computing power of commercial gateways in resource-constrained conditions.

Results: As shown in Table 1, LoRaSeek maintains a lightweight yet effective design for LoRa demodulation:

(1) **Model Efficiency:** LoRaSeek maintains minimal parameter count from 0.5 million (e.g., SF7) to 5.7 million (e.g., SF12)



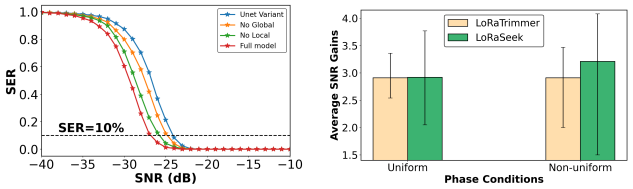
(a) SER-SNR Curves of incorporating different decoding methods (b) SER-SNR Curves of the dual attention skip connection

Figure 10: Ablation study on the impact of different decoding methods and the skip connection.

parameters. In stark contrast, NELoRa scales exponentially, reaching up to 2.4 billion parameters at SF12, 425 times larger than LoRaSeek. Compared to GLoRiPHY-Core, LoRaSeek has a significantly smaller footprint, approximately 1.7 to 6.4 times smaller, while inducing consistently higher demodulation accuracy. Furthermore, the compressed version further reduces the parameter count by 22.2% to 42.1%, maintaining efficiency without compromising performance.

(2) **Minimal Storage:** LoRaSeek requires only 2.1 MB (SF7) to 21.8 MB (SF12) storage on the disk, with a slight increase as SF increases. Notably, LoRaSeek's storage overhead at SF12 (21.8 MB) is 32% lower than that of NELoRa at SF7 (32.1 MB). In the compressed variant, LoRaSeek reduces storage requirement by 1.6-7.4× compared to GLoRiPHY-Core, enabling deployment on resource-constrained devices.

(3) **Low Inference Latency:** LoRaSeek exhibits the fastest inference times, introducing 3.8 ms at SF7 and 14.7 ms at SF12. This is substantially lower than other methods: from SF=7 to SF=12, NELoRa incurs 7.1 ms to 149.9 ms, and GLoRiPHY-Core requires 6.7 ms to 17.4 ms. The compressed version of LoRaSeek obtains from 21-57.9% speedup, up to 1.6× faster than GLoRiPHY-Core and 5× faster than NELoRa, making LoRaSeek an ideal solution for near real-time LoRa demodulation. Furthermore, though our method can run effectively on PC-based platforms, we further evaluate the performance on a resource-constrained device, the Raspberry Pi. As depicted in Table 2, our method incurs approximately 577 ms at SF7, which reduces to 390 ms after compression, achieving a 32% speedup. This highlight LoRaSeek's optimal balance between decoding performance and computational requirements for practical applications. Given its low computational footprint and minimal time consumption, LoRaSeek facilitates real-world deployments across diverse hardware and computational resource configurations.



(a) SER-SNR Curves of local and global feature extraction blocks (b) SNR gains between LoRaSeek and LoRaTrimmer.

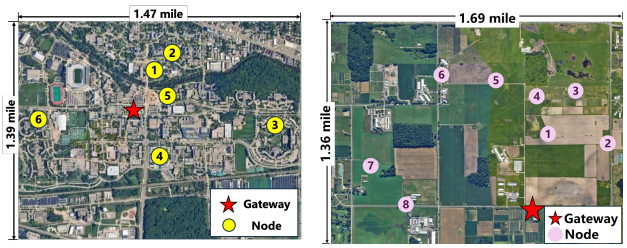
Figure 11: (a) Ablation study on local and global feature extraction blocks, (b) SNR gains comparison between LoRaSeek and LoRaTrimmer under dataset with uniform and non-uniform phase distribution.

5.4 Ablation Study

Different Decoding Methods: We evaluate LoRaSeek's performance with all possible LoRa demodulation methods: LoRaPHY [57] (LoRaSeek by default), LoRaTrimmer [11] (Trimmer), and NELoRa's neural decoder (DNN) [24]. Figure 10a illustrates the SER of LoRaSeek with different decoding techniques. LoRaSeek obtains the highest SNR gains when incorporated with the standard demodulation approach: 3.86 dB gains at SF7 and 2.80 dB gains at SF10 over LoRaPHY. In contrast, NELoRa's neural decoder achieves the lowest SNR gains: 3.14 dB gains at SF7 and 1.74 dB gains at SF10 over LoRaPHY. This reduction in performance could be due to the expensive computational overhead of the neural decoder and its limited ability to capture the chirp signal's features in infinite noise patterns. It is noteworthy that across all decoding methods, LoRaSeek outperforms other baseline methods.

Dual Attention Skip Connection Performance: Figure 10b shows the SNR gains of LoRaSeek over LoRaPHY with and without the dual attention skip connection. The dual attention skip connections result in SNR gains of 2.8 dB at SF10 and 3.86 dB at SF7 over LoRaPHY while removing them reduces SNR gains to 1.68 dB at SF10 and 2.75 dB at SF7. These results confirm the effectiveness of the skip connection in preserving distinct characteristics and mitigating information loss in chirp signals. Moreover, the dual attention skip connection incurs only a slight increase in the model complexity, achieving a 1.12 dB SNR gain at the cost of a mere 0.04-0.08 million parameter increase. Even without dual attention skip connection, LoRaSeek still outperforms other baseline methods, with SNR gains of 0.72 dB over NELoRa (e.g., SF7) and 0.62 dB over GLoRiPHY-Core (e.g., SF10).

Local and Global Feature Extraction: We conduct three ablation studies as follows: (1) removing local feature extraction blocks, (2) removing global feature extraction blocks, and (3) replacing global blocks with local ones and removing skip connections, making a conventional U-Net variant. Figure 11a shows that local and global blocks contribute 1.05 and 2 dB SNR gains, respectively. The U-Net model suffers the most serious performance degradation, even worse than the baseline LoRaPHY. These results confirm the effectiveness of our hybrid design in capturing the unique characteristics of LoRa chirps for robust denoising.



(a) campus-scale deployment (b) farm-scale deployment

Figure 12: Deployment of coverage experiments.

5.5 Comparison with LoRaTrimmer

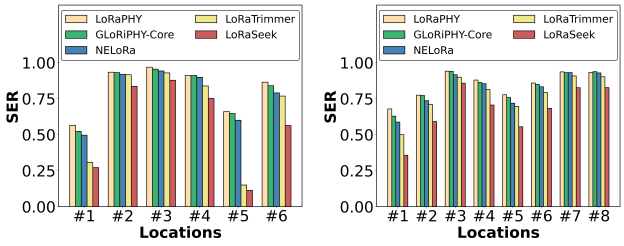
In LoRa demodulation, a frequency jump occurring within a symbol duration splits the symbol's energy into two segments, and the Sampling Time Offset (STO) causes a phase jump between these segments. At large STOs, the phase jump follows a uniform distribution; At small STOs, its distribution concentrates around zero.

LoRaTrimmer [11] is designed based on the assumption of a uniform phase jump distribution, but it may not be the case with smaller oscillator drifts and shorter packet lengths. To systematically evaluate the performance of LoRaSeek compared to LoRaTrimmer across different phase jump distributions, we conduct two experiments : (1) **Uniform Phase Distribution:** Signal phases are independently and uniformly sampled from the range $[-\pi, \pi]$. (2) **Non-Uniform Phase Distribution:** Signal phases exhibit a bias towards values close to zero, drawn from a normal distribution centered at zero with a standard deviation of 0.1.

Results: Figure 11b illustrates the results. Under non-uniform scenarios, LoRaSeek and LoRaTrimmer achieve comparable performance with an average SNR gain of 2.91 dB over LoRaPHY. On the other hand, under uniform phase conditions, LoRaSeek has an average SNR gain of 3.21 dB over LoRaPHY, while LoRaTrimmer achieves 2.91 dB. In fact, phase jump distributions are complex in real-world deployments depending on hardware and communication settings, and they can be further canceled by accurate CFO and STO calibration. LoRaTrimmer assumes that noise is evenly distributed AWGN and phase follows an even distribution over $[-\pi, \pi]$, both may deviate from the real-life distributions. LoRaSeek can learn the intricate patterns of real-life phase distributions and adapt to different settings, achieving a higher SNR gain.

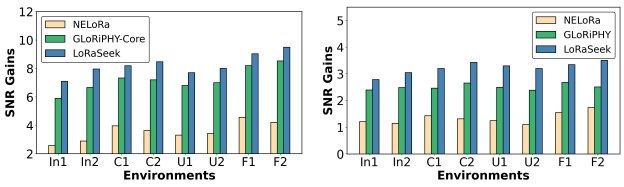
5.6 Real-world Coverage

Setup: To evaluate the real-world coverage of LoRaSeek, we conducted experiments in two environments: a 1.39×1.47-mile urban campus and a 1.36×1.69-mile rural farm, covering various landcover types (e.g., trees, buildings, roads, hills, and crops). We set SF at 10 and BW at 125 kHz. LoRa nodes were



(a) campus-scale testbed (b) farm-scale testbed

Figure 13: SER performance in real campus and farm.



(a) Unlicensed Band Interference (b) Fading and Multipath Channel

Figure 14: Robustness to (a) Unlicensed Interference and (b) Channel Fading in various environments.

deployed at six campus locations and eight farm locations, with each location transmitting 61 packets, each containing 36 chirp symbols. The deployment of the experiments is illustrated in Figure 12.

Results: All the packets transmitted across the 14 locations were successfully detected, validating the effectiveness of our packet detection method. Figure 13a presents the SER of five methods at 6 locations in an urban campus. Overall, as the distance between the gateway and the LoRa node increases, the SER tends to rise. However, despite its proximity to the gateway, location 5 experiences a high SER due to building obstructions. Compared with the baseline, the LoRaSeek decreases SER by 54.7% to 11.3%. Moreover, despite severe signal attenuation at location 6, LoRaTrimmer reduces the SER by 9.7% to 76.8%, while our proposed method, LoRaSeek, achieves an even greater reduction of 20.5%, lowering the SER to 56.3%. On campus, LoRaSeek lowers the SER by an average of 8.3%, and by up to 20.5% compared to LoRaTrimmer. Figure 13b shows the SER of eight locations in a farm. We can still observe that LoRaSeek outperforms other baseline methods. At location 1, which has the shortest distance, LoRaSeek achieves an SER of 59.1%, whereas LoRaTrimmer achieves 70.9%. At location 7, the farthest distance, LoRaSeek achieves an SER of 82.6%, while LoRaTrimmer achieves 90.8%. In the farm, LoRaSeek lowers the SER by an average of 10.2%, and by up to 14.11% compared to LoRaTrimmer.

Remark: This highlights LoRaSeek's advantage over previous methods in real scenarios, as it effectively captures multi-dimensional features under strong noise.

5.7 Robustness to Unlicensed Band Interference

The interference from other unlicensed band devices can be categorized into interference from other wireless protocols operating in the same band, and interference from other LoRa transmitters. For the former, we evaluate LoRaSeek with co-existence interference from Zigbee in the 902-928 MHz band [20]. We insert Zigbee BPSK modulated interference into high-SNR LoRa signals. We control Signal-to-Interference Ratio (SIR) by scaling Zigbee interference power to LoRa with SIR values from -20 dB to +20 dB. We also add AWGN at varying SNRs to simulate realistic channels. As shown in Figure 14a, LoRaSeek consistently outperforms LoRaPHY, NELoRa, and GLoRiPHY-Core with up to 9.49, 5.29, and 1.3 dB SNR gains. For LoRa interference, chirp signals may have the same or different SFs. In the first case, LoRaSeek utilizes the quasi-orthogonality of chirp time and frequency domains for effective decoding via chirp distinct multi-dimensional features. When multiple LoRa signals have the same SF, it is a collision decoding problem [25], which is not within our research scope. LoRaSeek is designed to effectively reconstruct and decode weak signals, so it is not optimized to resolve collisions, especially when interference is stronger than the signal and has high similarity in feature space. We acknowledge that ML–based LoRa collision decoding remains an open problem, requiring a novel architecture to maximize feature differences among multiple the same SF chirp signals.

5.8 Robustness to Fading and Multipath Channel

We evaluate LoRaSeek on a dataset with simulated Rayleigh fading and multipath effects. From high-SNR signal data, we use MATLAB's Communication Toolbox [35] to configure channel parameters: number of paths, average path gains, and Doppler shifts. AWGN is added at varying SNR levels to further assess robustness. As shown in Figure 14b, in different settings of office, campus, urban, and farm, LoRaSeek consistently outperforms LoRaPHY, NELoRa, and GLoRiPHY with up to 3.51, 2.11, and 1.0 dB SNR gains.

6 Related Work

MIMO-based Decoding: Recent studies have demonstrated that using multiple gateways or LoRa nodes can significantly enhance the SNR [2, 10, 14, 18]. For instance, Charm[10] coordinates multiple gateways to identify weak signals by detecting a combined energy peak. OPR [2] achieves better bit error recovery performance as the number of gateways increases. Choir [14] leverages signal correlation among nearby LoRa nodesChime [17] selects frequencies strategically across multiple gateways to mitigate multipath effects. MALoRa [18] and PCube [56] synchronize gateways using a shared clock, but this restricts the distance between gateways, resulting

in limited overall coverage. In contrast, LoRaSeek leverages state-of-the-art deep learning architecture to achieve additional SNR gain, seamlessly integrating with existing MIMO systems for improved decoding.

Single gateway Decoding: LoRaPHY[57] introduces a phase alignment strategy to enhance signal power. Chime[17] examines multiple wireless channel characteristics to identify optimal frequencies. Nephalai[28] uses a compressive-sensing cloud radio access network to enable efficient multi-channel LPWAN decoding. Falcon [49] employs selective interference with competing LoRa transmissions. Ostinato [58] enhances SNR by restructuring original packets into pseudo packets composed of repeated symbols. FerryLink [59] presents a two-hop forwarding approach to enhance unreliable LoRa connections. Demeter [41, 42] optimizes signal reception through polarization alignment. XCopy [55] achieves improved SNR by coherently integrating signals from retransmitted packets. LoRaTrimmer [11] optimizes decoding by trimming out noise components and merging power effectively. In contrast, LoRaSeek employs an ML-based hierarchical structure to denoise signals from multiscale features, improving the standard decoding technique's accuracy.

Deep Learning based wireless networks: Recent studies [3, 33, 34, 51] use neural networks to infer downlink from uplink channels. Amani et al. [1] explored radio frequency fingerprinting for authenticating LoRa. The DeepLoRa framework [29] leverages an LSTM network to accurately predict path loss in LoRa links. DeepSense [8] utilizes CNN and RNN architectures to facilitate random access and coexistence in LoRaWAN systems, effectively operating below the noise floor by converting complex signals into power spectrogram representations. NELoRa [12, 24] combines CNN and LSTM to decode LoRa signals received at a single gateway, particularly addressing low-SNR scenarios. SRLoRa [13] builds upon NeLoRa, enhancing the decoding performance when utilizing multiple LoRa gateways. ChirpTransformer [26, 38] modifies the encoder on the LoRa node and uses a DNN decoder to enhance the ultra-low SNR scenarios.

7 Discussion

Low-Power IoT Feasibility: A LoRa network consists of LoRa nodes, gateways, and cloud servers. While LoRa nodes operate under strict battery constraints, gateways are typically powered by cables or solar panels, providing more stable power and stronger computational capabilities. Gateways receive and decode packets from LoRa nodes and forward them to the cloud. LoRaSeek is deployed on LoRa gateways, allowing LoRa nodes to operate with standard LoRa protocols without compromising their original functionality. To further improve gateway energy efficiency, when backhaul bandwidth is sufficient, LoRa Cloud Radio Access Networks (CRAN) [48] can be used to offload packet decoding to the cloud. In this case, LoRaSeek is deployed on the cloud server, and raw LoRa signals captured by the gateways are transmitted to the cloud via CRAN.

End-to-End Computation Optimization: In addition to model compression used in LoRaSeek, we can selectively trigger full-layer LoRaSeek only for weak signals below a predefined SNR threshold. This SNR-aware triggering, trained through data-intensive methods, enables LoRaSeek to operate in a lightweight mode under high-SNR conditions. Beyond software optimization, computation can be accelerated through hardware–software co-design, leveraging low-power edge AI chips (e.g., Google Edge TPU [21]) to enhance LoRa gateway processing capabilities. The additional overhead for robust packet detection and offset estimation is approximately 12 ms per packet on a GPU server—comparable to chirp symbol decoding. With hardware–software co-design, this process can be further optimized by executing packet detection on FPGAs and decoding on edge AI chips, enabling pipelined processing to increase overall throughput. However, achieving optimal latency–throughput trade-offs through pipelined packet detection and decoding remains an open challenge.

Model Training and Enhancement: In a static environment, channel noise consists of both stable and dynamic components. Before deploying LoRaSeek in a given environment, noise-type-specific regularization can enhance denoising by learning noise-invariant features from preambles [45].

8 Conclusion

In this paper, we introduce LoRaSeek, a lightweight and reliable LoRa denoising framework that enhances chirp signal quality and decoding robustness. At its core, LoRaSeek features a hierarchical denoising architecture combining multi-stage CNN and Transformer blocks to extract local and global features, effectively mitigating multiscale distortions while maintaining linear scaling complexity. Additionally, a selective denoising mechanism with dual attention-based skip connections filters and integrates multistage information across both channel and spatial dimensions, ensuring high-quality signal reconstruction. We implement LoRaSeek with COTS LoRa end node and USRP N210. The evaluation results under various configurations and environments demonstrate that it surpasses state-of-the-art methods in both robustness and computational efficiency.

Acknowledgement

We sincerely thank the anonymous reviewers and our shepherd for their valuable feedback. This work was partially supported by NSF Awards NeTS-2312674, 2312675, and 2312676.

References

[1] Amani Al-Shawabka, Philip Pietraski, Sudhir B Pattar, Francesco Restuccia, and Tommaso Melodia. 2021. DeepLoRa: Fingerprinting LoRa devices at scale through deep learning and data augmentation. In *Proceedings of MobiHoc*.

[2] Artur Balanuta, Nuno Pereira, Swarun Kumar, and Anthony Rowe. 2020. A cloud-optimized link layer for low-power wide-area networks. In *Proceedings of ACM MobiSys*.

[3] Avishek Banerjee, Xingya Zhao, Vishnu Chhabra, Kannan Srinivasan, and Srinivasan Parthasarathy. 2024. HORCRUX: Accurate cross band channel prediction. In *Proceedings of ACM MobiCom*.

[4] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum learning. In *Proceedings of ICML*.

[5] Richard P Brent. 2013. *Algorithms for minimization without derivatives*. Courier Corporation.

[6] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. 2022. Swin-unet: Unet-like pure transformer for medical image segmentation. In *Proceedings of ECCVW*.

[7] Tusher Chakraborty, Heping Shi, Zerina Kapetanovic, Bodhi Priyantha, Deepak Vasisht, Binh Vu, Parag Pandit, Prasad Pillai, Yaswant Chabria, Andrew Nelson, Michael Daum, and Ranveer Chandra. 2022. Whisper: IoT in the TV White Space Spectrum. In *Proceedings of USENIX NSDI*.

[8] Justin Chan, Anran Wang, Arvind Krishnamurthy, and Shyamnath Gollakota. 2019. DeepSense: Enabling carrier sense in low-power wide area networks using deep learning. *arXiv preprint arXiv:1904.10607* (2019).

[9] ONNX Runtime developers. 2021. ONNX Runtime. https://onnxruntime.ai/. Retrieved by Aug 30 2025.

[10] Adwait Dongare, Revathy Narayanan, Akshay Gadre, Anh Luong, Artur Balanuta, Swarun Kumar, Bob Iannucci, and Anthony Rowe. 2018. Charm: Exploiting geographical diversity through coherent combining in low-power wide-area networks. In *Proceedings of ACM/IEEE IPSN*.

[11] Jialuo Du, Yunhao Liu, Yidong Ren, Li Liu, and Zhichao Cao. 2024. Loratrimmer: Optimal energy condensation with chirp trimming for lora weak signal decoding. In *Proceedings of ACM MobiCom*.

[12] Jialuo Du, Yidong Ren, Mi Zhang, Yunhao Liu, and Zhichao Cao. 2023. NELoRa-Bench: A Benchmark for Neural-enhanced LoRa Demodulation. *International Conference on Learning Representations (ICLR) Workshop on Machine Learning for IoT* (2023).

[13] Jialuo Du, Yidong Ren, Zhui Zhu, Chenning Li, Zhichao Cao, Qiang Ma, and Yunhao Liu. 2023. SRLoRa: Neural-enhanced LoRa Weak Signal Decoding with Multi-gateway Super Resolution. In *Proceedings of ACM MobiHoc*.

[14] Rashad Eletreby, Diana Zhang, Swarun Kumar, and Osman Yağan. 2017. Empowering Low-Power Wide Area Networks in Urban Settings. In *Proceedings of ACM SIGCOMM*.

[15] Chi-Mao Fan, Tsung-Jung Liu, and Kuan-Hsien Liu. 2022. SUNet: Swin transformer UNet for image denoising. In *Proceedings of ISCAS*.

[16] Biyi Fang, Xiao Zeng, and Mi Zhang. 2018. Nestdnn: Resource-aware multi-tenant on-device deep learning for continuous mobile vision. In *Proceedings of ACM MobiCom*.

[17] Akshay Gadre, Revathy Narayanan, Anh Luong, Anthony Rowe, Bob Iannucci, and Swarun Kumar. 2020. Frequency Configuration for Low-Power Wide-Area Networks in Heartbeat. In *Proceedings of USENIX NSDI*.

[18] Ningning Hou, Xianjin Xia, and Yuanqing Zheng. 2022. Don't Miss Weak Packets: Boosting LoRa Reception with Antenna Diversities. In *Proceedings of IEEE INFOCOM*.

[19] Jie Hu, Li Shen, and Gang Sun. 2018. Squeeze-and-excitation networks. In *Proceedings of the IEEE CVPR*.

[20] Digi International. 2025. Digi XBee-PRO 900HP RF Module. https://www.digi.com/products/embedded-systems/digi-xbee/rf-modules/sub-1-ghz-rf-modules/xbee-pro-900hp. Retrieved by Aug 30 2025.

[21] Norman P Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, et al. 2017. In-datacenter performance analysis of a tensor processing unit. In *Proceedings of ISCA*.

[22] M Krithika Alias AnbuDevi and K Suganthi. 2022. Review of semantic segmentation of medical images using modified architectures of UNET. *Diagnostics* 12, 12 (2022), 3064.

[23] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Proceedings of NIPS*.

[24] Chenning Li, Hanqing Guo, Shuai Tong, Xiao Zeng, Zhichao Cao, Mi Zhang, Qiben Yan, Li Xiao, Jiliang Wang, and Yunhao Liu. 2021. NELoRa: Towards Ultra-low SNR LoRa Communication with Neural-enhanced Demodulation. In *Proceedings of ACM SenSys*.

[25] Chenning Li, Xiuzhen Guo, Longfei Shangguan, Zhichao Cao, and Kyle Jamieson. 2022. CurvingLoRa to Boost LoRa Network Throughput via Concurrent Transmission. In *Proceedings of USENIX NSDI*.

[26] Chenning Li, Yidong Ren, Shuai Tong, Shakhrul Iman Siam, Mi Zhang, Jiliang Wang, Yunhao Liu, and Zhichao Cao. 2024. Chirptransformer: Versatile lora encoding for low-power wide-area iot. In *Proceedings of ACM MobiSys*.

[27] Yawei Li, Kai Zhang, Jiezhang Cao, Radu Timofte, and Luc Van Gool. 2021. Localvit: Bringing locality to vision transformers. *arXiv preprint arXiv:2104.05707* (2021).

[28] Jun Liu, Weitao Xu, Sanjay Jha, and Wen Hu. 2020. Nephalai: towards LPWAN C-RAN with physical layer compression. In *Proceedings ACM MobiCom*.

[29] Li Liu, Yuguang Yao, Zhichao Cao, and Mi Zhang. 2021. DeepLoRa: Learning Accurate Path Loss Model for Long Distance Links in LPWAN. In *Proceedings of IEEE INFOCOM*.

[30] Yimeng Liu, Maolin Gan, Gen Li, Younsuk Dong, and Zhichao Cao. 2025. Adonis: Neural-enhanced Fine-grained Leaf Wetness Sensing with Efficient mmWave Imaging. In *Proceedings of IEEE INFOCOM*.

[31] Yimeng Liu, Maolin Gan, Huaili Zeng, Li Liu, Younsuk Dong, and Zhichao Cao. 2024. Hydra: Accurate multi-modal leaf wetness sensing with mm-wave and camera fusion. In *Proceedings of ACM MobiCom*.

[32] Yimeng Liu, Maolin Gan, Huaili Zeng, Yidong Ren, Gen Li, Jingkai Lin, Younsuk Dong, Xiaobo Tan, and Zhichao Cao. 2025. Proteus: Enhanced mmWave Leaf Wetness Detection with Cross-Modality Knowledge Transfer. In *Proceedings of ACM SenSys*.

[33] Zikun Liu, Gagandeep Singh, Chenren Xu, and Deepak Vasisht. 2021. FIRE: enabling reciprocity for FDD MIMO systems. In *Proceedings of ACM MobiCom*.

[34] Zikun Liu, Changming Xu, Yuqing Xie, Emerson Sie, Fan Yang, Kevin Karwaski, Gagandeep Singh, Zhao Lucis Li, Yu Zhou, Deepak Vasisht, et al. 2023. Exploring practical vulnerabilities of machine learning-based wireless systems. In *Proceedins of USENIX NSDI*.

[35] MATLAB. 2025. MATLAB's Communication Toolbox. https://www.mathworks.com/products/communications.html. Retrieved by August 10, 2025.

[36] Keiron O'shea and Ryan Nash. 2015. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458* (2015).

[37] Yidong Ren, Amalinda Gamage, Li Liu, Mo Li, Shigang Chen, Younsuk Dong, and Zhichao Cao. 2024. Sateriot: High-performance ground-space networking for rural iot. In *Proceedings of ACM MobiCom*.

[38] Yidong Ren, Maolin Gan, Chenning Li, Shakhrul Iman Siam, Mi Zhang, Shigang Chen, and Zhichao Cao. 2025. Morph: ChirpTransformer-based Encoder-decoder Co-design for Reliable LoRa Communication.

arXiv preprint arXiv:2507.22851 (2025).

[39] Yidong Ren, Gen Li, Yimeng Liu, Younsuk Dong, and Zhichao Cao. 2025. Aeroecho: Towards agricultural low-power wide-area backscatter with aerial excitation source. In *Proceedings of IEEE INFOCOM*.

[40] Yidong Ren, Li Liu, Chenning Li, Zhichao Cao, and Shigang Chen. 2022. Is lorawan really wide? fine-grained lora link-level measurement in an urban environment. In *Proceedings of IEEE ICNP*.

[41] Yidong Ren, Wei Sun, Jialuo Du, Huaili Zeng, Younsuk Dong, Mi Zhang, Shigang Chen, Yunhao Liu, Tianxing Li, and Zhichao Cao. 2024. Demeter: Reliable Cross-soil LPWAN with Low-cost Signal Polarization Alignment. In *Proceedings of ACM MobiCom*.

[42] Yidong Ren, Yawen Wang, Younsuk Dong, Shigang Chen, Mi Zhang, Jiliang Tang, and Zhichao Cao. 2024. Demeter-demo: Demonstrating cross-soil lpwan with low-cost signal polarization alignment. In *Proceedings of ACM MobiCom*.

[43] Ettus Research. 2025. SBX 400-4400 datasheet. https://www.ettus.com/all-products/sbx/. Retrieved by Aug 30 2025.

[44] Ettus Research. 2025. USRP N210 datasheet. https://www.ettus.com/all-products/un210-kit/. Retrieved by Aug 30 2025.

[45] Kanav Sabharwal, Soundarya Ramesh, Jingxian Wang, Dinil Mon Divakaran, and Mun Choon Chan. 2024. Enhancing LoRa Reception with Generative Models: Channel-Aware Denoising of LoRaPHY Signals. In *Proceedings of ACM SenSys*.

[46] Semtech. 2025. LoRa by the Numbers. https://www.semtech.com/lora. Retrieved by Aug 30 2025.

[47] Semtech. Retrieved by Aug 30 2025. Semtech SX1276. https://www.semtech.com/products/wireless-rf/lora-connect/sx1276.

[48] Muhammad Osama Shahid, Daniel Koch, Jayaram Raghuram, Bhuvana Krishnaswamy, Krishna Chintalapudi, and Suman Banerjee. 2024. {Cloud-LoRa}: Enabling Cloud Radio Access {LoRa} Networks Using Reinforcement Learning Based {Bandwidth-Adaptive} Compression. In *Proceedings of USENIX NSDI*.

[49] Shuai Tong, Zilin Shen, Yunhao Liu, and Jiliang Wang. 2021. Combating link dynamics for reliable lora connection in urban settings. In *Proceedings of ACM MobiCom*.

[50] Deepak Vasisht, Zerina Kapetanovic, Jongho Won, Xinxin Jin, Ranveer Chandra, Sudipta Sinha, Ashish Kapoor, Madhusudhan Sudarshan, and Sean Stratman. 2017. FarmBeats: an IoT platform for Data-Driven agriculture. In *Proceedings of USENIX NSDI*.

[51] Deepak Vasisht, Swarun Kumar, Hariharan Rahul, and Dina Katabi. 2016. Eliminating channel feedback in next-generation cellular networks. In *Proceedings of ACM SIGCOMM*.

[52] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of NIPS*.

[53] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. 2022. Uformer: A general u-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF CVPR*.

[54] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. 2018. Cbam: Convolutional block attention module. In *Proceedings of ECCV*.

[55] Xianjin Xia, Qianwu Chen, Ningning Hou, Yuanqing Zheng, and Mo Li. 2023. XCopy: Boosting Weak Links for Reliable LoRa Communication. In *Proceedings of ACM MobiCom*.

[56] Xianjin Xia, Ningning Hou, Yuanqing Zheng, and Tao Gu. 2021. PCube: scaling LoRa concurrent transmissions with reception diversities. In *Proceedings of ACM MobiCom*.

[57] Zhenqiang Xu, Shuai Tong, Pengjin Xie, and Jiliang Wang. 2022. From Demodulation to Decoding: Toward Complete LoRa PHY Understanding and Implementation. *ACM Transactions on Sensor Networks* 18, 4 (2022), 64:1–64:27.

[58] Zhenqiang Xu, Pengjin Xie, Jiliang Wang, and Yunhao Liu. 2022. Ostinato: Combating lora weak links in real deployments. In *Proceedings of IEEE ICNP*.

[59] Jing Yang, Zhenqiang Xu, and Jiliang Wang. 2021. Ferrylink: Combating link degradation for practical lpwan deployments. In *Proceedings of IEEE ICPADS*.

[60] Kun Yuan, Shaopeng Guo, Ziwei Liu, Aojun Zhou, Fengwei Yu, and Wei Wu. 2021. Incorporating convolution designs into visual transformers. In *Proceedings of the IEEE/CVF ICCV*.

[61] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF CVPR*.