

HybNet: A Hybrid Deep Learning-Matched Filter Approach for IoT Signal Detection

KOSTA DAKIC¹ (Graduate Student Member, IEEE), BASSEL AL HOMSSI² (Member, IEEE), MARGARET LECH¹ (Member, IEEE), AND AKRAM AL-HOURANI¹ (Senior Member, IEEE)

¹School of Engineering, RMIT University, Melbourne, VIC 3000, Australia

²School of Engineering and Information Technology, UNSW Canberra, Canberra, ACT 2610, Australia

CORRESPONDING AUTHORS: K. DAKIC (kosta.dakic@ieee.org) AND A. AL-HOURANI (akram.hourani@rmit.edu.au)

This work was supported by the Australian Government RTP Stipend Scholarship (RSS).

ABSTRACT Random access schemes are widely used in IoT wireless access networks. They enable a reduced complexity and overcome power consumption constraints. Nevertheless, random access results in high packet losses which are caused by overlapping transmissions. Signal detection methods for digital modulation techniques are typically based on the well-established *matched filter*, which is proven as the optimal filter under additive white Gaussian noise for minimizing error probability. However, with the *colored* interference arising from the overlapping IoT transmissions, deep learning approaches are being considered a suitable alternative. In this paper, we present a hybrid framework, dubbed as *HybNet*, that alternates between deep learning and match filter pathways based on the perceived interference level. This helps the detector to work in a broader range of conditions, optimally leveraging both the matched filter and deep learning advantages. We compare the performance of several possible data modalities and detection architectures concerning the interference-to-noise ratio, demonstrating that leveraging domain knowledge by pre-processing the input data in conjunction with the proposed HybNet surpasses the complex conjugate matched filter performance under interference-limited scenarios.

INDEX TERMS Internet of Things, LoRa, deep learning, chirp spread spectrum, convolutional neural network, matched filter, signal detection, interference.

I. INTRODUCTION

INTERNET-OF-THINGS (IoT) wireless technologies play a significant role by facilitating the connectivity between sensor devices and Internet cloud services. According to a report from Ericsson, over 5 billion cellular IoT connections are anticipated by 2025 [1]. For wireless IoT devices, this increase would further congest the scarce radio spectrum and increase the level of interference. This is already evident in the license-free *industrial, scientific, and medical* (ISM) band [2] since this band is available to the public and does not require paid licensing.

In centrally-controlled networks, the base station, e.g., gNB in 5G, orchestrates the access to the spectral resources by employing a multiple-access technique such as time-division multiple access (TDMA), frequency-division multiple access (FDMA), code-division multiple access (CDMA), and orthogonal frequency-division multiple access (OFDMA). These access techniques are required to alleviate

co-channel interference. Nevertheless, such techniques mandate significant overhead, adding complexity to the protocol stack. Thus, typical license-free IoT networks employ random access techniques to simplify the protocol stack, thereby lowering energy consumption as well as device cost. Furthermore, a complex access technique would not be efficient in the first place given that the interference is originating from other co-existing systems in the same band.

Efficient signal detection is crucial for Low power wide area networks (LPWAN) technologies, which are receiving increasing attention due to the increasing use of IoT applications. LPWANs can achieve long-distance communication while maintaining low-power consumption at the cost of a reduced bit rate. In a license-free spectrum, LPWAN network performance is typically an interference-limited system in urban environments, while it becomes noise-limited in rural and remote locations. A prominent LPWAN technology is LoRaWAN, adopted by the LoRa Alliance [3]. LoRaWAN

typically uses an ALOHA-based protocol, allowing multiple IoT devices to transmit without coordination or prior signaling. However, the transmitted signals are prone to packet collisions which significantly reduces the throughput and in turn, the network performance [4], [5], [6], [7], [8]. LoRaWAN utilizes the LoRa modulation scheme, which is based on chirp spread spectrum (CSS) modulation spreading the signal energy over a wider bandwidth to combat narrowband interference [7]. Furthermore, the time-spreading of the transmission can be controlled using a parameter called *spreading factor* (SF), which increases the energy at the receiver without the need to transmit at a higher RF power. Aside from potential narrowband interference, LoRa-to-LoRa-induced interference significantly degrades the performance, especially when the interfering LoRa signal uses the same SF.

To achieve signal (frame) detection, receivers typically employ the matched filter architecture, where the signal is compared, or *matched*, to a known template(s). This method is proven to be optimal under additive white Gaussian noise (AWGN) conditions [9]. However, the rapid evolution of neural networks (NN), particularly deep learning (DL) methods, has shown great potential for signal detection [10], [11], [12], [13], [14], [15]. The main appeal of DL signal detection is its strength against non-linearities [16]. Notably, the Convolutional Neural Network (CNN) has been utilized to classify stochastic signals such as in image classification [17], which is not practically possible using deterministic methods such as the matched filter.

There seems to be little previous research work focusing on mitigating the issue of LoRa modulation with the same technology interference by utilizing DL-based detection. This paper presents a new framework, *HybNet*, that alternates between a (i) matched filter detector and (ii) a proposed DL detector, where the switching is automatically decided based on the interference level. Consequently, Our framework harnesses the benefits of both (i) the matched filter's ideal performance under noise-limited scenarios from one side, and (ii) the improved performance of the proposed DL detector in LoRa-to-LoRa interference on the other side. Additionally, we explore three different input data modalities for the DL-based detection and compare their performance; (i) time-domain (I/Q samples), (ii) time-frequency domain (spectrogram), and (iii) frequency domain (spectrum). Results show that the proposed DL-detector outperforms traditional noncoherent detection in non-Gaussian interference scenarios and without perfect phase synchronization. Also, the frequency domain DL detector shows a greater performance with a lower network complexity compared to the time-domain and time-frequency domain DL detectors. The lower complexity is due to the smaller input frame size compared to the I/Q frame and the 2D spectrogram. To evaluate the proposed framework at different interference-to-noise levels, we vary the interference-to-noise ratio (INR) and obtain the corresponding bit-error rate (BER). Accordingly, the overall BER performance results indicate that the proposed DL-based detectors can significantly improve LoRa detection perfor-

mance for LoRa symbols under high interference. The contribution of this work is summarized as follows.

- It presents a novel LoRa symbol detector, *HybNet*, which combines both matched filter detection and DL-enabled detection via an automatic switching mechanism. Through this combination, HybNet can capitalize on matched filter detection capabilities in noisy mediums while benefiting from DL detection under co-channel interference.
- It demonstrates the compatibility of the proposed neural network with conventionally deployed energy detectors to switch between DL and matched filter detection based on the measured interference level. This provides a trade-off in terms of complexity and detection performance.
- It assesses the performance of various data modalities and their corresponding neural network architectures. The data modalities assessed are time, frequency, and time-frequency.
- It proposes an experimental framework for systematically evaluating symbol detection efficiency using various input data modalities, AWGN impairments, and co-channel interference levels.

The rest of this paper is organized as follows. A literature review is presented in Section II. An overview of the systems model which utilizes LoRa modulation along with LoRa emulation and dataset creation is covered in Section III. Section IV discusses the detection architectures utilized in this paper. Section V discusses the results obtained from the experiments. Finally, Section VI concludes the paper.

II. BACKGROUND AND RELATED WORK

This section summarizes some of the previous research works. This includes those that addressed various ways to mitigate *colored* interference, which occurs due to the overlapping of the transmitted signals in time and frequency. Moreover, it includes works on various machine learning (ML) and DL techniques that have been applied to wireless communication for signal detection.

A. INTERFERENCE AND LORA

The next generation of wireless networks is bringing forth the necessity to support massive connectivity of devices. One such paradigm that requires massive connectivity is IoT. Apart from massive connectivity, this new generation of IoT devices will put a demand on low-power usage and wide coverage. However, massive connectivity promotes co-channel interference, which is a significant challenge that hinders adequate packet reception, particularly in shared radio frequency bands. In these bands, different users and systems are allowed to access the spectrum without prior resource coordination to simplify their operation, ultimately decreasing Spectral efficiency. To allow for efficient transmissions a traditional multiple-access technique could be incorporated. However, the massive connectivity would cause sizable scheduling

delays as well as typically short packet lengths to reduce access latency and receiver complexity, which would make the conventional access method difficult to implement [18]. Thus, we introduce a receiver that can alleviate the co-channel interference effect without the need for coordinating multiple access schemes between different users.

Extensive research work has been conducted on how to efficiently and dynamically utilize the available spectrum by relying on techniques such as cognitive radio (CR). CR relies on being able to sense the spectrum occupancy and devise a spectrum access plan to mitigate co-channel interference [19]. However, the efficiency of CR is subject to accuracy in estimating the occupancy in the spectrum. As such, random access networks severely reduce such capacity. Also, CR requires additional computational resources (and hence power) for spectrum sensing and prediction, which is not ideal for battery-operated devices [20]. Hence, CR has not been adopted in IoT networks due to these limitations.

Another technique to increase the capacity of a system with interference from another signal is to utilize successive interference cancellation (SIC) [21]. In SIC, signals coming from different users are successively decoded, with each decoded signal being subtracted from the total received signal. SIC is also utilized in NOMA methods for next-generation wireless communications. However, the channel gain and time synchronization information of all users should be known by the base station (BS) to achieve an accurate decomposition of the superimposed signals, which entails additional overhead and is often impossible to achieve in shared frequency bands [22].

As one of the widely adopted technologies, LoRa modulation is claimed to be robust to interference, whereby LoRa modulation provides a certain level of orthogonality in different SF transmissions, owing to the reduced cross-energy among the mismatched LoRa transmissions with different SFs [23]. However, collisions of inter-SF signals would in practice still cause packet losses, shown in [7], [23], [24], [25]. This is further aggravated in cases where the colliding signals have identical SFs, thus requiring a higher signal-to-interference plus noise ratio (SINR) to be successfully detected by the receiver [7]. The work in [26] proposes a methodology to allocate different SFs to users to mitigate this effect. However, while the proposed method improves the throughput, interference still increases the error probability due to the non-perfect orthogonality. Recent research has also shown how coherent detection methods enhance the performance in same-SF interference scenarios, demonstrated by theoretical approximations and Monte Carlo simulations as shown in [25]. However, the work focuses on same-SF interference and does not explore inter-SF scenarios. Additional work on investigating the theoretical BER performance of traditional LoRa receivers under same-SF interference is explored in [8], whereas [27] extends the work on same-SF interference to include chirp misalignment.

B. MACHINE AND DEEP LEARNING IN TELECOMMUNICATION

ML methods, with a focus on DL, have been recently investigated in the field of wireless communications, including works on automatic modulation recognition (AMR) [19], [28], occupancy detection in the license-free band [29], [30], and optimization of interference management algorithms [31]. Many more applications of ML in communications exist, which are reviewed in [16] for the physical layer of communications, and in [32] for higher layers. ML has also proven to be beneficial in the IoT application layer, enabling dynamic adaption to a change in user traffic conditions and providing advantages in terms of energy efficiency and resource control, as outlined in a survey of recent advances in IoT application layer protocols and ML research directions [33]. Our previous work in [10] explored the use of DL-based signal detection techniques with CNN models for the LoRa modulation scheme under AWGN, synchronization time offset, and carrier frequency offset impairments. The proposed neural network outperformed the conventionally used noncoherent detector. Another research work that utilizes DL sequence detector networks can be found in [12], constructing a sliding bidirectional recurrent neural network (SBRNN). The SBRNN can take in information from previous symbols, unlike symbol-by-symbol detectors, to combat inter-symbol interference (ISI) for non-LoRa signals. Another advantage of some DL-based detectors is that they do not need channel state information (CSI) which is further demonstrated in [13], which proposes a detection method based on the generative adversarial network (GAN).

Conventional DL-based detectors in past research literature rely on a data-driven approach, where the DL network learns from a large labeled dataset and mathematical models are not necessarily utilized to aid in training. However, DL-based detectors can be further optimized by leveraging domain knowledge for a model-based approach to significantly reduce the network complexity as well as the dataset size required for training [34]. An example of a model-based approach for signal detection in OFDM systems is presented in [35], where the research work achieves high performance relative to traditional methods and the model-driven DL approach converges quicker requiring fewer parameters compared to the data-driven DL signal detection receiver shown in [36]. Another work presenting model-based DL signal detection is shown in [14], which uses DL to learn the log-likelihoods and performs Viterbi signal detection. In addition to solid detection performance, the research work also proposes a method to train in real-time and adapt to time-varying channels. Essentially, leveraging some form of domain knowledge in the selection of the architecture by utilizing statistical models, preparing the input data, employing known mathematical processes, or in the selection of the DL architecture can increase the overall performance, reduce the computational complexity, lessen the dependency of a large

dataset, and speed up convergence in the training of a DL detector [37].

To combat non-Gaussian interference, ML can be used to realize efficient frequency band selection techniques for IoT devices [38]. ML for signal detection is also shown to effectively deal with non-Gaussian interference, such as in [36] for the OFDM scheme, demonstrating robustness to non-linear channel impairments, interference, and comparable BER performance to traditional detection methods. However, conventional detection methods outperform the DL methods in noise-limited scenarios. The HybNet architecture, on the other hand, leverages optimal detection in AWGN conditions with conventional detection, while also performing well in interference-limited conditions by using DL detection techniques. For detecting multiple-input and multiple-output (MIMO) signals with correlated interference, authors in [39] show a DL-based approach that uses a maximum likelihood detector first to detect the signal and then a DL network to predict and remove local correlation among the noise in different symbols. Extending the HybNet framework to MIMO signals would require a different architecture to accommodate a MIMO system. Though, DL is a promising tool for the suppression of correlated interference. Nonetheless, achieving efficient detection performance on signals impaired with co-channel interference with DL-detection is an open research question.

III. IOT SIGNAL MODEL

This section explains the general structure of LoRa modulation utilized to generate the synthetic data for training the DL neural network. It also details traditional LoRa detection methods utilized as a comparison benchmark with the proposed DL method.

A. LORA MODULATION

LoRa CSS modulation is based on linear cyclic chirping within a given bandwidth, denoted as B . Each chirp encodes a single symbol with a given duration, denoted as T_s . The bandwidth according to the LoRaWAN protocol can take one of the following values as $B \in \{125, 250, 500\}$ kHz, whereas the chirp rate of the symbols is controlled by the corresponding SF, where the SF $\in \{7, 8, 9, 10, 11, 12\}$ per the LoRaWAN protocol. Based on the SF the number of available symbols is given by $M = 2^{\text{SF}}$. We follow the same notations as in [6] and [10] to describe LoRa signal modulation where a single symbol (chirp) is represented as follows,

$$s_k(t) = \exp\left(j2\pi \int_0^t \left[(\beta x + \zeta_k)_{\text{mod } B} - \frac{B}{2}\right] dx\right), \quad (1)$$

where ζ_k is the shift frequency representing the symbol value as follows,

$$\zeta_k = m_k \delta_f, \quad (2)$$

where m_k is the data symbol value $m_k \in \{0, 1, \dots, M-1\}$, k is the symbol's index, and δ_f represents the frequency step between the shifts. The frequency step in LoRa is designed

to be equal to the symbol rate itself, i.e., $\delta_f = B/M$. On the other hand, β represents the chirp rate given by,

$$\beta = \frac{f_{\text{high}} - f_{\text{low}}}{T_s} = \frac{B}{T_s}, \quad (3)$$

where $f_{\text{low}} = f_c - \frac{B}{2}$ and $f_{\text{high}} = f_c + \frac{B}{2}$ are respectively the lower and upper-frequency bounds of the chirp around the carrier f_c . Accordingly, the sequence of symbols (message) can be expressed as follows,

$$x(t) = \sum_{k=1}^K s_k(t - kT_s), \quad (4)$$

where K is the total number of symbols contained within a message.

B. CONVENTIONAL LORA DETECTION

The detection of LoRa symbols is typically performed in two steps; (i) the symbol is dechirped with the same chirping rate to convert the received symbol into a single tone, accordingly, the sequence of symbols will manifest as a multiple frequency shift keying (MFSK) modulation signal, (ii) then in the second step the symbols are detected using a conventional frequency detector (matched filtering can be used find the peak given by the frequency tone).

To dechirp the signal, each symbol in the received LoRa waveform is multiplied by a synchronized and inverted chirp with zero frequency shift. The dechirping pulse train is represented as follows,

$$s^*(t) = \sum_{k=1}^K \exp\left(j\pi B(t - kT_s) - j2\pi\beta(t - kT_s)^2\right). \quad (5)$$

In an ideal channel, each dechirped LoRa symbol can then be expressed as follows,

$$y(t) = \sum_{k=1}^K \exp(j2\pi \zeta_k(t - kT_s)), \quad (6)$$

which is a sequence of single tones, each tone with a frequency offset ζ_k corresponding to the value of the symbol. Note that this is now a typical MFSK signal. Accordingly, we can use conventional MFSK matched filtering detection methods to detect the symbols. Two main approaches are typically utilized; (i) noncoherent detection when phase information is not available, and (ii) coherent detection when perfect phase information is known.

1) COHERENT DETECTION

Optimal detection of a LoRa symbol in AWGN environments can be achieved using coherent detection. Coherent detection is executed by correlating the dechirped signal with all permissible frequency shifts, then the estimated symbol value is based on the frequency shift that produces the largest real-valued output. Matched filtering of an MFSK LoRa symbol is shown as follows [9],

$$\hat{m}_{\text{coh}} = \underset{k}{\operatorname{argmax}} \Re \left[\int_0^\infty y(t) w_k(t - \tau) d\tau \right], \quad (7)$$

where $w_k(t) = \exp(j2\pi k\delta_f t)$ represents a set of harmonics and k is an integer representing the number of discrete frequency steps $k = \{0, 1, \dots, M - 1\}$. Coherent detection can be efficiently implemented with the FFT [40], as follows,

$$\hat{m}_{\text{coh}} = \left\lfloor \frac{1}{\delta_f} \operatorname{argmax}_f \Re\{Y(f)\} \right\rfloor, \quad (8)$$

where [9] provides an analytical expression for BER assuming a coherent detection and proves its optimality in an AWGN channel. However, this requires, the receiver to be phase-locked to the transmitter, by using a phase-locked loop (PLL) [40]. An alternative method is to use noncoherent detection to perform signal detection without phase synchronization.

2) NONCOHERENT DETECTION

Noncoherent detection follows the methodology of coherent detection. Whereby each LoRa symbol is dechirped and then correlated with permissible frequency shifts. However, the magnitude of the output is taken, which is shown as follows [9],

$$\hat{m}_{\text{ncoh}} = \operatorname{argmax}_k \left| \int_0^\infty y(t)w_k(t - \tau)d\tau \right|. \quad (9)$$

Equivalently, the FFT can be applied as follows [40],

$$\hat{m}_{\text{ncoh}} = \left\lfloor \frac{1}{\delta_f} \operatorname{argmax}_f |Y(f)| \right\rfloor, \quad (10)$$

where $Y(f) = \text{FFT}\{y(t)\}$ denotes the FFT of the received FSK signal and $\lfloor \cdot \rfloor$ represents the rounding function. This method achieves the same performance as in 9, but at less processing complexity. However, when compared to coherent detection, noncoherent methods come at the cost of higher BER in AWGN environments.

C. DATASET CREATION

To create a synthetic dataset of LoRa symbols, we use the open-source MATLAB emulator [6], [41] which was previously developed by our team. The target and interferer LoRa signals are generated as I/Q waveforms, with a $f_s = 1$ MHz, a SF = 7, and a $B = 125$ kHz. A randomly generated LoRa message consisting of the symbol sequence vector $M = \{m_1, m_2, \dots, m_n\}$ is utilized to create each waveform. For injecting interference to the original LoRa signal with a random phase shift, another LoRa signal, representing the interference, is added to the original signal which also has a random phase shift and delay. Finally, the resultant signal is resampled to Nyquist rate of $f_s = 125$ kHz. Superimposing the target signal with interference while the two signals are sampled at a higher sampling rate allows for a more realistic signal due to the higher resolution. For the training and detection process, a sequence is picked up arbitrarily to represent the *target* signal and is assigned a controlled power of p_s . In addition to the interference, the received target signal is further impaired with a Gaussian noise process having

a controlled power. We denote the controlled interference power as p_I and the noise power as σ^2 . The INR is denoted as α , and the SINR is denoted as γ ,

$$\alpha = \frac{p_I}{\sigma^2}, \quad \gamma = \frac{p_s}{p_I + \sigma^2}. \quad (11)$$

Since the performance is only related to the INR and SINR, the dataset consists of signals with a uniformly distributed SINR between $\{-15, 15\}$ dB and a uniformly distributed INR between $\{-30, 30\}$ dB. Additionally, we normalize the interference and noise power concerning the target LoRa signal power, i.e., $p_s = 1$. Accordingly, the stored emulated received waveform is comprised of three parts; (i) the target LoRa signal, (ii) the interference, (iii) the AWGN noise, as follows,

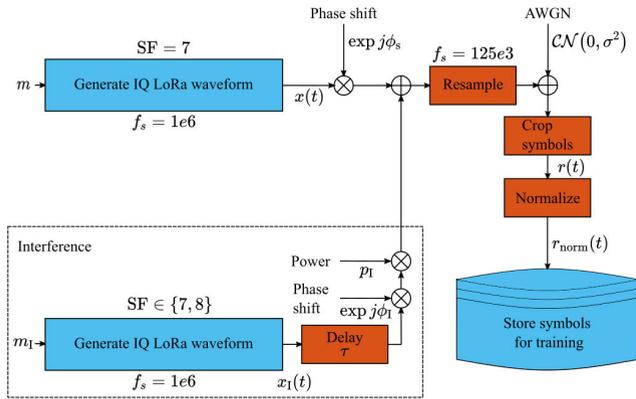
$$r(t) = \underbrace{x(t) \exp(j\phi_s)}_{\text{Signal}} + \underbrace{\sqrt{\frac{\alpha}{\gamma + \alpha\gamma}} x_I(t - \tau) \exp(j\phi_I)}_{\text{Interference}} + \underbrace{\sqrt{\frac{1}{\gamma + \alpha\gamma}} n(t)}_{\text{AWGN}}, \quad (12)$$

where $x(t)$ is the target LoRa baseband signal, $x_I(t)$ is the interfering baseband LoRa signal with a random time shift of τ , and $n(t) \sim \mathcal{CN}(0, \sigma^2)$ is the complex zero-mean AWGN. Both the target signal and the interfering signal have a uniformly distributed random phase shift between 0° and 360° denoted by ϕ_s and ϕ_I respectively. Then, to normalize the signal, the overall signal is divided by the target signal power as follows,

$$r_{\text{norm}}(t) = \frac{r(t)}{\sqrt{p_s}}. \quad (13)$$

Note that the reasoning behind normalizing the signal power to $p_s = 1$ is to keep the magnitude of the target signal power consistent while training the DL receiver. This is because normalization can give a clue to the DL receiver as to which signal is the interfering signal and which is the target signal. We assume every signal has an uncorrupted preamble which can be used to estimate the signal power.

The complex baseband signal is cropped into LoRa symbols, where each target symbol is synchronized in time, as such this implicitly assumes that a receiver is capable of performing preamble detection and synchronization with the target signal. Each symbol in the dataset is labeled according to the transmitted symbols. Since increasing the SF leads to more permissible values in a symbol (e.g. SF 7 means 128 different variations could be represented in a symbol, and SF 8 means 256, etc.), the training dataset size will increase due to the number of classes increasing according to the number of possible symbol values that a symbol can encode, which is equal to 2^{SF} . Consequently, the training time would be much longer, and the network complexity would also need to be increased. An illustration of the creation process of the dataset is depicted in Fig 1.


FIGURE 1. Dataset generation process for training each CNN.
TABLE 1. Notations and symbols

Parameter	Symbol	Value
Spreading factor	SF	7
Bandwidth	B	125 kHz
Sampling rate	f_s	125 kHz
Samples per symbol	N	128
Frequency offset	ζ	-
Chirp slope	β	-
Symbol time	T_s	1.024 ms
Noise power spectral density	N_o	-
Symbol value	m	-
Discrete frequency step	δ_f	976.56 Hz
Average noise power	σ^2	-
LoRa symbol	$s_k(t)$	-
Dechirping LoRa signal	$s^*(t)$	-
Dechirped LoRa signal	$y(t)$	-
FSK signal FFT	$Y(f)$	-
Noncoherent FSK symbol	$\hat{m}_{n\text{-coh}}$	-
Coherent FSK symbol	\hat{m}_{coh}	-
Interference to noise ratio (INR)	α	-
Signal to interference plus noise ratio (SINR)	γ	-
Target signal power	p_s	1
Interference signal power	p_I	-
Number of training frames	-	400,000
Number of validation frames	-	100,000
STFT window type	-	Hamming
Number of DFT points	N_{DFT}	128
Points per FFT window for the STFT	W	16
STFT window overlap	L	15
I/Q modality network input	R	-
Time-frequency modality network input	X	-
Spectrum modality network input	F	-

IV. DEEP LEARNING DETECTION

Following typical CNN architectures, the main elements that commonly makeup CNNs are as follows: (i) convolutional layers interlaced with (ii) max-pooling layers and (iii) batch

normalization layers. The convolutional layer performs the convolution operation on the input with a kernel (filter) with a certain dimension. The window of the kernel slides across the input data with the unity stride. The output is then multiplied by an activation function, where the Rectified Linear Unit (ReLU) is utilized. Note that the ReLU function is linear for positive inputs and gives zero for negative inputs. Typically, each convolutional layer is followed by a batch normalization layer, which normalizes the mean and variances of the convolutional layer output, which speeds up training. A max-pooling layer then follows, which down-samples the data by taking the maximum value in each max-pooling filter. The output is then passed into a fully connected layer that uses a Softmax activation function that outputs an l -length vector of scores summing to 1, where l is the number of classes. Finally, a classification layer assigns the classes according to the probabilities.

A. DATA MODALITIES

In this section, we explain the three different data modalities that could be used for detecting LoRa signals:

1) I/Q MODALITY

The first model is constructed with time-domain I/Q samples. The model takes in the complex input signal represented as the real part,

$$R_I = \Re\{r(1), r(2), \dots, r(N)\}, \quad (14)$$

and the imaginary part,

$$R_Q = \Im\{r(1), r(2), \dots, r(N)\}, \quad (15)$$

where for each time-domain symbol $r(\cdot)$ there are N temporal samples. The samples are then arranged into two 1D vectors to be used in the DL network as follows,

$$R = \begin{bmatrix} R_I \\ R_Q \end{bmatrix}, \quad (16)$$

where $R \in \mathbb{R}^{2 \times N}$.

2) TIME-FREQUENCY DOMAIN MODALITY

In the second modality, we convert the time domain samples into a spectrogram using the squared magnitude of the short-time Fourier transform (STFT). STFT works by taking segments of the time domain signal and converting each one using the FFT. After that, the FFT vectors are combined in a 2D matrix representing the spectral change across time. In such representation, a linear chirp, for example, will appear as a straight line. The calculation of a spectrogram matrix is achieved by performing the STFT and then taking the squared magnitude as follows,

$$X(\omega, p) = \left| \sum_{n=0}^{N_{\text{DFT}}-1} r(n)g(n-pO)e^{-j\omega n} \right|^2, \quad (17)$$

where $X(\omega, p) \in \mathbb{R}^{N_{\text{DFT}} \times \frac{N-L}{W-L}}$. $r(\cdot)$ denotes the captured sample of a cropped discrete-time LoRa baseband symbol.

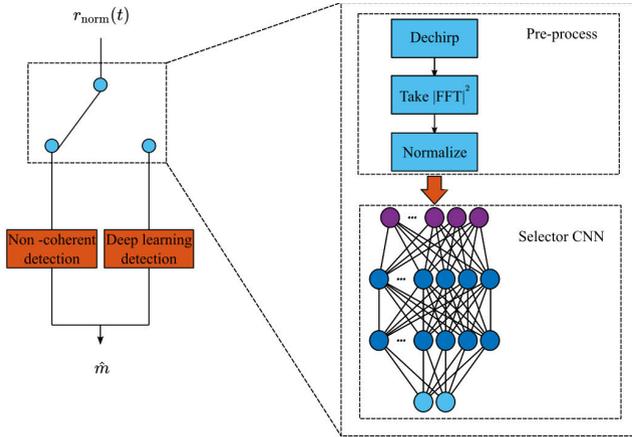


FIGURE 2. Illustration of the proposed HybNet architecture switching between a deep-learning branch and a matched filter-branch based on the interference-to-noise ratio.

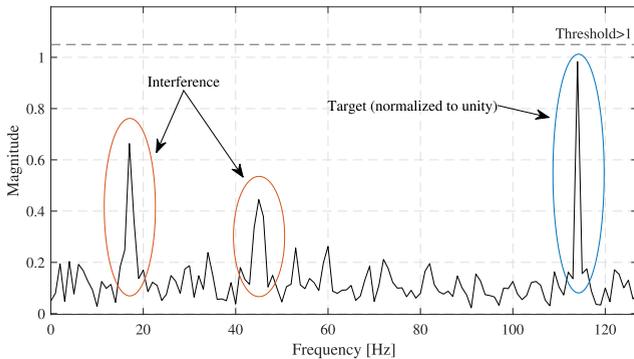


FIGURE 3. Plot of an absolute valued dechirped LoRa symbol with added interference from another LoRa signal with the same SF. The annotations show how the energy detector switching works.

Every p th column of the spectrogram matrix has $\frac{N-L}{W-L}$ values. A Hamming windowing function is denoted by $g(\cdot)$ with an overlap length between each Discrete Fourier Transform (DFT) of $O = W - L$. The parameters chosen for the STFT are outlined in Table 1.

3) FREQUENCY DOMAIN MODALITY

The third DL approach uses the frequency domain, where only the absolute of the FFT of a dechirped LoRa symbol is used. The input signal can be represented as follows,

$$F(\omega) = \left| \sum_{n=0}^{N-1} r(n)e^{-j\omega n/N} \right|, \quad (18)$$

where $r(\cdot)$ is the dechirped LoRa symbol and N is the number of temporal samples. This modality exploits domain knowledge by pre-processing the input data to simplify the detection process.

B. HYBRID ARCHITECTURE: HybNet

The proposed HybNet architecture is the core contribution of this paper, which in essence is a classifier that switches between two different detection branches; (i) a DL network (in our case we select the FFT-CNN as referred to in Fig. 6),

due to the performance characteristics which are shown in Section V, and (ii) a traditional matched detector (in our case we utilize noncoherent detection which is a type of matched filter detection). This architecture is developed to marry the advantages of both detectors, i.e. the optimal performance of matched filters in AWGN and the improved performance of CNN in co-channel interference conditions. In order to decide which detector branch to choose, a supervisory selector system is utilized to decide whether to pass the signal to the first branch or the second.

Note that this selector is trained with another CNN network, which we dub as the *Selector CNN* which utilizes the absolute valued dechirped FFT of a LoRa symbol as the input. A dataset consisting of 200,000 symbols with a uniformly distributed SINR ranging between $\{-15, 15\}$ dB and a uniformly distributed INR between $\{-30, 30\}$ dB is created to be trained on. Each LoRa symbol is labeled with either one of two possible classes; (i) *Traditional* or (ii) *Deep learning*. If the received symbol is identified to lack LoRa interference or the gain of the interfering LoRa signal is less than that of the target, the symbol is labeled as “Traditional”. Consequently, the received symbol is passed to be purely detected by the noncoherent detector (outlined in subsection IV-B.) On the other hand, if the gain of the interfering transmissions is larger than the gain of the target, the symbol is labeled as “Deep learning”. In that case, the received symbol is passed onto the DL detection branch. An illustration of the utilized switching architecture is shown in Fig. 2.

Instead of the Selector CNN, simple threshold switching can be used, whereby the simple switching method based on the energy level is shown in Fig. 3. If the peak magnitude of the FFT exceeds the threshold, then the signal is passed to be detected by the DL detection branch. On the other hand, if the peak magnitude of the FFT falls below the threshold, noncoherent detection is used.

To evaluate the efficacy of a simple threshold energy detector instead of a DL network for the switching logic, different threshold values are accessed and plotted against the switching accuracy, which is shown in Fig. 4. The figure shows that the optimum accuracy of the energy detector occurs at a threshold of magnitude 1.089. Note that the resulting FFT is normalized by $Y = \text{FFT}(y)/\text{length}(y)$. The performance of the energy detector is also compared to the Selector CNN switching network at varying INR levels in Fig. 5. For this test, a dataset consisting of 20,000 symbols with a uniformly distributed SINR between $\{-15, 15\}$ dB per INR point is used. From the figure, the Selector CNN consistently classifies whether noncoherent or DL detection should be used to reduce the error rate from the detection process. Also from the figure, the accuracy of the Selector CNN slightly dips at an INR between -15 dB and -5 dB since the magnitude of the interfering and target LoRa symbol in the dechirped FFT are similar, thus it becomes more difficult for the Selector CNN to correctly discern which is the interfering signal and which is the target. Finally, threshold switching performs worse than the Selector CNN at all INR values.

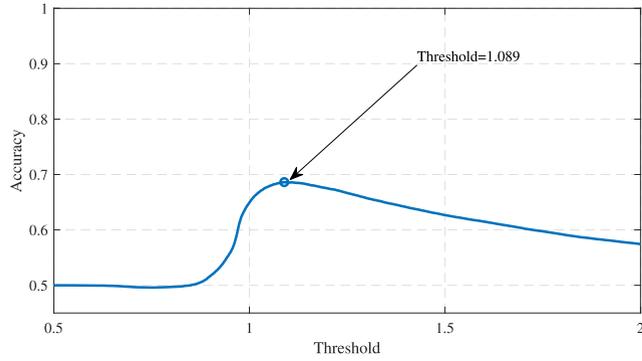


FIGURE 4. The plot of the overall accuracy for the energy detector against a varying threshold value.

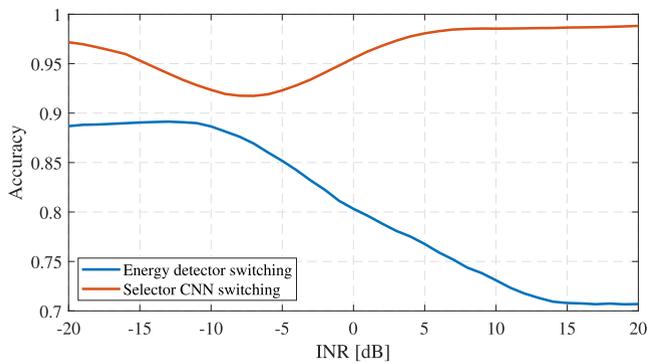


FIGURE 5. The plot of the overall accuracy against a varying INR value to compare the performance of a threshold energy detector and the Selector CNN.

C. CNN SETUP DESCRIPTION

Three different CNN networks were designed to cater to the three data modalities explained in Section IV-A, and this is because both input dimensions and contents are different. For each of the three networks, we utilize a Bayesian optimizer to tune the network’s hyperparameters. Bayesian optimization is a more efficient method for selecting hyperparameters compared to search methods such as brute-force, grid search, and random search [42]. The hyperparameters that were chosen for optimization are; (i) the number of convolutional layers, (ii) the convolutional filter size, (iii) the initial learning rate, and (iv) the dropout rate. The resulting network architectures are summarized in Table 2. The summary of the CNN architectures is also illustrated in Fig 6, showing the input format to each network.

Considering the simple classification task, the network architectures are shallow. The Selector network, on the other hand, is manually optimized for the hard switching architecture since high accuracy is achieved with a straightforward network. A simple two-layer CNN is used with a convolutional layer width of 16 using the ReLU activation function. All the networks were trained with stochastic gradient descent with momentum (SGDM) over 60 epochs. The training parameters for all the networks discussed in this paper are

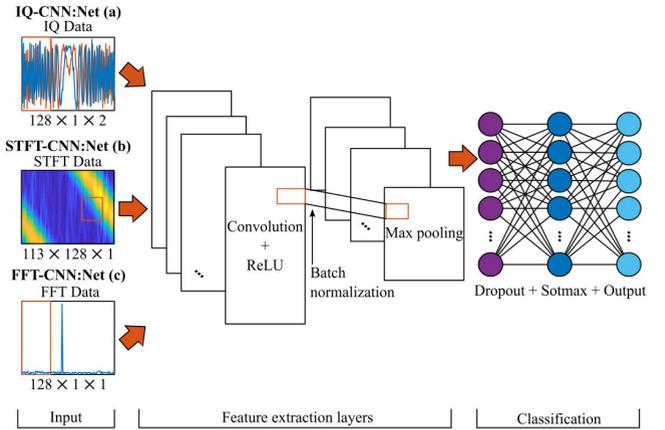


FIGURE 6. Illustration of all three DL-based detector models used in this paper.

outlined in Table 3. Also note that the re-training frequency of a DL network depends on the environment in which the DL receiver is deployed. If the behavior of the wireless channel is not represented in the training dataset, then the DL receiver requires transfer learning (incremental training). Incremental learning further trains the existing network with new data and updates the weights. Incremental learning does not need to train a network from the beginning.

V. SIMULATION RESULTS AND DISCUSSION

This section presents the performance benchmark of the different architectures based on the BER indicator. A Monte-Carlo simulation of both *same-SF* interference (SF7 on SF7) and *inter-SF* interference (SF8 on SF7) is performed. The performance is investigated for a variable level of INR and we also investigate a static INR with varying levels of SINR. A lower INR value indicates a noise-dominant scenario (noise-limited performance), while a higher value indicates an interference-dominant scenario (interference-limited performance). Note that this work investigates the detection performance for signals with a normalized signal power $p_s = 1$ unit power. We assume that the receiver is able to estimate the received power of the target signal based on an uncorrupted preamble. This is because the focus of this work is on the retrieval of the corrupted payload which has an order of magnitude longer duration than the preamble.

A. DETECTION PERFORMANCE

The BER performance for a *same-SF* LoRa interferer, with an SF7 in Fig. 7. Furthermore, the performance of the DL-based techniques is depicted in Fig. 8 showing the performance in an *inter-SF* interference scenario. The figures both show the BER of the noncoherent detector increases as the INR increases because the interfering signal masks the target signal, thus noncoherent detection is unable to correctly detect the symbol. Additionally, the FFT-CNN has the best performance under noise-limited scenarios. However, at high INR values in an inter-SF scenario, both the IQ-CNN and

TABLE 2. CNN network summary

Layer	Shape			Parameter
	IQ – CNN	STFT – CNN	FFT – CNN	
	Net(a)	Net(b)	Net(c)	
Input	128 × 1 × 2	113 × 128 × 1	128 × 1 × 1	-
Convolutional 1	57, 19 × 1	57, 7 × 7	57, 11 × 1	ReLU
Batch Normalization 1	-	-	-	-
Max Pooling 1	2 × 1	2 × 1	2 × 1	-
Convolutional 2	57, 19 × 1	57, 7 × 7	57, 11 × 1	ReLU
Batch Normalization 2	-	-	-	-
Max Pooling 2	2 × 1	2 × 1	2 × 1	-
Convolutional 3	57, 19 × 1	57, 7 × 7	57, 11 × 1	ReLU
Batch Normalization 3	-	-	-	-
Max Pooling 3	2 × 1	2 × 1	2 × 1	-
Convolutional 4	57, 19 × 1	57, 7 × 7	57, 11 × 1	ReLU
Batch Normalization 4	-	-	-	-
Max pooling 4	2 × 1	2 × 1	2 × 1	-
Convolutional 5	57, 19 × 1	57, 7 × 7	57, 11 × 1	ReLU
Batch Normalization 5	-	-	-	-
Max pooling 5	2 × 1	2 × 1	2 × 1	-
Dropout	0.43	0.37	0.50	-
Fully Connected	128	128	128	Softmax

TABLE 3. Training options

Parameter	Value
Optimizer	SGDM
Momentum	0.9
Epochs	60
Learning rate drop schedule	40
Learning rate drop factor	0.1
Mini-batch size	256
L2 regularization	0.0001

STFT-CNN slightly outperform the FFT-CNN. This is due to the FFT-CNN having no time information, unlike the IQ-CNN and STFT-CNN.

To present the performance improvement of the *HybNet* framework compared to the DL-based detectors, fixed INR plots with varying SINR are used. Traditional noncoherent methods (upper-bound) are plotted alongside the *HybNet* framework and FFT-CNN performance. The performance of *HybNet* is depicted in Fig. 9 and in Fig. 10 for SF7 and SF8 interference respectively, at a SINR at the value of $\gamma = -15$ dB. It can be clearly seen how the proposed *HybNet* architecture can effectively switch between the noncoherent detector and the FFT-CNN branches and thus follows the optimal performance in both noise-limited and interference-limited scenarios. The efficient switching indicates that the interference detector network can accurately select a detector depending on the interference level. Furthermore, the switching performance of the CNN compared to the energy detector

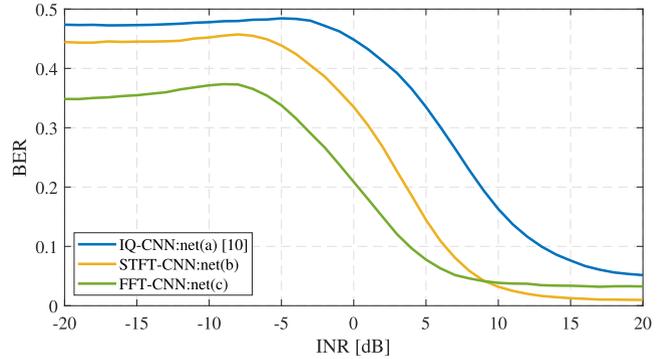


FIGURE 7. Detection performance for a target LoRa symbol with SF7, an interference LoRa signal with SF7, and a fixed SINR = -15 dB.

can be observed in Fig. 11. From the plot, it can be seen that the energy detector switches between the noncoherent and DL pathways at a lower efficacy compared to the CNN. Also to note, the efficacy of the energy threshold switching network further degrades when the fixed SINR is decreased while the CNN switching network continues to be efficient. The performance of the *HybNets* at a fixed SINR of $\gamma = -20$ dB can be observed in Fig. 12.

We further analyze the overall dataset accuracy, training time, and detection time of the three main architectures discussed in this paper, the IQ-CNN, STFT-CNN, and the FFT-CNN, in addition to the performance of *HybNet* and the *HybNet* architecture when the threshold switching is used (discussed in subsection IV-B). We utilize MATLAB to create

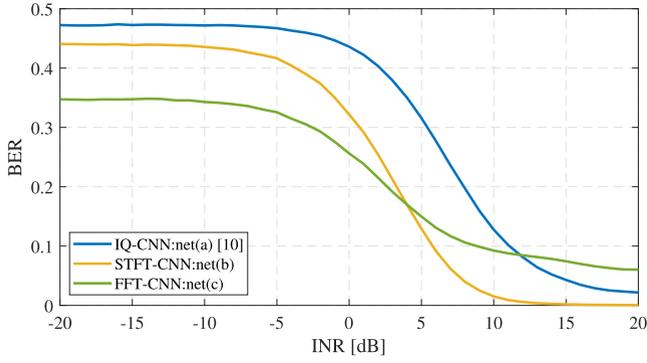


FIGURE 8. Detection performance for a target LoRa symbol with SF7, and an interference LoRa signal with SF8, and a fixed SINR = -15 dB.

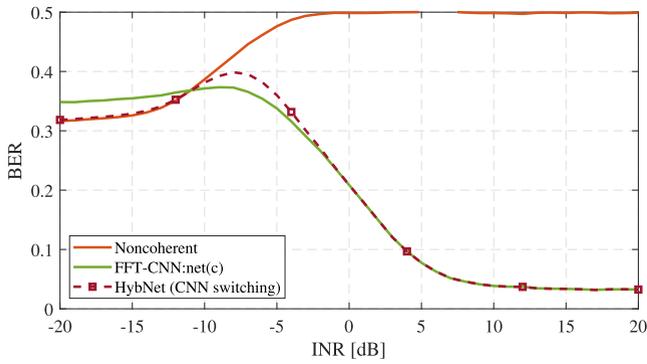


FIGURE 9. Detection performance of the switching architecture for a target LoRa symbol with SF = 7, an interference LoRa signal with SF = 7, and a fixed SINR = -15dB.

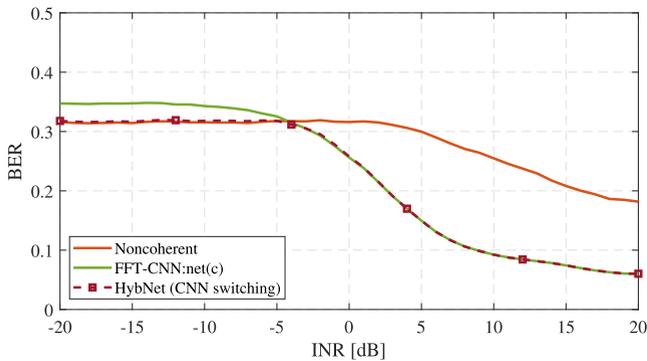


FIGURE 10. Detection performance of the switching architecture for a target LoRa symbol with SF = 7, an interference LoRa signal with SF = 8, and a fixed SINR = -15 dB.

the testing dataset which consists of 200,000 symbols with a uniformly distributed SINR between $\{-15, 15\}$ dB and a uniformly distributed INR between $\{-30, 30\}$ dB. MATLAB is also used for preprocessing and for DL. The system used for the experimentation has a 16-logical core Intel Xeon CPU at 3.2 GHz and an Nvidia Quadro 4000 GPU. Table 4 shows the additional performance metrics of the networks, including the

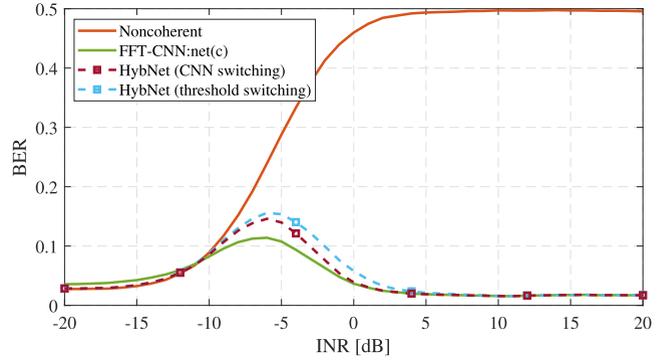


FIGURE 11. Detection performance of both the CNN switching architecture and the energy detection switching architecture for a target LoRa symbol with SF = 7, an interference LoRa signal with SF = 7, and a fixed SINR = -10 dB.

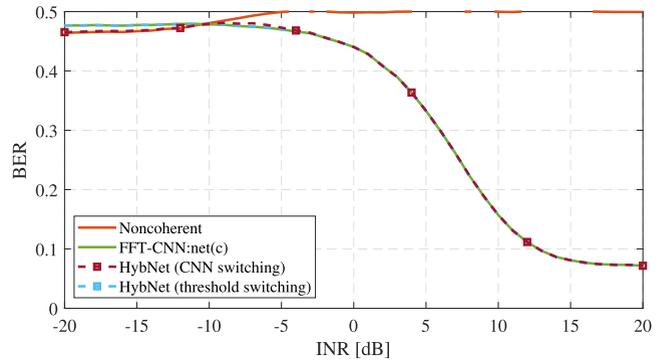


FIGURE 12. Detection performance of both the CNN switching architecture and the energy detection switching architecture for a target LoRa symbol with SF = 7, an interference LoRa signal with SF = 7, and a fixed SINR = -20 dB.

overall dataset accuracy, training time for each architecture, and detection time per symbol.

For the DL detectors, the IQ-CNN requires the least amount of pre-processing compared to the other networks and has a low training time, however, it has the lowest accuracy. On the other hand, the STFT-CNN has higher accuracy, nevertheless, it also exhibits the longest training time and time per symbol. Notwithstanding, the FFT-CNN exhibits the lowest complexity, the highest accuracy, and trains the fastest. The HybNet increases the dataset accuracy, however, the training time is slightly higher as it is the sum of FFT-CNN and the Selector CNN training times. To compromise between high performance and high complexity, the HybNet network can be used with the energy detector as the switching network instead of the Selector CNN. Using the energy detector for the switching mechanism exhibits an accuracy higher than the noncoherent detector and the time per symbol is significantly lower compared to the pure DL detectors. One shortcoming of the HybNet architecture is the Selector CNN does not perfectly switch between noncoherent detection and DL detection which is shown in the transition where the channel switches between noise dominant and interference dominant.

TABLE 4. Performance analysis

Network	Test accuracy	Train time	Time per symbol
Net(a)	82%	16.7 min	55 μ s
Net(b)	88%	111.6 min	312 μ s
Net(c)	89%	17.2 min	47 μ s
HybNet (CNN)	90%	24.5 min	24 μ s
HybNet (threshold)	89%	17.2 min	19 μ s
Noncoherent	75%	-	0.05 μ s

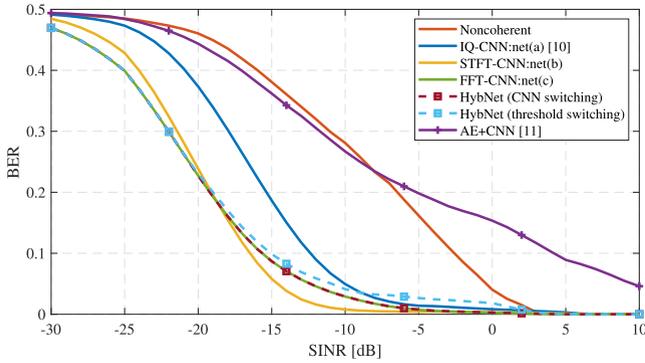


FIGURE 13. Detection performance of the proposed architectures compared with the noncoherent detector. The target LoRa symbol with SF = 7, an interference LoRa signal with a mixture of SF = 7&8, a fixed INR = 10 dB.

To make sure the Hybnet architecture as well as the DL detectors still have desirable BER performance in varying SINR regimes, we plot the BER with a constant INR value and sweep the SINR. The performance with a fixed INR of 10 dB is shown in Fig.13, which shows that the FFT-CNN outperforms the baseline noncoherent detector as the SINR is swept from -30 dB up until the BER reaches 0 at around 3 dB. In addition, the HybNet with Selector CNN switching consistently maintains the optimum detector selection. However, the STFT-CNN outperforms the FFT-CNN at a certain region which could be addressed by adapting the HybNet architecture to have three detection paths to switch between noncoherent, the FFT-CNN, and the STFT-CNN. Nonetheless, this significantly complicates the detector and could be addressed in future research.

We also investigate a noise-limited scenario in Fig.14. This plot further confirms that the HybNet with CNN switching outperforms the HybNet with threshold switching. Additionally, the HybNet with CNN maintains near-optimal performance.

Moreover, the error rate of the Hybnet architecture is shown to outperform the methods developed in the CNN-based detection architecture in [10] and the convolutional denoising autoencoder (AE) with a CNN detector in [11].

B. COMPLEXITY ANALYSIS

Firstly, the proposed deep learning method is supposed to run on the receiver side, which, in the massive IoT case, is mainly on the access point side (i.e. the majority of the

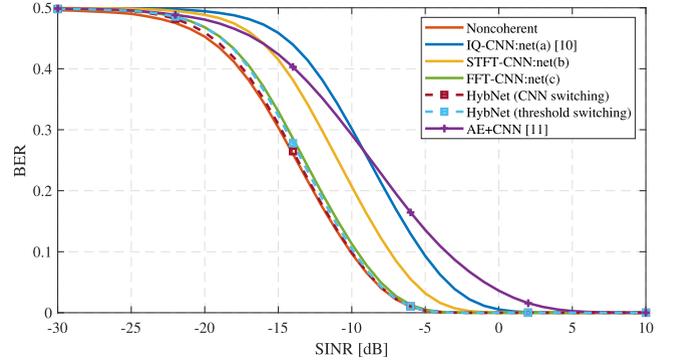


FIGURE 14. Detection performance of the proposed architectures compared to noncoherent detection. The target LoRa symbol has SF = 7, and the interference LoRa signal is a mixture of SF = 7&8. The interference-to-noise ratio is fixed at INR = -20 dB.

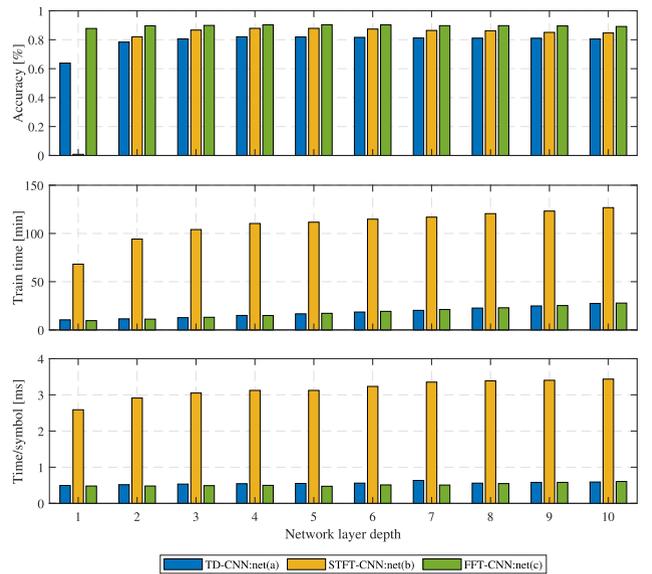


FIGURE 15. Complexity and detection accuracy characteristics as the number of layers increases.

traffic is in the uplink direction). IoT access points are typically well-resourced in terms of computational power and energy budget to uplink IoT end devices. The theoretical time complexity for all the networks can be expressed as $O\left(L \sum_{l=1}^M K_{l-1} F_l W_l K_l\right)$ [43], where l is the index of convolutional layers and M is the number of layers. L is the number of input symbols, and K_{l-1} denotes the number of input channels. F_l is the dimensions of the convolutional filter multiplied together. W_l is the number of filters per convolutional layer. Finally, K_l is the dimensions of the output multiplied together. The *max-pooling*, *dropout*, and *fully connected layers* have insignificant time complexity compared to the convolutional layers, so their complexity is not included in the calculation. From the theoretical expression, the time complexity linearly increases with several input symbols for all three networks. The theoretical expression confirms that

the time complexity is much higher for the STFT-CNN since the convolutional filters F_l as well as the output dimensions K_l are 2-dimensional.

We further investigate the performance of the proposed architecture as the number of convolutional layers in the CNNs is changed, the results are shown in Fig. 15. The plot confirms the complexity in terms of training time and detection time per symbol is much higher in the case of the STFT-CNN compared to the IQ-CNN and FFT-CNN. In terms of accuracy, as the number of layer increase, so does the accuracy up until the optimized number of layers is reached, which in our case is five. The optimized performance characteristics are shown in Table 4.

VI. CONCLUSION

This paper investigated deep learning approaches that rely on convolutional neural networks for the detection of LoRa symbols in the presence of AWGN and colored interference. This paper presented a new framework HybNet that switches between traditional matched filter detection and deep learning detection. As such the proposed architecture combines the merits of (i) the optimal detection in Gaussian noise based on the matched filter, with (ii) the improved performance of the deep learning detector under non-Gaussian interference. The paper tested different input data modalities for deep learning, namely; (i) I/Q-based, (ii) time-frequency-based, and (iii) spectrum-based. The spectrum-based deep detector showed the best detection performance in heavy interference conditions and the lowest time complexity compared to the I/Q-based and time-frequency-based networks. This improvement in the performance suggests that the proposed hybrid architecture would outperform stand-alone conventional detection methods in a broader range of applications, especially in random access IoT networks where the interference is caused by overlapping co-channel transmissions. Possible future work could include evaluating the performance of the HybNet architecture under multi-path channel conditions.

REFERENCES

- [1] *Ericsson Mobility Report*, Ericsson, Stockholm, Sweden, Jun. 2021.
- [2] B. Al Homssi et al., "Free spectrum for IoT: How much can it take?" in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.
- [3] LoRa Alliance. (2018). *Specification, LoRaWAN*. LoRaWAN. Accessed: Oct. 5, 2020. [Online]. Available: <https://lora-alliance.org/sites/default/files/2018-07/lorawan1.0.3.pdf>
- [4] B. Vejlgård, M. Lauridsen, H. Nguyen, I. Z. Kovacs, P. Mogensen, and M. Sorensen, "Interference impact on coverage and capacity for low power wide area IoT networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.
- [5] B. Manzoor, A. Al-Hourani, B. A. Homssi, K. Magowe, S. Kandeepan, and K. G. Chavez, "Evaluating coverage performance of NB-IoT in the ISM-band," in *Proc. 27th Int. Conf. Telecommun. (ICT)*, Oct. 2020, pp. 1–5.
- [6] B. Al Homssi, K. Dakic, S. Maselli, H. Wolf, S. Kandeepan, and A. Al-Hourani, "IoT network design using open-source LoRa coverage emulator," *IEEE Access*, vol. 9, pp. 53636–53646, 2021.
- [7] A. Mahmood, E. Sisinni, L. Guntupalli, R. Rondon, S. A. Hassan, and M. Gidlund, "Scalability analysis of a LoRa network under imperfect orthogonality," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1425–1436, Mar. 2019.
- [8] T. Elshabrawy and J. Robert, "Analysis of BER and coverage performance of LoRa modulation under same spreading factor interference," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1–6.
- [9] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*, vol. 95. Hoboken, NJ, USA: Wiley, 2005.
- [10] K. Dakic, B. Al Homssi, A. Al-Hourani, and M. Lech, "LoRa signal demodulation using deep learning, a time-domain approach," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–6.
- [11] A. A. Tesfay, S. Kharbech, E. P. Simon, and L. Clavier, "Signal denoising and detection for uplink in LoRa networks based on Bayesian-optimized deep neural networks," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 214–218, Jan. 2023.
- [12] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5663–5678, Nov. 2018.
- [13] L. Sun, Y. Wang, A. L. Swindlehurst, and X. Tang, "Generative-adversarial-network enabled signal detection for communication systems with unknown channel models," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 47–60, Jan. 2021.
- [14] N. Shlezinger, N. Farsad, Y. C. Eldar, and A. J. Goldsmith, "ViterbiNet: A deep learning based Viterbi algorithm for symbol detection," 2019, *arXiv:1905.10750*.
- [15] B. Al Homssi et al., "Artificial intelligence techniques for next-generation mega satellite networks," 2022, *arXiv:2207.00414*.
- [16] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
- [17] J. Gu et al., "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.
- [18] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.
- [19] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.
- [20] H. N. Abdullah and H. S. Abed, "Improvement of energy consumption in cognitive radio by reducing the number of sensed samples," in *Proc. Al-Sadeq Int. Conf. Multidisciplinary IT Commun. Sci. Appl. (AIC-MITCSA)*, May 2016, pp. 1–6.
- [21] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
- [22] M. Aldababsa, M. Toka, S. Gökçeli, G. K. Kurt, and O. Kucur, "A tutorial on nonorthogonal multiple access for 5G and beyond," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–24, Jun. 2018.
- [23] D. Croce, M. Gucciardo, S. Mangione, G. Santaromita, and I. Tinirello, "Impact of LoRa imperfect orthogonality: Analysis of link-level performance," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 796–799, Apr. 2018.
- [24] A. Waret, M. Kaneko, A. Guitton, and N. El Rachkidy, "LoRa throughput analysis with imperfect spreading factor orthogonality," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 408–411, Apr. 2019.
- [25] O. Afisiadis, S. Li, A. Burg, and A. Balatsoukas-Stimming, "On the advantage of coherent LoRa detection in the presence of interference," 2020, *arXiv:2010.00507*.
- [26] L. Amichi, M. Kaneko, E. H. Fukuda, N. El Rachkidy, and A. Guitton, "Joint allocation strategies of power and spreading factors with imperfect orthogonality in LoRa networks," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3750–3765, Jun. 2020.
- [27] O. Afisiadis, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "On the error rate of the LoRa modulation with interference," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1292–1304, Feb. 2020.
- [28] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [29] B. A. Homssi, A. Al-Hourani, Z. Krusevac, and W. S. T. Rowe, "Machine learning framework for sensing and modeling interference in IoT frequency bands," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4461–4471, Mar. 2021.
- [30] T. Gale, T. Solc, R.-A. Mosoi, M. Mohorcic, and C. Fortuna, "Automatic detection of wireless transmissions," *IEEE Access*, vol. 8, pp. 24370–24384, 2020.

[31] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.

[32] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.

[33] P. K. Donta, S. N. Srirama, T. Amgoth, and C. S. R. Annavarapu, "Survey on recent advances in IoT application layer protocols and machine learning scope for research directions," *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 727–744, Oct. 2022.

[34] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 77–83, Oct. 2019.

[35] X. Gao, S. Jin, C.-K. Wen, and G. Y. Li, "ComNet: Combination of deep learning and expert knowledge in OFDM receivers," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2627–2630, Dec. 2018.

[36] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[37] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning: Key approaches and design guidelines," in *Proc. IEEE Data Sci. Learn. Workshop (DSLW)*, Jun. 2021, pp. 1–6.

[38] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.

[39] J. Xia, K. He, W. Xu, S. Zhang, L. Fan, and G. K. Karagiannidis, "A MIMO detector with deep learning in the presence of correlated interference," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4492–4497, Apr. 2020.

[40] A. Marquet, N. Montavont, and G. Z. Papadopoulos, "Investigating theoretical performance and demodulation techniques for LoRa," in *Proc. IEEE 20th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2019, pp. 1–6.

[41] B. Al Homssi and A. Al-Hourani. (2020). *LoRa MATLAB Emulator*. Data Retrieved From AURIN. [Online]. Available: <https://au.mathworks.com/matlabcentral/fileexchange/81166-loramatlab>

[42] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 2951–2959.

[43] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5353–5360.



KOSTA DAKIC (Graduate Student Member, IEEE) received the B.Eng. degree in electrical and electronic engineering from RMIT University, Melbourne, Australia, in 2019, where he is currently pursuing the Ph.D. degree. His research interests include the IoT satellite communications, deep learning, and spiking neural networks for wireless communications.



BASSEL AL HOMSSI (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the American University of Sharjah, United Arab Emirates, in 2014 and 2016, respectively, the M.Eng. degree in telecommunications from The University of Melbourne, Australia, in 2017, and the Ph.D. degree from the School of Engineering, RMIT University, Melbourne, Australia, in 2020. He is currently a Lecturer of intelligent space systems with UNSW Canberra Space.

His research interests include the application of artificial intelligence in space, in particular edge-AI, distributed-AI, and neuromorphic computing for miniaturized satellite constellations.



MARGARET LECH (Member, IEEE) received the M.S. degree in physics from Maria Curie-Skłodowska University, Poland, the M.S. degree in biomedical engineering from the Warsaw University of Technology, Poland, and the Ph.D. degree in electrical engineering from The University of Melbourne, Australia. She is currently a Professor of signal processing and artificial intelligence with the School of Engineering, RMIT University, Australia. Her research interests include machine

learning applications in speech and image processing, system modeling, and optimization.



AKRAM AL-HOURANI (Senior Member, IEEE) received the Ph.D. degree from RMIT University, Melbourne, Australia, in 2016. Previously, he worked as a Research and Development Program Manager in the satellite and telecommunication industry, where he directed industry projects valued at over U.S. \$120 million. He is currently an Associate Professor and the Telecommunication Program Manager (Master's) with the School of Engineering, RMIT University. With his expertise,

he is currently engaged in multiple research and industry projects related to satellite communications, synthetic aperture radar, and AI in wireless communications. As the Lead Investigator, he played a crucial role in delivering the largest open Internet of Things (IoT) network in Victoria in collaboration with five local governments, the Northern Melbourne Smart Cities Network. This project won two prestigious industry awards, such as the Technology Awards for Excellence 2020 from the Municipal Association of Victoria (MAV) and the Smart Cities Award for 2020 from the IoT Alliance Australia (IoTAA). He has contributed to over 93 journal articles and conference proceedings. His research interests include satellite communications, radar systems, UAV communication systems, signal processing/machine learning, and stochastic geometry. He was awarded the IEEE Sensors Council Paper Award for his contribution in hand-gesture recognition using neural networks. He is currently ranked in the top 2% of scientists worldwide according to Stanford University's latest rankings, in both the Career-Long and Single Year Impact categories. He is currently an Associate Editor of multiple high-impact journals, including IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONICS SYSTEMS, IEEE OPEN JOURNAL OF THE COMMUNICATION SOCIETY, and *Frontiers in Space Technologies*. He is also an Editor of the *Remote Sensing* (MDPI) journal.