Recognition of Multifunction Radars Via Hierarchically Mining and Exploiting Pulse Group Patterns

ZHANG-MENG LIU
National University of Defense Technology, Changsha, China

Recognition of multifunction radar (MFR) is an open problem in the field of electronic intelligence. Parameters of MFR pulses are generally agile and difficult to distinguish statistically. A prospective way to realize credible MFR recognition is mining and exploiting more distinguishable high-dimensional patterns buried in pulse groups, which may be designed for implementing infrequently used radar modes such as target tracking. A high-dimensional pattern is defined according to the agile range and switching law of sequential pulse repetitive intervals within a pulse group. This article establishes deep recurrent neural networks (RNN) to discriminate and coarsely cluster different pulse groups hierarchically with respect to their sequential structures. Afterwards, RNN-based classifiers are trained to extract and exploit features within different pulse group clusters. Distinct degrees of confidence are then attached to these classifiers to indicate the discriminabilities of the corresponding pulse group clusters. The pulse group clustering and classifying models are finally cascaded to form an integrated classification model, which mines distinguishable patterns from sequentially arriving pulse groups of the same radar and accumulate them to realize MFR recognition. Simulation results demonstrate the much improved performance of the proposed method over existing counterparts in different scenarios.

Manuscript received June 22, 2019; revised March 7, 2020; released for publication May 19, 2020. Date of publication June 2, 2020; date of current version December 4, 2020.

DOI. No. 10.1109/TAES.2020.2999163

Refereeing of this contribution was handled by K.D. Pham.

This work was partially supported by Hunan Youth Talent Program under Grant 2019RS2026, Provincial Outstanding Youth Fund of Hunan and Provincial Innovation Group Program of Hunan.

Author Address: Zhang-Meng Liu is with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE), National University of Defense Technology, Changsha 410073, China, Email: (liuzhangmeng@nudt.edu.cn).

I. INTRODUCTION

Multifunction radars (MFR) are widely used in civil and military areas [1]–[4], they usually have multiple modes such as storm surveillance and tracking [5], target searching and tracking [6], [7]. Each radar mode is implemented with pulse groups that may have diversed discrete temporal structures [8]. MFRs have a high level of agility in transmitted waveform and antenna beam steering. This flexibility makes it possible for the radar to change operation strategies on the fly, as is shown in Fig. 1. MFRs can also modify their transmitted waveforms from burst to burst over a wide range of parameters, including carrier frequency, pulsewidth, and pulse repetitive interval (PRI). From the perspective of noncooperative receivers, the strong directionality of MFRs significantly reduces the probability of intercepting radar waveforms radiated to directions other than the receiver. Take the scenario in Fig. 1 for example, an airborne receiver can easily capture the radar signals used to search or track its carrier aircraft, but it will miss the signals radiated by the other two MFR beams with high probabilities. The miss probability depends on the sidelobe level of the oriented beams along the direction of the receiver. As a result, noncooperative observation data of agile MFRs are highly incomplete. They usually contain relatively isolated pulse groups with different amplitudes. Passive recognition of MFRs based on rapidly changing and severely incomplete observation data thus becomes a very demanding task in the field of electronic intelligence [9].

Two of the main MFR tasks are target searching and tracking, and some other tasks such as target identification [7], recognition [3], and weapon guidance [5], [6] may also be included. Each radar mode is implemented with pulse groups consisting of temporally discrete pulses [8]. The PRI pattern of each pulse group is task dependent. For example, equi-PRI pulse groups with PRIs falling within different ranges are exploited to perform target searching tasks at different altitudes in fire-control radars and airborne pulse-doppler radars [3], and stagger- or jitter-PRI pulse groups are used for ambiguity elimination and jamming deception [1]–[3]. Some tasks such as searching are regularly performed by MFR, the corresponding pulse groups with relatively simple patterns of equi-PRI make up a large part in noncooperatively observed signals. On the contrary, pulse groups of some rarely implemented tasks such as target tracking and weapon guidance take small proportions in observed signals [6], but these pulse groups contain more distinguishable patterns between different MFRs.

As MFRs usually vary their signal parameters over wide ranges, it is very difficult to recognize which class they belong to according to the statistical characteristics of their pulses. Canadian researchers proposed to model MFRs operating modes with a syntactic model, so that temporal patterns of successive pulses can be described in a more compact way [7], [10]. This is an early attempt to apply automatic technologies to process noncooperative radar signals. However, the establishment of the syntactic model requires a detailed description of MFR, which is

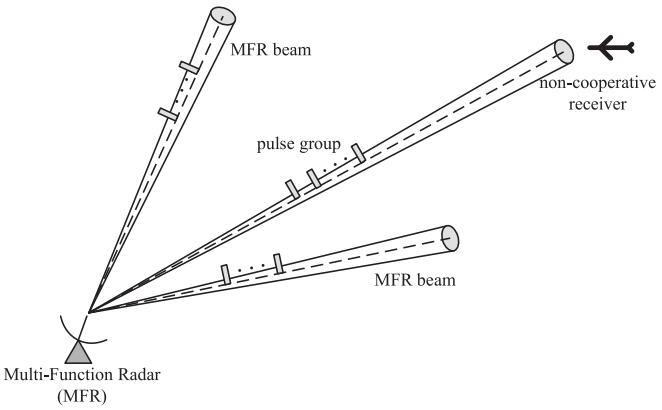

Fig. 1. MFR have high flexibilities of beam steering and waveform modulation.

hardly available from the perspective of noncooperative radar intelligence [11], [12].

Another major difficulty in MFR recognition is the temporal discreteness of pulse groups. Both the ordering of multiple PRIs and the value of each discrete PRI contribute to the pattern of a pulse group. However, existing machine learning techniques are seldom designed for extracting and exploiting patterns from multidimensional and analog-valued PRI sequences, which may be further contaminated by data noises such as losing pulses.

Deep learning techniques have attracted much research interest in various fields in the past few years [13]–[15]. The data in some of the fields has similar multidimensional and discrete forms as pulse groups, such as word sequences in sentiment analysis, machine translation, and text comprehension [16], [17]. Some recurrent neural networks (RNNs) have been specially designed for processing word sequences and they have achieved great successes in related applications. The other deep learning techniques, e.g., convolutional neural networks, have fixed-size frameworks, they do not well fit such tasks due to various reasons including the variable-length sequential patterns in data sequences.

The author of this article has introduced RNN to classify radar pulse trains in a previous work [18]. RNN extracts high-dimensional patterns buried in pulse trains to realize radar classification, and the method achieves satisfactory performances in cases of agile pulse parameters and significant observation noises [18]. But this method may not work for MFRs. That is because deep learning methods usually have strong statistical tendency, i.e., they incline to learn from and match predominant modes in the given dataset. However, as is described abovementioned, there is significant imbalance in observed MFR signals. The imbalance is defined between pulse groups of regularly used MFR modes such as target searching and that of infrequently used modes such as target tracking. The former signals take a large proportion in noncooperatively received dataset, but they may contribute very little to MFR recognition due to their lack of distinguishable patterns. On the contrary, target tracking pulse group take a small proportion but contains more abundant patterns available for MFR recognition. The contradiction between imbalanced MFR data and the statistical tendency of deep learning techniques blocks the application of the method in [18] in MFR recognition. It should be noted that, the classification problem based on imbalanced data is also an open problem in the field of computer science [19]–[22].

This article addresses the recognition problem of MFR by hierarchically mining and exploiting high-dimensional patterns of pulse groups associated with different radar modes. The recognition is implemented among more than one MFR classes, with radars belonging to each class having the same modes and different classes of radars have somewhat diversed modes. By considering that frequently and infrequently used modes of MFR may differ largely in discriminability and take much differentiated proportions in noncooperatively observed signals, this article trains a series of RNNs hierarchically to cluster the pulse groups blindly and extract sequential PRI patterns from them to realize MFR recognition. The clustering process helps to separate pulse groups in original datasets successively, so that pulse groups associated with infrequently used modes but containing more abundant distinguishable patterns will be unfolded to enhance the availability of recognition. A series of classifiers are then trained with respect to the pulse group clusters and cascaded to form an integrated classification model. During performance testing, the integrated model is exploited to extract and accumulate patterns from sequentially arriving pulse groups to recognize MFR.

The contributions of this article are mainly threefold.

1) The problem of MFR recognition is studied by taking data imbalance between different radar modes into account, which is common in practical electronic intelligence systems but has been rarely considered in previous literature.
2) Unsupervised clustering and supervised classification techniques are developed synthetically to mine and exploit multidimensional patterns from pulse groups of infrequently used radar modes, which are deeply hidden beneath mixing observations.
3) MFR recognition is realized by extracting patterns and accumulating recognition information from pulse streams instead of static datasets, thus, the proposed method meets practical online application requirements well.

The rest of the article consists of five parts. In Section II, a noncooperative observation model of MFR signals is presented, some closely related terms are defined, and the availability of deep learning techniques in MFR recognition is roughly discussed. In Section III, a new integrated recognition model is established via hierarchically mining and exploiting pulse group patterns. Section IV introduces a method for MFR recognition based on pulse group streams using the established model. Section V verifies the performance of the proposed method via simulations. Section VI summarizes the article.

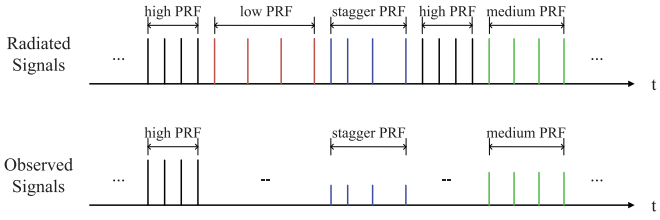

Fig. 2. Illustrative example of the relation and difference between radiated and observed signals.

II. OBSERVATION MODEL OF MFR SIGNALS

MFR sends pulse signals to space to perform various modes such as target searching and tracking. Each pulse has parameters of carrier frequency, pulsewidth and amplitude, and pulse-to-pulse intervals within a pulse group form a PRI sequence. A PRI is defined as the differential time-of-arrival (DTOA) between adjacent pulses. The ordering of PRIs contains sequential patterns of a pulse group, it is highly related with radar modes, e.g., the regularly used target searching mode is usually implemented with equal PRIs [3], whereas the infrequently used target tracking mode is implemented with stagger or jitter PRIs [1]–[3]. Moreover, pulse groups consisting of PRIs valued in different ranges are used to search or track targets on different ranges or altitudes [3].

In this article, agile radars, which perform sensing actions to either detect previously undetected targets or to update track estimates for known targets, are taken as MFR examples. The radiated signals of such radars form a sequence of pulse groups as follows:

$$\begin{aligned} &\ldots, \{pri_1^{(i)}, pri_2^{(i)}, \ldots, pri_{L_i}^{(i)}\}, \\ &\ldots, \{pri_1^{(j)}, pri_2^{(j)}, \ldots, pri_{L_j}^{(j)}\}, \ldots \end{aligned} \tag{1}$$

where the PRIs within each brace belong to the same pulse group, $pri_l^{(i)}$ represents the lth PRI of the ith pulse group, and L_i represents the total number of PRIs in the ith pulse group. In this model, the other pulse parameters, such as carrier frequency, pulsewidth, and amplitude, are ignored to simplify expression. These parameters can be included in the problem via straightforward model expansion as that in [18].

From the perspective of noncooperative systems, only part of MFR signals are received with diversed amplitudes due to the strong directivity of radar antenna, while the other pulse groups are lost as a whole as their amplitudes are lower than the receiver sensitivity. The time-of-arrival of each radar pulse is measured, and the DTOA between adjacent pulses within each pulse group is calculated and recorded as PRIs. In Fig. 2, five successive pulse groups are shown as an example to illustrate the relation and difference between radiated and observed signals. The MFR is assumed to search and track multiple targets by steering its beam electronically. Target searching tasks are performed with equi-PRI pulse groups, and the pulse repetition frequency (PRF, the reciprocal of PRI) takes values in multiple ranges for targets at different altitudes. Take a fire-control MFR for example, it uses low PRF pulse groups to look up, and medium or high PRF pulse groups to look down [3]. Target tracking tasks are performed with pulse groups consisting of stagger PRIs. In Fig. 2, five pulse groups are shown in the radiated pulse stream, they are associated with searching or tracking tasks at different altitudes. Two of the pulse groups, i.e., the second and fourth ones, are missed as they have very low level at the receiver, and the other three pulse groups are intercepted with different amplitudes. Interpulse PRIs of intercepted pulse groups are well preserved from radiated signals to observed data, and adjacent pulse groups are weakly related as they perform isolated tasks. The observed data are expressed as follows:

$$\begin{aligned} &\ldots, \{pri_1^{(i)}, pri_2^{(i)}, \ldots, pri_{L_i}^{(i)}\}, \\ &\{pri_1^{(j)}, pri_2^{(j)}, \ldots, pri_{L_j}^{(j)}\}, \ldots \end{aligned} \tag{2}$$

Compared with the radiated signal stream in (1), the observations of the ith and jth pulse groups in (2) assumes that the pulse groups between them are lost due to poor observation conditions. A high rate of pulse missing greatly weakens the temporal correlation between adjacent pulse group observations, and only sequential patterns within the pulse groups are retained. These roughly independent pulse groups can be separated from each other in the observations according to their parameters such as amplitude. This article studies how to use these relatively independent pulse groups to recognize MFRs.

The observed pulse groups can be categorized coarsely into two types according to their frequency of usage. They are denoted by $\boldsymbol{X}$ and $\boldsymbol{Y}$ as follows:

$$\begin{aligned} \boldsymbol{X} &= \{[pri_1^i, pri_2^{(i)}, \ldots, pri_{L_i}^{(i)}]\}_{i=1,\ldots,N_1} \\ \boldsymbol{Y} &= \{[pri_1^i, pri_2^{(i)}, \ldots, pri_{L_i}^{(i)}]\}_{i=N_1+1,\ldots,N_1+N_2} \end{aligned} \tag{3}$$

where N_1 and N_2 denote the number of pulse groups in the two categories, respectively. For many MFRs, the pulse groups in $\boldsymbol{X}$ correspond to regular target searching tasks and take a large proportion in observed signals, whereas these in $\boldsymbol{Y}$ correspond to infrequently used target tracking tasks and take a much smaller proportion in observed signals, i.e., $N_1 \gg N_2$.

In order to facilitate further discussions, some terms closely related to the concerned problem are explicitly defined as follows.

MFR mode: Predefined functions of MFR. Two ubiquitous modes, i.e., target searching and tracking, are usually contained in MFR mode list. Some other functions such as weapon guidance and jamming deception may also be included. In this article, only the two ubiquitous modes are considered.

Pulse group: Several successive pulses radiated by a MFR at a beam position to perform particular modes, such as target searching or tracking. Pulses within the same group have similar amplitudes, thus they are generally detected or missed as a whole by the receiver. Therefore, the observed data consists of a series of pulse groups with diversed intergroup DTOAs.

Pulse group pattern: The value range, ordering and switching law of agile PRIs within a pulse group, and the particular values of the PRIs are excluded in a broad definition of pulse group pattern. It is assumed that pulse groups corresponding to the same MFR mode share identical patterns, and pulse groups corresponding to different MFR modes have distinct patterns. The patterns of pulse groups containing more than one PRI is also termed as high-/multidimensional patterns.

Data imbalance: Pulse groups of regularly used modes such as target searching and these of infrequently used modes such as target tracking take much diversed proportions in observed data. The number of regular searching pulse groups is larger or even overwhelming over that of infrequent pulse groups. This imbalance greatly increases the difficulty of MFR recognition, since stagger or jitter PRI sequences of infrequent pulse groups generally contain more distinguishable patterns, but these patterns are concealed beneath the large number of regular and indistinguishable pulse groups used for target searching.

Sequential recognition: The noncooperatively received pulse groups are assumed to arrive one by one. The pulse stream of a MFR has been segregated from noisy signals according to emitter location, and it has been separated into pulse groups according to signal amplitudes. The observed pulse groups are not tagged with radar mode labels as they are not known at noncooperative receivers. The receiver can hardly recognize a MFR correctly based on a single pulse group, it extracts distinguishable features from sequential pulse groups and accumulate them to improve recognition probability gradually.

Noncooperative pulse observations of MFR usually have agile parameters and high pulse missing rates. It is very difficult to directly extract reliable statistical parameters from the observations to recognize MFRs. For example, if a MFR has two searching modes performed with equi-PRIs of AAA and BBB (A and B stand for particular PRIs), together with a tracking mode performed with stagger-PRIs of AAB, while another MFRs have the same searching modes but a slightly different tracking mode performed with stagger-PRIs of ABB. The two MFRs cannot be distinguished from each other from a statistical perspective unless the high-dimensional pattern in the tracking mode is mined and exploited.

As pulse group patterns are usually consist of multiple discrete PRIs, they are very difficult to be handled by machines, which prefer binary data and continuous vectors, in their original forms. RNNs have shown their superiority in processing discrete sequences. Typical applications include sentiment analysis, machine translation, and text comprehension [16], [17]. Words are basic information units in these applications, they have discrete one-hot formulations in a word dictionary. RNN first transform discrete words to continuous vectors, and then process phrases or sentences as a vector stream. Some specially designed neural networks are able to extract and remember high-dimensional patterns in data sequences [16], [17].

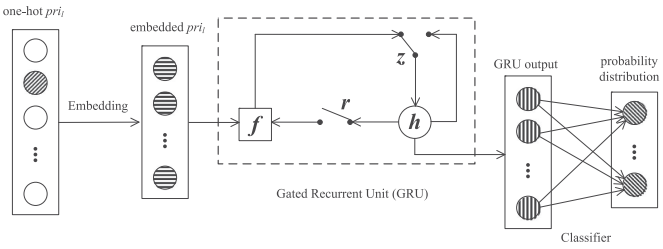

Fig. 3. Deep neural network structure used for radar classification in [18].

At first sight, the PRI sequences in pulse groups have similar formulations as word sequences. Both of them are consist of multiple discrete values. Therefore, the RNN is a prospective technique for processing radar pulse sequences. In [18], a RNN is established to extract high-dimensional temporal features of the pulses, and a deep learning-based method is proposed for radar recognition. The deep learning model used in [18] is shown in Fig. 3, which includes modules of PRI quantization, vectorization, recurrent processing, and classification. A gated recurrent unit (GRU), which has shown superior performances in many sequence processing applications [23], is used for recurrent processing of PRI sequences, so as to extract and exploit sequential patterns in radar pulse streams. The GRU model, instead of the long-short term memory one [15], is chosen because it has fewer parameters and performs comparably with the latter.

The model and method proposed in [18] can well extract high-dimensional sequential patterns hidden in pulse trains with agile parameters, and have shown satisfactory performances in recognizing radars that are hardly distinguishable according to their statistical parameters. They are also robust to various nonideal factors including data noises and small pulse numbers. However, the regularly used target searching mode of MFRs is usually implemented with equi-PRI pulse groups. Such PRI sequences are hardly distinguishable between different MFRs if their PRIs have the same or overlapped ranges. Therefore, it is very difficult to recognize MFRs according to their overall sequential patterns. Different from the regular modes, infrequently used radar modes such as target tracking are implemented with pulse groups having more distinguishable patterns, such as stagger or jitter PRI sequences [1]–[3]. These pulse groups are expected to contribute more largely to MFR recognition, but their patterns are difficult to extract as they take a much smaller proportion in observed pulse streams. As deep learning methods generally have statistical tendency, the strong imbalance between frequently and infrequently used pulse groups makes RNN-based MFR recognition a very demanding task.

Based on the abovementioned considerations, this article proposes to first cluster radar pulse groups according to their sequential patterns, so that pulse groups with diversed patterns are separated from each other. Distinguishable features of pulse groups belonging to different clusters

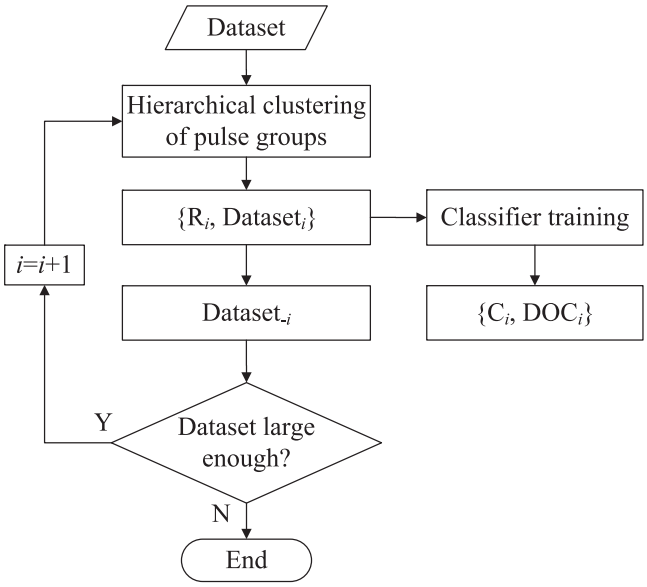

Fig. 4. Establishing process of the integrated classifier consisting of hierarchical clustering and classification models.

are then extracted and accumulated to realize robust radar recognition.

III. HIERARCHICAL CLUSTERING AND CLASSIFICATION OF PULSE GROUPS

A. Framework Establishment of Integrated Classifier

An integrated classifier will be established in this article for MFR recognition. It consists of hierarchical clustering and classification models trained on a large amount of pulse group observations. Each pulse group is tagged with a radar class label but no radar mode label. The clustering models and classification ones are built by mining and exploiting sequential patterns in pulse groups. The process for establishing the framework via hierarchical clustering is shown in Fig. 4. Three kinds of operations are included in the process. First, PRI sequence clustering models are trained with module "hierarchical clustering of pulse groups" based on a dataset including a large amount of pulse groups, and RNN models (R_i) that are able to cluster pulse groups according to their patterns will be obtained. Pulse groups whose patterns well match the model are separated from the original dataset to form Dataset$_i$, whereas the other pulse groups form Dataset$_{-i}$. Second, train a radar classifier C_i based on Dataset$_i$ for $i = 1, \ldots, I$, and output a classification confidence DOC$_i$ (DOC: degree of confidence) for each classifier, which indicates the reliability of the classification result outputted by the corresponding classifier. Third, select the next step according to the data amount in Dataset$_{-i}$. If the amount of remaining pulse groups is large, return to (1) to start a new iteration; otherwise, terminate the iterations.

The hierarchical processing framework is designed mainly to separate pulse groups corresponding to different radar modes, and then extract more distinctive patterns from infrequently used pulse groups. The simulation results in Section V will provide strong evidence of this hierarchical processing framework that, when compared with an indiscriminative process, it has significant advantages in pattern extracting in mixed pulse group datasets consisting of frequently and infrequently used modes.

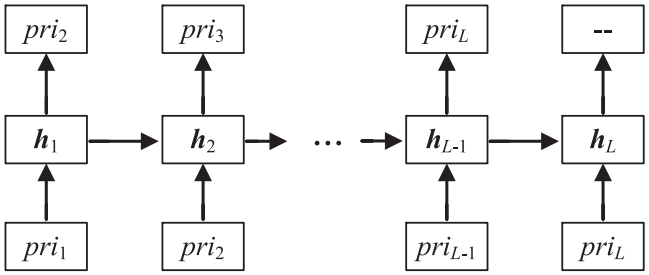

Fig. 5. Illustration of the prediction model used for pulse group clustering.

B. Pulse Group Clustering

Pulse groups performing different radar modes generally have diversed PRI patterns. They can be categorized by mining the discriminating features. Agile range and switching law of successive PRIs within a pulse group are important parts of the pattern. Therefore, diversed patterns can be distinguished according to a prediction model that can be formulated as follows:

$$pri_{l+1} = \arg\max_{\text{DTOA}} p_i(\text{DTOA}|\boldsymbol{h}_l) g_i(\boldsymbol{h}_l|pri_1, \ldots, pri_l)$$
$$l = 1, \ldots, L_i - 1 \tag{4}$$

where $g_i(\boldsymbol{h}_l|pri_1, \ldots, pri_l)$ is a sequence-to-vector function that extracts sequential features from the l preceding PRIs, and $p_i(\text{DTOA}|\boldsymbol{h}_l)$ is a projection function that predicts the upcoming PRI value. The subscript i in the two functions is identical with the index of radar mode, and it is used to indicate that diversed prediction models should be established for different radar modes. L_i stands for the number of PRIs in the pulse group.

The prediction model can be illustrated intuitively in Fig. 5. The PRIs $pri_1, \ldots, pri_L$ of a pulse group are read in by the model one by one. They are processed recurrently and their sequential pattern is embedded in a vector $\boldsymbol{h}_l$ when the lth pulse has been processed. A well-trained prediction model is expected to forecast the upcoming interpulse interval pri_{l+1} precisely. When the pulse group terminates, no prediction will be outputted.

The prediction model, shown in (4) and Fig. 5, is established to interpret pulse group patterns in more compact forms. As pulse groups associated with frequently and infrequently used radar modes have diversed patterns, they will be separated into different clusters by training a series of prediction models. This process can be deemed as unsupervised pulse group clustering. After pulse group separating, the difficulty of pattern mining from infrequent pulse groups caused by data imbalance will be greatly reduced.

The discreteness and analog-value of PRI sequences largely block machine processing. The prediction model

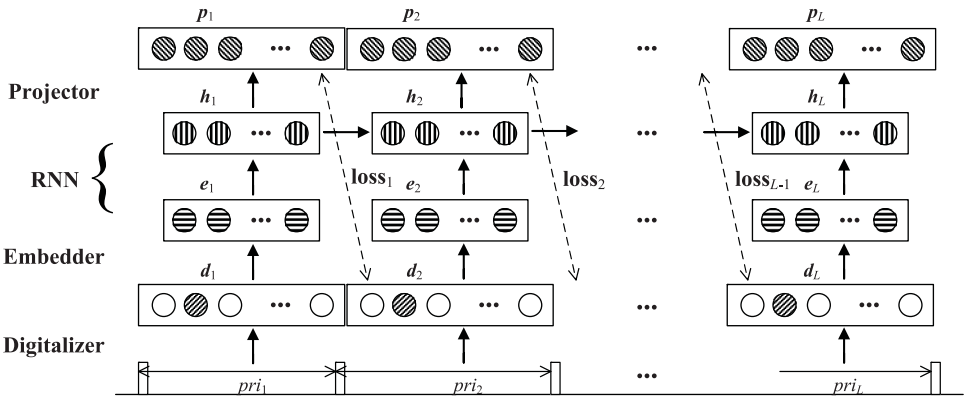

Fig. 6. Detailed structure of PRI prediction model.

in Fig. 5 should be modified to facilitate PRI forecasting. A major modification lies in the vectorization of the inputted and predicted PRIs. Based on the modification, a more detailed illustration of the PRI prediction model is established and shown in Fig. 6. The deep neural network processes PRIs within each pulse group sequentially, and it is trained by improving the accuracy of PRI predictions. Similar to the model in [18], this deep neural network consists of layers for PRI quantization, vectorization, recurrent processing, and prediction.

The PRI quantization layer reads in the DTOA between the current pulse and the previous one, and converts it to a one-hot vector $\boldsymbol{d}_l \in \mathbb{R}^{M\times 1}$ according to a certain quantization unit. The vector has only one nonzero element of 1 at a position according to the PRI value. The vectorization layer reduces the dimension of $\boldsymbol{d}_l$ via embedding to obtain a low-dimensional dense vector $\boldsymbol{e}_l \in \mathbb{R}^{m\times 1}$ with $m \ll M$ [24]

$$\boldsymbol{e}_l = \boldsymbol{E}\boldsymbol{d}_l \tag{5}$$

where $\boldsymbol{E} \in \mathbb{R}^{m\times M}$ is an embedding matrix. The recurrent processing layer uses a RNN to process multiple successive PRIs one-by-one [23], in order to extract sequential patterns implied in PRI sequences. The RNN unit of GRU has the same structure as that shown in Fig. 3. In addition to an input vector $\boldsymbol{e}_l$ and a state vector $\boldsymbol{h}_l \in \mathbb{R}^{h\times 1}$, the GRU also contains three intermediate vectors, namely, update vector $\boldsymbol{z}_l$, reset vector $\boldsymbol{r}_l$, and memory vector $\boldsymbol{f}_l$. The sequential updating procedures of the vectors are implemented as follows:

$$\boldsymbol{z}_l = \sigma\left(\boldsymbol{W}^{(z)}\boldsymbol{e}_l + \boldsymbol{U}^{(z)}\boldsymbol{h}_{l-1} + \boldsymbol{b}^{(z)}\right) \tag{6}$$

$$\boldsymbol{r}_l = \sigma\left(\boldsymbol{W}^{(r)}\boldsymbol{e}_l + \boldsymbol{U}^{(r)}\boldsymbol{h}_{l-1} + \boldsymbol{b}^{(r)}\right) \tag{7}$$

$$\boldsymbol{f}_l = \tanh\left(\boldsymbol{W}^{(f)}\boldsymbol{e}_l + \boldsymbol{r}_l \odot \boldsymbol{U}^{(f)}\boldsymbol{h}_{l-1} + \boldsymbol{b}^{(f)}\right) \tag{8}$$

$$\boldsymbol{h}_l = \boldsymbol{z}_l \odot \boldsymbol{f}_l + (1 - \boldsymbol{z}_l) \odot \boldsymbol{h}_{l-1}. \tag{9}$$

where $\boldsymbol{W}^{(z)}$, $\boldsymbol{W}^{(r)}$, $\boldsymbol{W}^{(f)}$ and $\boldsymbol{U}^{(z)}$, $\boldsymbol{U}^{(r)}$, and $\boldsymbol{U}^{(f)}$ are weighting matrices; $\boldsymbol{b}^{(z)}$, $\boldsymbol{b}^{(r)}$, and $\boldsymbol{b}^{(f)}$ are offset vectors, and their dimensions can be determined according to their contexts; $\sigma(\bullet)$ is an element-wise sigmoid function; $\tanh(\bullet)$ is an element-wise hyperbolic tangent function; and $\odot$ represents dot-product between vectors.

The recurrent processing module is expected to extract sequential patterns from the PRI sequence $[pri_1, \ldots, pri_l]$, and predict the value of pri_{l+1}. The prediction process is implemented by mapping $\boldsymbol{h}_l$ as follows:

$$\hat{\boldsymbol{p}}_l = s\left(\boldsymbol{W}^{(p)}\boldsymbol{h}_l + \boldsymbol{b}^{(p)}\right) \tag{10}$$

where $\boldsymbol{W}^{(p)}$ is a weighting matrix, $\boldsymbol{b}^{(p)}$ is an offset vector, and $s(\bullet)$ is a softmax function. The outputted normalized vector $\hat{\boldsymbol{p}}_l$ has the same dimension as $\boldsymbol{d}_{l+1}$, and each element in $\hat{\boldsymbol{p}}_l$ represents the probability that the value of pri_{l+1} falls within the corresponding quantization unit. In the ideal case, $\hat{\boldsymbol{p}}_l = \boldsymbol{d}_{l+1}$, but in practice, there is often an estimation error between them. This error can be used as a cost function to optimize the unknown parameters of the prediction model in Fig. 6.

As $\boldsymbol{d}_l$ has a one-hot form, the prediction error of $\boldsymbol{d}_l$ can be expressed as $-\log_2(\hat{\boldsymbol{p}}_{l-1}(d_l))$, where d_l represents the coordinate of the nonzero element in $\boldsymbol{d}_l$, and $\hat{\boldsymbol{p}}_{l-1}(d_l)$ represents the probability when $\boldsymbol{d}_l$ is accurately predicted. The prediction errors of the whole pulse group are averaged, and the overall prediction loss of a pulse group is

$$\text{loss}^{(\text{predictor})} = -\frac{1}{L-1}\sum_{l=2}^{L}\log_2\left(\hat{\boldsymbol{p}}_{l-1}(d_l)\right) \tag{11}$$

where L represents the number of PRIs in the pulse group, and the summation begins at $l = 2$ where the first prediction is available. At an early stage of model training, the parameters are randomly initialized, and the prediction error will be very large. By iteratively optimizing the parameters via back-propagation according to the loss, the prediction error can be continuously reduced [25], [26]. The parameter updating process is given as follows:

$$\alpha_{\text{new}} = \alpha_{\text{old}} - \eta\frac{\partial \text{loss}^{(\text{predictor})}}{\partial \alpha} \tag{12}$$

where α represents any unknown variable in (5)–(10), and η represents an iteration weight.

By optimizing the model parameters continuously, the prediction error of PRI sequences will become smaller and smaller. The iterative process is terminated when the prediction loss does not reduce any more, and the model

(denoted as R_i) is now able to predict the PRI sequences of pulse groups having similar patterns. The level of pattern matching is reflected by the PRI prediction error, and a clustering decision can be made by evaluating whether a pulse group matches the learned PRI prediction model in Fig. 6. The evaluation criterion is given as follows:

$$\text{loss}^{(\text{predictor})} < Thr \tag{13}$$

where Thr is a preset threshold.

Deep learning techniques generally prefer to learn patterns of data with larger amounts. Therefore, after sufficient training, the model in Fig. 6 can well describe the PRI pattern of pulse groups (Dataset_i in Fig. 4) corresponding to more frequently used radar modes, and it will produce large prediction errors for pulse groups of infrequently used radar modes. In this way, pulse groups are roughly clustered according to their associated radar modes in the absence of mode labels. After data clustering, pulse groups corresponding to radar modes with more distinguishable features but smaller amounts (Dataset_{-i} in Fig. 2) can be gradually revealed. By repeating the training process, a series of clustering models will be obtained hierarchically, and the original dataset will be separated into multiple clusters.

C. Pulse Group Classification

In the previous section, pulse groups of frequently and infrequently used radar modes are clustered and separated according to their diversed PRI patterns. However, no information for radar classification is obtained in this process. In order to realize MFR recognition, a classifier should be trained based on the clustered pulse groups in $Dataset_i$. The structure of the classifier is the same as that shown in Fig. 3, and the processing of the PRI sequences follows the guideline in (5)–(9). After outputting the final GRU state $\boldsymbol{h}_L$ corresponding to the last PRI of pri_L, the pulse group is classified via the fully connected layer on the right-side of Fig. 3. The classifying process is realized as follows:

$$\hat{\boldsymbol{c}} = s\left(\boldsymbol{W}^{(o)}\boldsymbol{h}_L + \boldsymbol{b}^{(o)}\right) \tag{14}$$

where $\boldsymbol{W}^{(o)}$ is a weight matrix, $\boldsymbol{b}^{(o)}$ is an offset vector, and $\hat{\boldsymbol{c}} = [\hat{c}_1, \ldots, \hat{c}_K]^T \in \mathbb{R}^{K\times 1}$ is the probability distribution vector of this pulse group along all radar classes, which has been normalized by a softmax function $s(\bullet)$.

Similar to the sequence prediction model in the previous section, the classifier in Fig. 3 also contains a large number of unknown parameters, which need to be optimized by taking the classification error as a cost function. The loss function is

$$\text{loss}^{(\text{classifier})} = -\frac{1}{K}\sum_{k=1}^{K}[c_k \log_2(\hat{c}_k) + (1-c_k)\log_2(1-\hat{c}_k)] \tag{15}$$

where K represents the number of radar classes, c_k and $\hat{c}_k$ represent, respectively the true and estimated probabilities when the pulse group is emitted by the kth radar. If the pulse group comes from the kth radar, $c_k = 1$; otherwise, $c_k = 0$. By minimizing the loss function, the parameters of the classification model can be optimized gradually, and a classifier with good classification performance can be obtained (denoted by C_i). The parameter optimization process is implemented as follows [25], [26]:

$$\alpha_{\text{new}} = \alpha_{\text{old}} - \eta\frac{\partial \text{loss}^{(\text{classifier})}}{\partial \alpha} \tag{16}$$

During the previous clustering process, pulse groups in the training dataset have been separated into several clusters. Pulse groups in different clusters have diversed patterns, which generally contribute differently to radar recognition. Although pulse groups corresponding to frequently used radar modes, such as target searching, have a large amount, they are generally performed with simple equi-PRI sequences and very few useful features can be extracted from them to support radar recognition. As a result, the loss function of the classifier in (15) has large values in these indistinguishable pulse group clusters. Pulse groups corresponding to infrequently used radar modes have smaller amounts, but they are generally more distinguishable from the perspective of MFR recognition, and the loss function of (15) is small. In order to measure the reliability of the recognition results of different classifiers, a DOC is defined for each classifier as follows

$$\text{DOC}_i = 1 - \text{loss}^{(\text{classifier})}. \tag{17}$$

When multiple pulse groups are received from a MFR, each pulse group will be categorized into one of the clusters according to the PRI prediction models, and the classifier associated with this cluster will produce a recognition result for the pulse group. As different pulse groups may have diversed distinguishabilities, the recognition results are weighted according to their DOCs defined in (17) and summed up to obtain overall recognition probabilities for each class, which finally indicate a radar recognition result.

D. Termination of Model Training

As the hierarchical clustering process continues, more and more pulse groups are separated from the original training dataset, and they form I clusters named $\{\text{Dataset}_1, \text{Dataset}_2, \ldots, \text{Dataset}_I\}$. The PRI prediction models associated with the clusters are denoted by $\{R_1, R_2, \ldots, R_I\}$. The number of remained pulse groups in $Dataset_{-I}$ becomes smaller and smaller, and the learning process is terminated when the number is smaller than a preset threshold. The threshold is set according to the scale of the PRI prediction model, and empirical values can be used to avoid insufficient training of new models.

Each of the clusters in $\{\text{Dataset}_1, \text{Dataset}_2, \ldots, \text{Dataset}_I\}$ contains a large number of pulse groups belonging to different MFR classes, and a classifier set of $\{C_1, C_2, \ldots, C_I\}$ is trained for the clusters with recognition confidences of $\{\text{DOC}_1, \text{DOC}_2, \ldots, \text{DOC}_I\}$. These pulse group prediction models, classifiers, and recognition confidences are obtained via offline training based on

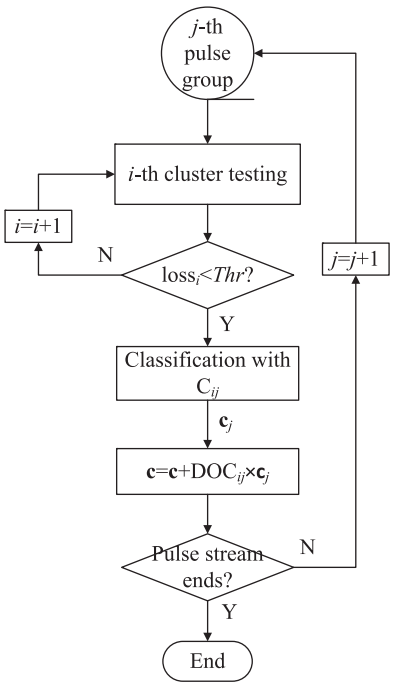

Fig. 7. Recognition process of MFR based on sequential pulse groups.

training dataset. They form a set $\{R_i, C_i, \mathrm{DOC}_i\}$ for $i = 1, \ldots, I$. If the PRI sequence of a pulse group can be well predicted by the prediction model R_i, the associated classifier C_i is used to classify this pulse group according to the K radar classes, and the recognition result has a DOC_i.

IV. RECOGNITION OF MULTIFUNCTION RADARS BASED ON PULSE STREAMS

Suppose that a series of pulse group prediction models, classifiers, and confidence degrees, which are denoted by $\{R_i, C_i, \mathrm{DOC}_i\}_{i=1,\ldots,I}$, have been obtained via offline training based on big data, they can then be exploited to recognize MFRs by processing sequentially arriving pulse groups. As a single pulse group may not contain sufficient information for MFR recognition, distinguishable features should be extracted from pulse streams continuously and accumulated together to improve recognition accuracy. In this section, a MFR recognition process based on sequential pulse groups is proposed.

The MFR recognition process is shown in Fig. 7. When a new pulse group (indexed by j) arrives, its processing can be decomposed into three steps.

1) Match the PRI sequence of the jth pulse group with prediction models in $\{R_1, R_2, \ldots, R_I\}$ iteratively. Once the prediction error falls belowmentioned a preset threshold, terminate the iteration and sort this pulse group to the corresponding cluster, with the cluster index denoted by i_j.
2) Use the associated classifier C_{ij} to classify the pulse group, which obtains a classification probability vector $\boldsymbol{c}_j$. The classification result is then weighted with DOC_{ij} and accumulated sequentially.
3) When all pulse groups have been received, the sequential clustering and classification process will be terminated and the recognition result will be outputted; otherwise, return to (1) to start a new iteration for processing the $(j+1)$th pulse group.

Specifically, the PRI prediction model of the jth pulse group is selected as follows:

$$i_j = \arg\min_i \left(\mathrm{loss}_i^{(\mathrm{predictor})} < Thr \right) \tag{18}$$

where $\mathrm{loss}_i^{(\mathrm{predictor})}$ is the prediction error when testing the PRI sequence of the pulse group with the prediction model R_i according to (11). Thr is a preset bias threshold for prediction model selection, which is identical with the threshold in (13). Suppose that the i_jth model is the first one whose prediction error is smaller than the threshold, the jth pulse group is sorted to the i_jth cluster.

After that, classifier C_{ij} is used to extract and exploit distinguishing patterns contained in the PRI sequence of this pulse group. The classification probability vector is denoted by $\hat{\boldsymbol{c}}_j$, which is weighted by DOC_{ij} and summed to the overall classification probability of the preceding $j-1$ pulse groups. When all J pulse groups have been processed, a final classification probability vector is outputted as

$$\hat{\boldsymbol{c}} = \sum_{j=1}^{J} \mathrm{DOC}_{ij} \times \hat{\boldsymbol{c}}_j. \tag{19}$$

The index of the maximal value in $\hat{\boldsymbol{c}}$ indicates the recognition result of the pulse group stream.

V. SIMULATION RESULTS AND ANALYSES

A. Parameter Settings

In this section, multiple simulations will be carried out to verify the performance of the proposed method. MFRs with agile parameters are considered, they perform sensing modes to either detect previously undetected targets or to update track estimates for known targets. The regularly used searching mode of all radars is performed with equi-PRI pulse groups, which account for a large proportion in the observed data. The PRI range of this mode of different radars is identical, so this type of pulse group does not contain distinguishable features for radar recognition. Infrequently used modes of the radars are implemented with stagger-PRI pulse groups, they have low ratios of usage, but their patterns are more distinguishable between different radar classes. In order to improve MFR recognition performance, the negative influence of indistinguishable pulse groups should be excluded as much as possible, and the more distinguishable pulse groups should be identified from the observations to extract and exploit their features fully. In the simulations, the method in [18] is selected for

performance comparison, and its RNN structure is set the same as the classifier in the proposed method. However, the process of pulse groups clustering is excluded in the method of [18], and the features extracted from multiple pulse groups are summed up directly. The proposed method and the method in [18] are labeled as *cascaded classifiers* and *over-all classifier*, respectively.

The PRI quantization unit in the proposed method is set to 2 us, and the PRI prediction range between adjacent pulses is [0, 1200] *us*, thus the dimension of the one-hot PRI vector $\boldsymbol{d}_l$ is $M = 601$. The dimension of the embedded PRI vector $\boldsymbol{e}_l$ is set to $m = 64$, and that of the RNN state vector $\boldsymbol{h}_l$ is $h = 64$. When training PRI prediction models for the pulse groups, the prediction vector $\hat{\boldsymbol{p}}_l$ has a dimension of $M = 601$, which equals that of $\boldsymbol{d}_{l+1}$. In the last layer of the classifiers, the probability vector of each radar, i.e., $\hat{\boldsymbol{c}}$, is calculated from $\boldsymbol{h}_L$, and its dimension is set to $K = 2$ or 3 according to the number of radar classes. In the PRI prediction model shown in Fig. 6, the dimensions of unknown variables can be inferred from their contexts, and the weighted matrices and offset vectors are initialized randomly.

The parameter optimization processes of the deep neural networks used for PRI prediction and pulse group classification are automatically implemented with PyTorch [27]. The number of batch size during model training is set to 64, and the weight coefficient of back propagation is $\mu = 0.3$. In (13) and (18), the threshold of PRI prediction error during pulse group clustering is set to $Thr = 0.1$.

Training datasets and test datasets are obtained independently via simulations. The numbers of training pulse group and test pulse group of each class of radar are 50 000 and 5000, respectively. Unless otherwise stated, the observation error of each PRI is zero-mean and Gaussian distributed with a standard deviation of $\sigma_{\text{PRI}} = 0.2\,u$s. The radar mode of each pulse group is determined randomly according to a preset proportion in the simulation process, and there may be a small deviation between the distributions of simulated pulse groups and the preset proportions due to sample finite. Radar recognition performance is tested based on pulse group streams. In each stream, several pulse groups are randomly selected from the 5000 test pulse groups of a radar class, and they are processed sequentially by the classifiers. Thirty thousands of batches, with each containing 64 pulse groups randomly chosen from the 50 000 training samples for each radar class, are used for training clustering models. Five thousands of batches, with each containing 64 pulse groups randomly chosen from the clustered pulse groups, are used for training each classifier. The statistical recognition performance of each type of radar is obtained by testing 6400 pulse group trains, and the correct recognition probability is obtained by averaging over all radar classes. The sequential training process of clustering and classification models in Fig. 4 is terminated when the number of remaining pulse groups is smaller than 1000. The classifier in [18] used for performance comparison has the same RNN structure as that of the integrated classifiers in the proposed method. It sums up the classification vector of each pulse group equally to obtain a final classification vector.

TABLE I
PRI Patterns of MFR Pulse Groups in Scenario 1 ($\tau \in [300\,u\text{s}, 400\,u\text{s}]$)

	mode 1	mode 2
class 1	$[\tau, \tau]$	$[\tau + 100, \tau + 220]$
class 2	$[\tau, \tau]$	$[\tau + 100, \tau + 240]$

TABLE II
Results of Hierarchical Clustering and Classification

		number of pulse groups		degree of confidence
		mode 1	mode 2	
1st iteration	total	90072	0	0.019
	1st class	45003	0	
	2nd class	45069	0	
2nd iteration	total	0	9363	0.984
	1st class	0	4889	
	2nd class	0	4474	

B. Scenario 1: Two Classes of MFR With Each Having Two Modes

This simulation tests the recognition performance between two MFR classes, with the MFR of each class having two modes. The PRI patterns of each mode are shown in Table I, where $\tau \in [300\,u\text{s}, 400\,u\text{s}]$. When generating a pulse groups, the value of τ is chosen randomly in the range of $[300 - 400\,u\text{s}]$ to bring in the factor of parameter agility, and then the PRI sequence in the pulse group is determined accordingly.

First, σ_{PRI} is set to 0 to eliminate the influence of measurement errors, and the regular mode of each radar, i.e., mode 1 in Table I, accounts for 90% in the training and testing datasets. Then, the hierarchical clustering and classification methods proposed in this article is implemented on the training dataset, and the number of clustered pulse groups, their distributions and degrees of classification confidence obtained during the first two iterations are shown in Table II. After two rounds of clustering, the number of remaining pulse groups is smaller than 1000, and the iteration is terminated.

As can be seen in Table II, the two-round clustering process has trained two PRI prediction models, which realizes the separation of pulse groups associated with the regular and infrequently used radar modes. The classifiers trained based on the two pulse group clusters have much diversed DOC. Pulse groups that fall in the first cluster have equi-PRI patterns, no matter which radar class they belong to. They also share the same PRI range and therefore contain negligible features for distinguishing the two classes of radars. Thus, the corresponding recognition confidence is close

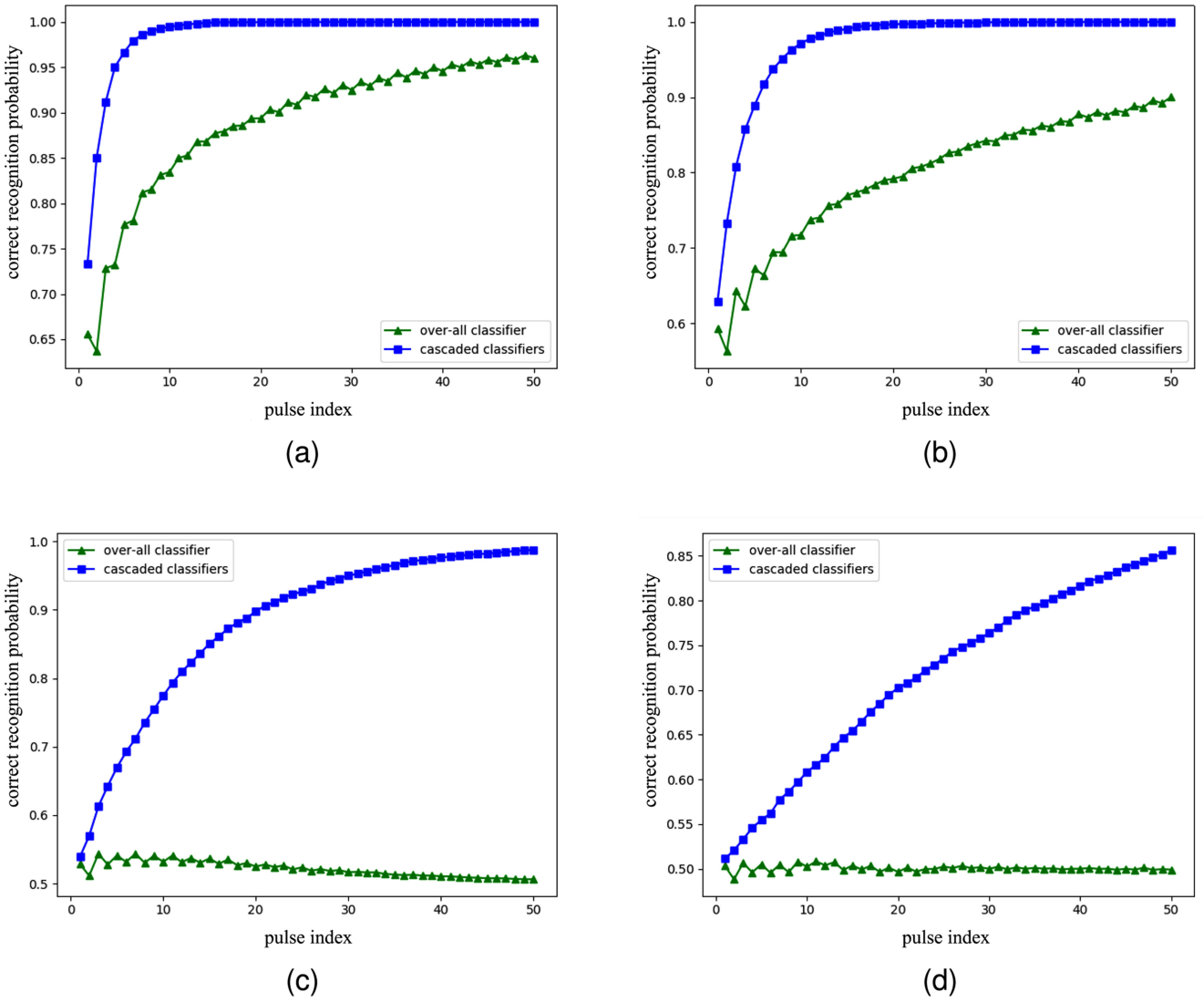

Fig. 8. Classification performances of two dual-mode MFR when mode-1 pulse groups take a proportion. (a) 50%. (b) 70%. (c) 90%. (d) 95%.

to 0. The second cluster of pulse groups has stagger-PRI patterns with different switching laws and agility ranges, which contain much more distinguishable features between the two classes of radars. The corresponding classification confidence is thus close to 1, which is much higher than that of the first classifier. The different confidence degrees indicate that, recognition results obtained from pulse groups belonging to the first cluster are not reliable, while that obtained from pulse groups belonging to the second cluster have higher reliability. The difference also supports the idea of weighting the recognition results of different pulse groups before accumulating them in (19) in this article. Such a weighting process is excluded in the previous counterpart in [18]. As pulse groups of regular radar modes take larger proportions but contain fewer distinguishable features, the weighting-excluded accumulation process generally leads to deteriorated recognition performances. This simulation partially supports the great potential of the proposed method in improving MFR recognition performance.

The PRI measurement errors are then included by setting $\sigma_{\text{PRI}} = 0.2\,us$ and the proportions of mode-1 pulse groups in Table I are set to 50%, 70%, 90%, and 95%, respectively. When the number of pulse groups in the testing pulse stream increases, the correct recognition probabilities of the proposed method and the method in [18] are shown in Fig. 8. As the proposed method separates pulse groups of different modes first, and attaches higher weights to pulse groups having abundant distinguishing patterns, its recognition performance is far better than the method in [18].

When the proportion of indistinguishable mode-1 pulse groups increases from 50% to 95%, the occurrence probability of mode-2 pulse groups, which have higher discrimination degrees, becomes lower and lower in the test data. As a result, the recognition performances of both methods decline significantly. With the increase of the number of sequential pulse groups, the proposed method gains greater recognition performance advantages over the method in [18]. Its recognition performance improves

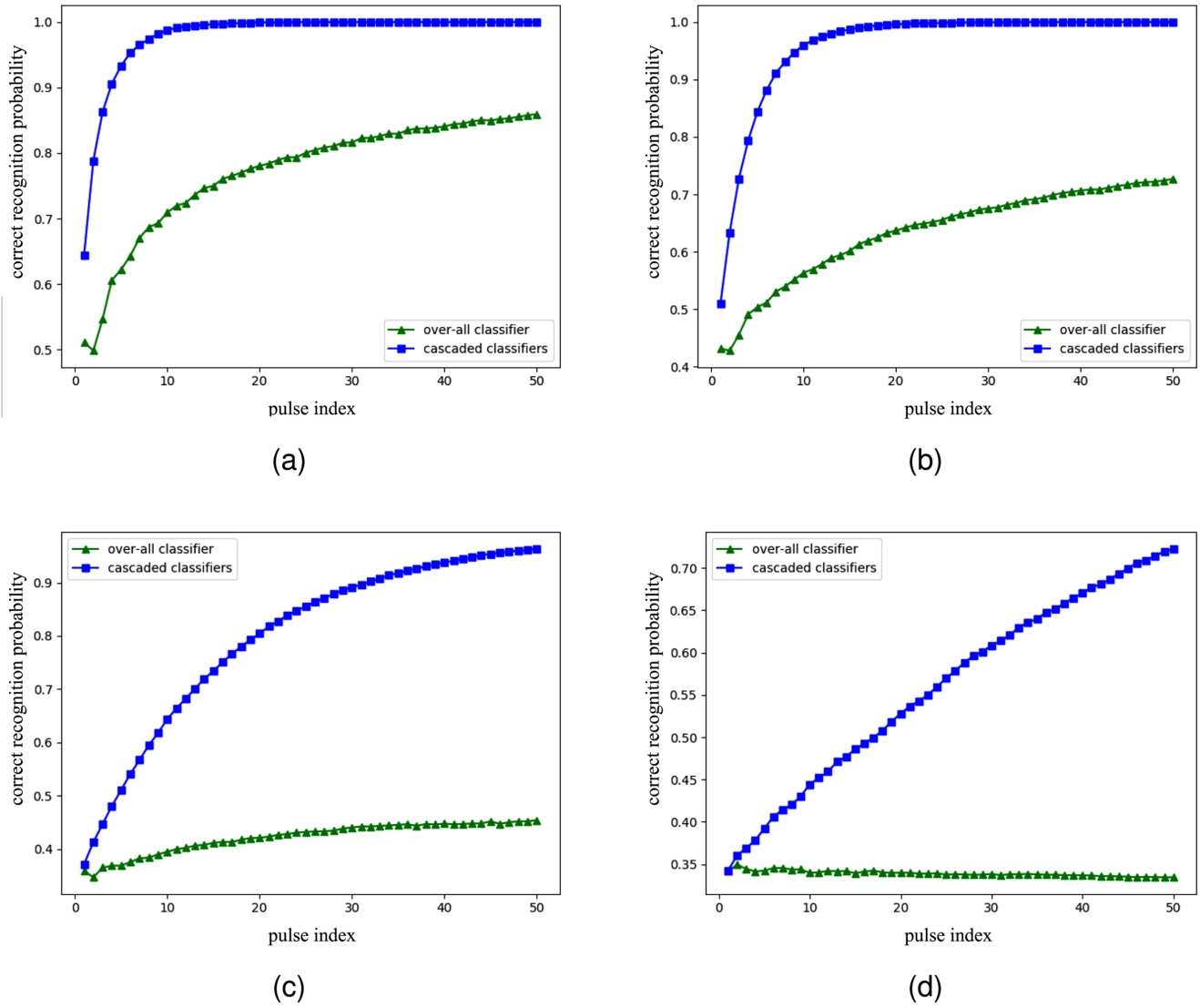

Fig. 9. Classification performances of three dual-mode MFR when mode-1 pulse groups take a proportion. (a) 50%. (b) 70%. (c) 90%. (d) 95%.

steadily even when the proportion of distinguishable pulse groups is as low as 5%. That is because, the proposed method can identify mode-2 pulse groups once they emerge, although they have very low proportions, and enhance the recognition results by attaching high weights to them. The highly weighted recognition results of mode-2 pulse groups greatly reduce the uncertainty in the accumulated classification results caused by indistinguishable mode-1 pulse groups. When more and more pulse groups arrive, the probability that at least one mode-2 pulse group is received increases continuously, and the probability of correct MFR recognition increases accordingly.

C. Scenario 2: Three Classes of MFR With Each Having Two Modes

Based on the scenario corresponding to Fig. 8, the number of radar classes is increased to three, and each class of radar still has two modes, with their PRI patterns listed in Table III. Pulse groups belonging to the first mode are indistinguishable between different radar classes. Mode-2 pulse groups have stagger PRI patterns, and their PRI ranges are very similar to that of mode-1 pulse groups. The other parameters remain unchanged from that in Scenario 1.

The proportion of mode-1 pulse groups is set to 50%, 70%, 90%, and 95%, respectively. The correct recognition probabilities of the proposed method and the method in [18] when the number of sequential pulse groups increases are shown in Fig. 9. It can be seen that the performance comparisons of the two methods are quite similar to Fig. 8. One of the differences is that, when the proportion of mode-1 pulse groups is as high as 95%, the correct recognition probability of the method in [18] drops from 50% in Fig. 8 to about 33% in Fig. 9 (both are $1/K$). This difference indicates that, as the distinguishable pulse groups have a very small amount and are concealed beneath the indistinguishable ones, the over-all classifier fails to extract effective features from pulse groups to realize reliable MFR recognition.

TABLE III
PRI Patterns of MFR Pulse Groups in Scenario 2 ($\tau \in [300\ us, 400\ us]$)

	mode 1	mode 2
class 1	$[\tau, \tau]$	$[\tau, \tau + 20]$
class 2	$[\tau, \tau]$	$[\tau, \tau + 30]$
class 3	$[\tau, \tau]$	$[\tau, \tau + 40]$

D. Scenario 3: Two Classes of MFR With Each Having Three Modes

Based on the simulation corresponding to Fig. 8, the number of MFR modes is increased to 3, whereas the number of radar class is fixed at 2. The PRI patterns of each mode are shown in Table IV. The proportions of pulse groups associated with the three modes are set to 70%, 20%, and 10%, respectively. The correct recognition probabilities of the proposed method and the method in [18] are shown in Fig. 10.

The performance comparisons of the two methods are similar to that in Figs. 8(b) and 9(b), where indistinguishable mode-1 pulse groups take the same proportion of 70%. The proposed method still shows significant performance advantages over the method in [18].

E. Scenario 4: Two Classes of Radars With Each Having Nine Modes

In the previous simulations, the number of MFR modes has been significantly reduced when compared with practical systems [3], [7]. Although both regular and infrequent modes are included, some other factors for radar waveform design have been ignored, such as target altitude [3]. The simplifications are introduced to make the simulation results more concise and intuitive.

In this group of simulations, target altitude is taken into consideration to subdivide MFR radar modes according to the design of fire-control radars in [3]. Each class of radar still has three major categories of modes, one equi-PRI mode and two stagger-PRI modes. The modes have patterns similar to that shown in Table IV, except that their PRI range is subdivided to three ranges of [200 us, 250 us], [300us, 350 us], and [400 us, 450 us] according to look-up or look-down radar beams. The three PRI ranges are set corresponding to high, medium, and low PRFs, respectively [3]. By subdividing the PRI range, each MFR now have nine modes altogether.

Pulse groups of different MFR modes take similar proportions as that in Scenario 3, except that each mode is further divided into three equal parts with low, medium, and high PRFs. Based on the settings, pulse groups of the mode with a PRI pattern of $[\tau, \tau]$ take a proportion of $1/3 \times 70\%$ for each τ range shown inTable V, that of the mode with a PRI pattern of $[\tau, \tau + 20]$ or $[\tau, \tau + 30]$ take a proportion of $1/3 \times 20\%$ for each τ range, and that of the mode with a PRI pattern of $[\tau, \tau + 40]$ or $[\tau, \tau + 60]$ take a proportion of $1/3 \times 10\%$ for each τ range.

In this scenario, the correct recognition probabilities of the proposed method and the method in [18] are shown in Fig. 11. The proposed method again exceeds its counterpart largely in this simulation, which indicates that it can be well applied to the recognition of MFR with more complicated modes.

TABLE IV
PRI Patterns of MFR Pulse Groups in Scenario 3 ($\tau \in [300\,us, 400\,us]$)

	mode 1	mode 2	mode 3
class 1	$[\tau, \tau]$	$[\tau, \tau + 20]$	$[\tau, \tau + 40]$
class 2	$[\tau, \tau]$	$[\tau, \tau + 30]$	$[\tau, \tau + 60]$

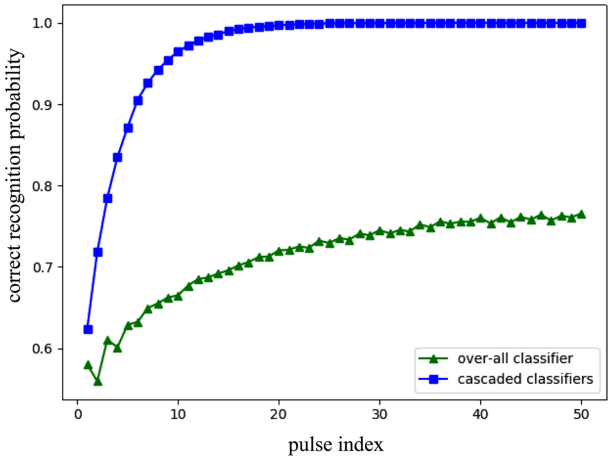

Fig. 10. Classification performances of two classes of three-mode MFR.

TABLE V
PRI Ranges of Pulse Groups With Different PRF

PRF	low	medium	high
τ	$[400us, 450us]$	$[300us, 350us]$	$[200us, 250us]$

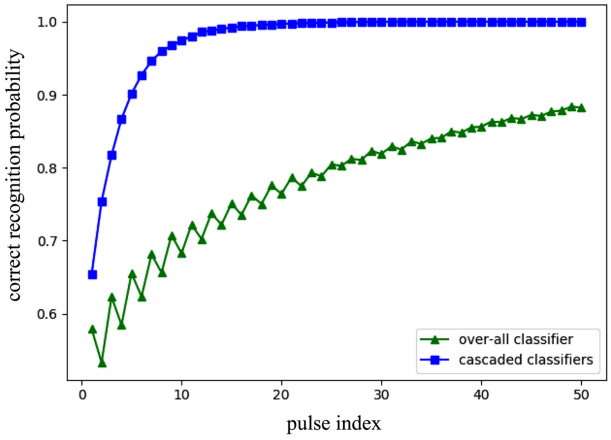

Fig. 11. Classification performances of two classes of nine-mode MFR.

VI. CONCLUSION

This article addresses the recognition problem of MFR, whose regular modes may be weakly distinguishable. An integrated classifier is established by synthesizing hierarchical clusters and classification models of pulse groups. Pulse group clustering is realized according to their sequential PRI patterns that are designed for different MFR modes, and different classifiers are trained for each cluster to extract distinguishable features from pulse groups. DOC are calculated for each classifier to indicate the reliability of its recognition results. In the MFR recognition process based on pulse group streams, each pulse group is sorted to one of the pretrained clusters, then the corresponding classifier is used to distinguish this pulse group, and the recognition result is weighted with the classifiers DOC. The overall recognition result is finally obtained by accumulating the weighted results of each pulse group. Simulation results show that the proposed method can well cluster pulse groups with different PRI patterns, and give them reasonable weights according to the significances of their patterns in MFR recognition. By handling sequential pulse groups with the integrated clustering and classification models, the proposed method obtains much better performances than its counterpart in the considered scenarios, and it can be well extended to the recognition of MFRs with more complicated modes.

REFERENCES

[1] M. A. Richards, J. A. Scheer, and W. A. Holm
Principles of Modern Radar: Basic Principles. Rijeka, Croatia: SCITECH, 2013.

[2] W. L. Melvin and J. A. Scheer
Principles of Modern Radar: Advanced Techniques. Rijeka, Croatia: SCITECH, 2013.

[3] W. L. Melvin and J. A. Scheer
Principles of Modern Radar: Radar Applications. Rijeka, Croatia: SCITECH, 2014.

[4] G. C. Tavik *et al.*
The advanced multifunction RF concept
IEEE Trans. Microw. Theory Technol., vol. 53, no. 3, pp. 1009–1020, Mar. 2005.

[5] R. Reinoso-Rondinel and T.-Y. Yu
Multifunction phased-array radar: Time balance scheduler for adaptive weather sensing
J. Atmospheric Ocean. Technol., vol. 27, pp. 1854–1867, 2010.

[6] A. J. Orman, C. N. Potts, A. K. Shahani, and A. R. Moore
Scheduling for a multifunction phased array radar system
Eur. J. Oper. Res., vol. 90, pp. 13–25, 1996.

[7] W. Alex and K. Vikram
Signal interpretation of multifunction radars: Modeling and statistical signal processing with stochastic context free grammar
IEEE Trans. Signal Process., vol. 56, no. 3, pp. 1106–1119, Mar. 2008.

[8] J.-P. Kauppi, K. Martikainen, and U. Ruotsalainen
Hierarchical classification of dynamically varying radar pulse repetition interval modulation patterns
Neural Netw., vol. 23, pp. 1226–1237, 2010.

[9] R. G. Wiley
ELINT: The Interception and Analysis of Radar Signals. Norwood, MA, USA: Artech House, 2006.

[10] N. Visnevski, V. Krishnamurthy, W. A., and S. Haykin
Syntactic modeling and signal processing of multifunction radars: A stochastic context-free grammar approach
Proc. IEEE, vol. 95, no. 5, pp. 1000–1025, May 2007.

[11] F. Digne, A. Baussard, A. Khenchaf, C. Cornu, and D. Jahan
Classification of radar pulses in a naval warfare context using Bezier curve modeling of the instantaneous frequency law
IEEE Trans. Aerosp. Electron. Syst., vol. 53, no. 3, pp. 1469–1480, Jun. 2017.

[12] T. R. Kishore and K. D. Rao
Automatic intrapulse modulation classification of advanced LPI radar waveforms
IEEE Trans. Aerosp. Electron. Syst., vol. 53, no. 2, pp. 901–914, Apr. 2017.

[13] I. Goodfellow, Y. Bengio, and A. Courville
Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[14] Y. LeCun, Y. Bengio, and G. Hinton
Deep learning
Nature, vol. 521, no. 7553, 2015, Art. no. 436C444.

[15] J. Schmidhuber
Deep learning in neural networks: An overview
Neural Netw., vol. 61, 2015, Art. no. 85C117.

[16] K. Cho *et al.*
Learning phrase representations using RNN encodercdecoder for statistical machine translation
in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Doha, Qatar, Oct. 2014, pp. 1724–1734.

[17] Z. Lin *et al.*
A structured self-attentive sentence embedding
In *Proc. Int. Conf. Learn. Representations*, 2017, pp. 1–15.

[18] Z.-M. Liu and P. S. Yu
Classification, denoising and deinterleaving of pulse streams with recurrent neural networks
IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 4, pp. 1624–1639, Aug. 2019.

[19] H. He and Y. Ma
Imbalanced Learning: Foundations, Algorithms, and Applications. Hoboken, NJ, USA: Wiley, 2013.

[20] A. Fernandez, S. Garcia, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera
Learning from Imbalanced Data Sets. Berlin, Germany: Springer, 2018.

[21] V. Lopez, A. Fernandez, S. Garcia, V. Palade, and F. Herrera
An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics
Inf. Sci., no. 11, pp. 113–141, 2013.

[22] B. Krawczyk
Learning from imbalanced data: Open challenges and future directions
Prog. Artif. Intell., no. 5, pp. 221–232, 2016.

[23] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio
Empirical evaluation of gated recurrent neural networks on sequence modeling
NIPS 2014 Workshop Deep Learn., pp. 8024–8035, 2014.

[24] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean
Distributed representations of words and phrases and their compositionality
In *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 3111–3119.

[25] D. Williams and G. Hinton
Learning representations by backpropagating errors
Nature, vol. 323, no. 6088, 1986, Art. no. 533C538.

[26] B. A. Pearlmutter
Gradient calculations for dynamic recurrent neural networks: A survey
IEEE Trans. Neural Netw., vol. 6, no. 5, pp. 1212–1228, Sep. 1995.

[27] A. Paszke *et al.*
PyTorch: An Imperative Style, High-Performance Deep Learning Library
Adv. Neural Inf. Process. Syst., pp. 8024–8035, 2017.

Zhang-Meng Liu received the Ph.D. degree in statistical signal processing from the National University of Defense Technology (NUDT), Changsha, China, in 2012.

He is currently an Associate Professor with NUDT, working in the interdiscipline of electronics engineering and computer science, especially electronic data mining. He has authored or coauthored more than 30 papers on signal processing, Bayesian learning, and data mining.