WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n3.05
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Deep Learning-Based Symbol-Frame Fusion Channel Estimation Network

Xiaojuan Bai, Yunqing Li *

College of Physics and Electronic Engineering, Northwest Normal University, Lanzhou, China
*Corresponding Author: Yunqing Li (Email: lyq20222452@163.com)

ABSTRACT

In recent years, the application of deep learning technology in vehicle communication has gradually increased, especially in channel estimation. The existing channel estimation methods based on deep learning can be divided into two categories: frame by frame (FBF) estimation and symbol by symbol (SBS) estimation. FBF estimation method can obtain high estimation accuracy by processing the data of the entire OFDM frame, but there is a large delay; Although the SBS estimation method can quickly respond to channel changes, it is vulnerable to noise and multipath effects in high dynamic scenarios, resulting in reduced estimation accuracy. This paper proposes a new Symbol Frame Fusion Channel Estimation Network (SFusNet) based on deep learning. First, the detail enhanced differential convolution (DeConv) is used to replace the traditional convolution layer for FBF channel estimation to improve the feature extraction ability of the convolution layer. Then, the FBF channel estimation results are converted into one-dimensional data, and GRU is used for SBS channel estimation. The simulation results show that SFusNet performs better in restoring real channel 2D images, and has significant advantages in BER and NMSE performance compared with other existing schemes.

KEYWORDS

Deep learning; Vehicle communication; Channel estimation; FBF; SBS; Detail enhanced convolution.

1. INTRODUCTION

With the rapid development of intelligent transportation system, the Internet of Vehicles technology has become one of the key technologies to achieve efficient, safe and reliable transportation. Especially in the high-speed mobile environment, inter vehicle communication (V2V) and vehicle infrastructure communication (V2I) put forward higher requirements for the accuracy of channel estimation [1]. In order to meet the increasingly complex traffic environment requirements, IEEE 802.11p standard has been developed as an international standard for vehicle communication. Its channel estimation performance directly affects the reliability and transmission quality of the system [2]. However, the IEEE 802.11p standard originated from the famous IEEE 802.11a standard, and did not take into account the impact of high mobility. Therefore, under the condition of dual selectivity of vehicle channel, especially in the vehicle channel environment with high mobility and fast channel change characteristics, the traditional IEEE 802.11p estimation scheme (Data Pilot Assisted (DPA)) can no longer provide satisfactory performance in such actual environment.

In order to solve this problem, several classical channel estimation schemes based on DPA estimation have been proposed. The basic idea of these schemes is to use previously detected data symbols as data pilots to estimate the current channel. [3] The spectrum time averaging (STA) scheme proposed

in takes averaging processing in the time domain and frequency domain as the operation after DPA estimates the channel. This scheme obtains significant performance advantages in low signal to noise ratio (SNR) regions, but presents significant error limits in high SNR regions. [4] The scheme of constructing data pilot (CDP) proposed in is better than the STA scheme in high SNR area, but it still cannot resist the influence of modulation order by analyzing the channel correlation characteristics between two adjacent symbols as the basis for whether to update the channel. [5] The time-domain reliable test frequency-domain interpolation (TRFI) scheme proposed in improves performance by using reliable data pilots to interpolate the channel estimation at unreliable data subcarriers in the frequency domain. However, the performance of these classical channel estimation schemes can no longer meet the vehicle channel environment with high mobility and fast channel change characteristics.

Therefore, in order to solve the problem that the classical channel estimation scheme cannot effectively restrain the influence of error, the rapid development of deep learning technology in recent years has attracted people's attention and has been widely used in the field of vehicle communication. The existing channel estimation methods based on deep learning can be divided into two categories: frame by frame (FBF) estimation and symbol by symbol (SBS) estimation. FBF estimation method can obtain high estimation accuracy by processing the data of the entire OFDM frame, but there is a large delay; Although the SBS estimation method can quickly respond to channel changes, it is vulnerable to noise and multipath effects in high dynamic scenarios, resulting in reduced estimation accuracy. In [6-9], the researchers applied deep learning technology to SBS channel estimation. [6] A scheme (STA-DNN) combining STA and deep neural network (DNN) is proposed in, STA and DNN are used as initial rough channel estimation and nonlinear post-processing, respectively, so that the performance is significantly improved. Subsequently, a TRFI-DNN scheme was proposed in [7], and the STA scheme in [6] was replaced by the TRFI scheme, which further improved the performance in the high SNR region. In order to solve the relatively simple architecture of DNN, an LSTM-MLP scheme [8] is proposed. This method combines the Long Short Memory Network (LSTM) and Multi Layer Perceptron (MLP) for error compensation. Although this scheme has better performance, it has higher computational complexity. Later in [9], it was found that the GRU unit is more suitable than the LSTM unit in processing channel estimation, and the complexity is significantly reduced after removing MLP processing. However, the initial GRU unit combined with DPA scheme still has room for improvement in error suppression. [10-12] The channel estimated in the whole frame is regarded as a two-dimensional low resolution noise image, and CNN based processing is applied as super-resolution and denoising technology to finally obtain the channel estimation result in the entire OFDM frame. In [10], the frequency response matrix of the fast fading channel is regarded as a low resolution two-dimensional image, which is input into ChannelNet to obtain a complete high-resolution channel frequency response image. ChannelNet is composed of SRCNN [13] and DnCNN [14], which respectively perform super-resolution reconstruction and denoising for channel response images. Both are based on convolutional neural network (CNN). Inspired by the image super-resolution technology, [11] proposed a time-frequency channel network based on deep learning to solve the problem of channel estimation. Specifically, first obtain all rough channel estimates in a frame through a DPA based method, and then use the proposed super-resolution convolution long and short-term memory network to correct the rough estimates. Taking advantage of the high correlation between frequencies of wireless channels, [12] proposes a SisRafNet scheme that circulates along a single slot of the frequency network. Using the sequence behavior of cross frequency channels, recursive neural network technology is used in a single OFDM slot, thus overcoming the delay and memory constraints usually associated with recursive based methods.

Existing research on channel estimation methods based on deep learning uses FBF estimation or SBS estimation alone [6-12], and does not integrate the advantages of both. Therefore, this paper proposes a Symbol Frame Fusion Channel Estimation Network (SFusNet) based on deep learning. Specifically, we first use DeConv to replace the traditional convolution layer for FBF channel estimation to

improve the estimation accuracy. As an improved convolution operation, DeConv can provide higher resolution and stronger feature extraction capability without increasing computational complexity. Then, the FBF channel estimation results are converted into one-dimensional data, and the GRU is used for SBS channel estimation. As an efficient recurrent neural network, GRU can effectively capture the temporal correlation between consecutive OFDM symbols, thus further improving the accuracy and robustness of channel estimation.

2. CORRELATION THEORY

2.1. System Model

According to the physical layer description of IEEE 802.11p standard, the standard adopts orthogonal frequency division multiplexing (OFDM) technology and operates at the carrier frequency of 5.9 GHz. As shown in Figure 1, the specific description of the frame structure of the physical layer protocol data unit under the IEEE 802.11p standard is provided. During transmission, the data packet consists of the preamble, signal domain and data domain. The preamble part includes short training symbols $(t_1 - t_{10})$ and two identical long training symbols (LTS_1, LTS_2). The former is used for synchronization and packet detection, and the latter is used for channel estimation. The signal domain contains information such as data rate and packet length. The data field is used for actual data transmission. The guard interval GI2 is used for long training sequences, while GI is used for the cyclic prefix (CP) guard interval of each OFDM symbol. These protection intervals are used to reduce inter symbol interference (ISI). Each OFDM symbol contains 64 subcarriers, the index set is $\mathcal{S}_N = \{-32, -31, \dots, 31\}$, which is divided into three parts, respectively, four pilot subcarriers $\mathcal{S}_P = \{-21, -7, 7, 21\}$, 12 empty subcarriers $\mathcal{S}_E = \{-32, \dots, -27, 0, 27, \dots, 31\}$, and 48 data subcarriers $\mathcal{S}_D = \mathcal{S}_N \bigcap (\mathcal{S}_P \cup \mathcal{S}_E)^C$, where $(\cdot)^C$ represents a complementary set.

Assuming perfect synchronization, ignoring the signal field, we will focus on a frame composed of K subcarriers and I OFDM data symbols. Let $\mathcal{S}_A = \mathcal{S}_P \cup \mathcal{S}_D$ be the set of active subcarriers, and the input-output relationship between transmitted and received OFDM frames $(\mathcal{S}_A \times I)$ can be expressed as:

$$\boldsymbol{Y}(k,i) = \boldsymbol{H}(k,i)X(k,i) + \boldsymbol{N}(k,i), k \in \mathcal{S}_A \tag{1}$$

Where $\boldsymbol{H}(k,i)$ above represents the time-varying frequency response of channels of all subcarriers in the transmitted OFDM frame. Where $\boldsymbol{X}(k,i)$, $\boldsymbol{Y}(k,i)$ and $\boldsymbol{N}(k,i)$ respectively represent the transmitted OFDM symbol, the received OFDM symbol and the additive white Gaussian noise (AWGN) of the k-th subcarrier in the i-th OFDM symbol.

Since sending a large number of pilot signals in the $K \times I$ dimension data grid will greatly limit the throughput of the communication link, only a few pre positioning points contain pilot signals, while other points contain data to be transmitted. At the receiving end, the LS based channel estimation is performed to obtain $\boldsymbol{H}_{LS} \in \mathbb{C}^{|\mathcal{S}_P| \times I}$, and then the interpolation method is used to estimate the channel corresponding to each element of $\boldsymbol{H}$. In this section, only the channel $\widehat{\boldsymbol{H}} \in \mathbb{C}^{|\mathcal{S}_A| \times I}$ estimated in the active subcarrier $\mathcal{S}_A$ is considered. The frequency domain response of channel $\widehat{\boldsymbol{H}}$ is regarded as a two-dimensional image, as shown in Figure 2.

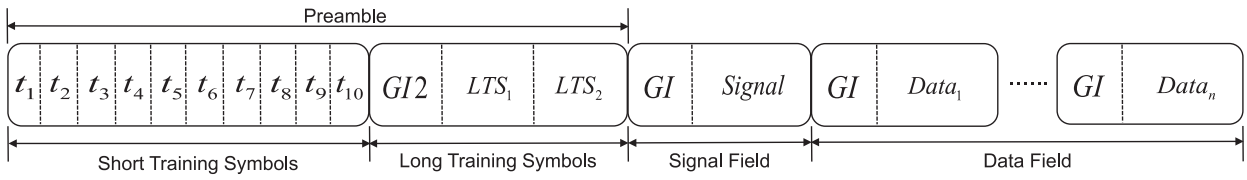

Figure 1. IEEE 802.11p transmission frame structure

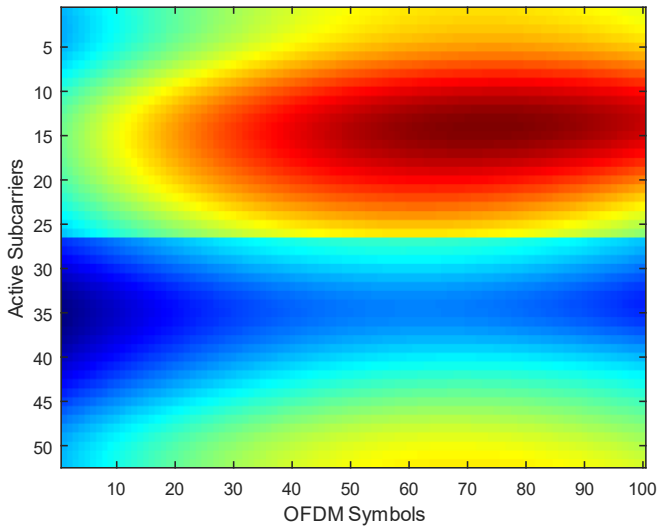

Figure 2. 2D image of real channel

2.2. Detail Enhanced Convolution

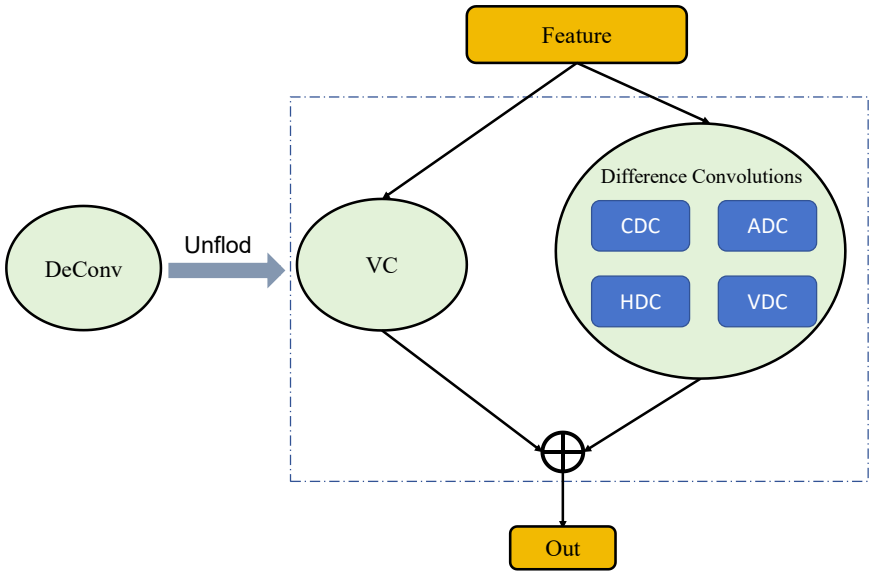

Figure 3. Unfolding structure of DeConv

Previous methods [15-17] usually use ordinary convolution layers for feature extraction and learning to process single image deblurring. However, the ordinary convolution layer searches a huge solution space without any constraints (even starting from random initialization), which limits the expression ability or modeling ability. Then researchers noticed that high-frequency information (such as edges and contours) is important for restoring images captured in blurred scenes. On this basis, some researchers used edge prior in the defogging model to help restore clearer contours [18-20]. Inspired by the work [19, 20], the author [21] designed a DeConv layer (Figure 3), which can integrate carefully designed prior information into ordinary convolution layers.

DeConv uses five convolutional layers deployed in parallel for feature extraction, four of which are differential convolutions (central differential convolution (CDC) and angular differential convolution (ADC), horizontal differential convolution (HDC) and vertical differential convolution (VDC)), and the other is ordinary convolution (VC). In difference convolution, a pixel pair difference calculation strategy can be designed to explicitly encode prior information into CNN. These convolution layers

are simplified into a standard convolution layer through the re parameterization technology, thus keeping the number of parameters unchanged and not adding additional calculation costs.

In DeConv, ordinary convolution is used to obtain intensity level information, while differential convolution is used to enhance gradient level information. If the learned features are simply added together, the output of DeConv can be obtained. However, deploying five parallel convolution layers for feature extraction will not ideally lead to an increase in parameters and reasoning time. Therefore, the additivity of the convolution layer is used to simplify the convolution of parallel deployment into a single standard convolution. The author of [45] noticed a useful attribute of convolution: if several 2D kernels have the same input and the same stride and fill and produce output with the same size operation, their output and final output can be obtained, then these kernels can obtain an equivalent kernel at the corresponding position to produce the same final output, and DeConv is completely suitable for this situation. Given the input feature F_{in}, DeConv can use the re parameterization technology to output F_{out} to the common convolution layer with the same calculation cost and reasoning time.

DeConv adds the results of all five convolutions to get the final output feature:

$$F_{out}^{DEConv} = F_{in} * W_{VC} + F_{in} * W_{CDC} + F_{in} * W_{ADC} + F_{in} * W_{HDC} + F_{in} * W_{VDC} \tag{2}$$

Specifically, the operation for each convolutional layer can be formulated as follows:

$$F_{out}^{DEConv} = \sum_{i=1}^{5} F_{in} * W_i \tag{3}$$

Where W_i represents the convolution core of VC, CDC, ADC, HDC and VDC respectively.

In order to reduce the number of parameters and speed up the training and testing process, the additive property of the convolution layer is used for re parameterization, as shown below:

$$F_{out}^{DeConv} = F_{in} * \left(\sum_{i=1}^{5} W_i\right) = F_{in} * W_{cvt} \tag{4}$$

Where W_{cvt} is the equivalent convolution kernel obtained by adding five convolution kernels by position.

Finally, the formula (4) is simplified to get:

$$F_{out}^{DeConv} = DeConv(F_{\text{in}}) \tag{5}$$

Where $DeConv(\cdot)$ represents the operation of DeConv.

2.3. Gate Recurrent Unit

Gate Recurrent Unit (GRU) is one of the two most classic variants of RNN (the other is LSTM) [10]. Compared with LSTM, GRU has a simpler gating mechanism. There are only two gating mechanisms: update gate and reset gate. As shown in Figure 4, it combines the forgetting gate and the input gate in LSTM into a single "update gate". It also merges the cell state and the hidden state, and makes some other changes. Compared with LSTM, it has fewer parameters and simpler calculation process. Its calculation formula is as follows:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{6}$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{7}$$

$$\widetilde{h_t} = tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{8}$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \widetilde{h_t} \tag{9}$$

Where r_t and z_t represent reset gate and update gate respectively, and W, U and b represent weight and offset parameters respectively. The reset gate r_t controls the combination of the current input x_t and the previous time step output h_{t-1}. A smaller r_t value will result in a smaller contribution from the previous output state. $\sigma(\cdot)$, $\tanh(\cdot)$ and $\odot$ represent sigmoid function, tanh function and hadamard product operation respectively.

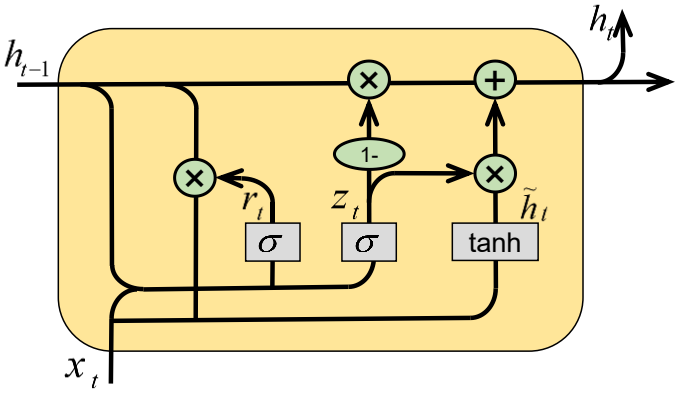

Figure 4. Structure of GRU

3. PROPOSED SCHEME

3.1. SFusNet Scheme

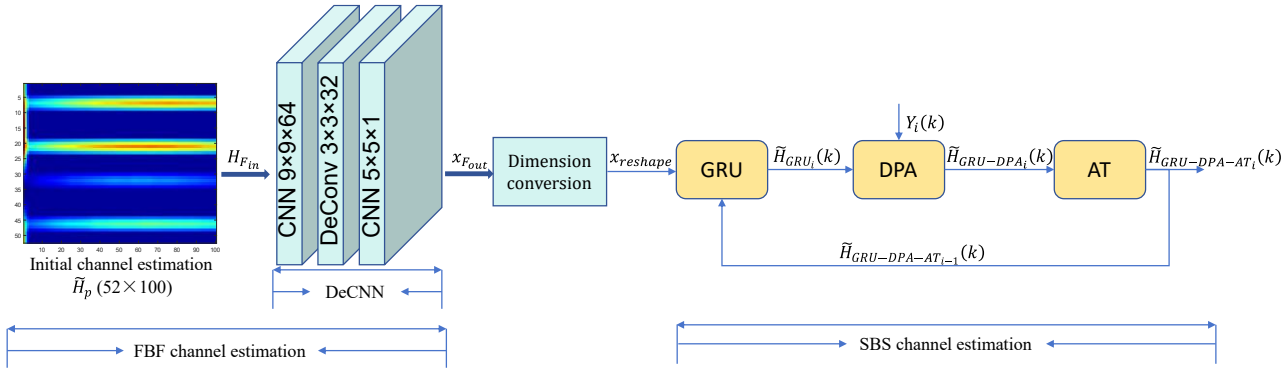

Figure 5. Process of the proposed SFusNet scheme

First, estimate the channel at the pilot position to obtain $\boldsymbol{H}_p \in \mathbb{C}^{|S_P| \times I}$, and then interpolate the estimated channel at the data symbol to obtain the initial channel estimate $\widetilde{\boldsymbol{H}}_p \in \mathbb{C}^{|S_A| \times I}$ (as shown in Figure 5). Then the real part and imaginary part are separated to obtain the channel estimation characteristic graph $\boldsymbol{H}_{F_{in}} \in \mathbb{R}^{2|S_A| \times I}$ used to input the convolutional neural network.

$$\boldsymbol{H}_{F_{in}}(i,k) = \mathrm{Re}\left(\widetilde{\boldsymbol{H}}_p(i,k)\right) + j \cdot Im\left(\widetilde{\boldsymbol{H}}_p(i,k)\right), k \in S_A \tag{10}$$

Input the channel estimation feature map $\boldsymbol{H}_{F_{in}}$ into the convolution layer. Since it is a single image input, the input of CNN can be expressed as $\boldsymbol{x}_{F_{in}} \in \mathbb{R}^{1 \times 2|S_A| \times I}$. The output $\boldsymbol{x}_{F_{out}} \in \mathbb{R}^{1 \times 2|S_A| \times I}$ is as follows:

$$\boldsymbol{x}_{F_{out}} = f_{DeCNN}\left(\theta_{DeCNN}; \boldsymbol{x}_{F_{in}}\right) \tag{11}$$

Where $f_{DeCNN}(\cdot)$ represents the operation process of DeCNN, and the DeCNN structure is composed of two CNN layers and one DeConv layer. The specific structure is shown in Figure 4. θ_{DeCNN} represents the parameter in DeCNN.

Next, we transform $\boldsymbol{x}_{F_{out}}$ from a two-dimensional image to a one-dimensional sequence $\boldsymbol{x}_{reshape} \in \mathbb{R}^{2|\mathcal{S}_A| \times I}$, which can also be regarded as the output channel estimation characteristic graph $\boldsymbol{H}_{F_{out}} \in \mathbb{R}^{2|\mathcal{S}_A| \times I}$. This process can be achieved through reshape operations:

$$\boldsymbol{x}_{reshape} = reshape\left(\boldsymbol{x}_{F_{out}}, (I, 2|\mathcal{S}_A|)\right) \tag{12}$$

In this way, the dimension of $\boldsymbol{x}_{GRU}$ becomes a dimension that can be used as GRU input data. Each time step corresponds to an OFDM symbol, and each feature vector contains the features extracted from DeCNN.

It is worth noting that the next channel processing becomes SBS channel estimation. Different from other SBS estimation, the initial input of GRU in this paper becomes $\boldsymbol{x}_{reshape}$ instead of LS based channel estimation.

Because the pilot subcarriers in the channel estimation after feature extraction by DeCNN are known information, in order to more accurately evaluate the performance of the proposed model, only the data subcarriers are estimated in this chapter. Therefore, the input $\widetilde{\boldsymbol{x}}_{GRU} \in \mathbb{R}^{2|\mathcal{S}_D| \times I}$ of GRU is expressed as:

$$\widetilde{\boldsymbol{x}}_{\boldsymbol{GRU_i}} = \widetilde{\boldsymbol{H}}_{GRU-DPA-AT_{i-1}}(k), k \in \mathcal{S}_D \tag{13}$$

Then $\widetilde{\boldsymbol{x}}_{GRU_i}$ is processed by GRU, and the following results are obtained:

$$\widetilde{\boldsymbol{H}}_{GRU_i}(k) = \Omega\left(\widetilde{\boldsymbol{x}}_{GRU_i}, \theta_{GRU}\right) \tag{14}$$

Where $\Omega(\cdot)$ is processed by GRU, and the overall weight is represented by θ_{GRU}.

Therefore, for DPA channel estimation, the $i-th$ received OFDM symbol is used for the following DPA channel estimation:

$$\boldsymbol{X}_{GRU-DPA_i}(k) = D\left[\frac{\boldsymbol{Y}_i(k)}{\widetilde{\boldsymbol{H}}_{GRU-DPA_{i-1}}(k)}\right], \widetilde{\boldsymbol{H}}_{GRU-DPA_0}(k) = \boldsymbol{x}_{reshape}(k) \tag{15}$$

Where $D[\cdot]$ refers to the operation of mapping the equalization symbol to the nearest constellation point during demodulation, and $\boldsymbol{X}_{GRU-DPA_i}(k)$ refers to the estimated transmission signal.

The final DPA channel estimation is as follows:

$$\widetilde{\boldsymbol{H}}_{GRU-DPA_i}(k) = \frac{\boldsymbol{Y}_i(k)}{\boldsymbol{X}_{GRU-DPA_i}(k)} \tag{16}$$

To mitigate the impact of AWGN, AT processing is applied to the DPA estimated channel. The results are as follows:

$$\widetilde{\boldsymbol{H}}_{GRU-DPA-AT_i}(k) = (1 - \alpha)\widetilde{\boldsymbol{H}}_{GRU-DPA-AT_{i-1}}(k) + \alpha\,\widetilde{\boldsymbol{H}}_{GRU-DPA_i}(k) \tag{17}$$

Where α is a dynamic parameter whose value is calculated according to the current SNR to ensure that α changes between 0 and 1. Specifically, the calculation method of α is:

$$\alpha = min\left(1, max\left(0, \frac{\text{SNR}}{30}\right)\right) \tag{18}$$

This means that when the SNR is low, α is close to 0, thus reducing the contribution of the new estimation to the current channel estimation; On the contrary, when the SNR is high, α approaches 1, increasing the contribution of the new estimation to the current channel estimation. This allows adaptive adjustment of channel estimation update speed with the change of SNR, effectively reducing the impact of AWGN.

3.2. Performance Evaluation

In the offline training phase of the scheme, the neural network is trained to capture more features of time-frequency correlation in the channel. In the training process, the loss function needs to be defined as the standard of neural network training. In FBF channel estimation, DeCNN is used to train the $Loss_{DeCNN}$ function of neural network as follows:

$$Loss_{DeCNN} = \frac{1}{|\mathcal{S}_A|I}\sum_{i=1}^{I}\sum_{k\in\mathcal{S}_A}||\widehat{\boldsymbol{H}}_i(k) - \boldsymbol{H}_{F_{out_i}}(k)||^2 \tag{19}$$

Where $\widehat{\boldsymbol{H}}_i(k)$ represents the real channel estimation at the active subcarrier, and $\boldsymbol{H}_{F_{out_i}}(k)$ represents the channel estimation processed by DeCNN.

In SBS channel estimation, GRU is used to train the $Loss_{GRU}$ function of the neural network as follows:

$$Loss_{GRU} = \frac{1}{|\mathcal{S}_D|I}\sum_{i=1}^{I}\sum_{k\in\mathcal{S}_D}||\widehat{\widetilde{\boldsymbol{H}}}_i(k) - \widetilde{\boldsymbol{H}}_{GRU_i}(k)||^2 \tag{20}$$

Where $\widehat{\widetilde{\boldsymbol{H}}}_i(k)$ represents the real channel estimation at the data subcarrier, and $\widetilde{\boldsymbol{H}}_{GRU_i}(k)$ represents the channel estimation after GRU processing.

4. SIMULATION ANALYSIS

4.1. Parameter Setting

The channel model considered in this paper is based on the TDL vehicle channel model proposed in [23]: RTV-Urban Canyon (RTV-UC): It captures roadside to vehicle communication in urban environment with severe multipath fading. This scenario simulates the communication between roadside infrastructure and vehicles in an urban canyon.

In order to compare with other schemes, the simulation parameters are set as follows: the number of OFDM symbols in each packet is set to I=100; The training set, verification set and test set of the neural network contain 10000, 2000 and 2000 samples respectively; The neuron parameters of STA-DNN and TRFI-DNN were set to 15-15-15; The hidden layer dimension of GRU is 128; The SisRafNet scheme consists of eight CNN layers and a two-way GRU with a hidden layer dimension of 140. The specific CNN is set in [12]; The DeCNN part of the proposed scheme SFusNet consists of 64 9 × 9 filter CNN layers, 32 3 × 3 filter DeConv layers and 1 5 × 5 filter CNN layer. The hidden layer dimension of GRU is 128. All the schemes are offline trained under the condition that the signal to noise ratio (SNR) is 30dB, and all of them are implemented in 16QAM modulation mode.

4.2. Performance Analysis

According to Figure 6, we can clearly see the comparison between the proposed SFusNet scheme and the single SBS channel estimation method. By comparing the channel estimated by SFusNet with the channel estimated by GRU-DPA, the channel estimated by SFusNet (on the left) is smoother and more detailed overall. This shows that SFusNet can effectively capture the subtle changes in the channel and show higher accuracy when processing OFDM symbols and data subcarriers. In contrast, the GRU-DPA estimated channel (on the right) can also capture some major trends, but it is rough in detail processing, especially in the edge area and complex changes of the channel. This shows that SFusNet can quickly respond to the short-term changes of the channel in a short time by combining the SBS and FBF methods, and improve the accuracy of the overall estimation by using the information between frames. This fusion method significantly improves the adaptability to complex dynamic environments. In addition, the performance is further improved by using the advantages of DeCNN method in restoring two-dimensional images.

Figure 7 shows the comparison of BER performance curves of various channel estimation schemes. It can be seen from the figure that the SFusNet scheme shows significant advantages in the whole SNR range. Especially after the SNR is 20dB, the BER of the SFusNet scheme starts to be significantly lower than that of other schemes. In the 10^{-5} phase, SFusNet has a gain of nearly 2.5dB compared with TS ChannelNet, and with the increase of SNR, the advantages of SFusNet are more obvious. This means that under the same SNR condition, SFusNet can provide lower BER, better recover the original signal, and thus ensure more reliable communication quality.

To further explore the accuracy of channel estimation in the SFusNet scheme, we compare the NMSE performance of other schemes in Figure 8. It can be seen from the figure that the SFusNet scheme shows significant advantages in the whole SNR range. The SFusNet scheme shows significantly better performance than other schemes in the RTV channel through the symbol and frame fusion method and the application of deep learning technology. SFusNet can provide lower NMSE under both low SNR and high SNR conditions, thus ensuring more accurate channel estimation.

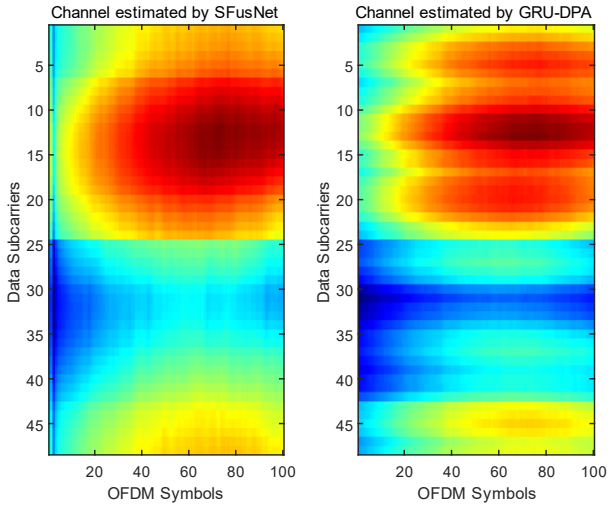

Figure 6. Performance of Different Methods for Restoring Real Channel 2D Images

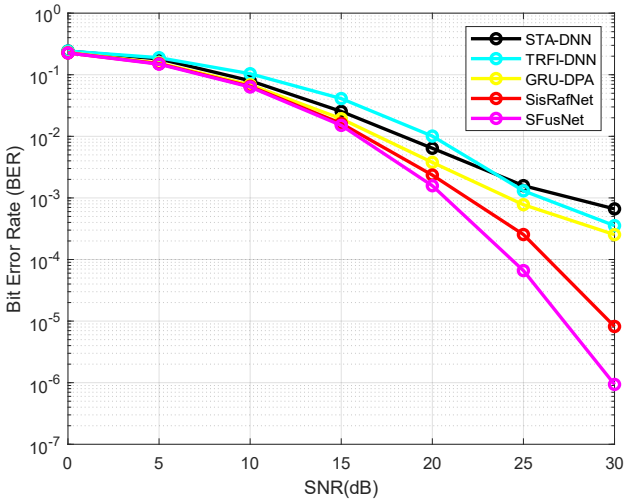

Figure 7. BER Performance Comparison of Channel Estimation Schemes

According to the performance comparison in Figure 7 and Figure 8, when the channel estimation scheme performs well on BER and NMSE, it means that the scheme can not only achieve high-precision channel estimation, but also ensure efficient and reliable data transmission. Low BER means that the error rate of data transmission is low, ensuring the accuracy of information transmission; Low NMSE means that the error between the channel estimation value and the true value is small, which improves the accuracy of signal recovery. In addition, these two excellent indicators also show that the system has strong anti noise ability and good adaptability, and can maintain stable performance even in complex or changing channel environments. In general, the dual excellence of BER and NMSE means that the scheme can provide high-quality services in a variety of application scenarios and effectively address complex communication challenges.

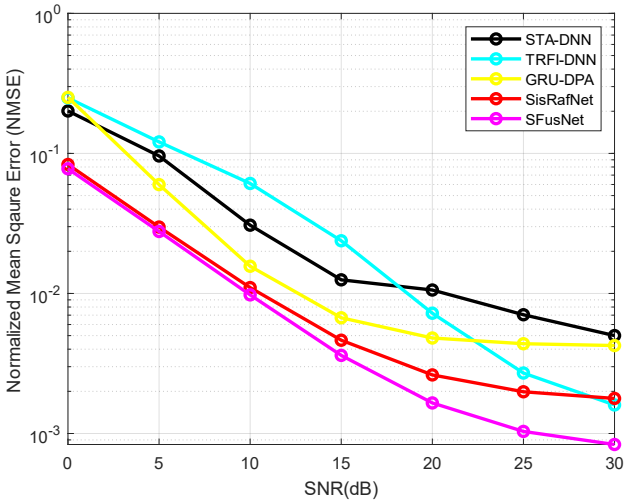

Figure 8. NMSE Performance Comparison of Channel Estimation Schemes

5. SUMMARY

This paper proposes a channel estimation method based on deep learning symbol and frame fusion, which is called SFusNet scheme. This scheme combines DeConv and GRU, and aims to improve the channel estimation performance in vehicle communication environment by using the advantages of FBF and SBS estimation methods. First, we use DeConv instead of the traditional convolution layer to estimate the FBF channel, to improve the accuracy of the estimation without increasing the computational complexity. As an improved convolution operation, DeConv can not only provide

higher resolution, but also extract feature information more effectively. Secondly, the results of FBF channel estimation are converted into one-dimensional data, and GRU is used for SBS channel estimation. As an efficient recurrent neural network, GRU can capture the temporal correlation between consecutive OFDM symbols, thus further improving the accuracy and robustness of channel estimation. The simulation results show that, compared with the SBS channel estimation method alone, the SFusNet scheme performs better in restoring the two-dimensional image of the real channel, especially in the details processing. In addition, the SFusNet scheme has significant advantages over other existing schemes in BER and NMSE performance.

REFERENCES

[1] J. Wang, C. Li, H. Li and Y. Wang. Key Technologies and Development Status of Internet of Vehicles. 2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2017, pp. 29-32.

[2] IEEE Standard for Information Technology-Local and Metropolitan Area Networks Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments [M]. New York: 2010.

[3] Joseph A. Fernandez, Kevin Borries, Lin Cheng, B. V. K. Vijaya Kumar, Daniel D. Stancil, and Fan Bai. Performance of the 802.11p physical layer in vehicle-to-vehicle environments [J]. IEEE Transactions on Vehicular Technology, 2011, 61(1): 3-14.

[4] Zijun Zhao, Xiang Cheng, Miaowen Wen, Bingli Jiao and Cheng-Xiang Wang. Channel estimation schemes for IEEE 802.11p standard [J]. IEEE Intelligent Transportation Systems Magazine, 2013, 5(4): 38-49.

[5] Yoon-Kyeong Kim, Jang-Mi Oh, Yoo-Ho Shin and Cheol Mun. Time and frequency domain channel estimation scheme for IEEE 802.11p [C].17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014: 1085-1090.

[6] Abdul Karim Gizzini, Marwa Chafii, Ahmad Nimr, and Gerhard Fettweis. Deep learning based channel estimation schemes for IEEE 802.11p standard [J]. IEEE Access, 2020, 8: 113751-113765.

[7] Gizzini A. K, Chafii M, Nimr A, et al. Joint TRFI and deep learning for vehicular channel estimation [C]. 2020 IEEE Globecom Workshops GC Wkshps, 2020, 1-6.

[8] Pan J, Shan H, Li R, et al. Channel estimation based on deep learning in vehicle-to-everything environments [J]. IEEE Communications Letters, 2021, 25(6): 1891-1895.

[9] J. Hou, H. Liu, Y. Zhang, W. Wang and J. Wang. GRU-Based Deep Learning Channel Estimation Scheme for the IEEE 802.11p Standard [J]. IEEE Wireless Communications Letters, 2023, 12(5): 764-768.

[10] Soltani M, Pourahmadi V, Mirzaei A, et al. Deep learning-based channel estimation [J]. IEEE Communications Letters, 2019, 23(4): 652-655.

[11] Xuchen Zhu, Zhichao Sheng, Yong Fang, and Denghong Guo. A deep learning-aided temporal spectral ChannelNet for IEEE 802.11p-based channel estimation in vehicular communications [J]. EURASIP Journal on Wireless Communications and Networking, 2020, 94: 1-15.

[12] A. S. M. M. Jameel, A. Malhotra, A. E. Gamal and S. Hamidi-Rad. Deep OFDM Channel Estimation: Capturing Frequency Recurrence [J]. IEEE Communications Letters, 2024, 28(3):562-566.

[13] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 38(2): 295-307.

[14] Zhang K, Zuo W, Chen Y, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising [J]. IEEE transactions on image processing, 2017, 26(7): 3142- 3155.

[15] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia. Ffa-net: Feature fusion attention network for single image dehazing [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020(34)7: 11908–11915.

[16] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma. Contrastive learning for compact single image dehazing [J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:10551–10560.

[17] H. Wu, J. Liu, Y. Xie, Y. Qu, and L. Ma. Knowledge Transfer Dehazing Network for NonHomogeneous Dehazing [J]. CVPR Workshop IEEE, 2020:1975–1983.

[18] H. Zhang and V. M. Patel. Densely connected pyramid dehazing network [J]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018:3194–3203.

[19] C. Wang, H.-Z. Shen, F. Fan, M.-W. Shao, C.-S. Yang, J.-C. Luo, and L.-J. Deng. Eaa-net: A novel edge assisted attention network for single image dehazing [J]. Knowledge-Based Systems, 2021(228):107279.

[20] H. Bai, J. Pan, X. Xiang, and J. Tang. Self-guided image dehazing using progressive feature fusion [J]. IEEE Transactions on Image Processing, 2022(31): 1217–1229.

[21] K. Cho et al., “Learning phrase representations using RNN encoder– decoder for statistical machine translation,” in Proc. Conf. Empir. Methods Nat. Lang. Process. (EMNLP), Oct. 2014, pp. 1724–1734.

[22] G. Acosta-Marum and M. A. Ingram, "Six time- and frequency- selective empirical channel models for vehicular wireless LANs," IEEE Vehicular Technology Magazine, vol. 2, no. 4, pp. 4-11, Dec. 2007.