

Multi-Feature Fusion Based Recognition and Relevance Analysis of Propagation Scenes for High-Speed Railway Channels

Tao Zhou^{1b}, Member, IEEE, Yingjie Wang^{1b}, Cheng-Xiang Wang^{1b}, Fellow, IEEE, Sana Salous^{2b}, Senior Member, IEEE, Liu Liu^{1b}, Member, IEEE, and Cheng Tao, Member, IEEE

Abstract—This paper proposes a multi-feature fusion based propagation scene recognition model for high-speed railway (HSR) channels and presents the channel relevance analysis of HSR scenes. Extensive field measurement data in typical HSR scenes, including rural, station, suburban and multi-link scenes, are collected with the assist of railway long-term evolution (LTE) networks. The datasets of space-time-frequency channel features, involving Ricean K-factor, root mean square delay spread, Doppler spread, and angle spread, are generated for the model training and testing as well as the relevance analysis. The proposed model merges a weighted score fusion scheme into the deep neural network (DNN) in order to adaptively determine the optimal weights for each feature stream. This weighted score fusion based DNN model is implemented and evaluated in terms of accuracy, confusion matrix, F-score, and receiver operating characteristic (ROC) curve, which exhibits better performance than other machine learning models like random forest, support vector machine (SVM), k -nearest neighbor (KNN), and weighted KNN. In addition, the channel

relevance of HSR scenes is analyzed from perspectives of high-dimensional distribution distance and joint correlation of multiple features. Two metrics, Wasserstein distance and correlation matrix collinearity, are used in the analysis. Statistical results are provided, which reveals the relatively strong channel relevance between the multi-link and suburban scenes.

Index Terms—Deep neural network, high-speed railway channels, multi-feature fusion, propagation scene recognition, relevance analysis.

I. INTRODUCTION

THE global spread of high-speed railway (HSR) has been accelerating in recent years, due to increased awareness of the global environment and construction of major transport infrastructure for economic growth. A HSR system consists of many intricate aspects, in which the communication system is an indispensable part for satisfying railway operational and passengers' requirements. To provide higher data rates, long-term evolution for railway (LTE-R) will be the next-generation railway-dedicated mobile communication system [1]. Furthermore, the forthcoming fifth-generation (5G) system aims to support a variety of high-mobility scenes, and thus will be deployed on HSR to make sure the demand of passengers for 5G services is guaranteed.

The high-speed train usually runs through multiple scenes along a HSR line [2]. The HSR scenes were specifically partitioned into 12 scenes in terms of radio propagation, such as viaduct, cutting, tunnel, station, hilly terrain, rural, suburban, urban, in-train, etc [3]. Authors in [4] also defined 6 scenes for HSR, where the rural, urban, and suburban were combined into an open space scene and others were the same as those in [3]. Besides, a special scene appeared in HSR communication networks, called multi-link scene, was mentioned in [5]. This scene is caused by the use of the cell combination technology and will be specifically introduced in the later section. Various propagation scenes lead to different propagation characteristics, which fundamentally affect the performance of communication systems [6]–[13]. If a wireless system is able to intelligently recognize its operating propagation scene, the system performance will be significantly improved by using some adaptive technologies, such as adaptive modulation and coding, and by achieving intelligent decisions, e.g., intelligent resource allocation and scheduling. Although it will be more effective and accurate to determine the modulation and coding modes depending on CIR

Manuscript received August 8, 2019; revised December 9, 2019 and March 21, 2020; accepted May 19, 2020. Date of publication June 2, 2020; date of current version August 13, 2020. This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1801101, in part by the National Natural Science Foundation of China under Grants 61701017 and 61960206006, in part by the Center of National Railway Intelligent Transportation System Engineering and Technology (Contract RITS2019KF01), China Academy of Railway Sciences, in part by the Frontiers Science Center for Mobile Information Communication and Security, in part by the High Level Innovation and Entrepreneurial Research Team Program in Jiangsu, in part by the High Level Innovation and Entrepreneurial Talent Introduction Program in Jiangsu, in part by the Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2020B01, in part by the Fundamental Research Funds for the Central Universities under Grant 2242020R30001, in part by the Huawei Cooperation Project, and in part by the EU H2020 RISE TESTBED2 Project under Grant 872172. The review of this article was coordinated by Dr. X. Dong. (Corresponding author: Cheng-Xiang Wang.)

Tao Zhou is with the Institute of Broadband Wireless Mobile Communications, Beijing Jiaotong University, Beijing 100044, China, and also with the Center of National Railway Intelligent Transportation System Engineering and Technology, China Academy of Railway Sciences, Beijing 100081, China (e-mail: taozhou@bjtu.edu.cn).

Yingjie Wang, Liu Liu, and Cheng Tao are with the Institute of Broadband Wireless Mobile Communications, Beijing Jiaotong University, Beijing 100044, China (e-mail: 17120131@bjtu.edu.cn; liuliu@bjtu.edu.cn; chtao@bjtu.edu.cn).

Cheng-Xiang Wang is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: chxwang@seu.edu.cn).

Sana Salous is with the School of Engineering and Computing Sciences, Durham University, DH1 3LE Durham, U.K. (e-mail: sana.salous@durham.ac.uk).

Digital Object Identifier 10.1109/TVT.2020.2999313

features, using the propagation scene identification could be simpler. This is because it is only required to choose the modulation and coding mode once for a propagation scene. Otherwise, the modulation and coding mode could be frequently changed if the determination is based on the time-variant CIR features. Thus, propagation scene recognition is of great importance for the design of high-performance HSR communication systems.

Moreover, propagation scene recognition is vital to achieve efficient mapping of measurement data and scenes in passive channel sounding. As for traditional positive measurements using channel sounders, the measured scene is usually known and fixed. The drawback of this method is the lower efficiency when it is applied to HSR measurements [14]. Recently, a high-efficiency passive channel sounding approach by use of railway networks has been widely employed. Extensive HSR channel measurements have been conducted, with the assistance of global system for mobile communications for railway (GSM-R), wideband code-division multiple access (WCDMA), or LTE [15]–[17]. Since channel data are collected in the networks along the whole HSR line, the measured scenes are unknown and frequently changed. A direct way to recognize the scenes is manual recognition, either by site inspection or based on electronic map and global positioning system (GPS) information. However, it is not appropriate for the recognition of massive scenes and the accuracy cannot be guaranteed. Thus, automatic and accurate propagation scene recognition is quite necessary for passive channel measurements on HSR.

Machine learning (ML) has been successfully applied in speech, image and video recognitions [18]–[20]. However, the ML was less used to implement the propagation scene recognition although it has been widely employed for multipath components (MPCs) clustering [21], [22] and channel characteristics predicting [23]–[28]. A semi-supervised ML method was introduced in [29] for classification of indoor and outdoor environments. Authors in [30] utilized k -nearest neighbor (KNN) and weighted KNN (WKNN) methods to recognize different indoor environments based on the radio frequency features such as channel transfer function and frequency correlation function. However, this method is not appropriate for scene recognition in outdoor time-variant environments. To the best of our knowledge, propagation scene recognition or classification for HSR based on ML is still missing.

In addition, the relevance in time, frequency and space domains, multiple links, and different scenes should be analyzed, which is essential to reveal the underlying channel characteristics. Most of studies focused on time-frequency correlations of HSR channels [31], [32]. Due to the lack of multi-antenna channel measurements, there were few works referring to the spatial correlation in realistic HSR scenes. Using a moving virtual antenna array scheme, the spatial correlation was investigated in viaduct and cutting scenes [33]. According to a geometry-based stochastic model, space-time-frequency correlation functions were theoretically derived and analyzed for non-stationary multi-antenna HSR channels [34]. Based on measurement data collected in dedicated LTE networks on HSR, the multi-link correlations of large-scale parameters and small-scale fading were studied in [5], [35]. The correlation between different links is due to the common scatterers or environmental similarity,

which was merged into channel modeling [36]–[38]. In fact, various propagation scenes could have the similarity in terms of joint channel features (joint consideration of different channel features) although its physical environments are diverse. However, there is still no study referring to the analysis of channel relevance in HSR propagation scenes.

To fill the aforementioned research gaps, this paper aims to investigate the propagation scene recognition model and achieve the relevance analysis of propagation scenes for HSR channels. The major contributions and novelties of this paper are as follows.

- 1) Extensive channel measurement data are collected for typical HSR scenes based on railway LTE networks, including rural, station, suburban, and multi-link scenes. The space-time-frequency dispersion features, such as Ricean K-factor (KF), root mean square (RMS) delay spread (DS), Doppler spread (DPS) and angle spread (AS), are extracted and the corresponding datasets are generated.
- 2) A novel weighted score fusion based deep neural network (DNN) model for HSR propagation scene recognition is proposed. The proposed model is implemented by activation function of parametric rectified linear units (PReLU), parameter initialization strategy of He-initializer and gradient-based optimization algorithm of adaptive moment estimation (Adam), and is comprehensively evaluated by three metrics, such as accuracy, F-score, and receiver operating characteristic (ROC) curve.
- 3) The relevance analysis of the four HSR propagation scenes is newly performed from the perspectives of high-dimensional distribution distance and joint correlation of multiple features, using Wasserstein distance (WD) and correlation matrix collinearity (CMC). The statistical WD and CMC results for different scenes are obtained and analyzed.

The remainder of this paper is outlined as follows. Section II describes the LTE network assisted HSR channel measurements. In Section III, the multi-feature fusion based DNN model for HSR propagation scene recognition is proposed. Then, the performance evaluation of the proposed model is presented in Section IV. In addition, relevance analysis of HSR propagation scenes is studied in Section V. Finally, conclusions are drawn in Section VI.

II. LTE NETWORK ASSISTED HSR CHANNEL MEASUREMENTS

A. Scene Description

Our measurements were performed on Beijing to Tianjin (BT) HSR in China, assisted by the LTE network deployed along the railway [14], [17]. The BT HSR has about 120 km distance in total, about 86% percent of which is built on viaduct, and supports the operating speed of up to 350 km/h. The high-speed train runs through the suburban areas (near the Beijing or Tianjin), rural areas (between Beijing and Tianjin), and three train stations (excluding departure and terminal stations) on the whole railway line. Meanwhile, there exists the multi-link propagation condition in the HSR network. Thus, four propagation scenes are involved on BT HSR line, i.e., rural, station, suburban, and multi-link, as shown in Fig. 1. Due to limited

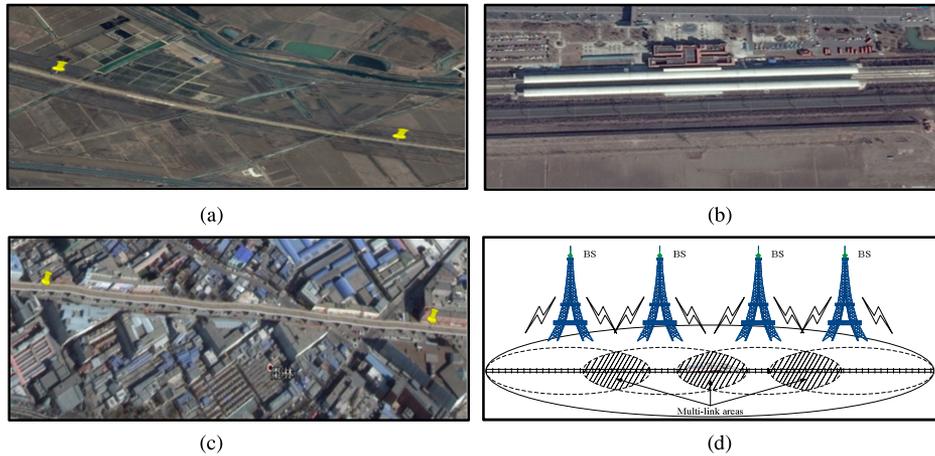


Fig. 1. The measured HSR propagation scenes: (a) Rural, (b) Station, (c) Suburban, and (d) Multi-link.

measurements, this paper only considers these four scenes. The detailed description for the scenes is introduced as follows.

- 1) *Rural*: The railway built on the viaduct is about 10 m above the ground and is higher than the surroundings including light forests and low buildings in most of cases. However, there still exist a few cases with higher and denser forests around the viaduct. The base station (BS) is about 10–30 m higher than the viaduct. The propagation in the HSR rural scene is the line-of-sight (LoS) or obstructed LoS (OLoS) dominance. Besides, non-LoS (NLoS) caused by the sparse scatterers will be identified as the train runs away from the BS after a certain distance.
- 2) *Station*: In the measurement the high-speed train runs through the three stations without stopping. These stations belong to a kind of open-type station with two awnings which only cover the platform supporting a clear free space over the railway [39]. However, the awnings can still yield some MPCs to complicate the fading behavior. The stations have similar geometrical size, with 400 m length, 15 m width of the awning, and 10 m width of the gap between the two awnings approximately.
- 3) *Suburban*: Suburban is a transition zone between the rural and urban areas. Compared with the rural and open-type station environments, the suburban environment has more reflectors and scatterers such as high buildings, which could produce richer MPCs. The density of the buildings in the suburban area is similar to that in the urban region, but the height of the buildings is lower. In fact, our measured suburban areas are close to the urban regions, which can be also regarded as a kind of urban scene.
- 4) *Multi-link*: It is a special propagation scene existing in the HSR network, due to the use of the cell combination technology [14]. To reduce the handover, several physical cells transmitting identical signals with the same frequency are combined into a logical cell. In the overlapping regions of the logical cell, multiple links from neighboring BSs exist simultaneously, which causes the echo channel effect (ECE) and yields additional MPCs. This region is regarded as the multi-link scene. The impact of ECE on propagation characteristics was analyzed in [40].

B. CIR Collection

The measurement was performed several times for return on BT HSR line using the experimental high-speed train with the maximum speed of 300 km/h. During the measurements, the BSs with average 1.2 km spacing send out cell-specific reference signals at 1890 MHz or 2605 MHz, and channel impulse responses (CIRs) are continuously collected by a customized LTE sounder. The measurement equipment is shown in Fig. 2 [14], [17]. Two cross-polarized directional antennas at BS side and two train-mounted omnidirectional antennas with the spacing of 1.2 m at train side are employed, which forms 2×2 multi-antenna measurement. In this paper, we regard the 2×2 multi-antenna measurement as four separate single-antenna measurements. Thus, four groups of CIR data were obtained for each measurement. The collected CIRs have 0.5 ms sample interval and 55.6 ns delay interval, and can support a maximum time delay of 11 μ s. The more detailed measurement parameters can be found in [14], [17].

C. Feature Extraction

In the paper, we aim to recognize the propagation scene based on the channel features derived from the CIRs. We extract channel feature parameters, including KF, RMS DS, RMS DPS, and RMS AS. These parameters comprehensively characterize the channel fading severity and dispersion in space-time-frequency domain, which can be beneficial for the propagation scene recognition. In the following, we briefly describe the extraction of the four feature parameters.

- 1) *KF*: KF is a measure of the fading severity, defined as the power ratio of the LoS component to the NLoS components. A traditional moment-based method is used to extract the KF. Note that the narrowband KF is considered here. The calculation of narrowband KF can be found in [41].
- 2) *RMS DS*: RMS DS is an important parameter used to characterize time dispersion of wireless channels. The RMS DS is estimated as the standard deviation of the second central moment of power delay profiles (PDPs) [5]. Here, a dynamic threshold is applied to the PDPs,

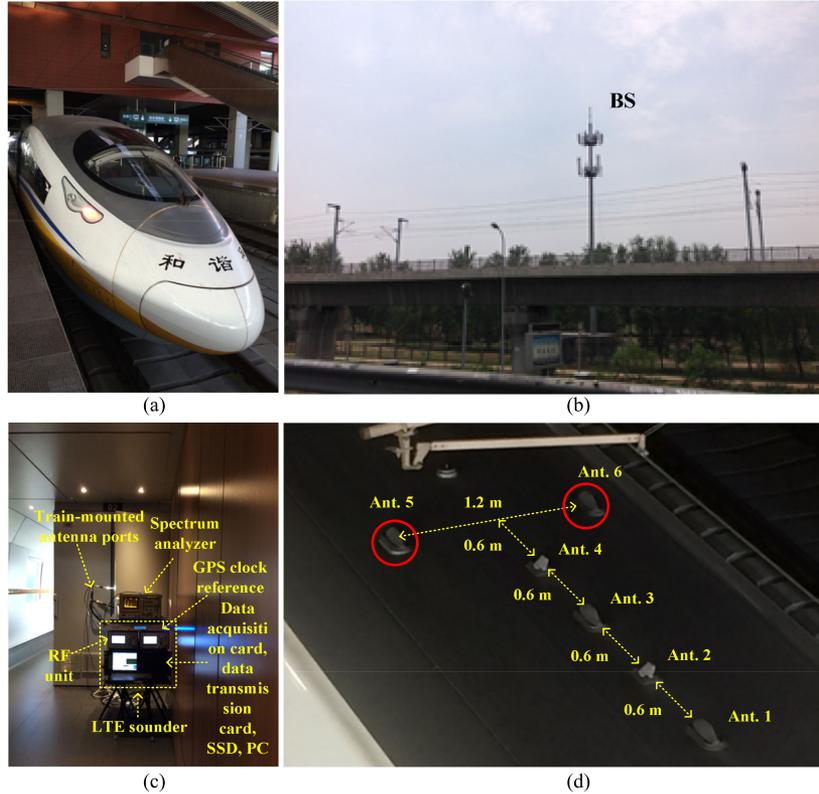


Fig. 2. Measurement equipment: (a) High-speed train, (b) LTE BS along the railway, (c) LTE sounder, and (d) Train-mounted antennas.

which can be used to capture the MPCs more effectively than the fixed threshold.

- 3) *RMS DPS*: RMS DPS is widely used to quantify the frequency dispersion of radio channels. The RMS DPS can be derived according to Doppler power spectral density (DPSD), using the similar calculation method as the RMS DS.
- 4) *RMS AS*: The severity of space dispersion is usually quantified by RMS AS, which can be computed by power angular spectrum (PAS) [5]. It is worth noting that the multi-antenna CIRs used for angle of arrival (AOA) estimation are obtained by the single-antenna CIRs, according to a moving virtual antenna array (MVAA) scheme. The detailed description of the MVAA scheme can be found in our previous work [33].

D. Dataset Generation

After the feature extraction, we generate feature datasets labelled by four scenes. One dataset contains the data for the four features, which are collected in the coverage area of one BS. In other words, one dataset belongs to one BS and the corresponding scene is labelled. Table I lists the amount of training and testing datasets for the four scenes. In this paper, 1528 effective datasets with four-dimensional channel features is obtained in total, which are further divided into 1028 datasets for training and 500 datasets for testing. It should be noticed that although 123 datasets for station scenes are obtained, these datasets do not correspond to 123 unique physical stations.

TABLE I
TRAINING AND TESTING DATASETS

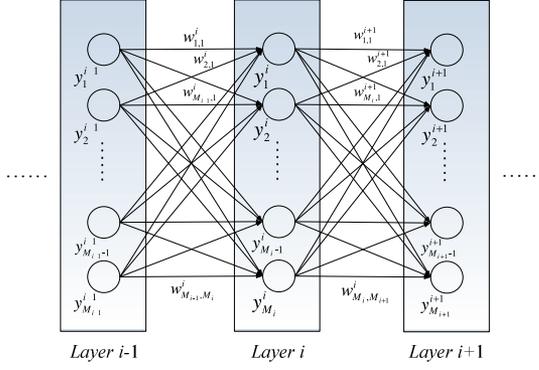
	Datasets for training	Datasets for testing
Rural	263	150
Station	73	50
Suburban	348	150
Multi-Link	344	150
Total	1028	500

III. MULTI-FEATURE FUSION BASED DNN MODEL FOR HSR PROPAGATION SCENE RECOGNITION

A. General Framework of DNN

The traditional DNN, also known as a multilayer feedforward neural network, comprises multiple layers of logistic regression models with continuous nonlinearities. The deeper architecture provides network with the possibility of extracting appropriate representations for classification or regression purpose [42]. Every hidden layer in deep architecture is fully connected to the adjacent layer with different weights and biases attached to the connection. Massive data are sent to the network system for training purpose to update the values of weights and biases, which can be regarded as the way how neural network understands input feature data.

Fig. 3 illustrates a general framework of N -layer DNN, defined as $i - 1$ th, i th and $i + 1$ th layers. The value of each neuron in i th layer is obtained through calculating weighted sum of previous layer's neurons. The weighted sum is then fed


 Fig. 3. A general framework of N -layer DNN.

into a non-linear activation function after adding a bias to each neuron in i th layer. The output of j th neuron in i th layer y_j^i is expressed as

$$y_j^i = f \left(\sum_{k=1}^{M_{i-1}} (w_{j,k}^i y_k^{i-1}) + b_j^i \right) \quad (1)$$

where $w_{j,k}^i$ denotes the weight that connects k th neuron in $i-1$ th layer and j th neuron in i th layer, b_j^i indicates the bias of j th neuron in i th layer, M_{i-1} denotes to total number of neurons in $i-1$ th layer, $f(\cdot)$ represents a non-linear activation function, e.g., sigmoid, rectified linear unit (ReLU), or softmax. The sigmoid or ReLU function is usually applied in hidden layers, while softmax is used in output layer.

Furthermore, the weights and biases need to be updated via gradient based algorithms. The back-propagation algorithm is commonly used for computing gradients in the DNN network [43]. We assume that the output will eventually be used to compute a scalar loss L . Given an input-output pair, $L = C(y, \hat{y})$, where $L = C(\cdot)$ denotes to categorical cross-entropy function. Each individual component of the gradient, $\partial L / \partial w_{j,k}^i$, can be computed by the chain rule. The back-propagation does not need to compute unnecessary intermediate values and is able to efficiently compute the gradient by avoiding duplicate calculations.

B. Multi-Feature Fusion Schemes

Intuitively, taking only one feature (KF, DS, DPS, or AS) into consideration for the propagation scene recognition task will have limited recognizing ability. Thus, we integrate multi-feature fusion schemes into the DNN, in order to take advantage of the four channel features and thus enhance the recognition performance. Motivated by fusion approaches that has been applied in RGB videos and human action recognition [44]–[46], three fusion schemes involving early fusion, feedforward fusion and score fusion are considered, as shown in Fig. 4. The principle of these fusion schemes is described as follows:

- 1) *Early fusion*: The regular scheme is early fusion, which makes use of all feature streams by designing a layer concatenating them together. After the concatenate layer, three fully connected or dense layers are employed, and then are connected to a softmax layer.
- 2) *Feedforward fusion*: Different from fusing feature streams at the input layer, the feedforward fusion aims to achieve

the fusion before the output layer. Each feature stream is as an independent input connected to a DNN. Then, the outputs of the four networks are concatenated together and fed to an extra fully connected layer before the softmax layer. This scheme considers the adaptive representation for different feature streams. However, it will lead to an over-fitting problem, which has been reported in [46].

- 3) *Score fusion*: Another fusion scheme is based on the softmax scores of the multiple DNN streams, where each stream outputs corresponding prediction scores of multiple classes. The scores will be fused to generate the final decision. A simple and widely used way of score fusion is to assign identical weights for each feature stream, called average score fusion. However, since different features may have unequal contribution to the final decision, they should have different weights. Thus, we consider a weighted score fusion scheme, which will be introduced next.

C. Weighted Score Fusion Based DNN Model

Combing the DNN and the score fusion scheme, we propose a weighted score fusion based DNN model to recognize the HSR propagation scenes, as shown in Fig. 5. We denote the confidence scores from the k th stream as $s^k \in \mathbb{R}^C$ ($k = 1, \dots, K$) with C being the number of classes and K being the number of streams, and let \hat{y} be the final confidence score vector. A straightforward way of late fusion is to compute the final prediction as $\hat{y} = f^t(s^1, \dots, s^K)$. Here, $f^t(\cdot)$ is a transition function, which can be a linear function, a logistic function, etc.

Different from the average score fusion method, we attempt to adaptively integrate the confidence scores of each class from multiple streams to determine the optimal weights for each feature stream. To this end, we first stack the multiple confidence score vectors of a training sample n as a coefficient vector, i.e.,

$$\mathbf{s}_n = [s_n^1, \dots, s_n^k, \dots, s_n^K] \in \mathbb{R}^{C \times K}. \quad (2)$$

Then, the stream-specific fusion weights $\alpha = [\alpha_1, \dots, \alpha_K] \in \mathbb{R}^K$ can be learned with an additional softmax layer that is connected to confidence scores. The objective of network training is to minimize the cross entropy in two steps.

The first step is to get the confidence score of every stream and the optimization target is written as

$$\mathbf{W} = \arg \min_{\mathbf{w}_1, \dots, \mathbf{w}_K} - \frac{1}{N} \sum_{n=1}^N \left(\sum_{i=1}^C \hat{y}_{n,i} \log \left(\frac{1}{K} \sum_{k=1}^K s_{n,i}^k \right) \right) \quad (3)$$

where $\frac{1}{K} \sum_{k=1}^K s_{n,i}^k$ denotes the i th class's average confidence score of a training sample n , $\hat{y}_{n,i}$ indicates the ground-truth label of the n th training sample, N represents to total number of samples.

The next step is to get the confidence score vector of weighted score fusion with an another softmax layer connected. The weighted score before the softmax layer is written as

$$\mathbf{o}_n = \sum_{k=1}^K \alpha_k s_n^k = \mathbf{s}_n \alpha^T \in \mathbb{R}^C. \quad (4)$$

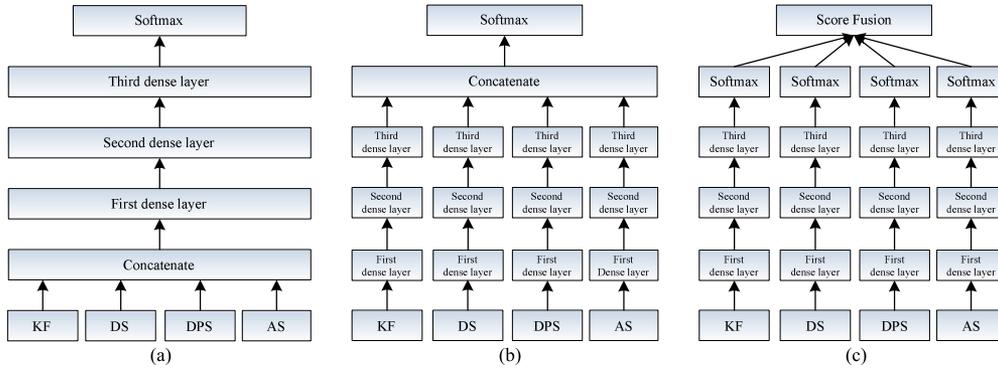


Fig. 4. Multi-feature fusion schemes: (a) Early fusion, (b) Feedforward fusion, and (c) Score fusion.

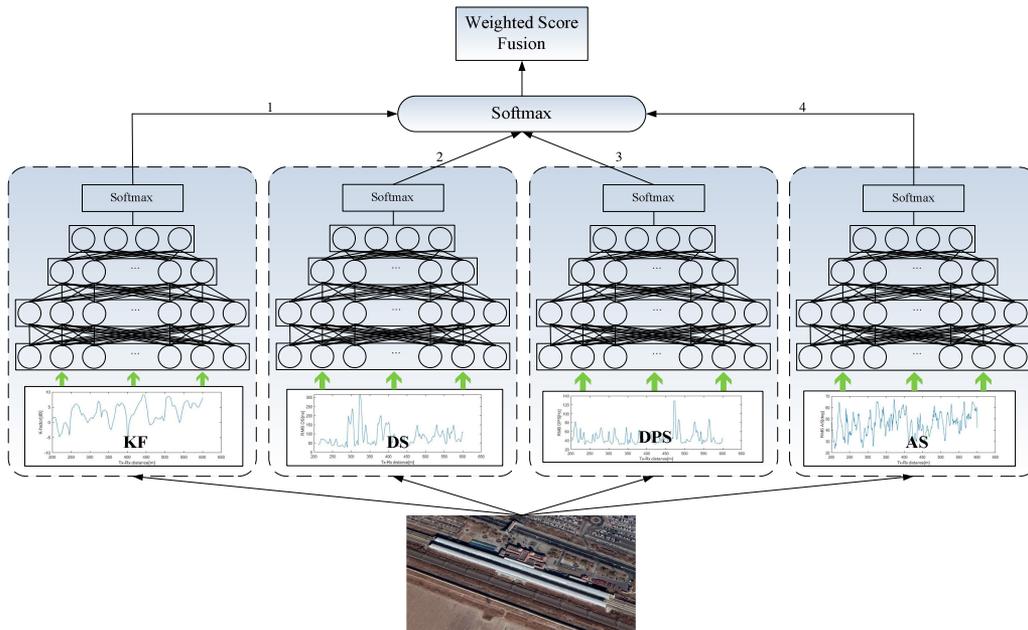


Fig. 5. A weighted score fusion based DNN model for propagation scene recognition.

Then, we optimize another objective function, expressed as

$$\alpha = \arg \min_{\alpha_1, \dots, \alpha_K} - \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^C \hat{y}_{n,i} \log(o_{n,i}). \quad (5)$$

It should be mentioned that the weights and biases of the network of each stream are frozen and they are not adjusted in this part of training, which can be seen as a straightforward way to alleviate the over-fitting problem.

D. Model Implementation

In the proposed model, we use PReLU as the activation function, instead of sigmoid or ReLU. The sigmoid function has the drawback of causing gradient vanishing, whereas ReLU is not suitable here as well for disabling a large number of neurons. The ReLU function tends to kill certain hidden neurons when the network is training by gradient-based algorithm, where the gradient becomes zero when the input is less than zero. The neuron will never be turned on again once it has been turned off (zero value). The PReLU avoids the occurrence of dead neurons

at a relative low price of growth of computation complexity [47]. A slight modification of the PReLU function is that it allows a non-zero value related to a trainable parameter when the input is less than zero.

A three-layer fully-connected network for each feature stream is used and the number of neurons at each layer in different fusion schemes is listed in Table II. We initialize the weights in each layer using He-initializer by taking PReLU into account [47]. This initialization method allows the models using PReLU to converge effectively, whereas the traditional Xavier initializer cannot. The weights that connect l th layer and $l + 1$ th layer need to fulfill $w \sim N\left(0, \sqrt{\frac{2}{N_l}}\right)$, where N_l is the number of neurons in l th layer, the biases in the network are initially set to zero.

The weights and biases are learnt by Adam which stores an exponentially decaying average m_t of past gradients and an exponentially decaying average of past squared gradients v_t , written as [48]

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (6)$$

TABLE II
 THE NUMBER OF NEURONS AT EACH LAYER IN DIFFERENT FUSION SCHEMES

	Single-feature	Multi-feature fusion		
		Early Fusion	Feedforward Fusion	Average/Weighted score fusion
First dense layer for each stream	64	16	16	16
Second dense layer for each stream	64	32	32	32
Third dense layer for each stream	32	64	64	64
Extra concatenate dense layer	-	-	128	-

and

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (7)$$

where t means the iteration index, m_t and v_t are the estimates of the first moment and the second moment of the gradients, respectively, β_1 and β_2 are the exponential decay rates for the first moment and the second moment of the gradients, respectively, g_t^2 indicates the elementwise square of gradient g_t . m_t and v_t are initialized as vectors of zero. Then, the bias-corrected estimates of \hat{m}_t and \hat{v}_t can be calculated as

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (8)$$

and

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}. \quad (9)$$

Finally, using these moment estimations updates the parameters θ_t to yield the Adam update rule, expressed as

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon} \hat{m}_t \quad (10)$$

where the learning rate η is initially set to 10^{-3} , β_1 and β_2 are set to 0.9 and 0.999, respectively, and ε is set to 10^{-8} in order to prevent zero denominator. Besides, a mini-batch of 64 samples is fed to the network for training purpose.

Stratified k -fold cross-validation procedure is enabled in the training process, which is an advanced validation strategy for model selection and optimal hyperparameter decision. The setting of $k = 5$ is employed and 20% of training datasets are split as validation datasets. Each validation datasets contains approximately the same percentage of samples of each target class. We find that the model has a similar performance when the number of hidden layers is more than three. Since more hidden layers will introduce extra computational complexity, three hidden layers are used in the proposed model in order to achieve the tradeoff between the performance and complexity.

The computational time of the proposed model is approximately 160 μ s, which corresponds to 0.013 m in case of 300 km/h. This means that the train only moves 0.013 m when we perform scene recognition per time. Therefore, the proposed model can be suitable for some real-time applications in HSR communications.

IV. PERFORMANCE EVALUATION OF THE PROPOSED MODEL

To comprehensively evaluate the performance of the proposed model for HSR propagation scene recognition, we focus on four metrics including accuracy, confusion matrix, F-score, and ROC curve.

 TABLE III
 ACCURACY OF THE DNN MODEL CONSIDERING SINGLE FEATURE

	KF	RMS DS	RMS DPS	RMS AS
Rural	67%	75%	73%	67%
Station	90%	68%	67%	84%
Suburban	62%	84%	81%	75%
Multi-Link	75%	83%	74%	87%
Total	70%	79%	73%	77%

A. Accuracy

Accuracy is the most commonly used measure to evaluate the recognition performance. We firstly compare the accuracy of the DNN model without using multi-feature fusion, which only considers a certain feature, as shown in Table III. It is found that a certain feature achieves a good performance for some scenes while it performs worse for other scenes. For instance, the KF has 90% accuracy for recognizing the station scene, whereas it is only 62% and 67% for the suburban and rural. This means that the KF is more suitable to distinguish the station scene. Similarly, we observe that RMS AS has the better performance on the multi-link scene recognition. These observations confirm that different features do not contribute equally to the final predictions and their weights used in multi-feature fusion should not be identical. Besides, it can be also seen that the RMS DS achieves the better recognition performance than other features, reaching 79% overall accuracy. However, it is still lower than the multi-feature fusion based methods, as shown in Table IV.

Table IV compares the recognition accuracy of ML models such as random forest (RF), support vector machine (SVM), KNN, WKNN, and DNN, using various multi-feature fusion schemes, including early fusion, feedforward fusion, average score fusion, and weighted score fusion. RF is a widely used ensemble learning method for classification tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees [49]. SVM was initially proposed in [50], which is a kind of supervised learning model that analyzes data used for classification and regression analysis. Input samples are mapped to a high-dimension feature space, and then the SVM constructs a hyperplane, which can be regarded as decision surface. It is observed from Table IV that the DNN model has the better accuracy than the other ML models when using the early fusion scheme. It can be also found that the accuracy of the feedforward fusion has better performance than the early fusion and average score fusion. This is because the feedforward fusion makes use of more discriminative features and suppressing somewhat less discriminative features at concatenation layer due to its adaptive

TABLE IV
PERFORMANCE OF THE ML MODELS USING MULTI-FEATURE FUSION SCHEMES

ML model	Multi-feature fusion scheme	Accuracy	F-1 score	AUC
RF	Early fusion	80.2%	0.80	0.90
KNN	Early fusion	80.8%	0.81	0.87
SVM	Early fusion	81.6%	0.81	0.90
WKNN	Early fusion	81.7%	0.82	0.90
DNN	Early fusion	83.4%	0.82	0.94
DNN	Average score fusion	85.4%	0.85	0.98
DNN	Feedforward fusion	86.6%	0.87	0.97
DNN	Weighted score fusion	90.8%	0.91	0.99

TABLE V
ACCURACY OF DIFFERENT FEATURE COMBINATIONS

Feature combination	Accuracy
KF + RMS DPS	74.3%
RMS DPS + RMS AS	77.7%
KF + RMS AS	78.2%
KF + RMS DS	79.8%
RMS DS + RMS DPS	80.9%
RMS DS + RMS AS	83.0%
KF + RMS DPS + RMS AS	79.0%
KF + RMS DS + RMS DPS	81.3%
KF + RMS DS + RMS AS	83.6%
RMS DS + RMS DPS + RMS AS	89.2%
KF + RMS DS + RMS DPS + RMS AS	90.8%

weights that connected to the softmax layer. However it may suffer the over-fitting problem by adding an extra fully-connected layer. Especially, our proposed weighted score fusion based DNN model shows the best performance. This is because this model adds a training process for the weights after scores of each feature stream and freezes the weights and biases in previous layers, which thus avoids the over-fitting problem.

To explore the contributions of different features, we compare the performance of different feature combinations, as listed in Table V. It can be found that the combination of RMS DS and RMS AS exhibits the highest accuracy when two-feature combination is considered. This means that RMS DS and RMS AS have the major contribution to scene recognition. It can be also seen that the combination of RMS DS, RMS DPS and RMS AS has almost similar accuracy to that of four features, which means that KF has less contribution than the other three features. These results are consistent with the previous findings as shown in Table III that RMS DS and RMS AS outperform the other two features and KF is the worst.

B. Confusion Matrix

Confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning algorithm. Each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class (or vice versa). The element in the confusion matrix is defined as

$$V_{i,j} = \frac{1}{|C_i|} \sum_{n \in C_i} \mathbf{1}_{\arg \max_C(\mathbf{o}_n) = C_j} \quad (11)$$

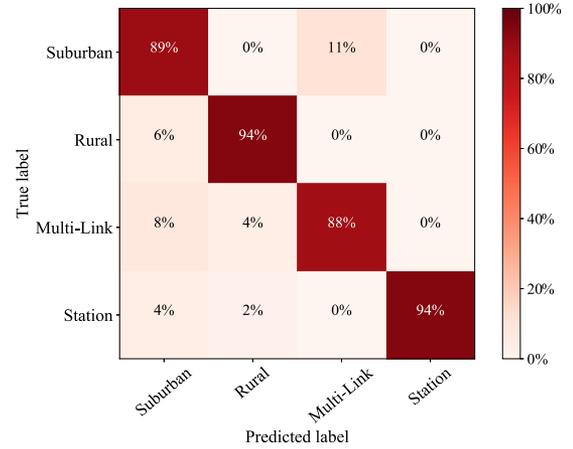


Fig. 6. Confusion matrix for the weighted score fusion based DNN model.

where $\mathbf{1}(\cdot)$ denotes the indicator function, C_i indicates the collection of testing samples that belongs to class i , $|\cdot|$ is the cardinality function, which represents the number of instances of C_i , $n \in C_i$ means the n th testing sample and also belongs to the class i , and $\arg \max_C(\mathbf{o}_n)$ is to get the label of the n th testing sample from the C dimensional confidence score vector \mathbf{o}_n obtained in (4). Here, each element $V_{i,j}$ is the i th column and j th row element in matrix which stands for the percentage of the samples with the ground-truth label of class C_i being wrongly classified into class C_j .

The confusion matrix of the proposed model is illustrated in Fig. 6. It can be found that the proposed model performs well on most of the scenes. However, the misclassification is not avoidable. It is worth noting that 11% data of suburban are misclassified to multi-link and 8% data of multi-link are misclassified to suburban, which means that there exists relatively great confusion between the suburban and multi-link scenes. The reasons for this confusion can be interpreted from two aspects: the recognition error of the model itself and the relevance between the two scenes. If two scenes have the strong channel relevance, it will deteriorate the performance of a classification model. The relevance between scenes will be analyzed in detail in the following section.

C. F-Score

The F-score can be interpreted as a weighted harmonic mean of the precision P and recall R . The recall is the ratio

$tp/(tp + fn)$, where tp is the numbers of true positives and fn is the number of false negatives. The precision is the ratio $tp/(tp + fp)$, where fp is the number of false positives. The F-score is defined as

$$F_\beta = \frac{(\beta^2 + 1)PR}{\beta^2 P + R} \quad (12)$$

where β is the trade-off between P and R . When $\beta = 1$, it means that the recall and the precision are equally important. F_β reaches its best value at 1 and its worst score at 0. Although F1 score is typically used for binary classifiers, it can also be extended to multi-class context, such as micro-F1 score, macro-F1 score and weighted macro-F1 score. Here, the micro-F1 score is considered. In order to obtain the micro-F1 score, tp , fn and fp in (12) should be calculated globally by counting the total true positives, false negatives and false positives of the four classes.

Table IV lists the results of F-1 score for different ML models using various multi-feature fusion schemes. It can be seen that the weighted score fusion based DNN model has the highest F-1 score, reaching 0.91. This confirms the better performance of the proposed model in terms of precision and recall.

D. ROC Curve

The ROC curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the fraction of true positives out of the positives (TPR = true positive rate) versus the fraction of false positives out of the negatives (FPR = false positive rate), at various threshold settings. Area under ROC curve (AUC) of a classifier C_i is the probability that C_i ranks a randomly drawn positive example higher than a randomly drawn negative example, expressed as

$$\text{auc}(C_i) = P [C_i(x^+) > C_i(x^-)]. \quad (13)$$

The higher the AUC is, the more likely a positive sample scores higher than a negative sample. AUC can be regarded as a measure of the robustness of the classifier. Similar to the F1-score, the ROC curve can also be used in multi-class classification based on two averaging strategies including one-vs-one (OvO) and one-vs-rest (OvR) algorithms [51]. Here, the OvR algorithm is used, where weighted average of the ROC for each class against all other classes is computed and the weights are decided by number of true samples of each class. The ROC curves for different multi-feature fusion based DNN models are depicted in Fig. 7 and corresponding AUC results of are shown in Table IV. It is found that the proposed model has the largest AUC with 0.99, which means that our model is more robust.

V. RELEVANCE ANALYSIS OF HSR PROPAGATION SCENES

The relevance of the four HSR propagation scenes will be investigated from macro and micro perspectives. On one hand, we analyze the high-dimensional distribution distance of multiple features in different scenes, which is regarded as the macro perspective. On the other hand, we analyze the joint correlation of multiple features from the micro perspective. In the paper, two metrics are used to perform the relevance analysis, including

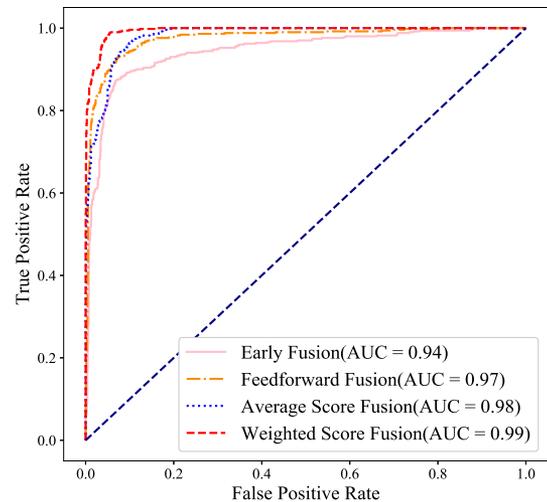


Fig. 7. ROC curves for different multi-feature fusion based DNN models.

Wasserstein distance (WD) and correlation matrix collinearity (CMC). These two metrics determine the channel relevance of propagation scenes together.

A. Wasserstein Distance

The most commonly used measure of distribution distance is Kullback-Leibler divergence (KLD). The KLD is a kind of asymmetric metric and can be infinite. In order to avoid these two defects, a new kind of divergence called Jensen-Shannon divergence (JSD) is introduced. The JSD is based on the KLD, with some notable modification, including that it is symmetric and it always has a finite value. However, The JSD will become a constant if the two distributions are so far apart that they don't overlap at all, which is not suitable for two non-overlapped distributions [52]. To solve this problem, a new measure known as WD is used. The WD is a distance function between two probability distributions on a given metric space. If the probability density function of each distribution is viewed as the piled up dirt, then the WD can be considered as the minimum cost of turning one pile into the other, defined as

$$W(P_1, P_2) = \inf_{\gamma \in \Pi(P_1, P_2)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (14)$$

where $\Pi(P_1, P_2)$ denotes the set of all joint distributions $\gamma(x, y)$ whose marginal are respectively P_1 and P_2 . And $\gamma(x, y)$ indicates how much mass of pile must be transported from x to y in order to transform the distribution P_1 into the distribution P_2 and $W(P_1, P_2)$ is the infimum of the cost of the optimal transport strategy. Although the infimum in (14) is highly intractable when dealing with extraordinary high dimensional distributions, only four-dimensional distributions are considered in our task.

The statistical results of WD for various HSR scenes are provided in terms of bar chart with error bar, as illustrated in Fig. 8. It is found that the mean value of WD between multi-link and rural is much larger than others. This means that these two scenes have less relevance from the view of distribution distance. It can be also seen that the multi-link and suburban scenes have the smallest mean value of WD, which implies that

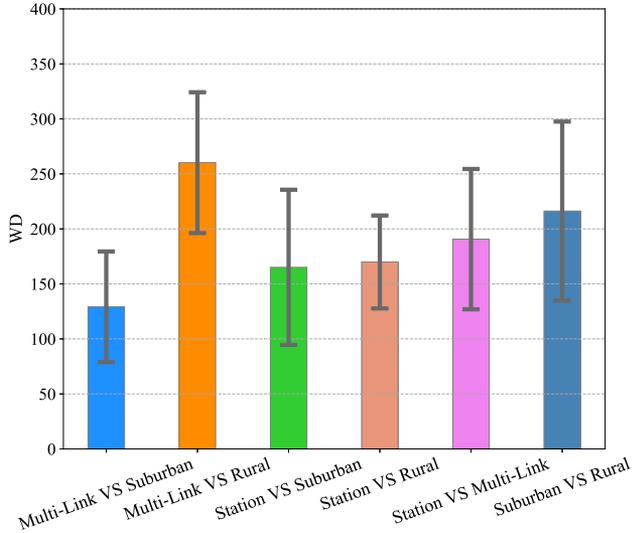


Fig. 8. Statistical results of WD for HSR scenes.

the multi-link and suburban have a certain degree of similarity. This is reasonable that both the scenes have rich MPCs, which causes the closer propagation dispersion.

B. Correlation Matrix Collinearity

To analyze the correlation between two propagation scenes, only focusing on the correlation of a certain feature is not appropriate. It is necessary to jointly consider the correlation of multiple features. The CMC is a parameter quantified the correlation of two matrices, which has been applied to evaluate the spatial correlation between MIMO channel matrices of different links [53]. Similarly, it is also suitable for the correlation analysis in different propagation scenes. The CMC coefficient can be calculated as

$$c(\mathbf{A}, \mathbf{B}) = \frac{|\text{tr}(\mathbf{A}\mathbf{B}^H)|}{\|\mathbf{A}\|_F \|\mathbf{B}\|_F} \quad (15)$$

where \mathbf{A} and \mathbf{B} are two matrices with the same dimension, $\|\cdot\|_F$ denotes the Frobenius norm of the matrix, and $(\cdot)^H$ represents the matrix conjugate transpose operation. The CMC coefficient reflects how similar the two matrices are. This coefficient ranges from zero (absolutely non-collinear, i.e. two matrices are orthogonal to each other) to one (fully collinear, i.e. two matrices are same).

We apply the feature matrices with four dimensions in different propagation scenes to the equation (15), the statistical CMC coefficient results can be obtained, as shown in Fig. 9. It is observed that the mean value of CMC coefficients between multi-link and suburban scenes is much higher than others. This infers that the multi-link and suburban scenes have a stronger similarity, while the other scenes show comparative low similarity. Since the multi-link and suburban scenes have both closer distribution distance and higher correlation, we can believe that there exists relatively stronger channel relevance between these two propagation scenes. Moreover, the relatively higher misclassification between multi-link and suburban scenes

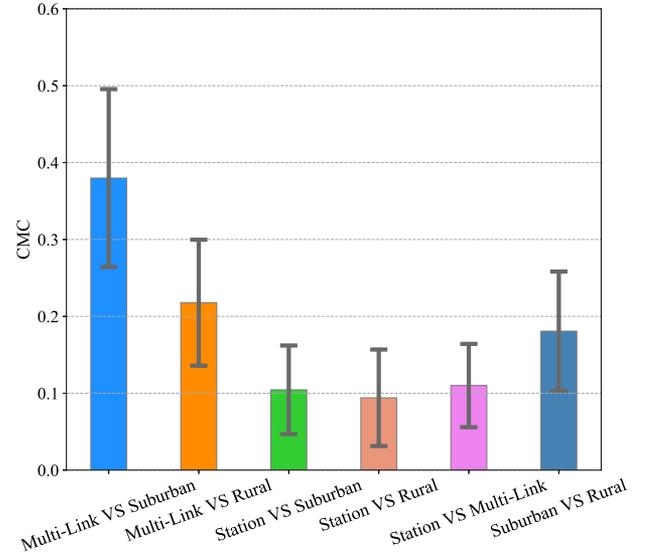


Fig. 9. Statistical results of CMC for various HSR scenes.

found in the confusion matrix result also confirms this relevance to some extent.

VI. CONCLUSION

In this paper, the propagation scene recognition using multi-feature fusion schemes and relevance analysis of propagation scenes have been investigated for HSR channels. Assisted by the railway LTE networks, CIR data have been collected in four typical HSR scenes, including rural, station, suburban, and multi-link. The corresponding datasets with four channel features involving KF, RMS DS, RMS DPS, and RMS AS have been generated. The DNN model using the weighted score fusion scheme has been proposed, implemented, and evaluated. It has been found that the proposed model reaches the accuracy of 90.8%, the F-1 score of 0.91, and the AUC value of 0.99, which outperforms the other mentioned recognition models. Furthermore, the relevance of HSR scenes has been analyzed based on high-dimensional distribution distance and joint correlation of multiple features. The statistical WD and CMC results have been derived, demonstrating that the multi-link and suburban scenes have the relatively stronger channel relevance than that of rural and station scenes.

REFERENCES

- [1] R. He, B. Ai, G. Wang, K. Guan, Z. Zhong, and A. F. Molisch, "High-speed railway communications: From GSM-R to LTE-R," *IEEE Veh. Technol. Mag.*, vol. 11, no. 3, pp. 1784–1793, Sep. 2016.
- [2] T. Zhou, H. Li, Y. Wang, L. Liu, and C. Tao, "Channel modeling for future high-speed railway communication systems: A survey," *IEEE Access*, vol. 7, no. 1, pp. 52818–52826, Apr. 2019.
- [3] B. Ai, R. He, Z. Zhong *et al.*, "Radio wave propagation scene partitioning for high-speed rails," *Int. J. Antennas Propag.*, vol. 2012, pp. 815232-1–815232-7, Sep. 2012.
- [4] C. X. Wang, A. Ghazal, B. Ai, Y. Liu, and P. Fan, "Channel measurements and models for high-speed train communication systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 974–987, May 2016.
- [5] T. Zhou, C. Tao, S. Salous, and L. Liu, "Joint channel characteristics in high-speed railway multi-link propagation scenarios: Measurement, analysis, and modeling," *IEEE Tran. Intell. Transp.*, vol. 20, no. 6, pp. 2367–2377, Jun. 2019.

- [6] A. Ghazal, C.-X. Wang, B. Ai, D. Yuan, and H. Haas, "A non-stationary wideband MIMO channel model for high-mobility intelligent transportation systems," *IEEE Tran. Intell. Transp.*, vol. 16, no. 2, pp. 885–897, Apr. 2015.
- [7] J. Bian *et al.*, "A WINNER+ based 3D non-stationary wideband MIMO channel model," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1755–1767, Mar. 2018.
- [8] Y. Liu, C.-X. Wang, C. F. Lopez, and X. Ge, "3D non-stationary wideband circular tunnel channel models for high-speed train wireless communication systems," *Sci. China Inf. Sci.*, vol. 60, no. 8, Aug. 2017, doi: [10.1007/s11432-016-9004-4](https://doi.org/10.1007/s11432-016-9004-4).
- [9] Y. Liu, A. Ghazal, C.-X. Wang, X. Ge, Y. Yang, and Y. Zhang, "Channel measurements and models for high-speed train wireless communication systems in tunnel scenarios: a survey," *Sci. China Inf. Sci.*, vol. 60, no. 8, Oct. 2017, doi: [10.1007/s11432-016-9014-3](https://doi.org/10.1007/s11432-016-9014-3).
- [10] Y. Liu, C.-X. Wang, J. Huang, J. Sun, and W. Zhang, "Novel 3-D non-stationary mmWave massive MIMO channel models for 5G high-speed train wireless communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2077–2086, Mar. 2019.
- [11] S. Wu, C.-X. Wang, H. Aggoune, M. M. Alwakeel, and X. You, "A general 3D non-stationary 5G wireless channel model," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3065–3078, Jul. 2018.
- [12] C.-X. Wang, J. Bian, J. Sun, W. Zhang, and M. Zhang, "A survey of 5G channel measurements and models," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3142–3168, Oct.–Dec. 2018.
- [13] F. Zhang and W. Fan, "Near-field ultra-wideband mmWave channel characterization using successive cancellation beamspace UCA algorithm," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7248–7259, Aug. 2019.
- [14] T. Zhou, C. Tao, S. Salous, L. Liu, and Z. Tan, "Channel sounding for high-speed railway communication systems," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 70–77, Oct. 2015.
- [15] R. He, Z. Zhong, B. Ai, and J. Ding, "An empirical path loss model and fading analysis for high-speed railway viaduct scenarios," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 808–812, Aug. 2011.
- [16] X. Yin, X. Cai, X. Cheng, J. Chen, and M. Tian, "Empirical geometry-based random-cluster model for high-speed-train channels in UMTS networks," *IEEE Tran. Intell. Transp.*, vol. 16, no. 5, pp. 2850–2861, Oct. 2015.
- [17] T. Zhou, C. Tao, S. Salous, L. Liu, and Z. Tan, "Implementation of an LTE-based channel measurement method for high-speed railway scenarios," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 25–36, Jan. 2016.
- [18] A. Graves and N. Jaitly, "Towards end-to-end speech recognition with recurrent neural networks," in *Proc. ICML*, Beijing, China, 2012, pp. 1764–1772.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. NIPS*, Lake Tahoe, USA, 2012, pp. 1097–1105.
- [20] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks" in *Proc. CVPR*, Columbus, USA, 2014, pp. 1725–1732.
- [21] R. He *et al.*, "On the clustering of radio channel impulse responses using sparsity-based methods," *IEEE Trans. Wireless Commun.*, vol. 64, no. 6, pp. 2465–2474, Jun. 2016.
- [22] R. He *et al.*, "A kernel-power-density-based algorithm for channel multipath components clustering," *IEEE Trans. Antennas Propag.*, vol. 16, no. 11, pp. 7138–7151, Nov. 2017.
- [23] E. Ostlin, H.-J. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, Jul. 2010.
- [24] J. Zhang, "The interdisciplinary research of big data and wireless channel: A cluster-nuclei based channel model," *China Commun.*, vol. 13, no. 2, pp. 14–26, 2016.
- [25] C. A. Oroza, Z. Zhang, T. Watteyne, and S. D. Glaser, "A machine learning-based connectivity model for complex terrain large-scale low-power wireless deployments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 576–584, Dec. 2017.
- [26] L. Bai, C.-X. Wang, Q. Xu, S. Ventouras, and G. Goussetis, "Prediction of channel excess attenuation for satellite communication systems at Q-band using artificial neural network," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2235–2239, Nov. 2019.
- [27] J. Huang *et al.*, "A big data enabled channel model for 5G wireless communication systems," *IEEE Trans. Big Data*, vol. 6, no. 2, pp. 211–222, Jun. 2020.
- [28] L. Bai *et al.*, "Predicting wireless mmWave massive MIMO channel characteristics using machine learning algorithms," *Wireless Commun. Mob. Com.*, vol. 2018, Aug. 2018, doi: [10.1155/2018/9783863](https://doi.org/10.1155/2018/9783863).
- [29] V. Radu, P. Katsikouli, R. Sarkar, and M. K. Marina, "A semi-supervised learning approach for robust indoor-outdoor detection with smartphones" in *Proc. ACM Conf. Embedded Netw. Sensor Syst.*, Memphis, USA, 2014, pp. 280–294.
- [30] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 12, pp. 2164–2168, Dec. 2018.
- [31] L. Liu *et al.*, "Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35GHz," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 834–845, May 2012.
- [32] Y. Zhang, Z. He, W. Zhang, L. Xiao, and S. Zhou, "Measurement-based delay and doppler characterizations for high-speed railway hilly scenario," *Int. J. Antennas Propag.*, vol. 2014, pp. 875345-1-875345-8, Apr. 2014.
- [33] T. Zhou, C. Tao, S. Salous, and L. Liu, "Measurements and analysis of angular characteristics and spatial correlation for high-speed railway channels," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 2, pp. 357–367, Feb. 2018.
- [34] A. Ghazal *et al.*, "A non-stationary IMT-A MIMO channel model for high-mobility wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2057–2068, Apr. 2017.
- [35] T. Zhou, C. Tao, and L. Liu, "LTE-assisted multi-link MIMO channel characterization for high-speed train communication systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2044–2051, Mar. 2019.
- [36] J. Poutanen, F. Tufvesson, K. Haneda, V.-M. Kolmonen, and P. Vainikainen, "Multi-link MIMO channel modeling using geometry-based approach," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 587–596, Feb. 2012.
- [37] X. Cheng *et al.*, "Cooperative MIMO channel modeling and multi-link spatial correlation properties," *IEEE J. Sel. Area. Commun.*, vol. 30, no. 2, pp. 388–396, Jan. 2012.
- [38] T. Zhou, C. Tao, S. Salous, and L. Liu, "Geometry-based multi-link channel modeling for high-speed train communication networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1229–1238, Mar. 2020.
- [39] K. Guan, Z. Zhong, B. Ai, and T. Kürner, "Propagation measurements and analysis for train stations of high-speed railway at 930 MHz," *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 3349–3516, Oct. 2014.
- [40] T. Zhou, C. Tao, S. Salous, and L. Liu, "Measurements and analysis of short-term fading behavior in high-speed railway communication networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 101–112, Jan. 2019.
- [41] T. Zhou, C. Tao, L. Liu and Z. Tan, "Ricean K-factor measurements and analysis for wideband high-speed railway channels at 2.35GHz," *Radioengineering*, vol. 23, no. 2, pp. 578–585, Jun. 2014
- [42] Q. Mao, F. Hu and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surv. Tuts.*, vol. 20, no. 4, pp. 2595–2621, Oct.–Dec. 2018.
- [43] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, Cambridge, MA, USA, 2016.
- [44] Z. Wu, Y.-G. Jiang, X. Wang, H. Ye, and X. Xue, "Multi-stream multiclass fusion of deep networks for video classification" in *Proc. ACM Multimedia*, New York, USA, 2016, pp. 791–800.
- [45] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos" in *Proc. NIPS*, Montreal, Canada, 2014, pp. 568–576.
- [46] S. Zhang *et al.*, "Fusing geometric features for skeleton-based action recognition using multilayer LSTM networks," *IEEE Trans. Multimedia*, vol. 20, no. 9, pp. 2330–2343, Sep. 2018.
- [47] K. He, X. Zhang, S. Ren and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification" in *Proc. ICCV*, Santiago, Chile, 2015, pp. 1026–1034.
- [48] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization" in *Proc. ICLR*, San Diego, USA, 2015, pp. 1–13.
- [49] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.
- [50] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
- [51] T. Fawcett, "Using rule sets to maximize ROC performance" in *Proc. ICDM*, San Jose, CA, USA, 2001, pp. 131–138.
- [52] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks" in *Proc. ICML*, Sydney, Australia, 2017, pp. 214–223.
- [53] N. Czink, B. Bandemer, G. Vazquez-Vilar, L. Jalloul, C. Oestges and A. Paulraj, "Spatial separation of multi-user MIMO channels" in *Proc. PIMRC*, Tokyo, Japan, 2009, pp. 1059–1063.



Tao Zhou (Member, IEEE) received the B.E. degree from the Changchun University of Science and Technology, Changchun, China, in 2009 and the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2016.

From 2014 to 2015, he was a Visiting Ph.D. student with the Centre for Communication Systems, School of Engineering and Computing Sciences, Durham University, U.K. Since 2016, he begins to work as an Associated Professor of the Institute of Broadband Wireless Mobile Communications, School of Electronics and Information Engineering, Beijing Jiaotong University. His current research interests include propagation channel characterization, channel sounding and modeling for high-speed railway communication systems.



Yingjie Wang received the B.S. degree from Ningbo University, Ningbo, China, in 2017. He is currently working toward the M.S. degree with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China. His research interests include wireless channel characterization and applications of machine learning in wireless communications.



Cheng-Xiang Wang (Fellow, IEEE) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2004.

He was a Research Assistant with the Hamburg University of Technology, Hamburg, Germany, from 2000 to 2001, a Visiting Researcher with Siemens AG Mobile Phones, Munich, Germany, in 2004, and a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005. He has been with Heriot-Watt University, Edinburgh, U.K., since 2005, where he was promoted to a Professor in 2011. In 2018, he joined Southeast University, China, as a Professor. He is also a part-time professor with the Purple Mountain Laboratories, Nanjing, China. He has authored four books, one book chapter, and more than 380 papers in refereed journals and conference proceedings, including 23 Highly Cited Papers. He has also delivered 20 Invited Keynote Speeches/Talks and 7 Tutorials in international conferences. His current research interests include wireless channel measurements and modeling, B5G wireless communication networks, and applying artificial intelligence to wireless communication networks.

Prof. Wang is a fellow of the IET, an IEEE Communications Society Distinguished Lecturer in 2019 and 2020, and a Highly-Cited Researcher recognized by Clarivate Analytics, in 2017-2019. He is currently an Executive Editorial Committee member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has served as an Editor for nine international journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2009, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2011 to 2017, and IEEE TRANSACTIONS ON COMMUNICATIONS from 2015 to 2017. He was a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Vehicular Communications and Networks (Lead Guest Editor), Special Issue on Spectrum and Energy Efficient Design of Wireless Communication Networks, and Special Issue on Airborne Communication Networks. He was also a Guest Editor for the IEEE TRANSACTIONS ON BIG DATA, Special Issue on Wireless Big Data, and is a Guest Editor for IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, Special Issue on Intelligent Resource Management for 5G and Beyond. He has served as a TPC Member, TPC Chair, and General Chair for over 80 international conferences. He received ten Best Paper Awards from IEEE GLOBECOM 2010, IEEE ICCT 2011, ITST 2012, IEEE VTC 2013-Spring, IWCMC 2015, IWCMC 2016, IEEE/CIC ICC 2016, WPMC 2016, and WOCC 2019.



Sana Salous (Senior Member, IEEE) received the B.E.E. degree from the American University of Beirut, Beirut, Lebanon, in 1978, and the M.Sc. and Ph.D. degrees from Birmingham University, Birmingham, U.K., in 1979 and 1984, respectively. She was an Assistant Professor with Yarmouk University, Irbid, Jordan, for four years followed by a one-year Research Fellowship at Liverpool University, Liverpool, U.K. She held a lectureship with the University of Manchester Institute of Science and Technology, Manchester, U.K., in 1989, where she was subsequently a Senior Lecturer and then a Reader. Since 2003, she has been the Chair of Communications Engineering, Durham University, Durham, U.K., where she is currently the Director of the Centre for Communication Systems. Her current research interests include radio channel characterization in various frequency bands ranging from skywave in the HF band to millimeter-wave bands, the design of radar waveforms, and novel radio channel sounders and radar systems for radio imaging.

Dr. Salous is a Fellow of the Institution of Engineering and Technology (FIET) and a Fellow of the International Union of Radio Science (IURSI). She was the Chair of Commission C on Radio Communication and Signal Processing Systems of IURSI, from 2014 to 2017. She is the Editor-in-Chief of the journal *Radio Science*, a publication of the American Geophysical Union.

Dr. Salous is a Fellow of the Institution of Engineering and Technology (FIET) and a Fellow of the International Union of Radio Science (IURSI). She was the Chair of Commission C on Radio Communication and Signal Processing Systems of IURSI, from 2014 to 2017. She is the Editor-in-Chief of the journal *Radio Science*, a publication of the American Geophysical Union.



Liu Liu (Member, IEEE) received the B.E. and Ph.D. degrees from Beijing Jiaotong University (BJTU), Beijing, China, in 2004 and 2010, respectively.

He was a Post Ph.D. Researcher with the Institute of Broadband Wireless Mobile Communications, School of Electronics and Information Engineering, BJTU, from 2010 to 2012. Since 2018, he has been an full Professor with the Institute of Broadband Wireless Mobile Communications, School of Electronics and Information Engineering, BJTU. His current research interests include channel measurement

and modeling for different propagation environments and signal processing of wireless communication in time-varying channel.



Cheng Tao (Member, IEEE) received the M.S. degree from Xidian University, Xian, China, in 1989, and the Ph.D. degree from Southeast University, Nanjing, China, in 1992, both in telecommunication and electronic system.

He was a Post-Doctoral Fellow with the School of Electronics and Information Engineering, Beijing Jiaotong University (BJTU), Beijing, China, from 1993 to 1995. From 1995 to 2006, he was with the Air Force Missile College and the Air Force Commander College. In 2006, he joined as an Academic Faculty

Member with BJTU, where he is currently a Full Professor and the Director of the Institute of Broadband Wireless Mobile Communications. He has authored over 50 papers and holds 20 patents. His current research interests include mobile communications, multiuser signal detection, radio channel measurement and modeling, and signal processing for communications.