A Deep Learning Approach for Universal NPRACH Detection with Inter-Cell Interference

Runhua Li, Yu Chu, Jiang Xue, *Senior Member, IEEE*, Jian Sun, *Member, IEEE* and Symeon Chatzinotas, *Fellow, IEEE*

***Abstract*—This paper works on the detection of physical random access channel (NPRACH) in Narrowband Internet of Things (NB-IoT) system. The frequency hopping preamble design and increasing number of IoT terminals lead to inter-cell interference among different cells, resulting in inevitable increase of false alarm rate. Due to the ambiguity between preamble and interference, it is a great challenge for NPRACH detection methods to achieve low false alarm rate when having strong interference. In this paper, we analyze the difference between preamble and interference in the propagation environments of NPRACH signals in the 2-dimensional Fast Fourier Transformation (2-D FFT) domain. Then we propose a deep learning-based NPRACH detection method, dubbed Mask Assisted Anti-Interference Universal Detection Scheme (MIUS), in the 2-D FFT domain for preamble detection with inter-cell interference in different repetition cases. In the proposed MIUS, the *Mask-ResNet Block* is designed as a building block to extract features distinguishing the preamble and interference based on masking operations. Our proposed MIUS utilizes the *Mask-ResNet Block* in a recurrent manner to detect the preambles in sequential repetitions across different repetition cases. Simulation results show that MIUS can simultaneously maintain the low false alarm rate and achieve high detection accuracy in low Signal to Interference and Noise Ratio (SINR) regime in all repetition cases.**

***Index Terms*—Deep learning, Fast Fourier Transformation, Internet of Things, Random access, Preamble detection.**

I. Introduction

WITH the development of communication technologies, the Internet of Things (IoT) has gradually become one of the hottest research topics because of its wide applications in recent years [1]. Driven by the demand of low power consumption, large coverage of base station (BS) and multiple terminal connections in IoT, the 3rd generation partnership project (3GPP) proposed the Narrow Band Internet of Things (NB-IoT) [2]–[5]. It is a new low power and wide area (LPWA) scheme relying on existing long term evolution (LTE) cellular networks [6]–[11]. The formulation of NB-IoT standard is completed in Release 14 in 2016 [12]–[19]. Compared with other LPWA schemes, it has a series of advantages such as global coverage, low interference, multiple connections and high level of security, etc. Therefore, NB-IoT stands out as the most widely accepted IoT standard. [20]

R. Li, J. Xue and J. Sun are with Xi'an Jiaotong University, Shaanxi, China.
Y. Chu is with Huazhong University of Science and Technology, Wuhan, China.
S. Chatzinotas is with University of Luxembourg, Luxembourg, UK.
Correspondence author is Jiang Xue (x.jiang@xjtu.edu.cn, 86-29-82665732)

Narrowband physical random access channel (NPRACH) is the latest uplink access technology of NB-IoT, which guarantees effective connection between BSs and terminals [11], [21]–[23]. In NPRACH, each terminal randomly selects one subcarrier in the allocated spectrum and transmits the preamble in the uplink channel. The BS conducts the preamble detection with the received signal using well designed NPRACH detection method, judges which terminal is active and makes the connection. Considering the narrow band characteristic and the demand of extended coverage of NB-IoT, a new single-tone frequency hopping preamble based on frequency hopping rules was designed for NPRACH, which has been included in the 3GPP standard [24]. To make the designed preamble adapt to terminals with different path losses, the entire cellular cell is divided into several coverage levels [25]. For different coverage levels, the preambles are sent with different repetition times (1,2,4,8,...,128) to ensure the preamble detection accuracy of the terminals while sacrificing the transmitting resource [26], [27]. Due to the characteristics of the preamble and the rising number of IoT terminals, the BS of one cell will also receive the preambles from neighborhood cells. These preambles occupy the same time-frequency resources with the original preambles in this cell. This phenomenon is the collision of inter-cell interference, which is called interference collision for simplicity in this paper. And it will cause significant increase of false alarm rate in the process of NPRACH detection. To the best of our knowledge, the problem of NPRACH preamble detection considering inter-cell interference has not been studied in the literature.

For NPRACH preamble detection problem, most of the existing methods are based on energy threshold. Lin *et al.* [24] proposed a method of searching for the time-of-arrival (ToA) and carrier-frequency-offset (CFO) based on two-dimensional Discrete Fourier Transformation (DFT), which can be implemented by a two-dimensional Fast Fourier Transformation (2-D FFT). Furthermore, in this work, the modulo operation was conducted to the transformed two dimensional signal. Then the maximum value was selected as the feature of energy. And the threshold was set by hand to make the detection. Hwang *et al.* [28] also proposed a detection algorithm based on energy threshold. The feature of energy was directly calculated as the sum of the energy of each signal element. And the threshold was determined through derivation according to Neyman-Pearson criterion, which aimed to control the false alarm rate under 0.1%. Due to the adaptability of the threshold, the performance of this method exceeded that of 2-D FFT based

detection method. Also in this work, a simplified ToA and CFO estimation algorithm was investigated, which had lower complexity than the 2-D FFT algorithm with the compensation of the estimation accuracy. In [29], the Neyman-Pearson criterion based method was promoted by introducing singular value decomposition (SVD) for calculating the feature of energy, and the detection performance was improved. In [30], a NPRACH detection algorithm with CFO elimination was proposed. The effect of CFO in NPRACH detection was eliminated through some mathematical operations, and the threshold was set to detect the preambles. However, inter-cell interference was not considered in the above researches. All these energy threshold-based methods concentrate on the energy related features and set threshold for detection. However, there is no difference between the preamble and interference from the energy point of view. Therefore, none of these algorithms could decrease the false alarm rate considering inter-cell interference.

As we know, deep learning (DL) has been widely used in wireless communication to enhance the system performance [31]–[41]. It has also been used in NPRACH detection problem to improve the detection accuracy [42], [43]. In these two works, the received frequency preambles of multiple users are aggregated together and fed into the convolutional neural networks (CNN) based detection algorithms. After extracting the two-dimensional features among multiple-users and different symbol groups, the detection result is obtained through the linear classifier. In the case with no inter-cell interference, the deep learning based methods achieve better performance than 2-D FFT energy threshold based methods. Besides, the similar network architecture is also used to estimate the ToA and the CFO of the received preamble and achieve better estimation accuracy. However, it is very hard to distinguish the preamble with inter-cell interference using the received frequency symbols directly. So the CNN based detection algorithms can't perform well in the time-frequency domain of the received symbols when inter-cell interference occurs.

There have already been several studies considering intra-cell interference in NPRACH. Kim *et al.* [44] considered the transmitting signal design and proposed the partial preamble transmission mechanism to decrease the probability of intra-cell interference collision. However, the available repetition times were also decreased, which would cause the decrease of preamble detection accuracy. The trade-off between mis-detection probability and the collision probability was investigated and an optimal resource utilization strategy was found in this work. Besides, Harwahyu *et al.* [45] presented an analytical method to estimate the success probability of access and average access delay of NPRACH. Then a joint optimization method was proposed to obtain the optimized parameters of the NPRACH system, which can reduce the collision probability of intra-cell interference. However, the interference collision caused by inter-cell interference still existed in the physical layer. And the false alarm rate still arose with the growth of the number of terminals.

For inter-cell interference, conducting the NPRACH detection directly with a well designed method is more straightforward instead of optimizing the NPRACH system to decrease the probability of interference collision. A good enough detection method can solve the interference problem entirely rather than decreasing its effect. However, this is a challenging way due to that the preamble and its collided interference are nearly the same in the time-frequency signal domain. Existing energy threshold methods and time-frequency domain based deep learning methods are unable to distinguish the interference with the preamble. Thus the key point is to find out the potential difference in one specific domain and extract them by a properly designed algorithm.

In this paper, a DL based method called Mask Assisted Anti-Interference Universal Detection Scheme (MIUS) is proposed to solve the challenging problem of preamble detection with inter-cell interference in all repetition cases. Specifically, on the basis of 2-D FFT transformation and modulo operation [24], the received signal is transformed to the two dimensional domain, which is called 2-D FFT domain and the transformed signal is defined as the 2-D FFT image. Instead of the energy based feature, more detailed differences between preamble and interference are found out in the 2-D FFT domain. And the differences are extracted through specially designed feature extractor, which is based on Residual Network (ResNet) [46]. To avoid the need for training separate networks for each repetition case, which would lead to significant resource occupancy, we propose a method in which a single network is trained to recurrently extract features for each 2-repetition signal in any repetition case. Finally, the detection procedure is conducted using well trained neural network (NN) based classifier. The major contributions of this paper are summarized as follows:

- The idea of conducting the NPRACH detection problem in the transformed 2-D FFT domain is proposed in this paper. The aim is to show the potential differences between preamble and inter-cell interference, which is caused by different propagation paths. Relevant theories are proposed and analyzed theoretically.
- *Mask-ResNet Block* is designed in MIUS based on ResNet to extract the potential differences between preamble and interference shown in the 2-D FFT images. By focusing on the characteristic of signals related to the propagation paths, *Mask-ResNet Block* is able to extract more detailed engineered features than traditional energy threshold methods.
- The recurrent operation is designed in this paper to make MIUS be able to deal with all repetition cases without re-training *Mask-ResNet Block* for each repetition case. On the basis of the signal design among repetitions, conducting the feature extraction of the divided signals recurrently with the shared *Mask-ResNet Block* can guarantee the detection performance. The resource occupancy and the number of training data can also be saved.
- Simulation results show that in real interference scenarios, the false alarm rate of MIUS can be maintained under 0.1% with strong interference, while the existing methods can't. Besides, the results for the preamble detection can be guaranteed in real interference scenario in all repetition cases. Furthermore, it is shown that the performance of MIUS exceeds that of the state-of-art deep learning based

methods and the backbone ResNet.

The rest of this paper is organized as follows. In Section II, we will describe the NPRACH system considering inter-cell interference. The transformation of the NPRACH signal and the analysis of the discrimination will be elaborated in Section III. The details of the detection method, MIUS, will be introduced in Section IV. Simulation results are presented in Section V. Finally in Section VI, we conclude the paper.

II. System Model

In this section, the system model of NPRACH with interference is illustrated. Firstly, the preamble design of NPRACH is elaborated, and then the received signal of NPRACH with interference is derived based on the preamble design. For interference, it is considered as the NPRACH preamble from neighboring cells with the assumption of ideally synchronized frames of neighboring cells and full frequency reuse.

A. NPRACH Preamble Design

NB-IoT is based on single-carrier frequency-division multiple access (SCFDMA) system in the uplink configured with W Hz frequency bandwidth and Δw Hz subcarrier bandwidth, therefore the number of subcarriers for NPRACH is $K = W/\Delta w$. The length of each symbol is denoted by N, and the length of a cyclic prefix (CP) portion is N_{cp}. Considering the NPRACH transmitting preamble, a complete random access preamble unit consists of 20 symbols, of which every 5 symbols constitute a symbol group (SG). Only one subcarrier is used for transmission in each SG. The frequency index of the subcarrier is called frequency hopping index. We assume the frequency hopping index of the n-th SG is $f(n)$. The frequency hopping indexes satisfy the frequency hopping rules designed by Ericsson [24]. The rules state that the frequency hopping indexes of the first two SGs differ by 1, those of the second and the third SGs differ by 6, and those of the third and the fourth SGs also differ by 1. According to the frequency hopping rules, we know that once its initial frequency hopping index is determined, the frequency hopping indexes of all other SGs are also determined in a complete random access preamble unit. The multiple users in one cell can be separated through the orthogonal preambles so that the multi-users NPRACH detection problem can be transformed to the single-user detection problem. In practical applications, in order to expand coverage, the random access preamble unit will be repeatedly sent for $N_{rep} = 2^j$ times, $j = 0, \cdots, 7$. The initial frequency hopping indexes of different repeated preamble units are determined by the pseudo-random sequence generated by the cell identity [47]. In one repetition unit of the NPRACH preamble, the value at the k-th subcarrier of the i-th symbol in the n-th SG satisfies

$$S_{n,i}(k) = 1, k = f(n). \tag{1}$$

The time and frequency representations of the whole preamble is shown in Fig. 1.

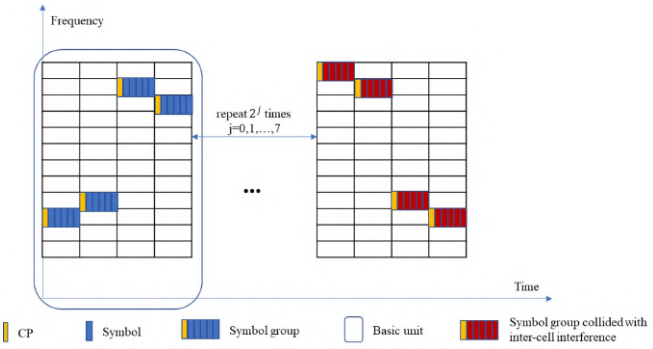


Fig. 1: Signal structure of NPRACH of NB-IoT in time-frequency domain with inter-cell interference.

B. Received Signal Model With Interference

As for the modeling of the received signal of NPRACH, the received preamble under the effect of only channel fading and noise in one repetition unit is derived firstly [28]. After that, the received signal including the received preamble and interference is derived.

Subjecting the i-th symbol in the n-th SG to an inverse Fast Fourier Transformation (IFFT), we can obtain the time domain transmitting symbol as

$$s_{n,i}(l) = \frac{\sqrt{P}}{N} \sum_{k=1}^{K} S_{n,i}(k) e^{\frac{j2\pi kl}{N}}, l = 0, \cdots, N-1, \tag{2}$$

where P denotes the transmitting power. The whole transmitting time domain symbol vector of the n-th SG is $\boldsymbol{s}_n = [\boldsymbol{s}_{cp}, \boldsymbol{s}_{n,1}, \cdots, \boldsymbol{s}_{n,5}]$, in which $\boldsymbol{s}_{cp}$ denotes CP with length of N_{cp}. And the length of the entire SG is $N_{cp} + 5N$. After passing through the wireless channel, the received time domain signal of n-th SG can be obtained as

$$\begin{aligned} y_n(l) =& h(n) s_n((i-1)N + l - D) e^{j2\pi\Delta f(l - D + n(N_{cp}+5N))} \\ &+ v_n(l), l = 0, \cdots, 5N-1, \end{aligned} \tag{3}$$

where $h(n) \in \mathcal{CN}(0,1)$ is the channel coefficient of n-th SG, which we assume as flat Rayleigh fading channel. And we also assume the channel coherence time is longer than the time to send a preamble unit. $D \in [0, N_{cp} - 1]$ is the ToA and $\Delta f \in [-200, 200] Hz$ denotes the CFO, which are both determined by the propagation environment. $v_n(l) \in \mathcal{CN}(0, \sigma^2)$ is the white Gaussian noise where σ^2 denotes the variance of the distribution of the white Gaussian noise and it is determined by the environment on the receiver side.

Furthermore, operating a FFT to the received time domain signal and sample at $f(n)$-th subcarrier in n-th SG. The received frequency domain signal is obtained as

$$Y_{n,i} = N S_N(\Delta f) h e^{\frac{-j2\pi D f(n)}{N}} e^{\frac{j2\pi\Delta f t(n,i)}{N}} e^{j(\theta+\psi)} + V_{n,i}, \tag{4}$$

where $S_N(\Delta f) = \frac{sin(\pi\Delta f)}{N sin(\pi\Delta f/N)}$, $t(n,i) = n(5N + N_{cp}) + iN$, θ is the carrier phase offset which satisfies uniform distribution $\mathcal{U}(0, 2\pi)$, $\phi = \frac{\pi\Delta f(N-1)}{N}$. $V_{n,i}$ is the noise term satisfying $\mathcal{CN}(0, \sigma^2)$.

The interference is actually the NPRACH preamble from the neighboring cells and it is defined as the inter-cell interference.

The interference collision means that some repetition units of the inter-cell interference occupy the same time frequency resource with the local cell preamble, which is shown in Fig. 1. This phenomenon leads to the increase of false alarm rate of NPRACH system, which causes that the user that employs this preamble can't access. It is assumed that each repetition unit of the received NPRACH signal has a certain probability p_0 to encounter an interference collision. It is known that the initial frequency hopping index of each repetition of a preamble is determined by the cell identity. So the probability that two preambles in different cells have the same frequency hopping indexes in more than one repetition unit is very low [47]. Therefore we assume that each interference will only cause the interference collision within a single repetition unit. With this assumption, the ToAs and CFOs of interference collided with different repetition units of the received signal are mostly different because the interference signals originate randomly from the adjacent cells.

According to the above assumptions, the signal on the $f(n)$-th subcarrier of the i-th symbol in the n-th SG of the g-th repetition unit is considered. And the received frequency signal containing both the preamble and interference at BS can be expressed as

$$Y_{n,i}^{S} = \alpha Y_{n,i}^{P} + I_g Y_{n,i}^{I} + V_{n,i}, \tag{5}$$

where $Y_{n,i}^{S}$ denotes the value of the i-th symbol in the n-th SG of the actually received signal at BS. α is the preamble indicator. If the local user equipment is active and the preamble exists, $\alpha = 1$, otherwise $\alpha = 0$. $Y_{n,i}^{P}$ and $Y_{n,i}^{I}$ denote the values of the i-th symbol in the n-th SG of preamble and interference with the same frequency hopping index respectively which can be obtained by equation (4). I_g is the interference indicator which is assumed to satisfy Bernoulli distribution in this paper and its probability density function is written as

$$p_{I_g}(x|p_0) = \begin{cases} p_0^x(1-p_0)^{1-x}, & x = 0, 1, \\ 0, & x \neq 0, 1, \end{cases}$$

where p_0 is the parameter of the distribution indicating the collision probability of inter-cell interference. It is determined by the cell coverage and number of terminals. $g = 0, \cdots, log_2(N_{rep})$ is the index corresponding to the signal repetition times, which indicates the 1-st to N_{rep}-th repetition unit of the received signal. Meanwhile, The signal to noise ratio (SNR) of interference signal should be defined as interference to noise ratio (INR) in this paper.

III. Signal Transformation and Analysis

To solve the tough NPRACH detection problem with interference, the key point is to find out the difference between the preamble and the interference. In this section, the received NPRACH signal will be studied from two dimensions, not one dimension as usual. More discriminative features between preamble and interference will be derived and analyzed.

A. Transformation of Received Signal

According to [24], each repetition unit of the received signals in frequency domain is filled to two dimensions according to the frequency hopping rule as

$$X_g[m, k] = Y_{n,i}^{S}, m = 6(n - 4g) + i, k = f(n), \tag{6}$$

where $X_g[m, k]$ is the element of the signal matrix $\boldsymbol{X}_g$ with size of $M_1 \times M_2$ in which M_1, M_2 are sizes of the 2-D FFT. Furthermore, we operate 2-D FFT and modulo operation to $\boldsymbol{X}_g$ as

$$W_g[p, q] = \sum_{m=0}^{M_1-1} \sum_{k=0}^{M_2-1} X_g[m, k] e^{\frac{-j2\pi mp}{M_1}} e^{\frac{-j2\pi kq}{M_2}}, \tag{7}$$

$$J_g[p, q] = |W_g[p, q]|^2, \tag{8}$$

where $p \in [0, M_1]$ and $q \in [0, M_2]$, $W_g[p, q]$ denotes the element of the complex matrix $\boldsymbol{W}_g$ after 2-D FFT. $\boldsymbol{J}_g$ is the real value matrix obtained through modulo operation to $\boldsymbol{W}_g$.

Thus, the transformation process is completed and the transformed signal $\boldsymbol{J} = \{\boldsymbol{J}_g\}_{g=0}^{N_{rep}}$ is obtained from the one dimensional received signal.

B. Analysis

Definition 1. *The three-dimensional signal $\boldsymbol{J}$ is called a 2-D FFT image when it consists of $\boldsymbol{J}_g$ given by eq. (8), with the g-th channel corresponding to $\boldsymbol{J}_g$.*

The properties of 2-D FFT image are summarized in Theorem I.

Theorem 1. *The maximum point of any channel of $\boldsymbol{J}$, which we refer as (p^*, q^*), satisfies*

$$p^* = \Delta f M_1 + \frac{M_1}{2}, \tag{9}$$

$$q^* = -\frac{D}{N} M_2, \tag{10}$$

where M_1, M_2 are sizes of the 2-D FFT and Δf, D represent the CFO and ToA of the corresponding repetition unit of the received signal respectively.

Proof: Let's assume there only exists preambles. The maximum energy point (p^*, q^*) of the g-th channel of the 2-D FFT image satisfies

$$(p^*, q^*) = \arg\max_{p,q} \left| \sum_{m=0}^{M_1-1} \sum_{k=0}^{M_2-1} X_g[m, k] e^{\frac{-j2\pi mp}{M_1}} e^{\frac{-j2\pi kq}{M_2}} \right|^2, \tag{11}$$

then expand $X_g[m, k]$ with equation (2), (5) and (6), the following expression can be obtained after simplifying:

$$(p^*, q^*) = \arg\max_{p,q} |f_g(p, q)|^2, \tag{12}$$

$$f_g(p, q) = \sum_{m=0}^{M_1-1} \sum_{k=0}^{M_2-1} e^{j2\pi(\Delta f m - \frac{m(p-M_1/2)}{M_1})} e^{-j2\pi(\frac{Dk}{N} + \frac{kq}{M_2})}. \tag{13}$$

According to the triangle inequality, we can scale $|f_g(p,q)|^2$ as

$$\begin{aligned}&\left|\sum_{m=0}^{M_1-1}\sum_{k=0}^{M_2-1} e^{j2\pi(\Delta f m-\frac{m(p-M_1/2)}{M_1})}e^{-j2\pi(\frac{Dk}{N}+\frac{kq}{M_2})}\right|^2\\ &\leq \sum_{m=0}^{M_1-1}\sum_{k=0}^{M_2-1}\left|e^{j2\pi(\Delta f m-\frac{m(p-M_1/2)}{M_1})}e^{-j2\pi(\frac{Dk}{N}+\frac{kq}{M_2})}\right|^2\\ &= M_1M_2, \end{aligned} \tag{14}$$

and the condition for equation is

$$\Delta f m-\frac{m}{M_1}(p-\frac{M_1}{2})=0, \tag{15}$$

$$\frac{Dk}{N}+\frac{k}{M_2}q=0. \tag{16}$$

We can derive the solution: $p^*=\Delta f M_1+\frac{M_1}{2}$, $q^*=-\frac{D}{N}M_2$. ■

From Theorem I, it is known that in each repetition, the maximum point of the 2-D FFT image is determined by the ToA and CFO of the received signal one by one. Due to that different repetition units of the preamble are transmitted together, passing through the same environment and reaching the receiver almost at the same time, so they have the same ToAs and CFOs. This means that the locations of the maximum points of different repetition units of the preamble are the same. However, the different repetition units of the received signal with interference come from different transmitters in different cells according to the assumption, so their ToAs and CFOs are different. And it leads to the difference of the locations of maximum points. However, due to the limitation of the FFT size, the positions of maximum points in different repetition units of the preamble have slight differences.

Because of the difference of the relative positions of maximum points among the channels, the patterns of 2-D FFT images with preamble and interference are obviously different. Take the two repetitions case for example and the 2-D FFT images of preamble and interference are shown in Fig. 2. The values in the first channel of the 2-D FFT image are shown in the red pattern and those in the second channel are shown in the green pattern. From the left image in which the 2-D FFT image of preamble is shown, we can see that the red and green patterns are almost overlaping. While in the right image where the signal collides with interference in both repetition units, the red and green patterns are separated. The further inference can be made that in the case with interference in only one repetition unit, there will only exists one red or green pattern. And finally in the case with only noise, there is no such pattern. The difference between preamble and interference has been observed in local area in 2-D FFT domain explicitly, like in Fig. 2. The intuitive way to extract the difference is to design CNN based deep learning algorithms to focus on the local area and extract the local two-dimensional features in the 2-D FFT domain. This is exactly the idea of our MIUS.

(a) 2-D FFT image of preamble.

(b) 2-D FFT image of interference.

Fig. 2: 2-D FFT images of preamble and interference.

IV. Deep Model for Anti-Interference NPRACH Detection

From the analysis in Section III, the difference between received signals of preamble and interference is shown in the relationship among the channels of the 2-D FFT image. MIUS is designed to extract the difference and make the detection based on the extracted features in this section. The whole procedure and the key points of MIUS are represented as below.

The whole procedure of MIUS consists of three parts, which are signal transformation part, recurrent feature extraction part and final detection part respectively. Firstly, the signals of NPRACH with any repetition time is transformed to the 2-D FFT image and divided into several 2-D FFT images with 2 repetition times. Secondly, the divided 2-repetitions 2-D FFT images undergo feature extraction recurrently. After that, the features are mixed together and taken as input to the classifier in the detection part, and the detection result can be obtained. The signal transformation part contains the 2-D FFT, module and dividing operations. *Mask-ResNet Block* is designed to extract the effective features of the 2-repetitions 2-D FFT images and the block is shared during the recurrent operation. Finally the classifier in the detection part contains a max-pooling layer following a multi-layer perceptron (MLP). The MLP consists of a linear layer and the soft-max activating function. Fig. 3 represents the procedure in detail.

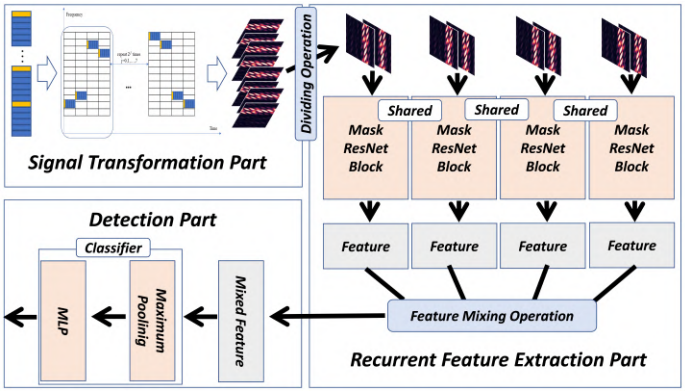


Fig. 3: Illustration of MIUS, which contains three parts respectively for signal transformation, feature extraction and final detection.

The first key point of MIUS is to adapt to all repetition cases. As we know, DL methods are data based techniques

and they can be used only when they are well trained by amount of data. Besides, a well trained neural network (NN) only supports fixed size inputs. Therefore, it is a challenge to make NN useful for all repetition cases with different input sizes. It is straightforward to consider to train different NNs for different repetition cases. However, it is not applicable because all the NNs need to be stored in the BS and that would engage considerable resource occupancy. In this paper, the recurrent feature extraction method of MIUS is designed to solve this challenging problem. For any repetition case, the NPRACH detection problem can be transformed to the 2-repetitions case. Therefore, the *Mask-ResNet Block* and the final classifier only need to be trained by the 2-repetitions NPRACH signals.

The second key point of MIUS shows in the design of *Mask-ResNet Block* and the max-pooling in the classifier. Due to the difference between preamble and interference exists among the channels of the 2-D FFT images, the *Mask-ResNet Block* is designed based on the ResNet and masking operation to extract the relationship among the channels. Max-pooling is used in the classifier. It is capable to keep the highlighted areas to make MIUS focus on the difference of the highlighted areas of preambles and interference.

The details on recurrent feature extraction, *Mask-ResNet Block* and the training and testing of MIUS will be illustrated in the following subsections.

A. Design of Recurrent Operation

Assuming that the repetition times of the signal is N, the 2-D FFT image of the signal is represented as $\{\boldsymbol{J}_i\}_{i=1}^{N}$. Each $\boldsymbol{J}_i$ is a real matrix with size of $M_1 \times M_2$. Considering the dividing method in the signal transformation part, the neighbouring two matrixes are put together as a 2-repetitions 2-D FFT image, the signal can be rewritten as $\{\boldsymbol{J}_{2i-1}\boldsymbol{J}_{2i}\}_{i=1}^{N/2}$. Assuming that there is already a *Mask-ResNet Block* for the 2-repetitions 2-D FFT images. Input all the 2-D FFT images $\{\boldsymbol{J}_{2i-1}\boldsymbol{J}_{2i}\}$ into the shared F. They are operated recurrently and $N/2$ extracted low dimensional features $\{\boldsymbol{z}_i\}_{i=1}^{N/2}$ can be obtained. After that, conduct average pooling operation to the extracted features as

$$\boldsymbol{z}_{avg} = \frac{1}{N}\sum_{i=1}^{N/2} \boldsymbol{z}_i, \tag{17}$$

where the averaged feature $\boldsymbol{z}_{avg}$ can be obtained. Feeding the averaged feature to the well trained classifier, the detection result can be gotten. The reason for using average pooling is that features can be normalized to the fixed size and the characteristics from different 2-repetitions parts can be mixed together effectively. Through this operation, the *Mask-ResNet Block* and the classifier only need to be trained in the 2-repetitions cases. The storage of MIUS in BS only needs to contain one *Mask-ResNet Block* and one classifier for 2-repetitions cases.

The intuition of this method is to consider detection problem of the N-repetitions NPRACH signal as that of the 2-repetitions NPRACH signal. Average pooling of the obtained features after recurrent feature extraction can be considered as taking the middle point of the divided 2-repetitions features. As long as the classifier performs well, the misclassified points will be averaged by the correctly classified points and the accurate probability of classification will increase for the middle point. In this way, the performance of MIUS in high repetition cases can be guaranteed.

B. Design of Mask-ResNet Block

As for the design of *Mask-ResNet Block*, only 2-repetitions NPRACH signals need to be considered. The structure of *Mask-ResNet Block* is shown in Fig. 4.

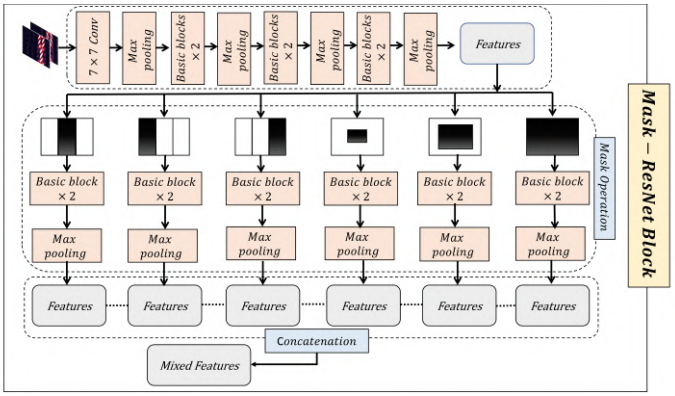


Fig. 4: Structure design of the Mask-ResNet Block in MIUS.

The structure is based on ResNet-18 [46] which consists of five parts. The first part contains a 7×7 convolutional layer, a batch norm layer, a ReLu activate function and a max-pooling layer. The following four parts are all composed of two basic blocks and a max-pooling layer. In the basic block, the input feature firstly undergoes the 3×3 convolution, the batch norm layer, the ReLu activate function and another 3×3 convolution with a batch norm layer. Then it is added to the feature which is obtained through sending the input feature into the 1×1 convolution with a batch norm layer. This kind of framework contributes to easier gradient back-propagation, which greatly improves the calculation speed and effectively solves the problem of gradient explosions. ResNet is applied as the backbone architecture so that MIUS has the significant basic ability to extract the feature of 2-D FFT image. In the design of the *Mask-ResNet Block*, masking operation is added to the backbone architecture to make *Mask-ResNet Block* pay more attention to the highlight areas which are capable of showing the difference. The detailed procedure of mask-ResNet block is as follows. Firstly, input the 2-D FFT image to the first four parts of the backbone ResNet, and the pre-extracted features can be obtained. Then the features undergo six different masking operations and the six masked features are further taken as input to the six basic blocks following max-pooling operation. And finally, the six outputs are concatenated together and the mixed output features can be obtained. The details on masking operation will be elaborated as follows:

From the conclusion in Section III, we know that the difference between preamble and interference is reflected in the relationship of the locations of the highlights in neighboring channels of 2-D FFT image. Considering that we put the 2-D FFT image into the backbone ResNet, the relationship

of the locations of the highlights in different channels is reflected in the highlight area of the feature map after several steps of convolutions and max-poolings. The feature map of preamble has one strong highlight area while the feature map of interference has one or two weak highlight areas. To make MIUS focus on this difference, we conduct masking operation to the feature map. Masking operation is to retain the values of the areas we want and set all other areas to zero. In this work, we consider the feature map output from the fourth part of the backbone ResNet and design six masks for different areas of the feature map. After six masking operations in six different areas, further feature extraction is conducted by six basic blocks and max-pooling layers. Then six output feature maps are concatenated together for classification. In this way, the network can learn to compare the difference of different areas of the feature map. And it will notice the strong highlight area of the preamble or weak highlight area of interference through back-propagation during the training. Six masks we designed in this work are shown in Fig. 4. Considering the size of each channel of the output feature map is 16×16, firstly, we divide the feature map into three disjoint parts horizontally and the width of each part are $5, 6$ and 5 respectively. These three parts lead to the first three masks. Then, considering different levels of central part of the feature map, we design three masks focusing on different scales of the central area, which have the size of 6×6, 12×12 and 16×16 respectively.

C. Network Training and Testing

As for the training part, only 2-repetitions NPRACH signals are needed. Conduct 2-D FFT and modulo operation to the received 2-repetitions NPRACH signals obtained under the assumption of Rayleigh fading channels in the training set, which involves preambles with different SNRs, interference with different interference collision ratios and INRs, and received noise. Signals with preambles are labeled as 1 while others are labeled as 0. Thus we can get a set of 2-D FFT images. The number of training epoches is set as 100, the batch-size is set as 40 and the learning rate is set as 0.01. In each training batch, input the 2-D FFT images to the *Mask-ResNet Block* and the classifier, calculate the cross entropy loss between the outputs and the labels as

$$L = E[-(o_1 \log(p_1) + o_2 \log(p_2))], \tag{18}$$

where (p_1, p_2) is the normalized output probability vector and $o_i, i = 1, 2$ denotes the label. $E[\cdot]$ is to get the mean of the results and it is operated over the training batch. Back-propagate the loss and optimize the parameters of *Mask-ResNet Block* and the classifier using ADAM algorithm. Cycle the above steps for all training batches and epoches. After finishing the training procedure, the validation set is used to choose the most generalized model for testing.

As for the testing part, all repetition cases are considered. After the NPRACH signals are received, they firstly undergo the signal transformation part, in which they are transformed to 2-D FFT images and then divided into several 2-repetitions 2-D FFT images. After that, the 2-D FFT images are sent to the shared *Mask-ResNet Block* to be operated recurrently. The output features are mixed together through feature mixing operation and the mixed feature is obtained. Finally the mixed feature is taken as input to the classifier and the detection result is obtained.

V. Simulation

In this section, we will verify the effectiveness of MIUS by simulation experiments. Since the users in the same cell have been separated by the orthogonal preambles, the problem of multi-user detection can be transformed to that of single-user detection. Therefore, in the simulation part, we only consider the detection of single user and we would pay more attention to the 2-repetitions case, which is the hardest one under the interference. Firstly, we will verify the false alarm rate performance of MIUS in the presence of interference scenarios. False alarm rate is defined as the ratio between the number of misjudged interference or noise samples and their total number. It is hard to be controlled in a low level with the presence of inter-cell interference. Secondly, we consider the metric of detection accuracy of preambles. The simulations are conducted in scenarios where only preamble is available as well as in scenarios where both preamble and interference are present across all repetition cases.

A. Simulation Setting

As for the simulation of NPRACH in NB-IoT, there are some existing simulation models proposed in literature based on OPNET or the 3-rd Generation Partnership Project (3GPP) protocol [5], [43]. However, only one cell is considered in these well-built simulation models and the inter-cell interference has not been considered. So in our simulations, the system link is constructed by ourselves with Matlab considering the simulation of inter-cell interference, and the system settings follow that of [28]. The specific parameter settings of the system are presented as follows: The bandwidth for NPRACH is set as 180 kHz, and that of a single subcarrier is 3.75 kHz. The FFT size and the length of CP are set as 512 and 128 respectively. The numbers of antennas on the transmitting side and receiving side are 1 and 2, respectively. Flat Rayleigh fading channel is considered following [28] and the size of 2-D FFT is 256×256.

There are totally six methods for comparison with MIUS in our simulation, which are described below:

- *2-D FFT threshold method:* 2-D FFT threshold method proposed in [24] is one of the existing energy threshold methods. In the simulation of this method, We firstly transform the received time-frequency signals with 2-D FFT and modulo operation. Then the maximum value of the matrix is selected as the energy value. The threshold selection scheme is not clarified in [24]. We magnify the energy value by 100 times and choose 0.37 as the energy threshold. This choice makes sure that the false alarm rate with only noise is just below 0.1%, which follows the requirement of the protocol.
- *Neyman Pearson threshold method:* Neyman Pearson threshold method in [28] is also an energy threshold method with adaptive thresholds. In our simulation, the

energy value is computed through adding up all energy of the received time-frequency symbols. As for threshold selection, we choose 0.312 as the threshold for the 2-repetition case according to the Neyman-Pearson criterion, which is shown in Fig. 3 of [28].

- *Multi-Layer Perceptron (MLP):* MLP is one of the basic neural networks in deep learning fields. It is selected as the comparison method to show 1) it is hard to control the false alarm in time-frequency domain of the signals with deep learning facing with inter-cell interference and 2) the effectiveness of the network design of MIUS. In the simulation, the received time-frequency symbols are directly input to the MLP and the output is the detection result. As for structure, the number of layers is set as 4 and the ReLU activate function is selected.
- *Long Short-Term Memory (LSTM):* The correlation among the received symbols in different repetition units is not considered in the MLP. LSTM can be seen as the enhanced version of MLP to detect the time-correlated symbol sequence in the time-frequency domain. The reason for selecting LSTM is the same as that of MLP. In the simulation, the multiple-repetition received time-frequency signal is directly input to the multi-cell of the LSTM and the detection result is the output. As for the network structure, the number of layers of LSTM is set as 2 and the bi-direction is considered.
- *NPRACH-CNN:* Considering the state-of-art work in [42], [43], CNN based deep learning methods are considered for comparison with our MIUS, which is called NPRACH-CNN in this paper. Since the difference between preamble and interference is more explicit in 2-D FFT domain than time-frequency domain, 2-D FFT image of the received signal is also considered as the input of NPRACH-CNN for fair comparison. The network structure of NPRACH-CNN follows that of [43].
- *ResNet:* As the improved version of basic CNN, ResNet is the backbone method of our MIUS. Selecting ResNet of the comparison method is to show the effectiveness of the mask design of MIUS. In the simulation, the input of ResNet is also the 2-D FFT image of the received time-frequency signals. 18-layers ResNet is selected as the network structure, which is the same as that in MIUS.

In conclusion, 2-D FFT threshold method and Neyman Pearson threshold method are considered as the representations of energy threshold methods. MLP, LSTM are selected as the time-frequency domain deep learning methods. NPRACH-CNN is the state-of-art method for NPRACH detection and it is considered in 2-D FFT domain to make the comparison fair enough. Finally, the backbone network ResNet is also selected for comparison as the ablation study to show the effectiveness of the network design of MIUS. When facing with high repetition cases for NPRACH-CNN and ResNet, which is conducted in the 2-D FFT domain, the recurrent operation method are utilized for test. The two comparison methods are also trained only in the 2-repetitions case for fair comparison with MIUS. The hyper-parameters of the above DL methods are all set the same as MIUS, which is illustrated in Section IV-C. All DL methods are trained on the same dataset. All experiments are conducted on an individual computer with i5-9400 CPU, RTX-2080 GPU and RAM 16GB. All DL methods are conducted with Pytorch 1.2.0 in Python 3.7 on Windows 10.

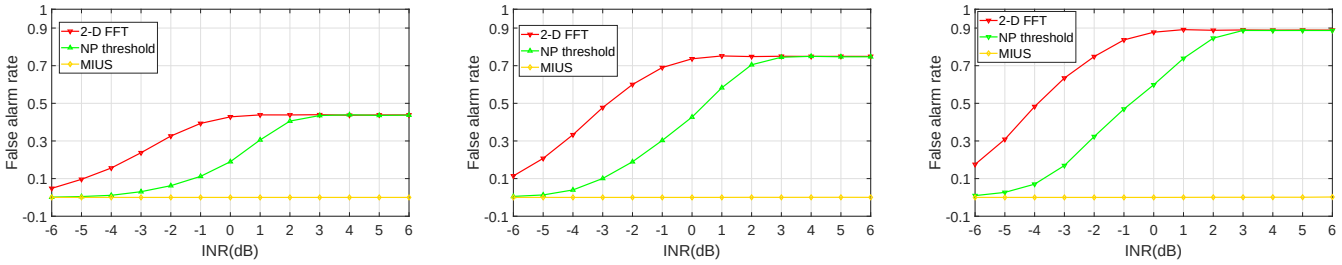
The training data-set consists of received signals in three scenarios, which are scenarios with preamble and noise, interference and noise and finally preamble, interference and noise. In the scenario with preamble and noise, we set the range of received SNR as $6 \sim 10$ dB, and generate 3000 signal samples randomly in total and set their labels as 1. To make the DL methods traverse all kinds of interference scenarios during training, the collision probability of interference is considered random in the training data-set. So in the scenario with interference and noise, the number of repetitions with interference collision is set as the integer random variable satisfying $\mathcal{U}(0, N_{rep})$ for simplicity. N_{rep} is the total number of repetitions of the signal. The range of INR is chosen as $-6 \sim 6$ dB. We generate 6000 samples randomly with label 0 in total. And finally in the scenario with preamble, interference and noise, the range of SNR and INR are respectively set as $6 \sim 20$ dB and $-6 \sim 6$ dB. The interference collision number is still assumed as an integer random variable obeying $\mathcal{U}(0, N_{rep})$. 3000 samples are generated randomly with label 1 in this case. The validation set consists of 1000 samples in interference and noise case and 1000 samples in preamble and noise case. The samples in interference and noise case is generated with random interference collision and 6 dB INR. And those in preamble and noise case are generated with 6 dB SNR. After we finish training the network, the validation set helps us to select the model that has the best generalized capability, and the selected model will get involved in the following test process.

B. Simulation Results of False Alarm Rate

In this subsection, the tests in interference scenarios are conducted. The tests are separated as two parts. In the first part, fixed collision probability scenario is considered to evaluate the effect of MIUS with comparison of existing energy threshold based methods and the backbone ResNet. In the second part, random collision probability scenario and all repetition cases are considered. The performance of MIUS is evaluated and compared with the deep learning methods.

Firstly, considering different fixed collision probabilities, the performance of false alarm rate is evaluated in the 2-repetitions case. The methods for comparison are respectively 2-D FFT threshold methods, Neyman Pearson threshold methods and the backbone ResNet. For the simplicity of experiment, $3/12, 6/12$ and $8/12$ are selected as the representative collision probabilities in the test. In each collision probability scenario, 5000 samples with interference and noise are generated for each fixed INR ranging from -6 dB to 6 dB. The performance comparison between MIUS and two existing energy threshold methods are shown in Fig. 5.

In these three figures, the red line represents the performance of 2-D FFT threshold detection method, the green line represents that of NP threshold detection method and the



(a) 3/12 collision probability (b) 6/12 collision probability (c) 8/12 collision probability

Fig. 5: Performance comparison with energy threshold based methods in the case of two repetitions in $3/12, 6/12$ and $8/12$ interference collision probability scenarios.

performance of MIUS is shown in the yellow line. Obviously, the false alarm rate of MIUS always stays about 0 with INR ranging in $[-6, 6]$ dB while that of two existing energy threshold methods is higher. It can be seen that under all collision probabilities, the false alarm rate of the 2-D FFT threshold detection algorithm is at least above 5%, which has exceeded the requirements of the NPRACH system. NP threshold method only works well under -6 dB INR in three interference collision levels, but when INR increases, the false alarm rate of this method increases rapidly. Finally, the false alarm rates of both two threshold detection algorithms stabilise at a similar level which are roughly $44\%, 75\%$ and 90% under $3/12, 6/12$ and $8/12$ collision probability respectively. Obviously, the NPRACH system can hardly work with these high false alarm rates. One of the reasons for the results is that interference has no difference with preambles in the view of received energy, so it is impossible to distinguish interference with preamble only by received energy. Furthermore, there is no other prior information about the interference when setting the judging threshold. The threshold only has an impact with the noise. This is also one reason for high false alarm rate of two existing energy threshold detection methods.

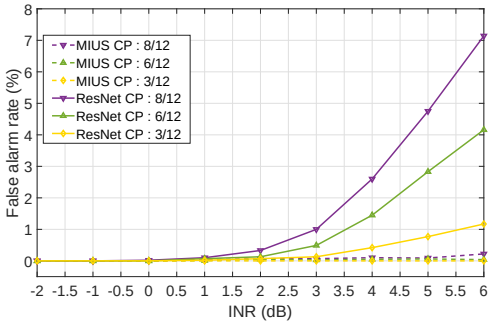
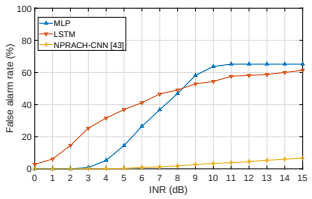


Fig. 6: False alarm performance comparison between ResNet and MIUS with 3/12, 6/12 and 8/12 collision probabilities respectively. CP in the figure represents the collision probability for simplicity.

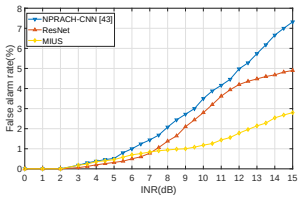
The comparison of the performance of false alarm rate between the backbone ResNet and MIUS is presented in Fig. 6. The solid line represents the performance of ResNet and the dotted line represents the result of MIUS. The false alarm rates in three different interference collision levels $3/12, 6/12, 8/12$ are represented by yellow, green and purple colors respectively. When the INR is below -2 dB, the false alarm rates of both methods are zero. So we only present the results with INR from -2 dB to 6 dB in the figure. We can see from the figure that MIUS outperforms ResNet in all three interference collision levels with the increasing of INR. Especially when the INR is 6 dB, the false alarm rates of ResNet are beyond $1\%, 4\%$ and 7% in three different collision levels, while the false alarm rate of MIUS are controlled under 0.2% in all simulation scenarios. This result confirms that the extraordinary ability of controlling the false alarm rate of MIUS comes from the masking operation and the max-pooling function which is one of our main contributions in this work. From Fig. 6, We can see that the false alarm rate of MIUS can always stay below 0.1% except for the scenario with $8/12$ collision probability and 6 dB INR. Even in this scenario, its false alarm rate can stay around 0.2%. From the above results, MIUS can meet the requirement of false alarm rate and ensure the NPRACH system work effectively even with strong enough interference. This consequence verifies the effectiveness of the discrimination features between preambles and interference. Meanwhile, MIUS extracts the detailed features of the relationship among different repetitions of received signal to make a correct judge between preambles and interference.

The above experiments verify the effectiveness of MIUS with the assumption of fixed collision probability in the 2-repetitions case. To test the performance of MIUS in more realistic scenario, the parameter settings are changed in the following. More repetition cases are considered, including $2, 4, 8, 16, 32, 64$. For each repetition case, 5000 interference and noise signals are generated through the data generators. As for the interference, the number of repetitions that be collided with interference satisfied $\mathcal{U}(0, N_{rep})$. As for the INR of the interference, $[0, 15]$ dB is selected for the 2-repetitions case and $[0, 100]$ dB is selected for the other repetition cases. Only DL based methods are considered for comparison for simplicity due to that existing energy threshold based methods have been proved unable to deal with the interference collision according to the above experiments. The test results for the 2-repetitions case are shown in Fig. 7.

The performance of MLP, LSTM and NPRACH-CNN is shown in the left side of Fig. 7. From the figure, it can be



(a) Performance of MLP, LSTM and NPRACH-CNN.



(b) Performance of MIUS, ResNet and NPRACH-CNN.

Fig. 7: False alarm performance of MIUS in realistic two repetition case with comparison of MLP, LSTM, NPRACH-CNN and backbone ResNet.

seen clearly that the performance of MLP and the LSTM is much worse than that of NPRACH-CNN. The false alarm rate of NPRACH-CNN can be maintained under 10% till 15 dB INR. Otherwise for MLP and LSTM, 4.5 dB INR and 1.5 dB INR may cause 10% false alarm rate and the rate can even exceeds 60% when INR exceeds 10 dB. This indicates that our MLP and LSTM fail to control the false alarm rate with high INR interference in the time-frequency domain, but NPRACH-CNN achieves much better performance in the 2-D FFT domain. This result can prove to some extent that detecting the preamble from interference in 2-D FFT domain is more effective than time-frequency domain. After that, the performance of MIUS is shown in the right side of Fig. 7 with comparison of the backbone ResNet and NPRACH-CNN. It can be seen that the false alarm rate of MIUS can be maintained under 2% till 15 dB INR while that of ResNet and NPRACH-CNN are respectively 6.5% and 4.5%. When the INR is low, the performance of the three methods are nearly the same. From about 6 dB INR, the false alarm rates of NPRACH-CNN and ResNet grow rapidly and that of MIUS can be maintained in a low level. This experiment verifies the effectiveness of MIUS to control the false alarm rate in more realistic scenarios, especially with high INR when interference is strong. It is clearly shown that the performance of MIUS is better than the state-of-art NPRACH-CNN in interference scenario. The comparison with ResNet also shows the effectiveness of the network design.

TABLE I: False alarm rates of different methods under 100 dB INR in the $4, 8, 16, 32, 64$ repetition cases.

False alarm rate (%) Repetition case / Methods	4	8	16	32	64
NPRACH-CNN [43]	8.0	3.5	2.3	1.4	0.5
ResNet	2.8	0.4	0.0	0.0	0.0
MIUS	1.8	0.2	0.0	0.0	0.0

As for the high repetition cases, the testing results are shown in Table I. Only NPRACH-CNN and ResNet are used for comparison according to the above experiments, which indicate that MLP and LSTM in our simulation can not perform well with interference. Since normal interference level cannot arise any false alarm rate in high repetition cases with the deep learning methods in 2-D FFT domain. So an extreme interference scenario with 100 dB INR is considered to show the difference of ability of these methods. It can be seen from the results that the performance of ResNet and MIUS exceeds that of NPRACH-CNN. For the 4-repetitions case, the false alarm rate reaches 1.8% for MIUS, 2.8% for ResNet and 8% for CNN uner 100 dB INR. As for the 8-repetitions case, MIUS can control the false alarm rate under 0.2% till 100 dB INR, while those of ResNet and NPRACH-CNN are respectively 0.4% and 3.5%. For the case of $16, 32$ and 64-repetitions cases, MIUS and ResNet can all stay 0% false alarm rate till 100 dB INR, while NPRACH-CNN does not have this ability, the false alarm rates for $16, 32$ and 64 repetition cases are respectively 2.3%, 1.4% and 0.5%. The above results can prove the effectiveness of recurrent operation methods. The neural network trained using only 2-repetitions signals can be implemented to any repetition case without any other training process. And the performance of false alarm rate is also satisfied. Besides, the advantage on performance of MIUS is also shown form the results, compared with the baseline ResNet and NPRACH-CNN.

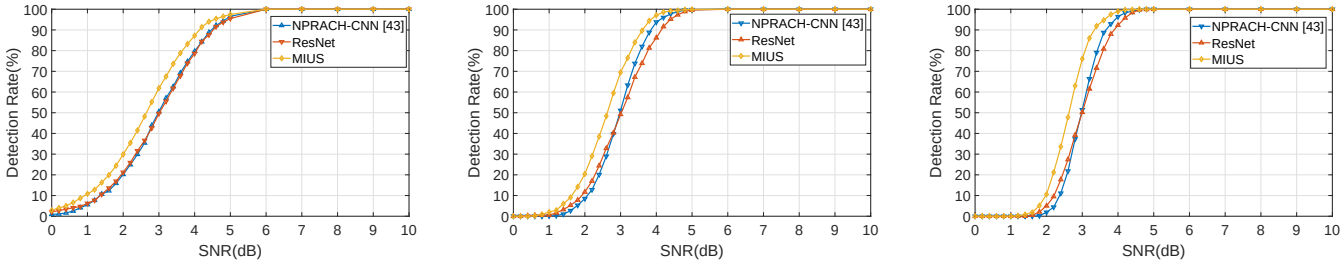
C. Performance of Preamble Detection Accuracy

After verifying the effectiveness of MIUS to control the false alarm rate in the interference scenario, the performance of preamble detection accuracy is evaluated in this subsection. The simulations are divided into two parts. In the first part, the scenario with only preambles across all repetition cases are considered. In the second part, only 2-repetitions cases with preambles and different levels of interference are under consideration.

Firstly, considering the performance of MIUS for all repetition cases with only preamble signals. 5000 preamble and noise signal samples are generated for each repetition case and SNR level, which ranges in $[0, 10]$ dB. Due to the limitation of the space for writing, three repetition cases are selected for the test, in which the numbers of repetitions are respectively $2, 8, 32$. According to the above results of false alarm rates, energy threshold based methods, MLP and LSTM can not control the false alarm rate in a low level in the interference scenario, so only NPRACH-CNN and ResNet are chosen as the comparison methods. And the results are summarized in Fig. 8.

From the three figures, we can conclude that the performance of MIUS exceeds those of ResNet and NPRACH-CNN at every SNR level, which verifies the effectiveness of the design of MIUS for preamble detection. For the 2-repetitions case, the detection accuracy of MIUS is nearly 0 at 0 dB SNR. With SNR increasing, the detection accuracy grows rapidly and reaches 100% at about 6 dB SNR. The trends of the curves in the 8 repetition and 32 repetition case are the same. MIUS reaches 100% detection accuracy at 4.8 and 4.2 dB respectively. The performance in 8 repetition case and 32 repetition case show that MIUS can deal with high repetition cases through recurrent operation methods, and it only need to be trained in the 2-repetitions case.

Besides, the ability of MIUS for preamble detection with interference is explored. The interference collision probability is also modeled as a random variable as before. Three levels of interference are considered, which are respectively no



(a) 2-repetitions case. (b) 8-repetitions case. (c) 32-repetitions case.

Fig. 8: Preamble detection performance of MIUS in $2, 8, 32$-repetition cases with comparison of NPRACH-CNN and ResNet.

interference, 0 dB INR interference and 6 dB INR interference. $[0, 15]$ dB SNR is considered and 5000 samples are generated for each SNR and interference level. The results are shown in Fig. 9.

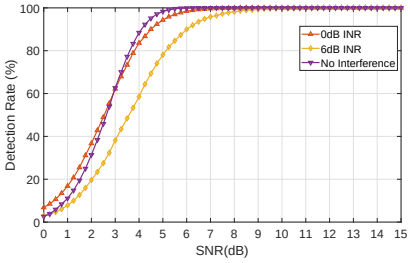


Fig. 9: Detection performance of MIUS in the case of two repetitions in different interference levels.

From the figure, it can be seen that when SNR is above 3 dB, the detection accuracy of MIUS decreases with the increase of interference level. However, when SNR is under 3 dB, the detection accuracy would be better with weak interference and worse facing strong interference. This is because when the INR of interference and the SNR of preamble are both low, the difference of the pattern between interference and preamble is not obvious, which may cause the mistake of MIUS. And the added interference would be recognised as preamble by MIUS, which may improve the detection accuracy mistakenly. When the SNR is high, the pattern of the preamble is obvious and MIUS can recognise it easily. The appearance of interference would affect the judgement of MIUS no matter the INR is low or high. If the interference is weak, the effect may be small, while it may be large if the interference is strong. So the order of the detection accuracy from small to large is 6 dB interference scenario, 0 dB interference scenario and no-interference scenario when the SNR is high. The results above verify that MIUS has the ability to work effectively in the interference scenarios, even in low SINR regime.

From the above test performance, we can see that in the interference scenarios, MIUS can work effectively. But the performance of detection accuracy is worse compared with the existing methods in previous works in which the interference is not considered. With the interference becomes stronger and stronger, the detection accuracy would decrease at the same SNR level. This phenomenon is reasonable. Comparing with the existing algorithms, we add many restrictions related to the features of difference between preamble and interference to the method to classify interference and preambles. This will make the determination of preamble more strict. So the detection accuracy will decrease and this can be seen as the trade-off of controlling low false alarm rate in the interference scenarios.

D. Complexity Analysis

To evaluate the complexity of MIUS, the number of parameters and flops are calculated and shown in this subsection. The number of parameters of MIUS is 9805056 and its flops is calculated as 1.7393×10^{10}. Furthermore to make the complexity more intuitive, the model size and the computation time of MIUS in our simulation device are presented. Those of ResNet are also evaluated for comparison. The computation time is defined as the time of detecting one signal. And the results are listed in Table II and Table III.

TABLE II: Model size and computation time of ResNet in the $2, 4, 8, 16, 32, 64$ repetition cases.

Measurement \ Repetition case	2	4	8	16	32	64
Model size (M)	43.73	43.73	43.73	43.73	43.73	43.73
Computation time (ms)	1.75	2.59	3.97	5.95	10.69	19.84

TABLE III: Model size and computation time of MIUS in the $2, 4, 8, 16, 32, 64$ repetition cases.

Measurement \ Repetition case	2	4	8	16	32	64
Model size (M)	43.75	43.75	43.75	43.75	43.75	43.75
Computation time (ms)	4.13	13.15	22.88	34.60	42.96	71.33

From the tables above, it can be seen that the model size of MIUS is nearly the same as that of ResNet, and for all repetitions, the model sizes are the same due to that only the network trained in 2 repetitions case needs to be stored. As for the computation time, MIUS is longer than that of ResNet for all repetitions. It is because that the feature output by the 3rd basic block goes through 6 masking operations in the MIUS, which cost more computation time than that of ResNet.

VI. Conclusion

In this paper, the NPRACH detection problem with inter-cell interference is considered. Firstly, the difference between

preamble and interference is studied theoretically, which is analyzed to exist in the relationship among channels of the transformed 2-D FFT images caused by different propagation paths. Then, a deep learning based method MIUS is proposed based on the theoretical analysis. In MIUS, recurrent operation method is utilized to make it suitable for all repetition cases. As for the design of the *Mask-ResNet Block* in MIUS, six masks are introduced to the backbone network ResNet, so that MIUS can concentrate on the difference of preambles and interference, and extracts efficient features about it. Simulations show that MIUS can work in all repetition cases and control the false alarm rate in a low level. The effectiveness of preamble detection with interference of MIUS can also be ensured. Besides, MIUS outperforms the existing energy threshold based methods and state-of-art deep learning based methods, which shows the effectiveness of the difference observed and the superiority of the structure design.

REFERENCES

[1] S. Andreev, O. Galinina, A. Pyattaev, M. Gerasimenko, T. Tirronen, J. Torsner, J. Sachs, M. Dohler, and Y. Koucheryavy, "Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 32–40, 2015.

[2] A. Lombardo, S. Parrino, G. Peruzzi, and A. Pozzebon, "Lorawan versus NB-IoT: Transmission performance analysis within critical environments," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1068–1081, 2021.

[3] O. Elgarhy, L. Reggiani, H. Malik, M. M. Alam, and M. A. Imran, "Rate-latency optimization for NB-IoT with adaptive resource unit configuration in uplink transmission," *IEEE Systems Journal*, vol. 15, no. 1, pp. 265–276, 2020.

[4] J. Zhang, D. Xie, and X. Wang, "TARA: An efficient random access mechanism for NB-IoT by exploiting TA value difference in collided preambles," *IEEE Transactions on Mobile Computing*, 2020.

[5] Y. Miao, W. Li, D. Tian, M. S. Hossain, and M. F. Alhamid, "Narrowband internet of things: Simulation and modeling," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2304–2314, 2017.

[6] C.-P. Lee, P. Lin, and Y. C. Sung, "The module switching mechanism for power-constrained devices in LTE and NB-IoT interworking networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 2018–2023, 2021.

[7] C.-F. Li, J.-K. Hwang, C. Ma, and C.-J. Lin, "Software defined radio implementation of LTE R13 NB-IoT downlink vector signal generator," in *2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. IEEE, 2017, pp. 69–70.

[8] J.-B. Seo and V. C. Leung, "Design and analysis of backoff algorithms for random access channels in UMTS-LTE and IEEE 802.16 systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 8, pp. 3975–3989, 2011.

[9] J. J. Nielsen, D. M. Kim, G. C. Madueno, N. K. Pratas, and P. Popovski, "A tractable model of the LTE access reservation procedure for machine-type communications," in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[10] C.-H. Wei, R.-G. Cheng, and S.-L. Tsao, "Performance analysis of group paging for machine-type communications in LTE networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3371–3382, 2013.

[11] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the random access channel of LTE and LTE-A suitable for M2M communications? a survey of alternatives," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 4–16, 2013.

[12] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band internet of things," *IEEE access*, vol. 5, pp. 20 557–20 577, 2017.

[13] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[14] C. Zhong, H. A. Suraweera, G. Zheng, I. Krikidis, and Z. Zhang, "Wireless information and power transfer with full duplex relaying," *IEEE Transactions on Communications*, vol. 62, no. 10, pp. 3447–3461, 2014.

[15] J. Schlienz and D. Raddino, "Narrowband internet of things whitepaper," *White Paper, Rohde&Schwarz*, pp. 1–42, 2016.

[16] G. Zhu, C. Zhong, H. A. Suraweera, G. K. Karagiannidis, Z. Zhang, and T. A. Tsiftsis, "Wireless information and power transfer in relay systems with multiple antennas and interference," *IEEE Transactions on Communications*, vol. 63, no. 4, pp. 1400–1418, 2015.

[17] I. Yaqoob, E. Ahmed, I. A. T. Hashem, A. I. A. Ahmed, A. Gani, M. Imran, and M. Guizani, "Internet of things architecture: Recent advances, taxonomy, requirements, and open challenges," *IEEE wireless communications*, vol. 24, no. 3, pp. 10–16, 2017.

[18] S.-M. Oh and J. Shin, "An efficient small data transmission scheme in the 3GPP NB-IoT system," *IEEE Communications Letters*, vol. 21, no. 3, pp. 660–663, 2016.

[19] S. Guirguis, K. R. Chaudhuri, A. Halder, and N. Mullaguru, "M2M communications: Enablement in 4G LTE, deployment considerations and evolution path to 5G," in *2017 International Conference on Wireless Networks and Mobile Communications (WINCOM)*. IEEE, 2017, pp. 1–8.

[20] J. Xu, J. Yao, L. Wang, Z. Ming, K. Wu, and L. Chen, "Narrowband internet of things: Evolutions, technologies, and open issues," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1449–1462, 2017.

[21] H. S. Jang, S. M. Kim, K. S. Ko, J. Cha, and D. K. Sung, "Spatial group based random access for M2M communications," *IEEE Communications Letters*, vol. 18, no. 6, pp. 961–964, 2014.

[22] Y. Wen, W. Huang, and Z. Zhang, "CAZAC sequence and its application in LTE random access," in *2006 IEEE Information Theory Workshop-ITW'06 Chengdu*. IEEE, 2006, pp. 544–547.

[23] C.-Y. Oh, D. Hwang, and T.-J. Lee, "Joint access control and resource allocation for concurrent and massive access of M2M devices," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4182–4192, 2015.

[24] X. Lin, A. Adhikary, and Y.-P. E. Wang, "Random access preamble design and detection for 3GPP narrowband IoT systems," *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 640–643, 2016.

[25] A. Adhikary, X. Lin, and Y.-P. E. Wang, "Performance evaluation of NB-IoT coverage," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2016, pp. 1–5.

[26] A. T. Abusabaha, M. A. Rahman, R. Oliveirab, and A. Flizikowski, "The importance of repetitions in ultra-dense NB-IoT networks," *IEEE Communications Letters*, 2022.

[27] N. Jiang, Y. Deng, M. Condoluci, W. Guo, A. Nallanathan, and M. Dohler, "RACH preamble repetition in NB-IoT network," *IEEE Communications Letters*, vol. 22, no. 6, pp. 1244–1247, 2018.

[28] J.-K. Hwang, C.-F. Li, and C. Ma, "Efficient detection and synchronization of superimposed NB-IoT NPRACH preambles," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 1173–1182, 2018.

[29] Q. Wu, P. Wu, W. Wen, T. Yang, and M. Xia, "An efficient NPRACH receiver design for NB-IoT systems," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10 418–10 426, 2020.

[30] H. Chougrani, S. Kisseleff, and S. Chatzinotas, "Efficient preamble detection and time-of-arrival estimation for single-tone frequency hopping random access in NB-IoT," *IEEE Internet of Things Journal*, vol. 8, no. 9, pp. 7437–7449, 2020.

[31] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Transactions on communications*, vol. 46, no. 4, pp. 431–436, 1998.

[32] T. OShea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, 2017.

[33] T. Gruber, S. Cammerer, J. Hoydis, and S. ten Brink, "On deep learning-based channel decoding," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2017, pp. 1–6.

[34] E. Nachmani, Y. Be'ery, and D. Burshtein, "Learning to decode linear codes using deep learning," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2016, pp. 341–346.

[35] E. Nachmani, E. Marciano, D. Burshtein, and Y. Be'ery, "RNN decoding of linear block codes," *arXiv preprint arXiv:1702.07560*, 2017.

[36] I. Dimnik and Y. Be'ery, "Improved random redundant iterative HDPC decoding," *IEEE Transactions on Communications*, vol. 57, no. 7, pp. 1982–1985, 2009.

[37] S. Cammerer, T. Gruber, J. Hoydis, and S. Ten Brink, "Scaling deep learning-based decoding of polar codes via partitioning," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[38] N. Samuel, T. Diskin, and A. Wiesel, "Learning to detect," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2554–2564, 2019.

[39] N. Farsad and A. Goldsmith, "Detection algorithms for communication systems using deep learning," *arXiv preprint arXiv:1705.08044*, 2017.
[40] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2017.
[41] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *arXiv preprint arXiv:1707.07980*, 2017.
[42] M. H. Jespersen, M. Pajovic, T. Koike-Akino, Y. Wang, P. Popovski, and P. V. Orlik, "Deep learning for synchronization and channel estimation in NB-IoT random access channel," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–7.
[43] Y. R. Kumar and N. M. Balasubramanya, "Deep learning based random access preamble detection for 3GPP NB-IoT systems," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2022, pp. 1689–1694.
[44] T. Kim, D. M. Kim, N. Pratas, P. Popovski, and D. K. Sung, "An enhanced access reservation protocol with a partial preamble transmission mechanism in NB-IoT systems," *IEEE Communications Letters*, vol. 21, no. 10, pp. 2270–2273, 2017.
[45] R. Harwahyu, R.-G. Cheng, C.-H. Wei, and R. F. Sari, "Optimization of random access channel in NB-IoT," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 391–402, 2017.
[46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
[47] 3GPP, "Technical specification group radio access network; evolved universal terrestrial radio access; physical channels and modulation," Tech. Rep. 36.211, 2018, v15.1.0.