Zhu *et al. EURASIP Journal on Wireless Communications and Networking* (2020) 2020:94
https://doi.org/10.1186/s13638-020-01714-4

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

A deep learning-aided temporal spectral ChannelNet for IEEE 802.11p-based channel estimation in vehicular communications

Xuchen Zhu, Zhichao Sheng, Yong Fang* and Denghong Guo

*Correspondence:
yfang@staff.shu.edu.cn
Shanghai Institute for Advanced Communication and Data Science, Key laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, 200444 Shanghai, China

Abstract

In vehicular communications using IEEE 802.11p, estimating channel frequency response (CFR) is a remarkably challenging task. The challenge for channel estimation (CE) lies in tracking variations of CFR due to the extremely fast time-varying characteristic of channel and low density pilot. To tackle such problem, inspired by image super-resolution (ISR) techniques, a deep learning-based temporal spectral channel network (TS-ChannelNet) is proposed. Following the process of ISR, an average decision-directed estimation with time truncation (ADD-TT) is first presented to extend pilot values into tentative CFR, thus tracking coarsely variations. Then, to make tentative CFR values accurate, a super resolution convolutional long short-term memory (SR-ConvLSTM) is utilized to track channel extreme variations by extracting sufficiently temporal spectral correlation of data symbols. Three representative vehicular environments are investigated to demonstrate the performance of our proposed TS-ChannelNet in terms of normalized mean square error (NMSE) and bit error rate (BER). The proposed method has an evident performance gain over existing methods, reaching about 84.5% improvements at some high signal-noise-ratio (SNR) regions.

Keywords: Vehicular communications, Channel estimation, IEEE 802.11p, Deep learning, Image super resolution

1 Introduction

Vehicular communications, which form a network to support vehicle-to-vehicle (VTV) and vehicle-to-infrastructure (VTI) communications, are essential techniques of intelligent transportation system (ITS). In recent years, lots of attention has been drawn to develop multiple applications in vehicular communications such as automatic selection of routing protocol[1]. To realize such high-speed mobile communications, the IEEE 802.11p standard [2], that defines the physical layers (PHY) and the medium-access layers (MAC), has been officially applied in 2010. The IEEE 802.11p is a modified version of 802.11a [3]. The main difference between them is that 802.11p facilitates the half of frequency bandwidth of 802.11a, thus making signals more robust against fading and

multipath propagation effects in vehicular environments[4]. What is more, it can support lower latency, realize higher data rate, and enhance security compared to other standards [5].

Channel estimation (CE) schemes play a crucial role in the performance of vehicular communication using 802.11p. The estimated channel response (CR) significantly affects the subsequent equalization, demodulation, and decoding. In general, the accuracy of CE determines the performance of the whole system. However, in PHY layer, the 802.11p protocol utilizes four pilot subcarriers per one OFDM symbol. The pilot positions are too loose to adequately track variations of channel frequency response (CFR). In addition, due to the fact that CR varies greatly in vehicular communications, coupled with no restrictions of the packet length in 802.11p, channel estimation (CE) keeps easily outdated during the entire packet.

A lot of work has been proposed to track channel variations over the frame duration for vehicular communication under the IEEE 802.11p standard. The current method focuses on data pilot-aided successive (DPAS) whose key is to consider the demapped data signals as aided pilot [6–9]. The performance gain, however, is not evident especially at high signal-noise-ratio (SNR) region because of the error propagation caused by accumulated noise in the iterative process. Recently, deep learning (DL) has shown impressively promising prospects. DL enables to extract inherent characteristic of signals and is applied for channel estimation [10–14]. However, due to deep fading caused by high Doppler shift under vehicular environment, above DL-based approaches degrades in the accuracy of CE. The main objective of this paper is to estimate precisely CFR by integrating DPAS with DL.

In this paper, a deep learning-aided temporal spectral channel network (TS-ChannelNet) for 802.11p-based channel estimation under high-speed vehicular scenarios is proposed to track variations of CFR. In general, the pilot taken as low-resolution (LR) version of CR is utilized to recover high-resolution (HR) version of CR by TS-ChannelNet. Our presented TS-ChannelNet consists of two phases. Initially, coarse CR is restored via pilot by leveraging averaging decision-directed with time truncation (ADD-TT). By averaging both in time and frequency domains, ADD-TT handles few of the impacts of error propagation caused by time truncation based on decision feedback. Afterwards, a neural network (NN) architecture named super resolution convolutional long short-term memory (SR-ConvLSTM) is introduced to make estimated CR more accurate. SR-ConvLSTM utilizes the power of convolutional long short-term memory (ConvLSTM) that exploits temporal spectral correlation to combat deep fast fading under the extremely time-varying environments. The obtained CR is tailored for vehicular communications. Simulation results demonstrate that our method is competent over previous methods under representative vehicular scenarios. Our contributions of this paper are summarized as follows.

- CFR is modeled as an image. The pilot is considered as a LR version of the image. The estimated CFR is viewed as a HR version of the image. Then, the TS-ChannelNet, which includes pilot-based interpolation and DL-based restoration, is presented to obtain HR version of CFR.
- An improved interpolation based on DD-TT called ADD-TT is taken to extend pilot into reasonable initial coarse CFR. ADD-TT handles few impacts of error

propagation by time truncation based on decision feedback and further improves the performance of the follow-up SR-ConvLSTM.
- The new super resolution technique-based architecture named SR-ConvLSTM is designed. It restores HR version of CFR by reflecting highly variations of channel.
- The extensive ablation experiment is conducted to verify that SR-ConvLSTM powerfully extract temporal spectral correlation of signal to track the variations of channel.

The rest of this paper is organized as follows. Section 2 illustrates related work in details. Section 3 introduces the system model, channel model, and benchmark algorithm. Section 4 presents our temporal spectral deep learning-based channel estimation scheme. Section 5 verifies the full advantage of TS-ChannelNet by simulation results. Section 6 concludes the paper.

2 Related work

In this section, the existing work of CE under vehicular communications using 802.11p standard is first elaborated. The downside of present work is then introduced. Furthermore, DL applied in the communication field, as a promising prospect, is investigated.

In few years, mobile ad hoc network (MANET) has successfully applied in amounts of field, such as health care [15, 16], broadcast encryption [17], vehicular streaming service [18], and urban management [19]. CE has been investigated actively because it decides the performance of the system in PHY layer [7]. The current CE focuses on DPAS method, such as STA [6] and CDP [7]. The key part of these algorithms is to consider the demapped data signals as aided pilot. Then, the estimated CR is iteratively used to construct data pilot in the follow-up orthogonal frequency division multiplexing (OFDM) symbol. Mehrabi [9] introduced decoded data bits into DPAS to suppress noise caused by demodulation, but the performance gain is still marginal at high SNR region. To further improve accuracy of CFR, Awad [8] transformed CFR into time domain and performed truncation operation, thus removing demodulation errors. However, because iterative accumulated noise is not eliminated completely, these schemes still suffer from error propagation especially in the rapid time-varying vehicular channels.

Compared to the conventional schemes, DL has been shown to extract powerfully the inherent characteristic of signals [20] and thus has been qualified when overcoming multiple problems in wireless communications field [21–25]. FCNN was utilized into channel estimation and pilot design [11, 12]. It initially demonstrates the powerful ability of DL to increase improvement in the accuracy of CE. However, the scheme is not fit for vehicular communications using 802.11p. Because the unlimited packet length in 802.11p leads to increase rapidly in the number of neurons and thus FCNN tends to overfit. Neumann [13] modeled channel as conditionally Gaussian distribution given a set of random hyperparameters. Those hyperparameters are learned via convolutional neural network (CNN). Soltani [14] viewed channel estimation as an image super resolution problem where the pilot was a low-resolution sampled version of the channel and time-frequency CR was the image to be recovered. But the performance of the method still degrades under fast time-variant environment. The goal of this paper is to integrate DPAS with DL to track variations of channel, thus estimating CFR precisely.

3 System model

In this section, the structure of IEEE 802.11p under vehicular communications is first presented. Then, the channel model for vehicular wireless environment employed in this paper is briefly introduced. Subsequently, ChannelNet applied as benchmark algorithm is elaborated.

3.1 Structure of IEEE 802.11p

IEEE 802.11p physical layer is based on OFDM which boosts spectrum utilization by turning serial large data streams into parallel data streams on orthogonal subcarriers.

In 802.11p, the received signal is turned into parallel data for fast Fourier transformation (FFT) input, thus obtaining follow-up output in the frequency domain.

$$Y(t,k) = H(t,k)X(t,k) + Z(t,k). \tag{1}$$

where $Y(t,k)$ and $X(t,k)$ represents received, transmitted OFDM data symbols using FFT respectively, $H(t,k)$ represents the CFR of the wireless channel, and $Z(t,k)$ is added white Gaussian noise (AWGN). t represents the index of length per frame with $1 \leq t \leq T$. T is the number of length per frame. k denotes the index of subcarriers per frame with $1 \leq k \leq K$. K is the number of subcarrier per frame. How to estimate H more accurately is the goal of this paper.

IEEE 802.11p defines 75 MHz band at 5.9 GHz. The 75 MHz bandwidth is divided into 7 channels including one control channel (CCH) and six service channels. Safety messages are transmitted through CCH when emergent events happen [26]. IEEE 802.11p standard defines that pilot tones for channel estimation is comb structure. It is located on subcarriers -21, -7, 7, and 21 as Fig. 1 is shown. The initial channel estimation is enabled by utilizing the known training symbols transmitted of the preamble. Due to the highly time-varying channel in vehicular environments and the fact that the frame length is unlimited in the IEEE 802.11p standard, the channel estimation for each packet outdates easily over

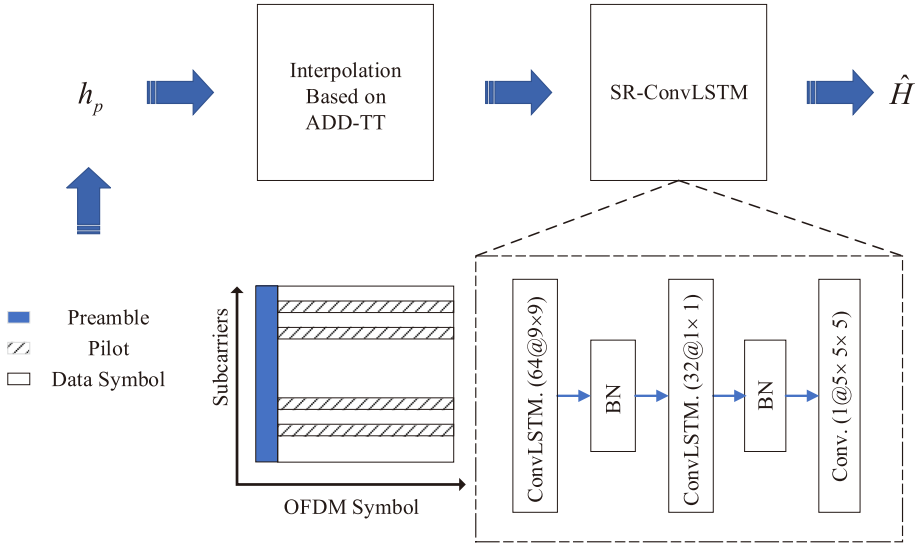

Fig. 1 The structure of TS-ChannelNet

the entire packet duration. Therefore, how to design a channel estimation scheme to track variations under vehicular channel is a challenging problem.

3.2 Channel model for vehicular communications

Due to the relative motion of the transmitter and receiver, a Doppler spectral spread or broadening appears under vehicular communication. The relatively high velocity causes fast time-varying CR. To capture the joint Doppler-delay characteristics of vehicular communications, the tapped-delay line (TDL) model is adopted following the parameter of [27]. In [27], taps are characterized by Doppler power spectral density due to Rayleigh fading. The channel impulse response is calculated as (2)

$$h(t,\tau) = \sum_{l=1}^{L} \phi_l(t)\delta(\tau - \tau_l(t)). \tag{2}$$

where $\phi_l(t)$ represents the fading coefficient, L denotes resolution multipath, δ is impulse function, and $\tau_l(t)$ denotes time delay in lth path.

In this paper, three representative models are given as in [27], i.e., VTV Expressway Oncoming (VTVEO), VTV Urban Canyon (VTVUC) Oncoming, and RTV Expressway (RTVE). In the VTV Expressway Oncoming scenario, the moving speed of the receiver and the transmitter is the highest compared to the other scenarios. Its speed is 100km/h and its Doppler shift is about 1200Hz. Then, the VTV Urban Oncoming is the medium challenging environment for channel estimation. Its Doppler shift is 400–500 Hz with about 32 km/h moving velocity. In conclusion, the models presented are typical standard vehicular environments that consist of different velocities (low velocity/high velocity), and a Doppler shift ranged from 400 to 1200 Hz.

3.3 Benchmark algorithm: ChannelNet

In [14], a deep learning-based channel estimation scheme named ChannelNet was implemented for the short length of frame in slow time-varying environment. By viewing CR as images, the pilot values were utilized via image super resolution technique to restore (estimate) CR.

The process of ChannelNet consists of two phases. On the one hand, the isolated pilot values are extended to initial CR via Gaussian interpolation. On the other hand, CR values as input are fed into super resolution neural network (SRCNN) [28] followed by denoising convolutional neural network (DnCNN) [29]. The NN generates the estimated CR. The authors investigate the performance of ChannelNet in relatively slow time-varying environment. In our experimental trial, ChannelNet furthermore degrades for high-velocity mobile communications. This is owing to the unreliability of initial interpolation method, coupled with the fact that CNN does not have enough capacity to uncover temporal spectral correlation of the CR, thereby keeping CR outdated over the frame duration.

4 Proposed method

In this section, we first describe the pre-process of TS-ChannelNet. It utilizes interpolation scheme based on ADD-TT via pilot values. Then, the NN architecture named SR-ConvLSTM is presented to track variations of vehicular channel. Afterwards, the training process of TS-ChannelNet that is made up of ADD-TT following by SR-ConvLSTM is illustrated.

4.1 Interpolation based on ADD-TT

In this subsection, interpolation based on pilot via ADD-TT is implemented to obtain coarse CR. It extends few pilot values to initial CR values that are taken as IR images.

Usually, least squares (LS) estimation utilizes two identical preambles which are sent at the beginning of received packet in IEEE 802.11p to estimate tentative CR. $Y(1,k)$, $Y(2,k)$ are the first two long training symbols. $X(1,k)$, $X(2,k)$ are identical and two transmitted predefined long symbols in the frequency domain. To obtain CFR for all subcarriers, the received $Y(1,k)$ and $Y(2,k)$ are divided by $X(1,k)$ as

$$\hat{H}_{LS}(1,k) = \frac{Y(1,k) + Y(2,k)}{2X(1,k)}, \tag{3}$$

where $\hat{H}_{LS}(1,k)$ represents the LS channel estimate at the 1th time slot on the kth subcarrier. LS estimation assumes the channel is stationary. However, vehicular channel varies fast and the performance of LS estimation degrades significantly.

Then, decision-directed channel estimation is presented. It is based on correlation of adjacent symbols. The symbols are equalized by previous channel estimation as follows

$$\hat{S}(t,k) = \frac{Y(t,k)}{\hat{H}(t-1,k)}, \tag{4}$$

where $\hat{S}(t,k)$ denotes equalized symbol at the tth time slot on the kth subcarrier and $\hat{H}(t-1,k)$ is the previous channel estimation. Based on the high correlation between adjacent data symbols, the current tth CFR is assumed to be unchanged with respect to the previous. The errors caused by such assumption are alleviated by the subsequent demodulation. Hence, the previous $\hat{H}(t-1,k)$ is utilized to estimate. Note that the first estimated CR is LS channel estimation using (2). Then, the decision feedback is used to update channel estimate according to (5)

$$\hat{H}(t,k) = \frac{Y(t,k)}{\hat{X}(t,k)}, \tag{5}$$

where $\hat{X}(t,k)$ represents the demodulated OFDM data symbol that stems from $\hat{S}(t,k)$. The errors of estimated CFR are alleviated by demapping $\hat{S}(t,k)$ to the corresponding constellation point $\hat{X}(t,k)$. Thus, data symbols can provide useful channel information to construct data pilot.

However, $\hat{H}(t,k)$ still cannot eliminate completely noise and accumulate error in iterative process, caused by error propagation, especially at low SNR region. The error propagation happens because the data symbols may be incorrectly demapped and thus the error is gradually accumulated during the iteration. To reduce such negative impact on decision-directed channel estimation, an average method based on time-domain truncation loop approach is applied. The scaled version of FFT matrix V is firstly calculated following by

$$V = \sqrt{M} F_M(:, 1 : L+1), \tag{6}$$

where M represents the modulation order of the signal, F_M is the FFT matrix, and L is the number of reserved time domain taps. $\hat{H}(t,k)$ is converted to $h(t,k)$ in the time domain using inverse fast Fourier transformation (IFFT). To curb noise, time truncation is operated. V is the scaled matrix that works for converting $\hat{h}(t,k)$ to frequency domain. Then, $\hat{H}(t,k)$ is turned into time domain to remove the time domain taps containing most of noise as follows

$$\hat{H}_f(t,k) = V\hat{h}(t,1:L), \tag{7}$$

where $\hat{h}(t,1:L)$ represents the reserved CR in time domain at the tth time slot and $\hat{H}_f(k)$ denotes scaled version of CR in frequency domain. The demodulation errors are equivalent to adding noise into $\hat{h}$. Converting from frequency domain to time domain, noise is uniformly distributed across the different taps. $\hat{h}$ is truncated in the time domain to alleviate the effect of noise caused by demodulation errors. Even though the time truncation is employed, there are some cumulative errors in $\hat{H}_f(k)$ caused by minor noise. Thus, the average of $\hat{H}_f(k)$ in time and frequency domains to smooth CR according to (8), (9)

$$\hat{H}_s(t,k) = \sum_{\lambda=-\beta}^{\lambda=\beta} \frac{1}{2\beta+1}\hat{H}_f(t,k+\lambda), \tag{8}$$

where $2\beta+1$ represents the number of averaged subcarriers. The high correlation between adjacent subcarriers $\hat{H}_f(t,k+\lambda)$ can be introduced to further improve the accuracy of the estimates. Then, averaging in time domain is calculated as follows,

$$\hat{H}_f(t,k) = (1-\alpha)\hat{H}(t-1,k)_f + \alpha\hat{H}_s(t,k). \tag{9}$$

where $\hat{H}_f(t,k)$ denotes the output of ADD-TT scheme at the tth time slot on the kth subcarrier, and α is coefficient parameter to update CR. Based on the high correlation across successive OFDM symbols, the weighted summation of previous and current estimated CFR can improve the performance. α, β are parameters related to knowledge of the vehicular environments. However, it is impossible to obtain such information in practice. It is observed in [6] that the best performance of averaging in time and frequency domain is achieved with $\alpha = 0.5$ and $\beta = 2$. Thus, α is fixed to 0.5 and β is set to 2 in this paper.

4.2 The architecture of SR-ConvLSTM

ChannelNet based on CNN is inept in uncovering the inherent characteristics of temporal spectral correlation, thus a NN architecture SR-ConvLSTM based on ConvLSTM is proposed. It models temporal spectral correlation of adjacent symbols to estimate CR and is suitable for non-stationary scenarios.

Channel estimation of vehicular communications using IEEE 802.11p is viewed as super resolution problem. Considering the time-variant channel, LSTM that enables to extract time correlation of series is introduced to tackle super resolution problem. In [30], the authors prove LSTM successfully handles channel state information (CSI) feedback for time-varying communications. Adding a convolution operation to the LSTM composes of ConvLSTM. ConvLSTM is more effective for feature extraction when the time series data are images. The ConvLSTM [31] originates from LSTM. The difference is that after adding the convolution operation which not only obtain the timing relationship, but also to extract features such as convolution layers. In this way, we obtain the temporal spectral characteristics via SR-ConvLSTM based on ConvLSTM.

The details of proposed SR-ConvLSTM are presented. SR-ConvLSTM is composed of five layers including ConvLSTM and batch normalization (BN). Since this paper views channel estimation as an image super resolution problem, inspired by the architecture of [28], the structure of ConvLSTM following by BN is chosen and such structure is repeated to track high variations of CFR. ConvLSTM works for capturing temporal spectral correlation between adjacent data symbols and BN enables SR-ConvLSTM to converge. The specific structure is seen in Table 1. The first layer applies 64 filters of size 9×9 of

Table 1 Architecture of the SR-ConvLSTM

Layer	Type	Number of filters	Size of filters	Activation function
1	ConvLSTM(Input)	64	9×9	ReLU
2	BN	–	–	–
3	ConvLSTM	32	1×1	ReLU
4	BN	–	–	–
5	ConvLSTM(Output)	1	5×5×5	ReLU

ConvLSTM following by rectified linear units (ReLU) activation following by

$$R(x) = \max(0, x), \tag{10}$$

where x is input of the ConvLSTM. When the activation value of the neuron enters the negative half region, the gradient is 0. That means this neuron is trained to keep sparsity. The second layer is BN. BN is able to solve the problem when the neural network is training with slow convergence speed or exploding gradients. In fact, we find out if BN is removed from SR-ConvLSTM, the network cannot be converged. The reason may lie in the complex distribution of channel that needs BN operation. In addition, BN is added to speed up the training speed and improve the accuracy of the model. The third layer uses 32 filters of size 1×1 of ConvLSTM following by ReLU activation. The fifth layer is BN. The last layer is 1 filter of size 5×5×5 to reconstruct the output. Notably, to strike balance between performance and complexity, TS-ChannelNet removes DnCNN compared to ChannelNet.

The relationship between input and output of proposed SR-ConvLSTM is represented as

$$\hat{H} = f\left(\theta; \hat{H}_{\text{seq}}\right). \tag{11}$$

where θ denotes the parameters of SR-ConvLSTM, $\hat{H}$ is the final estimated CR, and f means nonlinear function that is determined by θ.

The architecture of ChannelNet must be revised if the frame lengths are changed. What is worse, the whole ChannelNet should be trained from scratch, which is non-trivial in practice. In SR-ConvLSTM, the CR is divided into blocks that contain n data symbols. Hence, SR-ConvLSTM fits for the arbitrary frame length without amending the input shape of NN. In conclusion, SR-ConvLSTM is more robust than SRCNN that is building blocks of ChannelNet.

4.3 Training of TS-ChannelNet

In this paper, estimating CFR at the receiver is viewed as a super resolution problem which includes pilot-based interpolation and DL-based restoration [14]. Thus, the proposed TS-ChannelNet is composed of ADD-TT and SR-ConvLSTM. In the first phase, pilot values h_p are extended into the coarse CFR whose dimension is identical to estimated $\hat{H}$. In this second phase, SR-ConvLSTM parameterized by θ is utilized to make coarse CFR become HR version via DL. The relationship between the input and output of TS-ChannelNet can be represented by this equation:

$$\hat{H} = f(\theta; h_p) = f_\theta(f_{ADD-TT}(h_p); \theta) \tag{12}$$

where f_θ and f_{ADD-TT} are the network and interpolated functions, respectively.

ADD-TT in the first phase comprises decision-direction, time truncation, and weighted average. Firstly, decision-direction assumes that the tth CFR is highly correlated with the previous and thus $\hat{H}(t-1,k)$ is used as pseudo $\hat{H}(t,k)$ to calculate data pilot. The errors caused by such iterative operation are alleviated via demapping data pilot to constellation point. Secondly, accumulate errors by wrong demodulation are equal to adding noise. Noise is uniformly distributed across the different taps from frequency domain to time domain [8] and operate truncation to curb it. Thirdly, to make use of pilot, averaging $\hat{H}(t,k)$ in the frequency and time domain is taken into account. In general, ADD-TT utilizes average decision-directed time truncation to make pilot become coarse $\hat{H}$.

SR-ConvLSTM in the second phase is introduced to restore HR version of $\hat{H}$. Initially, training SR-ConvLSTM needs to extract real and imaginary part of $\hat{H}$ and stack them. Then, the stacked $\hat{H}$ is divided into several blocks to make SR-ConvLSTM reveal temporal spectral correlation. SR-ConvLSTM has impressive power to achieve intrinsic correlation of signal in a end-to-end manner. The stacked $\hat{H}$ is divided into several blocks. Finally, the output of SR-ConvLSTM is concatenated to obtain final estimated CFR. In addition, the optimization algorithm Adam [11] is chosen to make SR-ConvLSTM converge. To measure the accuracy of estimates, the normalized mean square error (NMSE) between $\hat{H}$ and H is utilized.

$$\mathrm{NMSE} = \frac{1}{N}\sum_{i=1}^{N}\frac{E\left[\left|H-\hat{H}\right|^2\right]}{E\left[|H|^2\right]} \tag{13}$$

where N is the frame length. Besides, bit error rate (BER) is also chosen to demonstrate the performance of TS-ChannelNet. The algorithm of TS-ChannelNet is summarized in Algorithm 1.

5 Simulation and results

In this section, we first introduce the settings of the simulation, which includes the parameters of IEEE 802.11p and DL-based model. Then, the simulation results demonstrate the strength of our proposed TS-ChannelNet.

5.1 Simulation setup and parameters

The IEEE 802.11p end-to-end PHY is implemented for simulation. The NMSE and BER are taken as the performance measurement of the scheme. The range of SNRs for simulation is from 0 to 30 dB with 4 quadrature amplitude modulation (QAM). The velocities range from 32 to 104 km/h. Frame length with 60 blocks is chosen.

Table 2 The parameters of simulation

Parameter	Value	Parameter	Value
Epoch	60	Modulation mode	QAM
Dropout	0.2	Number of symbols	60
Batch size	128	Number of subcarriers	52
Train dataset size	32000	Bandwidth	10 MHz
Validate dataset size	8000	Carrier frequency	5.2 GHz
Test dataset size	4000	FFT size	64

Algorithm 1 Proposed TS-ChannelNet Algorithm.

Input: Training epochs E, Frame length T
Initialize train dataset (the received data symbols Y_{train}, the label of CFR H_{label})
Initialize test dataset (Y_{test})
Initialize SR-ConvLSTM parameters θ with random values,
Calculate $H_{train}(0,k)$, $H_{test}(0,k)$ as described in (3)
for $t = 1 : T$ **do**
  Obtain ($\hat{H}_{train}(t,k)$, $\hat{H}_{test}(t,k)$) via decision-direction by (4), (5)
  Update ($\hat{H}_{train}(t,k)$, $\hat{H}_{test}(t,k)$) via time truncation as in (6), (7)
  Update ($\hat{H}_{train}(t,k)$, $\hat{H}_{test}(t,k)$) via averaging in frequency and time domain in (8), (9)
end for
($\hat{H}_{train}$, $\hat{H}_{test}$) ← Stack real and imaginary part of ($\hat{H}_{train}$, $\hat{H}_{test}$)
($\hat{H}_{train}$, $\hat{H}_{test}$) ← Split ($\hat{H}_{train}$, $\hat{H}_{test}$) into sequences
for *epoch* $= 1 : E$ **do**
  θ ← Train SR-ConvLSRM with ($\hat{H}_{train}, H_{label}$) via optimization algorithm Adam [11]
  if *NMSE* calculated with ($\hat{H}_{train}, H_{label}$) in (13) converges **then**
    break
  end if
end for
$\hat{H}_{out}$ ← Feed $\hat{H}_{test}$ into the trained SR-ConvLSTM
Concatenate $\hat{H}_{out}$
Output: $\hat{H}_{test}$

Tensorflow using graphics processing unit (GPU) is employed for our approach. The learning rate is 0.001 and the dropout is 0.2. The batch size is 128 and epochs are 60. The training size, validation size, and test size are 32000, 8000, and 4000 respectively. The two models are trained at the SNR values of 22 dB with above hyperparameters with respect to three different environments. The specific parameters of simulation are described in Table 2..

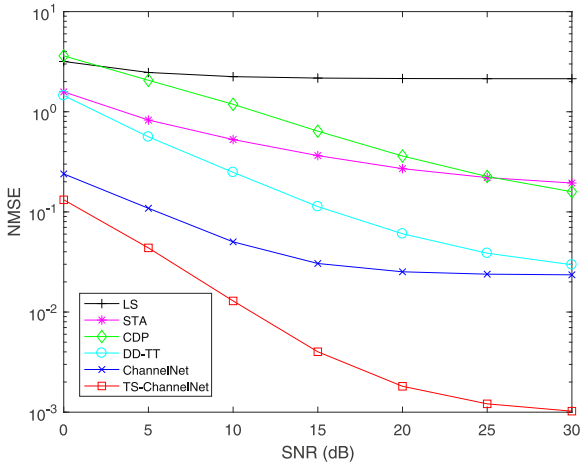

Fig. 2 NMSE with 4QAM modulation, maximum Doppler = 1200 Hz, and 60 OFDM symbols (VTV Expressway Oncoming)

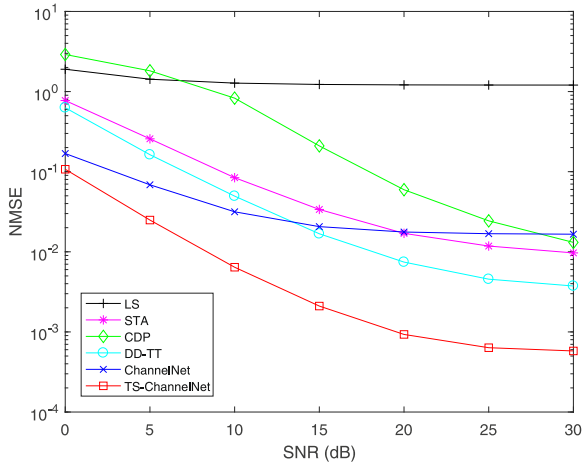

Fig. 3 NMSE with 4QAM modulation, maximum Doppler = 300 Hz, and 60 OFDM symbols (RTV Expressway)

5.2 Results and discussion

Figures 2, 3, and 4 compare the performance of TS-ChannelNet and other schemes with maximum Doppler shift ranged from 300 to 1200 Hz. It is seen that the DD-TT outperforms ChannelNet at the high SNR region. Our presented scheme consistently has a better performance advantage than other approaches. This is because our proposed scheme estimates CR by integrating pilot knowledge, data knowledge, and the correlation of adjacent symbol. TS-ChannelNet is competent under high-velocity communication, which is challenging for real vehicular communication.

In Fig. 5, ideal BER is illustrated. Ideal BER is obtained with known of CR without noise. It is seen that the performance of our method is approaching the ideal situation, which means TS-ChannelNet can nearly accurately recover CR. It is obvious that TS-ChannelNet has a better performance as deep fading for vehicular communications become severer. Through the performance under representative vehicular models, we

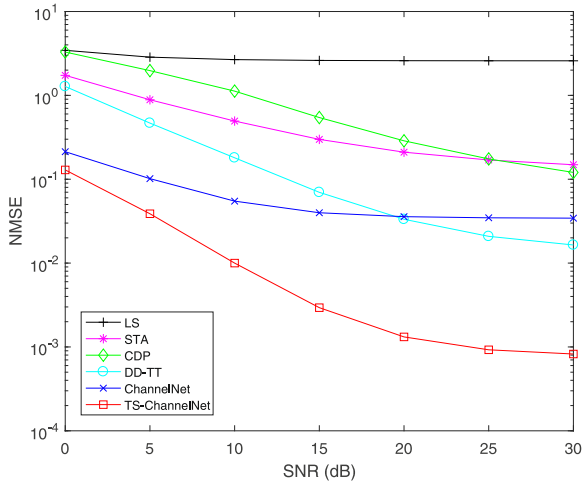

Fig. 4 NMSE with 4QAM modulation, maximum Doppler = 500 Hz, and 60 OFDM symbols (VTV Urban Canyon)

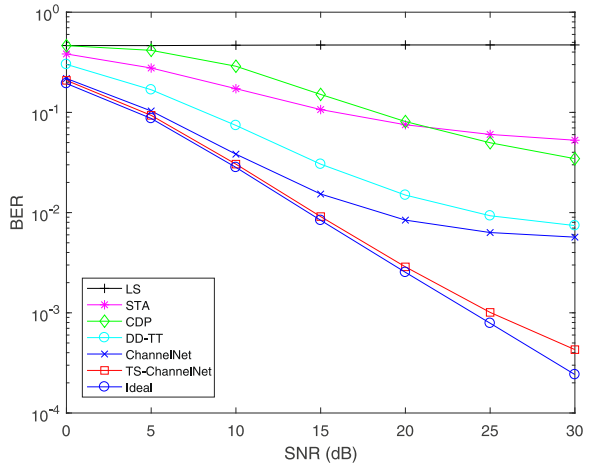

Fig. 5 BER with 4QAM modulation, maximum Doppler = 500 Hz, and 60 OFDM symbols (VTV Urban Canyon)

demonstrate our TS-ChannelNet is robust and has a evident performance in terms of BER or NMSE.

To further investigate our proposed method, an ablation analysis for fast time-varying environment is introduced. Due to the fact that Gaussian interpolation (GI) is utilized in ChannelNet, we take GI, DD-TT, and ADD-TT as interpolation methods in the first phase of TS-ChannelNet respectively while SR-ConvLSTM remains. We refer these approaches as GI-(SR-ConvLSTM), DD-TT-(SR-ConvLSTM), and ADD-TT-(SR-ConvLSTM). Besides, ChannelNet is taken as benchmark algorithm.

Figure 6 plots the NMSE of TS-ChannelNet with different interpolation methods under high mobility scenario while ChannelNet is considered as a reference. It is clearly seen that TS-ChannelNet with GI outperforms ChannelNet with GI. It suggests that our proposed SR-ConvLSTM has better capacity to extract temporal spectral correlation of data

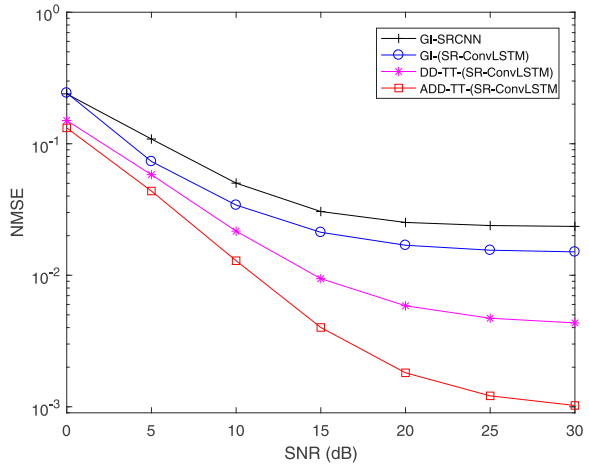

Fig. 6 NMSE for ablation analysis with 4QAM modulation, maximum Doppler = 1200 Hz and 60 OFDM symbols (VTV Expressway Oncoming)

Table 3 The improved percentage of TS-ChannelNet with respect to the compared methods

	LS (%)	STA (%)	CDP (%)	DD-TT (%)	ChannelNet (%)
RTVE	99.9	94.0	96.5	95.5	84.5
VTVEO	99.9	99.4	99.3	96.5	95.6
VTVUC	99.9	99.4	99.3	94.9	97.6

symbol than NN structure of ChannelNet. It is also observed that the different interpolation methods have effect on the performance of following SR-ConvLSTM. It proves that our proposed ADD-TT outperforms DD-TT, especially at high SNR values.

With respect to the compared methods, the improved result of our method in percentage is also presented in Table 3. Under three representative channel models, this percentage is obtained in terms of NMSE with SNR=30 dB. The three representative channel models are RTV Expressway (RTVE), VTV Expressway Oncoming (VTVEO), and VTV Urban Canyon (VTVUC) as mentioned before. It is obvious that our proposed method delivers fairly performance gain. Besides, the gain increases as the maximum Doppler shift grows. It demonstrates our proposed method can track more adequately variations of CFR with respect to the compared methods.

6 Conclusions

Because CFR in vehicular communications varies highly, it is difficult to track variations of channel. The current DPAS method suffers from error propagation caused by accumulative noise. In this paper, a TS-ChannelNet-based channel estimation method for the fast time-varying scenario using IEEE 802.11p is proposed. In this scheme, CR is taken as images and apply TS-ChannelNet to estimate the CR leveraging pilot. TS-ChannelNet is made up of two phases. Pilot values are first extended to coarse tentative CR via interpolation based on ADD-TT. Note that the estimated CR is divided into sequences that contain n adjacent symbols. Afterwards, the SR-ConvLSTM takes divided CR as input and generates recovered CR. Simulation results demonstrate that our proposed method enables prominent performance over previous schemes under high-sped scenarios. Further experiments verify the two building blocks of TS-ChannelNet have all evident performances in channel estimation accuracy. The proposed TS-ChannelNet sheds light on how DL can be successfully applied for CE under high velocity environments.

In this paper, the NN is trained separately with respect to the correspoding representive environments. Hence, the generalization ability of network needs to be further improved. How to use transfer learning to overcome this problem will be our future work.

Abbreviations

IEEE: Institute of Electrical and Electronics Engineers ISR : Image super resolution CR : Channel response CFR : Channel frequency response SNR : Signal-noise-ratio VTV : Vehicle-to-vehicle VTI : Vehicle-to-infrastructure PHY : Physical layer MAC : Medium-access Layer CE : Channel estimation STA : Spectral temporal averaging CDP : Constructed data pilot DL : Deep learning NN : Neural network FCNN: Fully connected neural network CNN : Convolutional neural network LSTM: Long short-Term ITS : Intelligent transportation system LR : Low resolution HR : High resolution OFDM: Orthogonal frequency division multiplexing AWGN: Added white Gaussian noise NMSE: Normalized mean square error CSI : Channel state information ReLU: Rectified linear units BN : Batch normalization BER : Bit error rate QAM : Quadrature amplitude modulation GPU :Graphics processing unit SRCNN: Super resolution convolutional neural network DnCNN: Denoiser convolutional neural network DD-TT: Decision-directed estimation with time truncation SR-ConvLSTM: Super resolution convolutional long short-term ADD-TT: Average decision-directed estimation with time truncation TS-ChannelNet: Temporal spectral channel network

Acknowledgements

Not applicable.

Authors' contributions
The study is conceived and designed by ZX. ZX performed the experiment. The manuscript is reviewed and revised by FY, SZ, and GD. All authors read and approved this manuscript.

Authors' information
Not applicable.

Funding
This work is supported by the National Natural Science Foundation of China (No.61673253, 61901254).

Availability of data and materials
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 27 December 2019 Accepted: 22 April 2020
Published online: 11 May 2020

References

1. I. Wahid, A. U. A. Ikram, M. Ahmad, F. Ullah, An improved supervisory protocol for automatic selection of routing protocols in environment-aware vehicular ad hoc networks. Int. J. Distrib. Sensor Netw. **14**(11) (2018)
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. IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009), 1–51 (2010). IEEE Xplore
3. 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: Further Higher Data Rate Extension in the 2.4 GHz Band. IEEE Std 802.11g-2003 (Amendment to IEEE Std 802.11, 1999 Edn. (Reaff 2003) as amended by IEEE Stds 802.11a-1999, 802.11b-1999, 802.11b-1999/Cor 1-2001, and 802.11d-2001), 1–104 (2003). IEEE Xplore
4. W. Lin, M. Li, K. Lan, C. Hsu, A comparison of 802.11a and 802.11p for V-to-I communication: a measurement study. Int. Conf. Heterog. Netw. Qual. Reliab. Secur. Robustness, 559–570 (2010)
5. S. Benkirane, M. Benaziz, in *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, Performance evaluation of IEEE 802.11p and IEEE 802.16e for vehicular ad hoc networks using simulation tools (Marrakech, 2018), pp. 573–577
6. J. A. Fernandez, K. Borries, L. Cheng, B. V. K. Vijaya Kumar, D. D. Stancil, F. Bai, Performance of the 802.11p physical layer in vehicle-to-vehicle environments. IEEE Trans. Veh. Tech. **61**(1), 3–14 (2012)
7. Z. Zhao, X. Cheng, M. Wen, L. Yang, B. Jiao, Constructed data pilot-assisted channel estimators for mobile environments. IEEE Trans. Intell. Transp. Syst. **16**(2), 947–957 (2015)
8. M. M. Awad, K. G. Seddik, A. Elezabi, Low-complexity semi-blind channel estimation algorithms for vehicular communications using the IEEE 802.11p standard. IEEE Trans. Intell. Transp. Syst. **20**(5), 1739–1748 (2019)
9. S. Baek, I. Lee, C. Song, A new data pilot-aided channel estimation scheme for fast time-varying channels in IEEE 802.11p systems. IEEE Trans. Veh. Tech. **68**(5), 5169–5172 (2019)
10. D. Gündüz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. R. Murthy, M. van der Schaar, Machine learning in the air. IEEE J. Sel. Areas Commun. **37**(10), 2184–2199 (2019)
11. Y. Yang, F. Gao, X. Ma, S. Zhang, Deep learning-based channel estimation for doubly selective fading channels. IEEE Access. **7**, 36579–36589 (2019)
12. H. Ye, G. Y. Li, B. Juang, Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wirel. Commun. Lett. **7**(1), 114–117 (2018)
13. D. Neumann, T. Wiese, W. Utschick, Learning the MMSE channel estimator. IEEE Trans. Signal Process. **66**(11), 2905–2917 (2018)
14. M. Soltani, V. Pourahmadi, A. Mirzaei, H. Sheikhzadeh, Deep learning-based channel estimation. IEEE Commun. Lett. **23**(4), 652–655 (2019)
15. F. Ullah, A. H. Abdullah, O. Kaiwartya, J. Lloret, M. M. Arshad, EETP-MAC: energy efficient traffic prioritization for medium access control in wireless body area networks. Telecommun. Syst. (2017)
16. F. Ullah, Z. Ullah, S. Ahmad, I. U. Islam, S. U. Rehman, J. Iqbal, Traffic priority based delay-aware and energy efficient path allocation routing protocol for wireless body area network. J. Ambient. Intell. Humanized Comput. (2019)
17. Y. Yang, Broadcast encryption based non-interactive key distribution in MANETs. J. Comput. Syst. Sci. **80**(3), 533–545 (2014)
18. H. Kung, C. Chen, M. Lin, T. Wu, Traffic priority based delay-aware and energy efficient path allocation routing protocol for wireless body area network. J. Int. Tech. **20**(7), 2083–2097 (2019)
19. Y. Chen, S. Weng, W. Guo, N. Xiong, A game theory algorithm for intra-cluster data aggregation in a vehicular ad hoc network. Sensors (Basel). **16**(2), 245 (2016)
20. X. Ke, J. Zou, Y. Niu, End-to-end automatic image annotation based on deep CNN and multi-label data augmentation. IEEE Trans. Multimedia. **21**(8), 2093–2106 (2019)
21. H. Cheng, Z. Xie, L. Wu, Z. Yu, R. Li, Data prediction model in wireless sensor networks based on bidirectional LSTM. EURASIP J. Wirel. Commun. Netw. **2019**, 203 (2019)
22. S. Zhong, C. Jia, K. Chen, P. Dai, A novel steganalysis method with deep learning for different texture complexity images. Multimedia Tools Appl. **78**, 8017–8039 (2019)

23. L. Liang, H. Ye, G. Yu, G. Y. Li, Deep-learning-based wireless resource allocation with application to vehicular networks. Proc. IEEE. **108**(2), 341–356 (2020)
24. T. Fu, C. Wang, N. Cheng, IEEE Internet Things J., Deep learning based joint optimization of renewable energy storage and routing in vehicular energy network, 1–1 (2020)
25. S. Khan Tayyaba, H. A. Khattak, A. Almogren, M. A. Shah, I. Ud Din, I. Alkhalifa, M. Guizani, 5G vehicular network resource management for improving radio access through machine learning. IEEE Access. **8**, 6792–6800 (2020)
26. I. Chu, P. Chen, W. Chen, in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, An IEEE 802.11p based distributed channel assignment scheme considering emergency message dissemination, (Yokohama, 2012), pp. 1–5
27. G. Acosta-Marum, M. A. Ingram, Six time- and frequency- selective empirical channel models for vehicular wireless lans. IEEE Veh. Tech. Mag. **2**(4), 4–11 (2007)
28. C. Dong, C. C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks. IEEE Trans. Patt. Anal. Mach. Intell. **38**(2), 295–307 (2016)
29. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans. Image Process. **26**(7), 3142–3155 (2017)
30. T. Wang, C. Wen, S. Jin, G. Y. Li, Deep learning-based CSI feedback approach for time-varying massive MIMO channels. IEEE Wirel. Commun. Lett. **8**(2), 416–419 (2019)
31. X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, W.-c. Woo, Convolutional LSTM network: a machine learning approach for precipitation nowcasting. ArXiv **abs/1506.04214** (2015)

Publisher's Note