Manuscript version: Author’s Accepted Manuscript
The version presented in WRAP is the author’s accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/141134

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
Abstract—In industrial sensor networks, complex industrial environments may be encountered leading to a mix of signals of different types. Complicated interference caused by mixed signals on industrial equipments may significantly degrade the classification rate of signals, which may result in a long training time in order to extract features. In addition, with limited channel resources, it is difficult to make the global optimal decision in industrial distributed wireless sensor networks (IDWSN). To address this problem, a signal classification method using feature fusion is proposed for industrial Internet of things (IIoT) in this paper. In the proposed method, the received signals of nodes are processed by frequency reduction and sampling pretreatment, based on which intelligent representations of signals are obtained. Using federated learning, the data samples are trained with the feature fusion network. Moreover, the trained deep learning network is used on each sensor node to classify signals, the results of which will be transmitted to aggregation center. In the aggregation center, the improved evidence theory method is used to aggregate the recognition results of each sensor node to achieve the final classification. Simulation shows that the proposed method has excellent classification performances. Notably, it is not required for the proposed method to transmit signals from nodes to the aggregation center, which could effectively protect the privacy of industrial information.

Index Terms—Industrial Internet of things, industrial distributed sensor networks, signal classification, feature fusion, federated learning, alpha-stable noise

I. INTRODUCTION

Industrial Internet of things (IIoT), integrate Internet, new-generation information technology and industrial system. In IIoT, a variety of entities and front-ends, such as machines, raw materials, control systems, information systems, products and people, are interconnected. A number of advanced technologies are needed to facilitate efficient and reliable IIoT networks, including the comprehensive deep perception of industrial data, real-time transmission and exchange, rapid computing process and modeling analysis for intelligent control, operation optimization and production organization mode change [1]-[2]. The birth of industrial distributed wireless sensor networks (IDWSN) technology has further promoted the development of IIoT. IDWSN is able to overcome many problems in traditional field bus technologies and industrial Ethernet technologies. Furthermore, the significant advantages of IDWSN, such as massive nodes, the wide-range distribution, multiple data transmission paths, and simple network layout, result in a wide range of applications [3]-[5].

One key and enabling technology of IDWSN is wireless signal classification, which is used to identify the modulation information of wireless signals. Wireless signal classification is essential in many applications, such as signal demodulation, suspicious transmission monitoring, anomaly detection, and interference localization [6]-[7]. In a distributed network, propagation and radio environments vary fast with great randomness. Even for the same signals, transmissions at different times may generate different signals observed at a receiving sensor. Moreover, since all users operate in a distributed fashion without centralized control, interference in such an environment will become uncontrollable and unpredictable. The artificial noise, such as alpha-stable noise, will worsen this issue by causing considerable interference to industrial signals. To achieve the global optimal hypothesis testing, it is required that the complete observation data are transmitted to the primary node without loss. However, the limitation of channel capacity in sensor networks makes this not possible [8]. To cope with this local signal processing can be utilized in a complex industrial environment for the purpose of effective signal classification out of mixed signals with various modulation methods, frequent parameter changes, and strong interference. Artificial intelligence has been widely used in signal classification [9]-[10]. The signal classification leveraging artificial intelligence has been proven effective to address the problems of signal classification.

Deep learning, as a powerful and efficient artificial intelligence technology, has unique technical advantages and could bring in significant performance gains in many application and fields, such as accuracy, efficiency, and robustness [11]-[13]. With deep learning, complex machine learning algorithms can be efficiently realized with the preferable performance. Recently, deep learning has been utilized to develop electro-
magnetic signal recognition methods. In [14], the bispectrum
estimation of the electromagnetic signal and the sparse self
encoder were applied in identifying signals. However, this
method cannot guarantee the global optimization in the consid-
ered model. In [15], a signal recognition method was designed
based on the wavelet fuzzy neural network method to address
the problem caused by various types of noise. Unfortunately,
this method does not perform well in convergence speed and
is easy to end in local extremum. In [16], the performance
of different deep reinforcement algorithms in resource allo-
cations is discussed. Most existing classification algorithms
depend on a strong assumption of the Gaussian white noise,
which is normally unrealistic in industrial applications with
complicated and mixed signals. With alpha-stable noise, the
modulation classification of signal is very challenging. In
[17], a joint estimation algorithm based on the generalized
cyclic spectrum is proposed. In [18], a new method based
on explicit countless cost function and global optimization is
designed. The authors in [19] put forward a modulation type
classification method using sparse signal decomposition (SSD)
of additive mixture Gaussian noise and impulse noise with an
over complete mixture dictionary.

In an environment of industrial distributed networks, the
main challenge exists in the training of learning model, where
nodes have to transmit data for global optimal classification
performance. To support distributed learning system, enor-
mous resources are required for signal processing and data
exchange. However, due to the scarcity of channel resources,
how to use limited resources to effectively execute distributed
learning is a key issue. Different from the existing methods,
where data need to be delivered from nodes to a main server
for training and learning, we will use distributed data that
collected and stored on multiple edge nodes to train machine
learning model as federated learning. With federated learning,
the privacy of users’ personal data can be protected, and
the signal classification could be improved by deep learning
technologies. In view of these tremendous benefits, federated
learning has been deemed as a promising technology in
IDWSN signal modulation classification, vehicular Internet of
Things and other fields [20].

In this paper, a novel wireless signal classification framew-
work leveraging federated learning is proposed. The main
contributions of this paper are summarized as follows:

- We propose a feature fusion network, where the deep
information of intelligent representation is fused with the
shallow information at different levels. The deep feature
fusion structure will also be used as the input of the
shallow feature fusion structure, expanding the amount of
feature information. The shallow feature can encompass
depart of the deep feature information for the recognition
accuracy improvement.

- The neural network based on feature fusion is developed
for the signal detection of distributed sensors. Compared
with existing methods, this structure can automatically
learn and extract features. Moreover, the classification
performance can be improved without affecting the trans-
mission performance.

- Federated learning is employed for learning and recogni-
tion in distributed networks, which solves the problem of
bandwidth limitation and protects the privacy of data. D-
S evidence theory is adopted to aggregate the recognition
for classification performance enhancement.

The rest of this paper is organized as follows. The system
model of signal classification for IDWSN is shown in the
Section II. In Section III, the intelligent representation of
electromagnetic signals is presented. Federated learning on
distributed sensors is developed in Section IV. In section
V, specific experiments are given to verify the classification
performance. Finally, Section VI show the main research
findings of this paper.

II. SYSTEM MODEL

As shown in Fig. 1, system model of signal classification in
IDWSN-based IIoT is considered in this paper, which consists
of one transmitter and multiple parallel sensor nodes. In an
IDWSN, wireless signals are broadcast by transmitters and
sent on parallel channels that experience independent channel
noise. In order to reduce channel resource consumption and
protect privacy, signals need to be processed and classified
locally on sensor nodes, and the wireless signals $s(k)$ received
on distributed sensor nodes can be expressed as

$$s(k) = x(k) + e(k), k = 0, 1, ..., N - 1,$$

(1)

where $e(k)$ represents artificial noise in IDWSN, which can
be described as the alpha-stable noise. Alpha-stable noise is
expressed as characteristic function:

$$\varphi(t) = \exp\{j\delta t - \gamma|t|^{\alpha}[1 + j\beta \text{sgn}(t)\omega(t, \alpha)]\}$$

(2)

where

$$\text{sgn}(t) = \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0, \end{cases}$$

(3)

$$\omega(t, \alpha) = \begin{cases} \tan(\alpha \pi/2), & \alpha \neq 1, \\ (2/\pi) \log |t|, & \alpha = 1, \end{cases}$$

(4)

$0 < \alpha \leq 2$ stands for the characteristic index, $\gamma \geq 0$ is the dis-
ersion parameter and $-1 \leq \beta \leq 1$ is the index of skewness,
$\delta$ represents the location parameter. $x(k)$ are wireless signals,
which are comprised of amplitude modulation (AM) signal,
frequency modulation (FM) signals, binary phase shift keying
(BPSK) signals, quadrature phase shift keying (QPSK) signals,
8 phase shift keying (8PSK) signals, 2 amplitude shift keying
(2ASK) signals, 4 amplitude shift keying (4ASK) signals, 2

Fig. 1: System model of signal classification in IDWSN-based
IIoT.
frequency shift keying (2FSK) signals and 4 frequency shift keying (4FSK) signals.

In the analog modulation type, the modulation signal can be expressed as

\[
s(t) = A \cos[2\pi f_c t + \phi(t) + \phi_0],
\]

(5)

where \( A \) represents the instantaneous amplitude of the signal, \( f_c \) stands for the carrier frequency, \( \phi_0 \) denotes the modulation phase and \( \theta \) is the carrier initial phase. AM modulation signal can be expressed as

\[
A = m_0 + m_t,
\]

(6)

where \( m_t \) denotes the baseband modulation signal and \( m_0 \) represents the DC component. FM modulation signal is expressed as

\[
A = 1.
\]

(7)

For digital modulation signal, whose baseband waveform can be expressed as follows

\[
s(t) = \sum_n a_n g(t - kT),
\]

(8)

where \( a_n \) is the symbol sequence sent by the transmitter, \( g(t) \) stands for the equivalent filter including shaping filter, channel filter and matching filter. Different types of modulation have the different symbol sequence renderings. MPSK signals can be expressed as

\[
s(t) = A e^{j(2\pi f_c t + \phi_i)},
\]

(9)

where \( \phi_i \) denotes the phase modulation function, which is given by

\[
\phi_i = \frac{2\pi i}{M}, \quad i = 0, 1, \ldots, M - 1,
\]

(10)

MASK signals are expressed as follows

\[
s(t) = A e^{j(2\pi f_c t)},
\]

(11)

MFSK signals can be expressed as

\[
s(t) = e^{j(w_c t + 2\pi f_i t)},
\]

(12)

where \( f_i \) denotes the modulation frequency.

III. INTELLIGENT REPRESENTATION OF WIRELESS SIGNALS

A. Generalized Envelope Square Spectrum

Since the alpha-stable noise does not have the second-order or higher-order statistics, the signal disturbed by alpha-stable noise do not have the effective envelope square spectrum. The main reason is that there is a large impulse pulse, resulting in a large amplitude in the disturbed signal, so it is necessary to preprocess the signal by

\[
f(x) = \frac{\left(1 + e^{-|x|\alpha} H(x(t))\right) - 1}{|x(t)| + j H(x(t))},
\]

(13)

where \( H(.) \) denotes the Hilbert transform. The signal of any point can be written as \( x(t) = r \cos \theta \), so we can obtain

\[
f(r \cos \theta) = \frac{\left(1 + e^{-\frac{r}{|e^{\theta}|}}\right) - 1}{|r e^{\theta}|} = \left(\frac{2}{1 + e^{-r}} - 1\right) \cos \theta,
\]

(14)

where the value range of \( \left(\frac{2}{1 + e^{-r}} - 1\right) \) is \([-1, 1]\). Therefore, the above functions can map the amplitude of the processed signal to a range of \([-1, 1]\), and do not change the phase information of the signal.

The square of the processed signal envelope can be expressed

\[
u(t) = a^2(t) = f(x)^2 + H[f(x)]^2.
\]

(15)

Using classical spectrum estimation to estimate the power spectrum of the signal as follows

\[
P(\omega) = \frac{1}{N} |U_N(e^{j\omega})|^2,
\]

(16)

where \( U_N(e^{j\omega}) \) is Fourier transform of the \( N \)-point observation data of \( U(t) \).

B. Fractional Lower Order Cyclic Spectrum

Fractional low-order moment is a powerful tool for analyzing and processing non-Gaussian signals. If the characteristic index of random signal is alpha, the fraction low-order moment of the signal \( x(t) \) with alpha-stable noise is defined as

\[
E[|X|^p] = C(p, \alpha) \gamma^{p/\alpha},
\]

(17)

where \( p \) represents the fractional factor, whose value range is \( 0 < p < \alpha \leq 2 \). \( C(p, \alpha) \) is a constant related to \( p \) and \( \alpha \). There are two methods of calculation for signal cycle spectrum based on low-order fractional moments as follows: one is based on covariance and the other is based on the low-order covariance of fractions. On the basis of fractional lower moment, the definition of \( p \)-order covariance of \( x(t) \) can be expressed as

\[
COV_{x,p}(t, \tau) = E[x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})^{(p-1)}],
\]

(18)

where \( \tau \) denotes the time delay, \( p \) stands for the order factor, and its value range is \([1, \alpha]\). \( \alpha \) is the characteristic index of alpha-stable noise, and its value range is \([1, 2]\). If \( COV_{x,p}(t, \tau) \) is a periodic function of \( t \), it is expanded into a Fourier series, and the coefficient of the Fourier series is the fractional low-order cyclic autocorrelation function of the signal, which is expressed as

\[
R^{\varepsilon}_{x,p}(\tau) = \left(\frac{x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})^{(p-1)} e^{-j2\pi \varepsilon \tau}}{2}\right),
\]

(19)

where \( \varepsilon \) represents the cyclic frequency of order \( p \). By performing the Fourier transformation of \( R^{\varepsilon}_{x,p}(\tau) \), we can obtain
the fractional low-order cyclic spectral density function of signal $x(t)$ as
\[ S_x^r(p) \triangleq \int_{-\infty}^{+\infty} R_{x,p}(\tau)e^{-j2\pi f\tau}d\tau, \]
where $f$ stands for the normal frequency. Obviously, when the order factor $p = 2$, the fractional lower order cyclic spectral density function is the second order cyclic spectral density function.

The fractional lower order cyclic spectrum of MASK signal can be expressed as follows
\[
S_x^r = \begin{cases} 
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})] + Q(f - f_0 + \frac{\varepsilon}{2})Q^*(f - f_0 - \frac{\varepsilon}{2}), & \varepsilon = m/T \\
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})], & \varepsilon = -2f_0 + m/T \\
\frac{E^p}{T^p}[Q(f - f_0 + \frac{\varepsilon}{2})Q^*(f - f_0 - \frac{\varepsilon}{2})], & \varepsilon = 2f_0 + m/T \\
0, & \text{other}
\end{cases}
\]
and the fractional lower order cyclic spectrum of MFSK signal can be expressed as
\[
S_x^r = \begin{cases} 
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})] + Q(f - f_0 + \frac{\varepsilon}{2})Q^*(f - f_0 - \frac{\varepsilon}{2}), & \varepsilon = m/2T \\
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})], & \varepsilon = \pm 2f_0 + m/2T \\
Q^*(f + f_0 + \frac{\varepsilon}{2})e^{-j2\pi(\alpha+2f_0)}, & \varepsilon = \Delta = \pm 2f_0 + m/2T \\
0, & \text{other}
\end{cases}
\]

and the fractional lower order cyclic spectrum of MPSK ($M \geq 4$) signals can be written as
\[
S_x^r = \begin{cases} 
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})] + Q(f - f_0 + \frac{\varepsilon}{2})Q^*(f - f_0 - \frac{\varepsilon}{2})e^{-2j\phi_0}, & \varepsilon = \pm 2f_0 + m/T \\
\frac{E^p}{T^p}[Q(f + f_0 + \frac{\varepsilon}{2})Q^*(f + f_0 - \frac{\varepsilon}{2})] + Q(f - f_0 + \frac{\varepsilon}{2})Q^*(f - f_0 - \frac{\varepsilon}{2})e^{2j\phi_0}, & \varepsilon = \Delta = \pm 2f_0 + m/T \\
0, & \text{other}
\end{cases}
\]

IV. SIGNAL CLASSIFICATION BASED ON FEATURE FUSION AND FEDERATED LEARNING IN IDWSN

A. Feature Fusion Based on DenseNet

In traditional CNN layers, input $x$ is mapped to $F(x)$, and then $F(x)$ is used to fit the target distribution. But ResNet connects the input directly to the output layer through a bypass between the input and output, so that the object to be fitted by the layer changes from $F(x)$ to $G(x)$. In abstract, such a connection makes the layer only need to fit 0, which greatly reduces the problem of gradient disappearance of the deep network in the learning process. DenseNet goes further and lead into the concept of Densenlock. In a DenseBlock module, the input of each layer comes from the output of all previous layers of this layer. The basic structure of DenseNet is shown in the Fig.6.
The feature matrix of size $\times 224 \times 224 \times 1$ is put into the feature fusion network, and set the number of convolution kernel channels K, convolution layers N in DenseBlock and compression factor $\theta$ in Transition Layer;

2: Carry on $7 \times 7$ conv with step 2 and $3 \times 3$ maxpool with step 2;

3: The output features $x$ obtained in the second step are through the structure of DenseBlock1-Transition Layer1-DenseBlock2-Transition Layer2-DenseBlock3-Transition Layer3-DenseBlock4. The output of the $ith$ Transition Layer is recorded as $x_i$;

4: Fuse $x$ and $x_3$. Deconvolution the fusion results to the same size with $x_2$ and fuse them. Then deconvolution the fusion results with $x_1$ and fuse them to get the final fusion features;

5: $[7 \times 7]$ global average pool and 1000D full-connected;

6: Use SoftMax classifier to get finally result.

**B. Signal Classification Based on Federated Learning**

In a distributed sensor network, the unknown transmission data sequence is broadcast by the transmitter and transmitted on a parallel channel experiencing independent channel noise.
It is assumed that each sensor has the same number of observations and all the sensors in the network collect and process the noisy data series at the same time. Due to the variations of propagation and transmission environments, even if the transmitter sends the same signal, different signals may be observed on the receiving sensor. To reach the global optimization of hypothesis testing, the complete sensor observation data are required to be collected by the master node. This is a strong and unrealistic assumption, as limitation of channel capacity in sensor networks makes the master node impossible to obtain a complete set of original observations. In order to reduce the requirement of channel bandwidth, local sensors are deployed to have the relevant chip, so that the signal processing can be completed locally. The following details the federated learning on the distributed sensor network, which is divided into training and testing. The structure of federated learning process in IDWSN is shown in Fig.10.

The data is sent out by the transmitter and reaches the sensor in the network through the parallel channel. Each sensor node has set up the neural networks but has not been trained. After each sensor receives the signal, its intelligent representation is calculated and put into network training. Different from the traditional training, after getting the loss function of this training, we do not rush to the next step of gradient optimization but send the loss function of each node to the aggregator. Then aggregate the results according to some rules and return them to each node, and then each node uses the global loss function for gradient optimization to achieve the goal of global aggregation. Then it repeats until the global loss function reaches an ideal state. This can not only improve the quality of the model, reduce the channel resource occupation, but also effectively protect the privacy of the data.

The loss function and the batch size of all sensors are sent to the aggregator for aggregation and the aggregator receives the data of $N$ sensors. According to the weighted average, the loss function after aggregation can be expressed as follows:

$$L_{\text{loss}} = \frac{\sum_{i=1}^{N} L_{\text{loss},i} \times b_i}{\sum_{i=1}^{N} b_i},$$  

where $L_{\text{loss}}$ is the loss function after aggregation, $L_{\text{loss},i}$ is the loss function of the $ith$ sensor and $b_i$ is the batch size of the $ith$ sensor. These results are sent back to each sensor. After receiving the aggregator’s return, each sensor uses the global loss function to continue to train the local network, and repeats until the loss function tends to be stable or the accuracy reaches the ideal value.

Each sensor uses gradient descent to update the network locally, then the local update in the $ith$ sensor is carried out as follows:

$$w_i(t) = w_i(t-1) - \eta \nabla F_i(w_i(t-1)),$$

where $\eta > 0$ is the learning step, and the model parameter $W$ is updated along the negative gradient direction following $F_i$. In [22], this method has been proved to have global convergence and good convergence performance for convex optimization problems.

From the above analysis, the training process of feature learning are given by algorithm 2.

**Algorithm 2** Training process of feature learning.

1: The received signal is intelligently represented and obtain the two channel characteristic matrix;
2: The feature matrix is used to train the neural network locally. After get the cross entropy loss function of each batch, send the loss function to the master node;
3: The main node merges the loss function from each node and returns the result to each node;
4: The node uses the received loss function for gradient descent training to complete the network training.
so their reliability matrix is

\[
\chi = \begin{bmatrix}
1 & \chi_{12} & \cdots & \chi_{1N} \\
\chi_{21} & 1 & \cdots & \chi_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\chi_{N1} & \chi_{N2} & \cdots & 1
\end{bmatrix}.
\]  

(30)

From the above matrix, we can calculate that the credibility of the evidence of sensor node \( i \) to other nodes is

\[
Sup(M_i) = \sum_{j=1, j\neq i}^{n} \chi_{ij}.
\]  

(31)

Normalize the credibility one by one, then we can get their respective weights as follows

\[
\omega_i = \frac{Sup(M_i)}{\sum_{j=1}^{N} Sup(M_j)}.
\]  

(32)

Finally, using the above weight to fuse the evidence, the probability \( m(\theta_i) \) of each classification is

\[
m(\theta_i) = \sum_{j=1}^{N} \omega_j \times m_j(\theta_i)
\]  

(33)

where \( M = \{m(\theta_1), m(\theta_2), ..., m(\theta_{12})\} \) denotes the final classification result.

The signal classification algorithm based on federated learning in IDWSN-based IIoT is summarized in Algorithm 3.

**Algorithm 3 Signal classification based on federated learning in IDWSN.**

1. The received signal is intelligently represented, and obtain the two channel characteristic matrix;
2. The feature matrix is input into the trained feature fusion network for local recognition. Use SoftMax classifier to obtain the set vector of basic probability \( M_i = \{m_i(\theta_1), m_i(\theta_2), ..., m_i(\theta_{12})\} \), and send it to the main node;
3. After the master node receives the set vector of basic probability distribution from each node, use D-S evidence theory to get the weight of each node vector;
4. Fuse set vector of basic probability distribution by weight and obtain the classification results.

The details of Algorithm 3 can be discussed as follows: main node is to receive the set vector \( M_i = \{m_i(\theta_1), m_i(\theta_2), ..., m_i(\theta_{12})\} \) of independent basic probability distribution of evidence generated from the classification results of each sensor, and then use (28) to obtain the trust degree between each node and form the reliability matrix shown in (29), and combine (30) and (31) to obtain the weight of each result to get the final classification results.

V. SIMULATION RESULTS AND ANALYSIS

To demonstrate the effectiveness and superiority of the proposed method, simulation experiments are conducted in this section. Nine types modulation signals are considered, including AM, FM, BPSK, QPSK, 8PSK, 2ASK, 4ASK,
TABLE I: Classification performance with the different number of DenseBlocks and revolution layers.

<table>
<thead>
<tr>
<th>DenseBlocks</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>82.87%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>83.5%</td>
<td>85.4%</td>
<td>86.29%</td>
<td>88.9%</td>
<td>88.99%</td>
<td>88.47%</td>
</tr>
<tr>
<td>5</td>
<td>83.56%</td>
<td>85.57%</td>
<td>86.46%</td>
<td>88.21%</td>
<td>88.24%</td>
<td>88.77%</td>
</tr>
</tbody>
</table>

2FSK and 4FSK. The simulation parameters of nine types of wireless signals used in this simulation are: frequencies of 2FSK signals are 20MHz and 40MHz, frequencies of 4FSK signal are 15MHz, 25MHz, 35MHz and 45MHz, respectively. The carrier frequency of other signals is 30MHz. The pulse width is set to 10 and the sampling frequency is 120MHz. The number of samples used in training and testing of each type of signals are 10000 and 1000, respectively.

Tab. 1 shows the classification performance of signals under different DenseBlocks and revolution layers. It should be noted that the table shows the number of layers in the first DenseBlock, and the number of layers in the second, third, fourth and fifth DenseBlocks is 2, 6, 9 and 8 times of the first. From Tab. 1, when the number of DenseBlock is 4 and there are 6 layers in the first one, the classification performance tends to be stable. Increasing the number of neurons will not significantly increase the performance, or even slightly decrease.

Using the network structure with the best performance obtained above, this neural network has four DenseBlocks and the number of convolution layers of each DenseBlock is 6, 12, 36 and 54. Fig. 12 shows the classification performance of different types wireless signals versus different GSNRs. According to the Fig. 12, the average classification rates are greater than 85% when the GSNR is greater than 5 dB. When the GSNR is more than 10 dB, the classification rates can reach more than 90%, so the proposed signal classification method is effective and feasible.

Fig. 13 shows the classification performance with different alpha-stable noise characteristic parameter $\alpha$ values. From the Fig. 13, the classification performance will gradually improve with the increase of $\alpha$. However, the classification performance improvement is exceedingly small when $\alpha$ is greater than 1.5. The proposed method has better classification performance for
non-Gaussian noise and is also robust to the characteristic parameter $\alpha$.

Fig. 14 shows the classification performance with different $\beta$ and $\delta$. From Fig. 14, $\delta$ has a great impact on classification performance. With the increasing of $\delta$, the accuracy of classification is decreasing, and the decreasing speed is faster. We also carried out the classification performance with different roll-off factors, and the results are shown in Fig. 15. From Fig. 15, we can see that the roll-off factor has a great impact on the classification performance under low GSNR, but the impact is smaller with the increasing of GSNR. When GSNR is greater than 10 dB, the effect of roll-off coefficient has been reduced to less than 5%.

Under the same simulation environment and parameter settings, we compare the classification performance of the proposed method in Gaussian noise environment with that of the method based on ENN in [21] and the comparison results are shown in Fig. 16. From Fig. 16, we can see that the proposed method has better performance than that in [21] when the SNR is less than 5 dB. In addition, the proposed method is suitable for signal classification in Gaussian noise environment.

In the same simulation environment and the same signal parameter settings, feature fusion network based on Dense-Blocks is compared with the method in [19] and DenseNet without feature fusion with $\alpha = 1.2$, and the comparison results are shown in Fig. 17. The average classification rate of the proposed method is significantly higher than other methods when the GSNR is less than 10 dB. The computational complexity of the proposed method as follows: the complexity of intelligent representation is $O(KN \log N)$, where $K$ is the number of data segmentation; the complexity of federated learning is $O(KM)$, where $K$ denotes the total number of global aggregation executions and $M$ represents the number of nodes, respectively; the complexity of the feature fusion network is $O \left( \sum_{l=1}^{D} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l \right)$, where $D$ stands for the depth of the network, $M_l$ denotes the length of the output feature side of the $l$th layer, $K_l$ represents the size of the convolution core of the $l$th layer, and $C_l$ is the number of output channels of the $l$th layer.

VI. Conclusion

To cope with complicated wireless signals in industrial Internet of things, an intelligent signal classification framework is developed in this paper. To extract accurate features that can fully represent the characteristics of wireless signals, the generalized envelope square spectrum and the fractional low-order cyclic spectrum are obtained through the intelligent representations of wireless signals. Then, the feature fusion neural network based on DenseNet is used to classify the signals locally on sensors receiving signals. Furthermore, the federated learning is used to fuse the learning process and recognition results of each sensor node in order to recognize the modulation recognition of signals. Finally, extensive simulation studies have been carried out to verify the effectiveness of the proposed framework. Simulation results show that the proposed framework possesses an excellent signal classification performance in industrial distributed wireless sensor networks, which significantly outperform that of the existing methods using sparse signal decomposition. In this paper, the aggregation frequency has optimization space. It can be considered to dynamically adjust the aggregation frequency according to some rules, to further reduce resource consumption without affecting performance.

REFERENCES


