

Dual-grained directional representation for infectious disease case prediction

Peisong Zhang^{a,b}, Zhijin Wang^{a,*}, Yaohui Huang^{a,c}, Mingzhai Wang^d

^a College of Computer Engineering, Jimei University, Yinjiang Road 185, Xiamen 361021, China

^b School of Science, Jimei University, Yinjiang Road 185, Xiamen 361021, China

^c College of Electronic Information, Guangxi Minzu University, Daxue East Road 188, Nanning 530006, China

^d Xiamen Center for Disease Control and Prevention, Shengguang Road 681, Xiamen 361021, China

ARTICLE INFO

Article history:

Received 12 October 2021

Received in revised form 23 August 2022

Accepted 27 August 2022

Available online 2 September 2022

Keywords:

Infectious disease

Prediction

Dual-grained time series

Directional representation

ABSTRACT

The uncertain infection transmission causes challenges in accurate disease prediction. Numerous methods have been proposed to capture the temporal pictures from past observations within equal time intervals, which are called single-grained time series. However, these methods are not suitable for capturing uncertain temporal dynamics from infectious disease time series, since the infectious diseases may propagate in the incubation period. To address this issue, this paper proposes a Dual-Grained Directional Representation (DGDR) to generate predictions, via consolidating the representations of an equal-grained time series and several fine-grained time series. Firstly, the proposed DGDR learns a transformed segmentation into three kinds of representations. And then those representations from both equal-grained data and fine-grained data are temporally consolidated to connect with outputs. Extensive experiments on two real infectious disease datasets are done to validate the proposed DGDR. Compared with the other twelve methods, MAE value is decreased by 31.5%, RMSE value is decreased by 29.9%, and R^2 value is improved by 87.6%.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Infectious diseases have caused hundreds of billions of dollars in losses, and hundreds of millions of people suffer from them annually [1]. The infectious disease surveillance system in China had been built up since 2003 [2]. In the past two decades, the surveillance system is getting powerful, which has prevented and controlled several disease outbreaks [3]. The infectious disease prediction tools are significant in supporting decision-making [4]. They effectively help the government and medical-related agencies allocate healthcare resources, which would help prevent and control disease transmission as well. It is also the core component of early warning techniques.

Numerous studies have been proposed for infectious disease prediction [5–7]. From the data perspective, these methods are mainly based on single-grained time series data. A single-grained time series means that the time intervals of past observations and the time intervals of prediction targets share a common length. For infectious disease datasets, the predictions based on single-grained time series commonly ignore effects on the incubation

periods of infectious disease. To overcome this problem, *fine-grained time series* are considered and modeled [8]. When the time interval length of the time series is smaller than the time interval length of the prediction target, it is called fine-grained time series. For the purpose of distinguish, when they have the same time interval length, we call them *equal-grained time series*.

The infectious disease prediction is often regarded as a problem of time series prediction [9,10]. Recently, representation learning achieves great success in time series prediction [11]. These methods mainly focus on presentation learning with respect to time dimensional. CNN-variant methods [12,13], RNN-variant methods [14], and attention mechanisms [15,16] have been widely applied to this dimension. However, the representation learning on infectious disease prediction has not been well studied. When considering both equal-grained time series and fine-grained time series, learning representations only on the time dimension are not enough [11].

There are two major challenges in accurate disease case prediction: (1) how to organize the time series data of two granularities? (2) how to efficiently and effectively learn the temporal dynamics from the dual-grained time series data?

To address the above issues for accurate disease prediction, the Dual-grained Directional Representation (DGDR) model is proposed. The proposed DGDR consists of five stages: (1) time series alignment; (2) time series normalization; (3) directional

* Corresponding author.

E-mail addresses: pencil007123@gmail.com (P. Zhang), zhijinecnu@gmail.com (Z. Wang), yhuang5212@gmail.com (Y. Huang), xmcdcmwz@163.com (M. Wang).

representation; (4) temporal fusion; and (5) de-normalization and prediction generation. The core part of the DGDR is the directional representation (DR) component. It learns representations of the transformed time series tensor from three directions, which is different from the previous methods.

The major contributions are summarized below:

- (1) The alignment of the time series of two granularities is firstly applied to general disease prediction. This brings the possibility on observing the disease incubation period in terms of outpatient visit counts.
- (2) The essential and feasibility of learning representations from different directions of time series have been validated.
- (3) The proposed DGDR can learn temporal dynamics from dual-grained time series with better performance. Technically, the RMSE value is decreased 29.9% at most.
- (4) Extensive experiments of twelve methods on two real disease datasets in Xiamen have been done.

The remains of this paper are organized as follows. Section 2 addresses this work. Section 3 gives problem definition. Section 4 illustrates the proposed DGDR. Section 5 gives datasets, evaluation metrics, and benchmark methods. Section 6 conducts experiments and analyses. Finally, a conclusion is drawn in Section 7.

2. Related work

This section addresses this research by review infectious disease prediction methods and temporal representation learning methods.

2.1. Infectious disease prediction methods

Infectious disease prediction has received increasing attention from epidemiologists and researchers in recent years. These methods can be divided into three categories: mechanical methods, statistical methods, and deep learning methods.

Mechanical methods capture disease transmission dynamics by utilizing disease progression mechanisms [17]. Susceptible–Infectious–Recovered (SIR) [18] and Susceptible–Exposed–Infected–Recovered (SEIR) [19] are famous mechanical predictive methods. These methods incorporate the causation of disease population in the inference process and are extensively applied in the field of public health. However, the performance is much dependent on the accurate initial parameters. So it is difficult to adapt to the dynamics of disease progression in the real world.

Statistical methods trace the upcoming trends of infectious disease based on statistical theories, such as probabilistic distribution and Bayesian theorem [20]. The ARIMA and seasonal ARIMA (SARIMA) [21] are two popular approaches to learning stable time series. Due to the unstable nature of disease transmission, these methods are not applicable in most cases. Some other methods, such as Kalman filtering and Bayesian inference, have also been used for epidemic forecasting [9]. But these methods rely on some statistical assumptions of the data. It inevitably limits these methods in achieving better prediction accuracy.

Deep learning methods have superiority in capturing the complex representation information. Recent studies have leveraged these methods to make predictions about the spread of infectious diseases [22,23]. However, these methods commonly utilize single-grained time series for prediction. This limits the predictive performance by ignoring the effect of incubation periods.

2.2. Temporal representation learning methods

Recently, numerous temporal representation learning methods have been proposed to do time series prediction. These methods can be divided into two categories: traditional deep learning methods and attention-based methods.

Traditional deep learning methods include some widely used approaches, such as convolutional neural networks (CNN) and recurrent neural networks. Temporal convolutional network (TCN) [12] is proposed for time series prediction based on the CNN component, which has the superiority in capturing local feature information. However, there are many long-term temporal dependencies are ignored by this method, which significantly affects the prediction accuracy. The RNNs methods are proposed to solve the above problem. Dual-grained Representation (DGR) [8] is a well-designed temporal representation learning method based on recurrent layers. This method fuses dual-grained epidemic time series and shows acceptable performance. However, DGR merely considers the temporal dynamics between different time steps, neglecting the relationships between different periods.

Attention-based methods have achieved excellent performance in dealing with time series data. Typically, LSTNet [24] and DARNN [14] are two well-known attention-based methods that show good performance in extracting time dimension representation from strong periodic time series. However, due to the weak periodicity and instability of epidemic time series data, attention-based methods still have room for improvement in infectious disease case prediction.

In this paper, we proposed a novel temporal representation learning method DGDR. From the perspective of the data, DGDR integrates fine-grained time series to enhance the ability to discover mechanisms of disease progression. From the perspective of the method, DGDR learns the temporal representations from different directions, which captures more internal transmission patterns from epidemic time series.

3. Problem definition

This section gives notations and problem definition.

A *time series* is defined as a series of observed values with identical time intervals. An *equal-grained time series* has the same length of time intervals as the prediction target. For example, if the predictive target is the outpatient cases of the upcoming week, then the time series of weekly outpatient cases can be called equal-grained time series. A equal-grained time series with length K is denoted by symbol $\mathbf{E} \in \mathbb{R}^{K \times 1}$. A *fine-grained time series* has a smaller length of time intervals than the prediction target. For example, if the prediction target is the outpatient cases of the upcoming week, then the time series of daily outpatient cases can be called fine-grained time series. A fine-grained time series with length L is denoted by symbol $\mathbf{F} \in \mathbb{R}^{L \times 1}$.

A *look-back segment* observes a part of historical values. Let d_2 be the length of a look-back segment. Symbol $\mathbf{E}_{1:d_2,1}$ denotes a segment of the equal-grained time series. Two segments of a time series may have overlapping values.

Commonly, infectious disease prediction is regarded as a problem of equal-grained time series prediction, and it is formulated as:

$$\hat{E}_{d_2+1,1} = f(\mathbf{E}_{1:d_2,1}), \quad (1)$$

where $\hat{E}_{d_2+1,1} \in \mathbb{R}$ is the predicted value in the upcoming time interval, and $f(\cdot)$ is a mapping.

The dual-grained time series prediction leverages both equal-grained time series and fine-grained time series to predict the upcoming values. Because the length of equal-grained time series is smaller than fine-grained time series, the fine-grained time

series is aligned with equal-grained time series according to their specific time intervals. This can also avoid the problem of timeline chaos in different time series. Given that one time interval of equal-grained time series consists of d_3 time intervals of fine-grained time series, symbol $\mathbf{F} \in \mathbb{R}^{K \times d_3}$ denotes the aligned fine-grained time series. The problem of dual-grained time series prediction is formulated as:

$$\hat{E}_{d_2+1,1} = f(\mathbf{E}_{1:d_2,1}, \mathbf{F}_{1:d_2,\cdot}), \quad s.t. \quad E_{i,1} = \sum_{j=1}^{d_3} F_{i,j}, \quad (2)$$

where $\hat{E}_{d_2+1,1} \in \mathbb{R}$ is the predicted value, $\mathbf{F}_{1:d_2,\cdot} \in \mathbb{R}^{d_2 \times d_3}$ is the observed fine-grained values, and $f(\cdot)$ is a mapping. The $\mathbf{E}_{1:d_2,1}$ has the same time span as $\mathbf{F}_{1:d_2,\cdot}$.

The consideration of a single predictive point would greatly limit the model scope and cause training biases. Technically, lots of small target values usually constrain the model's abilities in predicting outbreak transmission events. Hence, the consecutive predictions of dual-grained time series are considered and formulated as:

$$\hat{\mathbf{E}}_{d_2+1:d_2+d_1,1} = f(\mathbf{E}_{1:d_2,1}, \mathbf{E}_{2:d_2+1,1}, \dots, \mathbf{E}_{d_1:d_2+d_1-1,1}, \mathbf{F}_{1:d_2,\cdot}, \mathbf{F}_{2:d_2+1,\cdot}, \dots, \mathbf{F}_{d_1:d_2+d_1-1,\cdot}), \quad (3)$$

where $\hat{\mathbf{E}}_{d_2+1:d_2+d_1,1} \in \mathbb{R}^{d_1 \times 1}$ denotes consecutive predicted values with size d_1 , $[\mathbf{E}_{1:d_2,1}, \mathbf{E}_{2:d_2+1,1}, \dots, \mathbf{E}_{d_1:d_2+d_1-1,1}] \in \mathbb{R}^{d_1 \times d_2 \times 1}$ denotes segments of equal-grained time series with size d_1 , and $[\mathbf{F}_{1:d_2,d_3}, \mathbf{F}_{2:d_2+1,d_3}, \dots, \mathbf{F}_{d_1:d_2+d_1-1,d_3}] \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ denotes segments of fine-grained time series with size d_1 , and $f(\cdot)$ is a mapping. For easy representation, symbol $\mathbf{T}^E \in \mathbb{R}^{d_1 \times d_2 \times 1}$ denotes segments of equal-grained time series, symbol $\mathbf{T}^F \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ denotes segments of fine-grained time series, and symbol $\hat{\mathbf{T}}^E \in \mathbb{R}^{d_1 \times 1}$ denotes consecutive predicted values.

The representation learning techniques can be leveraged to learn the temporal dynamics from tensor \mathbf{T}^E and tensor \mathbf{T}^F . We believe that the tensors can be represented from three directions. These directions are defined as:

Direction 1 observes the associations between past observations and the consecutive upcoming values. It can be learned from the first dimension of \mathbf{T}^F or \mathbf{T}^E .

Direction 2 observes the associations between different time steps of a look-back segment. It can be learned from the second dimension of \mathbf{T}^F or \mathbf{T}^E .

Direction 3 observes the associations between different time series. It can be learned from the third dimension of \mathbf{T}^F or \mathbf{T}^E .

In summary, the problem of dual-grained time series prediction via directional representation is formulated as:

$$\hat{\mathbf{T}}^E = f(r_1(\mathbf{T}^E), r_2(\mathbf{T}^E), r_1(\mathbf{T}^F), r_2(\mathbf{T}^F), r_3(\mathbf{T}^F)), \quad (4)$$

where $r_1(\cdot), r_2(\cdot), r_3(\cdot)$ are the functions of three directions, respectively.

The main notations are listed in Table 1.

4. The proposed DGDR

This section illustrates the proposed DGDR. The graphical descriptions of the proposed DGDR are plotted in Fig. 1. The DGDR consists of the five stages below.

The first stage is the alignment of fine-grained time series and equal-grained time series, see the upper left part of Fig. 1(a). Given a set of specific outpatient records, they are counted in different lengths of time intervals to get different granularities of time series. The fine-grained time series are aligned according to a specific time interval of a coarse-grained time series. This makes the observation of periodic events and consecutive events become possible and easy.

Table 1
Notations and meanings.

Notation	Meaning
K	The length of the equal-grained time series
d_1	The number of look-back segments
d_2	The time steps of a look-back segment
d_3	The length of an aligned period
\mathbf{E}	Equal-grained time series $\mathbf{Z} \in \mathbb{R}^{K \times 1}$
\mathbf{F}	Aligned fine-grained time series $\mathbf{F} \in \mathbb{R}^{K \times d_3}$
$\mathbf{E}_{1:d_2,1}$	A look-back segment of \mathbf{E}
$\mathbf{F}_{1:d_2,d_3}$	A look-back segment of \mathbf{F}
$\hat{E}_{d_2+1,1}$	A predicted value
$\hat{\mathbf{E}}_{d_2+1:d_2+d_1,1}$	Consecutive predicted values
\mathcal{X}^E	Equal-grained tensor $\mathcal{X}^E \in \mathbb{R}^{(K-d_2+1) \times d_2 \times 1}$
\mathcal{X}^F	Fine-grained tensor $\mathcal{X}^F \in \mathbb{R}^{(K-d_2+1) \times d_2 \times d_3}$
\mathcal{Y}	Output tensor $\mathcal{Y} \in \mathbb{R}^{(K-d_2+1) \times 1}$
\mathbf{X}^E	Input equal-grained data $\mathbf{X}^E \in \mathbb{R}^{d_1 \times d_2 \times 1}$
\mathbf{X}^F	Input fine-grained data $\mathbf{X}^F \in \mathbb{R}^{d_1 \times d_2 \times d_3}$
\mathbf{Y}	Target data $\mathbf{Y} \in \mathbb{R}^{d_1 \times 1}$
$r_1(\cdot)$	The representation function on direction 1
$r_2(\cdot)$	The representation function on direction 2
$r_3(\cdot)$	The representation function on direction 3
$f(\cdot)$	A mapping
$[\cdot]$ or $[\cdot]$	Concatenation operation

The second stage is time series normalization. The values between the two granularities of time series are quite different. The normalization procedure makes the values of all the time series have a common or close order of magnitude. It works on the whole-length dual-grained time series.

The third stage is time series transformation. Since supervised models cannot directly work well on time series data, it needs to be transformed into supervised data. The "one-step-forward split" [6] technique is adopted to transform time series data to tensor data.

The fourth stage is the representations of transformed tensor data, the detailed representation procedures of one direction are displayed in Fig. 1(b). To comprehensively capture the temporal dynamics of an input tensor, three directional representation (DR) components are applied to it.

The final stage is temporal fusion and prediction generation. The procedures of temporal fusion is in Fig. 1(c). The three represented tensors of fine-grained tensor and two of equal-grained tensor are concatenated and pass to a temporal fusion layer, which makes connection with the outputs. The predicted values is the de-normalized model outputs.

4.1. Time series data processing

Time series alignment. To ensure the same length of dual-grained time series, we reordered the sequence of fine-grained time series, and arranged them in chronological order. Let symbol $\mathbf{F} \in \mathbb{R}^{K \times d_3}$ represent the aligned fine-grained time series.

The Min-Max normalization is used to compress time series into the range [0, 1]. The normalization and de-normalization are mathematically formulated below:

$$\mathbf{d}' = \frac{\mathbf{d} - \min(\mathbf{d})}{\max(\mathbf{d}) - \min(\mathbf{d})}, \quad (5)$$

$$\mathbf{d} = \mathbf{d}' * (\max(\mathbf{d}) - \min(\mathbf{d})) + \min(\mathbf{d}), \quad (6)$$

where $\mathbf{d} \in \mathbb{R}^K$ denotes all of the observed samples, $\mathbf{d}' \in \mathbb{R}^K$ is the normalized samples, $\max(\mathbf{d})$ is the maximum value of \mathbf{d} , and $\min(\mathbf{d})$ is the minimum value of \mathbf{d} . The de-normalization formula is applied to the model outputs.

The one-step-forward split is used to transform time series data into supervised data. For a equal-grained time series and

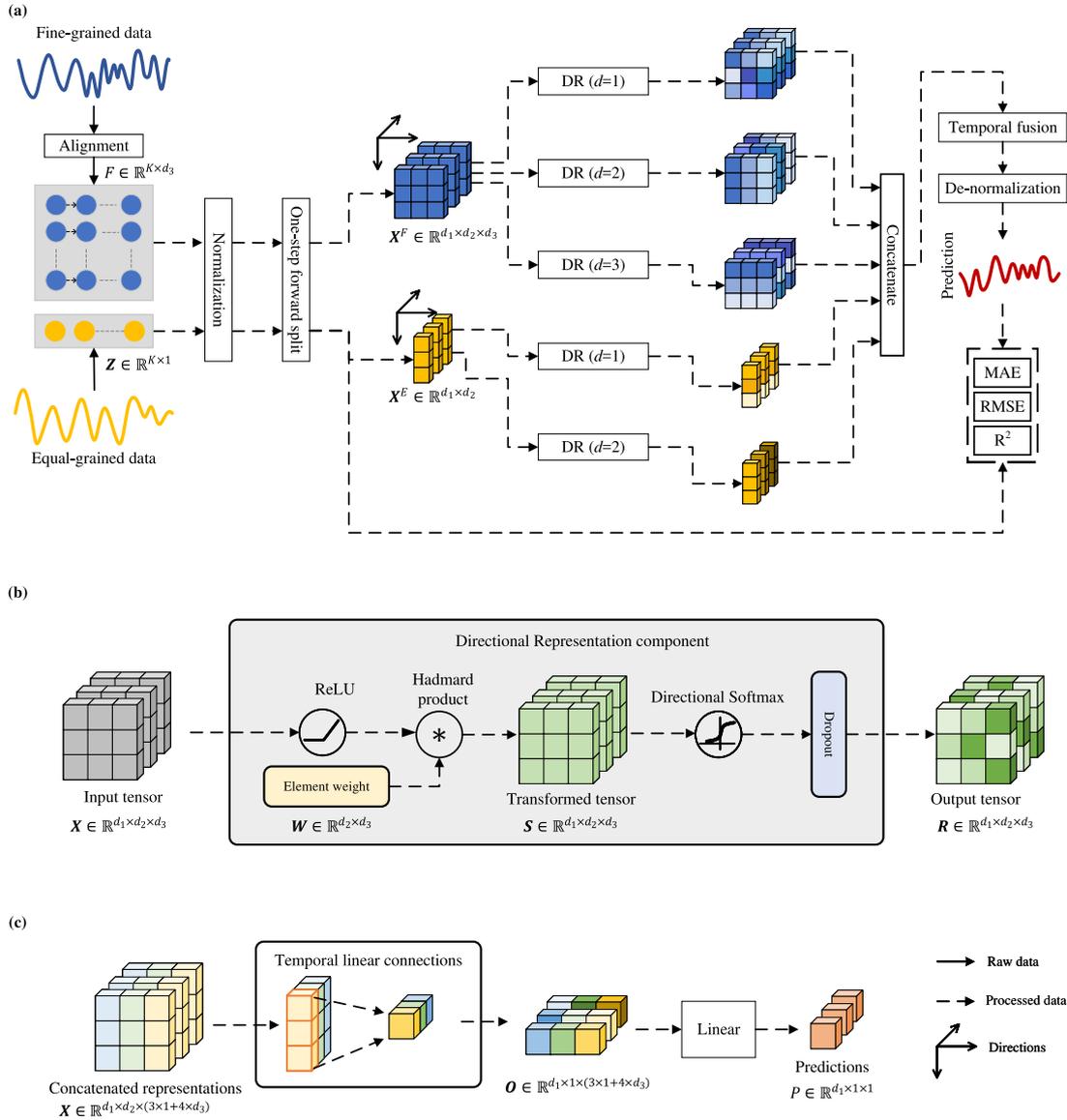


Fig. 1. The schematic illustration of the proposed Dual-grained Directional Representation (DGDR). (a) Workflow. (b) Directional representation (DR) component. (c) Temporal fusion layer.

a fine-grained time series, the transformation results are formulated as below:

$$\begin{bmatrix} E_{1,1} & \cdots & E_{d_2,1} & \mathbf{F}_{1,d_3} & \cdots & \mathbf{F}_{d_2,d_3} \\ E_{2,1} & \cdots & E_{d_2+1,1} & \mathbf{F}_{2,d_3} & \cdots & \mathbf{F}_{d_2+1,d_3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ E_{K-d_2-1,1} & \cdots & E_{K-1,1} & \mathbf{F}_{K-d_2-1,d_3} & \cdots & \mathbf{F}_{K-1,d_3} \end{bmatrix} \rightarrow \begin{bmatrix} E_{d_2+1,1} \\ E_{d_2+2,1} \\ \cdots \\ E_{K,1} \end{bmatrix},$$

where the left part is the inputs, and the right part is the outputs. For a lucid presentation, let symbol $\mathcal{X}^E \in \mathbb{R}^{(K-d_2-1) \times d_2 \times 1}$, symbol $\mathcal{X}^F \in \mathbb{R}^{(K-d_2-1) \times d_2 \times d_3}$ and symbol $\mathcal{Y} \in \mathbb{R}^{(K-d_2-1) \times 1}$ represent equal-grained inputs, fine-grained inputs and outputs respectively. Several consecutive samples in $(\mathcal{X}^E, \mathcal{X}^F, \mathcal{Y})$ are denoted by $(\mathbf{X}^E \in \mathbb{R}^{d_1 \times d_2 \times 1}, \mathbf{X}^F \in \mathbb{R}^{d_1 \times d_2 \times d_3}, \mathbf{Y} \in \mathbb{R}^{d_1 \times 1})$. $\mathbf{E}_{T+1:T+B,1}$ is denoted by \mathbf{Y} , and $[E_{1:d_2,1}, E_{2:d_2+1,1}, \dots, E_{d_1:d_2+d_1-1,1}]$ is denoted by \mathbf{X}^E , $[F_{1:d_2,d_3}, F_{2:d_2+1,d_3}, \dots, F_{d_1:d_2+d_1-1,d_3}]$ is denoted by \mathbf{X}^F .

4.2. Directional Representation (DR) component

The directional representation (DR) component represents the transformed tensor $\mathbf{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ by four procedures. The graphical procedures of DR component are shown in Fig. 1(b).

The first two procedures can be formulated as:

$$\mathbf{S} = \mathbf{W} * \text{ReLU}(\mathbf{X}), \quad (7)$$

where \mathbf{S} is element-wise weighted output matrix, $\mathbf{X} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the input matrix, $\mathbf{W} \in \mathbb{R}^{d_2 \times d_3}$ is the weight matrix of \mathbf{X} , $*$ is the Hadamard product between \mathbf{X} and \mathbf{W} , and $\text{ReLU}(\cdot)$ is the rectified linear unit function. The \mathbf{X} can be tensors from equal-grained data or fine-grained data.

The third procedure highlight the key element in three directions. The three dimensions are formulated as:

$$R_{i,j,k} = \frac{\exp(S_{i,j,k})}{\sum_{v=1}^{d_1} \exp(S_{v,j,k})}, \quad \text{if } d = 1 \quad (8)$$

$$R_{i,j,k} = \frac{\exp(S_{i,j,k})}{\sum_{v=1}^{d_2} \exp(S_{i,v,k})}, \quad \text{if } d = 2 \quad (9)$$

$$R_{i,j,k} = \frac{\exp(S_{i,j,k})}{\sum_{v=1}^{d_3} \exp(S_{i,j,v})}, \quad \text{if } d = 3 \quad (10)$$

where $\mathbf{R} \in [0, 1]$ is the represented tensor, and $\mathbf{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the element-wise transformed matrix. A dropout layer is added in the last procedure to avoid overfitting.

4.3. Fusion on directional representations

The *temporal fusion layer* fuses the transformed tensor and represented tensors to connect with the model output. The graphical procedures of the temporal fusion layer is displayed in Fig. 1(c). The fusion process consists of a tensor concatenation procedure and a temporal connection procedure.

For both equal-grained tensors and fine-grained tensors, their representations are symbolized as:

$$\mathbf{R}_j^E = DR(\mathbf{X}^E, d = j), \quad (11)$$

$$\mathbf{R}_j^F = DR(\mathbf{X}^F, d = j), \quad (12)$$

where $\mathbf{X}^E \in \mathbb{R}^{d_1 \times d_2 \times 1}$ is the equal-grained tensor, $\mathbf{X}^F \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the fine-grained tensor, $\mathbf{R}_j^E \in \mathbb{R}^{d_1 \times d_2 \times 1}$ is the represented tensor of \mathbf{X}^E in j th direction, $\mathbf{R}_j^F \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the represented tensor of \mathbf{X}^F in j th direction, and $DR(\cdot)$ is the directional representation component. All the values of \mathbf{R}_j^E and \mathbf{R}_j^F are in range $[0, 1]$. The j is in $\{1, 2, 3\}$.

The transformed tensors and represented tensors are concatenated as follows:

$$\mathcal{X} = [\mathbf{X}^E; \mathbf{R}_1^E; \mathbf{R}_2^E; \mathbf{X}^F; \mathbf{R}_1^F; \mathbf{R}_2^F; \mathbf{R}_3^F], \quad (13)$$

where $\mathcal{X} \in \mathbb{R}^{d_1 \times d_2 \times (1 \times 3 + d_3 \times 4)}$ is the concatenated representation outputs tensor, and $[\cdot]$ is the concatenation operation.

To reduce the model complexity and highlight the impact of time step dimension, the global autoregression (GAR) [5] is employed. The linear outputs is formulated as:

$$\mathcal{O} = (\mathcal{W}_g \mathcal{X}^T + B_g)^T, \quad (14)$$

where $\mathcal{O} \in \mathbb{R}^{d_1 \times 1 \times (1 \times 3 + d_3 \times 4)}$ is the global auto-regressive outputs on the concatenated outputs \mathcal{X} , $\mathcal{W}_g \in \mathbb{R}^{1 \times d_2}$ is the weighted matrix, and $B_g \in \mathbb{R}$ is bias term.

Finally, a linear layer fuses the above outputs and generate predictions. The model outputs are formulated as:

$$\hat{\mathcal{P}} = \mathcal{W}_o \mathcal{O} + B_o, \quad (15)$$

where the $\hat{\mathcal{P}} \in \mathbb{R}^{d_1 \times 1 \times 1}$ is the consecutive predictions, $\mathcal{W}_o \in \mathbb{R}^{1 \times (3 \times 1 + 4 \times d_3)}$, and $B_o \in \mathbb{R}$ is bias term.

Although some other methods can be used to combine and fuse all the immediate tensors, such as CNN, skip network, and highway network, the simplest method is adopted in this paper.

5. Experimental settings

This section gives data preparation, performance metrics, and benchmark methods.

5.1. Data preparation

Owing to not public dual-grained time series data that can be directly used to validate the methods. The raw outpatient records were collected and shared by the Xiamen City Center for Disease Control and Prevention (XMCCDC). These records were collected in the duration from January 1, 2011 to December 31, 2020, for a total of 3652 days. These records were cleaned, extracted, and transformed into time series data.

Two types of infectious diseases have been collected: hand, foot, and mouth disease (HFMD) and hepatitis beta (HB). For each disease, the daily outpatient visit cases and weekly outpatient visit cases are counted according to the diagnosed disease onset time. The weekly time series of outpatient cases are plotted in Fig. 2(a) and Fig. 2(i), respectively. The correlations between periodic daily time series and weekly time series are plotted in subfigures.

The prediction target is set the outpatient case prediction in the upcoming week. The weekly outpatient time series can be regarded as equal-grained time series. The daily outpatient time series can be regarded as fine-grained time series.

5.1.1. HFMD datasets

68 730 HFMD outpatient visit records were collected. Some observations from visualized analyses of Figs. 2(b)–2(h) are summarized as follows:

- (1) The activities on Friday have the strongest correlations with the activities of the whole week than any other day of the week. The activities on Saturday have second place.
- (2) The activities on Monday have the weakest correlation with the activities of the whole week.
- (3) For HFMD datasets, all days of the week are strongly correlated with the activities of the whole week. All the Pearson Correlation Coefficient (PCC) values are great than 0.9, and their corresponding p -values are less than 0.001.
- (4) HFMD frequently occurs in children under five years old [25]. Since the social activities on weekdays are frequent, adults are possibly exposed to the virus and spread it to children after home.
- (5) Adults infected with HFMD typically appear milder symptoms, so they may intend to seek medical on weekends.

5.1.2. HB datasets

75 891 HB outpatient visit records were collected. The HB cases are reported when the transaminase exceeds the twice standard. Some main observations from visualized analyses of Figs. 2(j)–2(p) are summarized as follows:

- (1) The activities on weekdays were significantly correlated with the activities of the whole week. Most patients with HB have no obvious symptoms, and about 76% outpatient cases were diagnosed through physical examination [26].
- (2) The activities on weekends were weakly correlated with the activities of the whole week. The reason is that HB diagnosis requires blood testings, and the hospitals have not this exam on weekends. Hence, the reported number of HB outpatient cases on weekends is relatively small.

5.2. Metrics

To evaluate the model performance, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), correlation coefficient (R^2) are combined to validate the models. They are formulated as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Z_i - \hat{Z}_i|, \quad (16)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_i - \hat{Z}_i)^2}, \quad (17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Z_i - \hat{Z}_i)^2}{\sum_{i=1}^n (Z_i - \bar{Z})^2}, \quad (18)$$

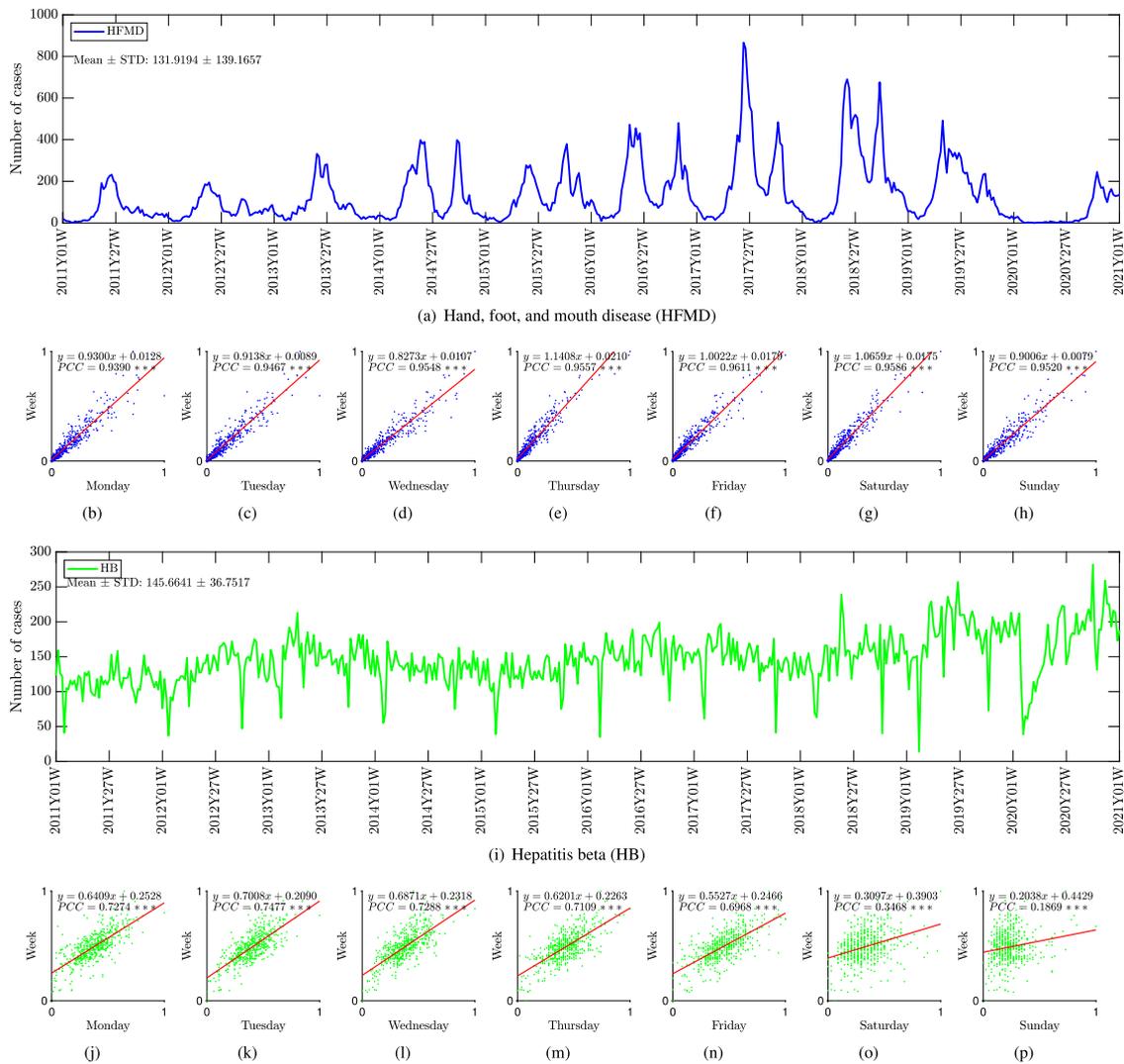


Fig. 2. The weekly distributions of hand, foot, and mouth disease (HFMD) outpatient counts and hepatitis beta (HB) outpatient counts. The visualized correlations between weekly variables and daily variables. For HB diagnosis, only the transaminase exceeding the standard twice will be reported.

where Z_i, \hat{Z}_i denotes real values and predictive values, respectively. Z is defined as the mean value of the test set, n represents the number of time steps in the test samples.

The lower MAE, RMSE values, and the higher R^2 value are the better performance.

5.3. Benchmarks

Some benchmarks are implemented to investigate the effectiveness of the proposed DGDR. These methods are:

- (1) Multiple Linear Regression (MLR) [27] models the linear relationship between inputs and outputs.
- (2) Long and short term memory (LSTM) [28] is a variation of the RNN, which equips various gate units to overcome RNN's limitation of the long-term dependency.
- (3) Gate Recurrent Unit (GRU) [29] is a variation of the LSTM, which replaces the hidden gates and cell gates by an updating gate.
- (4) Encoder–decoder (ED) [30] is composed of two RNN components which are separated in the encoder stage and the decoder stage to capture the non-linear dependencies.
- (5) Convolutional Neural Network (CNN) [31] extracts temporal patterns of sequential data, and uses a fully-connected layer to generate predictions.

- (6) CNNRNN [32] extracts the local temporal pattern with CNN, and uses a RNN to learn the temporal dynamics from CNN outputs.
- (7) DGR [8] extract temporal patterns from both equal-grained and fine-grained data with GRU, and uses a fully-connected layer to generate predictions.

6. Results

This section conducts extensive experiments to validate datasets and models. These experiments intend to address the following questions:

- (1) How the parameter d_1 (i.e., model scope) affects the prediction performance?
- (2) How the past observations affect the prediction performance?
- (3) How the directional representations affect the prediction performance?
- (4) Can the methods benefit from fine-grained time series data?
- (5) Could the DGDR outperforms other benchmark methods?

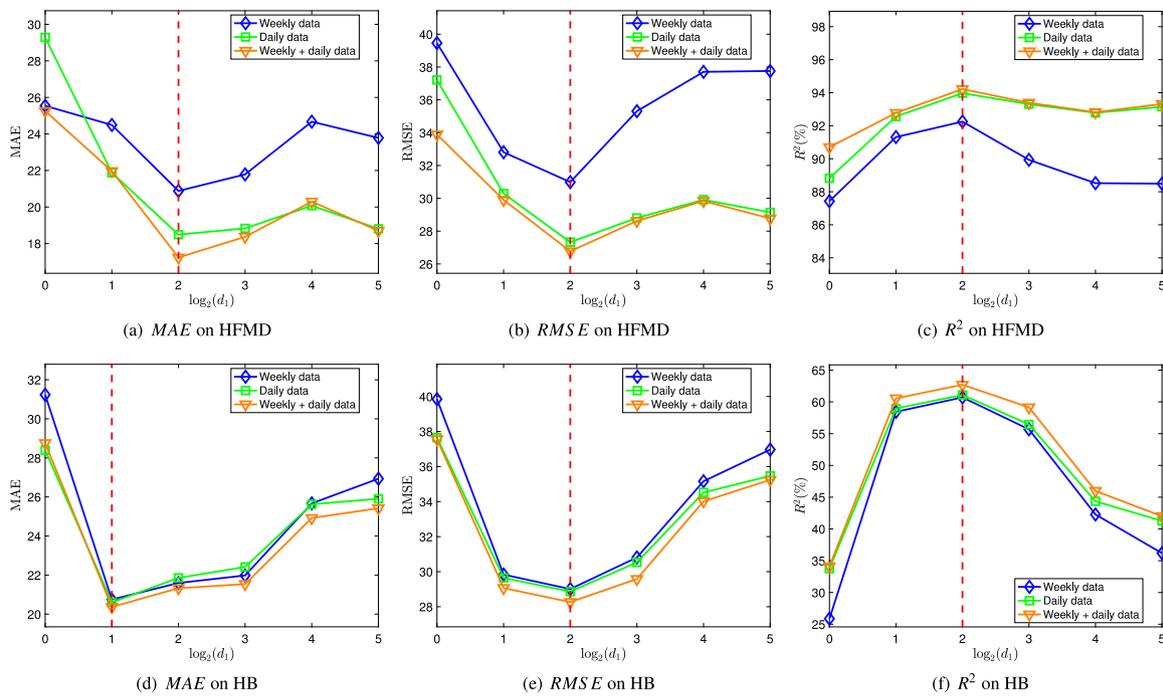


Fig. 3. The DGDR performance with varying d_1 in terms of MAE, RMSE, and R^2 . The optimal values are found at red dash lines. d_2 is fixed at 4.

6.1. Parameter d_1

The constant parameters of the DGDR are d_1, d_2, d_3 . When observing the prediction performance, the target parameter is varied while holding the others. For weekly data and daily data, d_3 is always fixed at 1 and 7, respectively. For HFMD datasets, parameter d_2 is held at 11. For HB datasets, parameter d_2 is held at 3.

The investigation of parameter d_1 on the both disease datasets are plotted in Fig. 3. The major observations are summarized below:

- (1) For HFMD datasets, the optimal values of MAE, RMSE, and R^2 are found at $d_1 = 4$.
- (2) For HB datasets, the optimal value of MAE is found at $d_1 = 2$, and the optimal values of RMSE and R^2 are found at $d_1 = 4$.
- (3) For all the experiments on parameter d_1 , the experiments on weekly data and daily data have the best performance than only one kind of data.
- (4) For the experiments solely based on weekly data, they usually have the worst performance.
- (5) The performance degrades when d_1 is less or greater than 4.
- (6) The poorest performance was found at $d_1 = 1$. $d_1 = 1$ means that the first directional representation has not been used.

For HFMD datasets, see Figs. 3(a)–3(c), the optimal performance are found at $d_1 = 4$. Parameter d_1 is the number of look-back segments. The adult-to-child transmission is a major cause of children infected with HFMD. Adults do not be infected by HFMD due to strong resistance, but they are carriers of the virus [33]. This period is similar to the period that Enterovirus (EV) continues to survive in the body.

For HB datasets, see Figs. 3(d)–3(f), the optimal RMSE value, optimal R^2 value, and the second optimal MAE value are found at $d_1 = 4$.

For HFMD datasets and HB datasets, the DGDR has the worst performance when d_1 is set to 1. The first directional representation is not working when $d_1 = 1$. In reality, the model scope is limited at a single data point when $d_1 = 1$. Parameter d_1 controls the model scope, which affects the learning ability of the first direction. This also reflects the effectiveness the first directional representations.

6.2. Parameter d_2

To investigate the effects on parameter d_2 , parameter d_1 is held at 4 for HFMD datasets and HB datasets. The experimental results are shown in Fig. 4. The main observations are summarized below:

- (1) For HFMD datasets, the optimal MAE value, RMSE value, and R^2 value are found at $d_2 = 11$.
- (2) For HFMD datasets, the experiments based on both weekly data and daily data (see the yellow line with triangle marks) have the best performance when compared with experiments on other datasets.
- (3) For HFMD datasets, the experiments based on weekly data (see the blue line with diamond marks) have the worst performance when compared with experiments on daily data or experiments on both.
- (4) For HFMD datasets, the performance exists fluctuations, among which using weekly data fluctuates the most.
- (5) For HB datasets, the optimal MAE value, RMSE value, and R^2 value are found at $d_2 = 2$.
- (6) For HB datasets, the experiments based on daily data (see the green line with square marks) usually have the worst performance than other experiments on other datasets.
- (7) For HB datasets, the experiments based on weekly data show a steady performance.

The experimental results on HFMD datasets are plotted in Figs. 4(a)–4(c). The optimal performance are found at $d_2 = 11$. HFMD has been proved to be seasonal trends [34]. The Chinese Center for Disease Control and Prevention (CDC) pointed out that

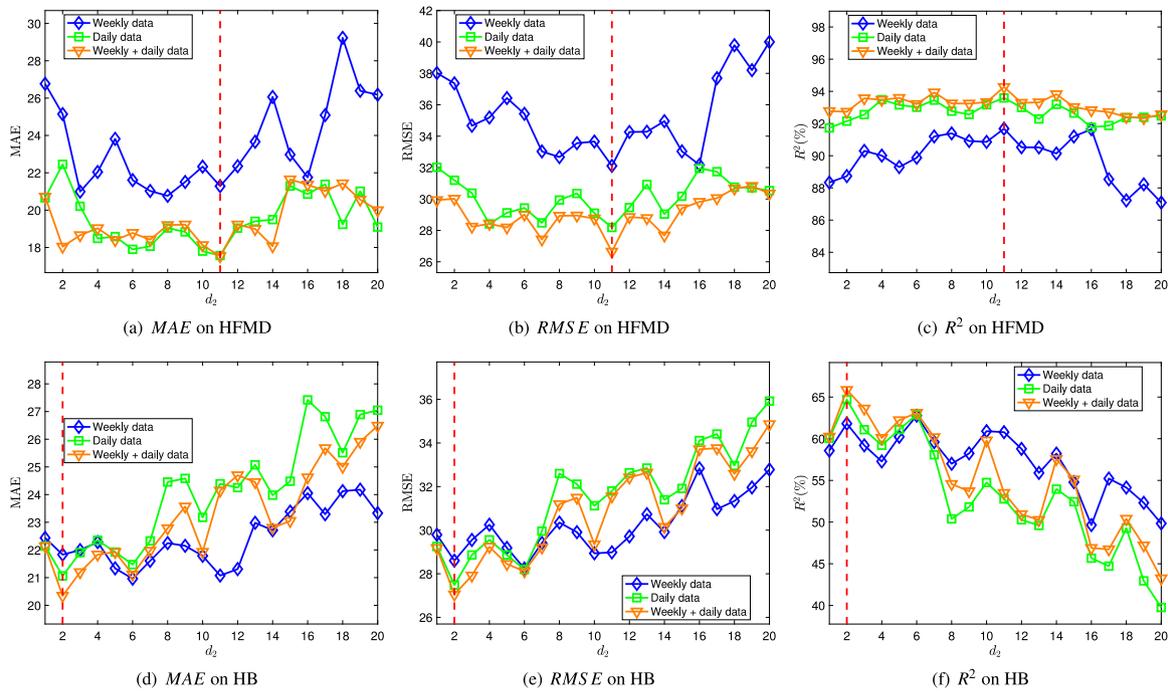


Fig. 4. The DGDR performance with varying d_2 in terms of MAE, RMSE, and R^2 . The optimal value is marked using red dash line. d_1 is fixed at 4.

the high incidence of HFMD usually occurs in southern China from April to June, with a small peak in autumn [35]. The second-best and third-best performance are found at $d_2 = 7$ and $d_2 = 14$, respectively. This means that the coming values have strong correlations with the observations in the past seven weeks and fourteen weeks. In reality, seven weeks and fourteen weeks are two HFMD high-incidence periods in Xiamen.

Figs. 4(d)–4(f) show the experiments on HB datasets. The optimal values of MAE, RMSE and R^2 are found at $d_2 = 2$. This means that the activities of the upcoming week have the strongest correlations with the activities of the past two weeks. HB disease is a type of chronic disease, which means the healthcare strategies are not urgent. The incubation period of HB may be long or very short, such as disease with birth.

When d_2 is set to, the second directional representations do not work, the performance would greatly degrade for HFMD datasets and HB datasets. This reveals the effectiveness of the second directional representations.

6.3. DR consolidation

To investigate the effects on consolidations of directional representations, parameters d_1, d_2 are both fixed while measuring the combinations of directional representations. For HFMD datasets, parameter d_1 is fixed at 4 and parameter d_2 is fixed at 11. For HB datasets, parameter d_1 is fixed at 4 and parameter d_2 is fixed at 2.

The experimental results on combinations of directional representations are plotted in Fig. 5. Some results are summarized below:

- (1) For HFMD datasets and HB datasets, when observing the experiments on the first or second directional representations, the performance of daily data is better than the performance of weekly data.
- (2) For HFMD datasets and HB datasets, the combination of all directional representations have the optimal performance in terms of three metrics.

- (3) When observing the performance on different directional representations, the first direction have better performance than the performance of second direction or third direction.

When observing the experiments in a single direction, the first direction achieves the best prediction accuracy. For the experiments of first directional representation on HFMD datasets, the RMSE value is decreased by 15.3% at most. For the experiments of first directional representation on HB datasets, the RMSE value is decreased by 27.1% at most. This means that the highlight of data points may be more important than the highlight of time steps. Past time steps are relative values to the upcoming time step. The highlight of some time steps usually does not have good results.

The second directional representation and third directional representation affect prediction accuracy by highlighting time steps and a dimension of exogenous time series. For the experiments of representation combinations on HFMD datasets, the RMSE value is decreased by 5.9% at most compared with the first representation. For the experiments of representation combinations on HB datasets, the RMSE value is decreased by 1.9% at most compared with the first representation. Although the second direction and the third direction show poor performance when working alone, they hold some temporal dynamics that the first direction does not have.

The experiments based on daily data outperforms the experiments based on weekly data. For the experiments on the same directional representations on HFMD datasets, the methods based on daily data decrease the RMSE value by 18.3% at most. For the experiments on the same directional representations on HB datasets, the methods based on daily data decrease the RMSE value by 5.8% at most. This shows that the DGDR learns the temporal dynamics from fine-grained data more efficiently than equal-grained data.

The consolidation of dual-grained data on three kinds of directional representations achieves the optimal performance. This reveals that the temporal dynamics captured from equal-grained and fine-grained data both have their unique part. It also shows the effectiveness of the proposed DGDR.

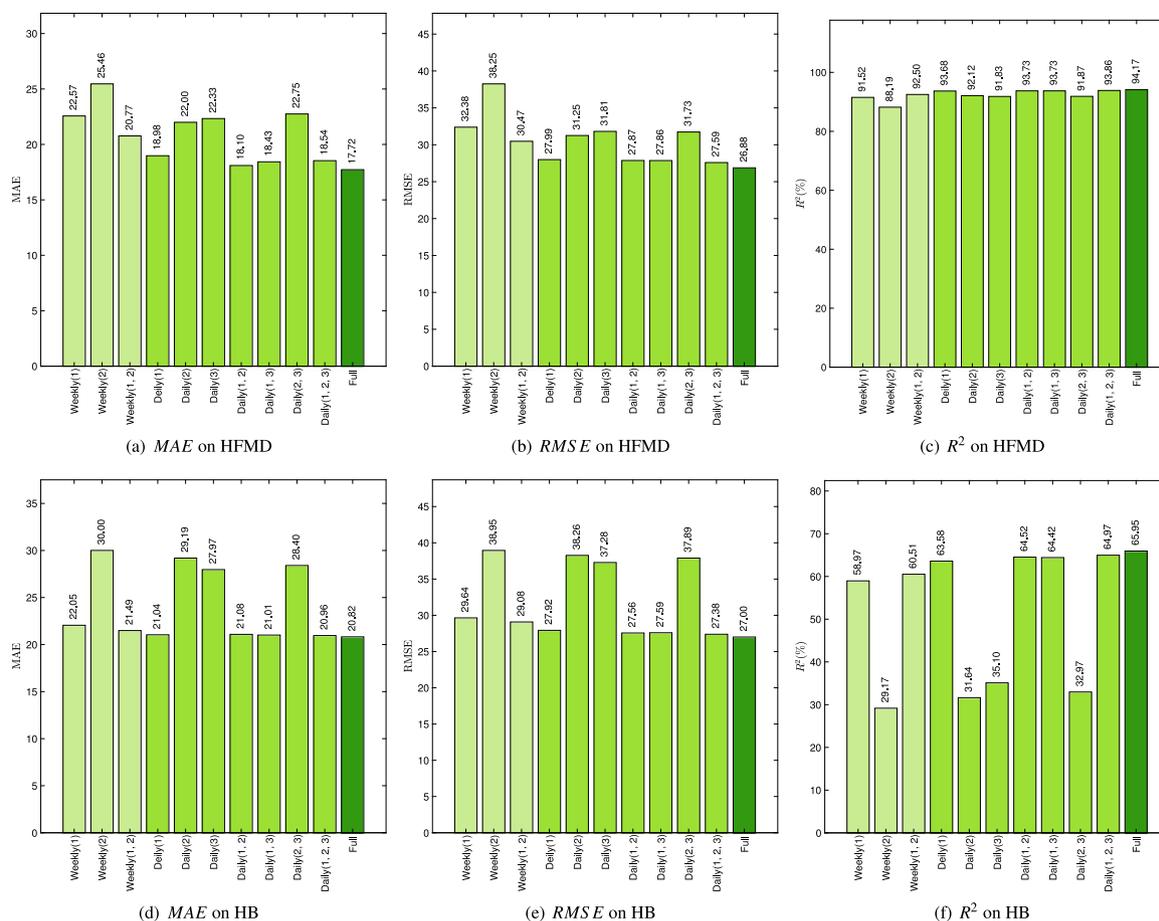


Fig. 5. The performance of directional representations and their combinations in terms of MAE, RMSE, and R². For HFMD dataset, d₁ is fixed at 4, and d₂ is fixed at 11. For HB datasets, d₁ is fixed at 4, and d₂ is fixed at 2.

6.4. Comparison

To validate the prediction performance of all the mentioned methods, extensive experiments of twelve methods are conducted on two disease datasets. For these methods, d₁ and d₂ are set to be the same, and the parameters of benchmark methods are tuned at the optimal performance.

The experimental results are displayed in Fig. 6. Some major observed results are listed below:

- (1) The DGDR significantly outperforms the other methods.
- (2) For two disease datasets, CNN1D obtain the worst performance among benchmark methods.
- (3) For RNN-variant methods, the number of hidden neurons has slight effects on prediction performance.
- (4) The methods based on daily data significantly outperform the methods based on weekly data.
- (5) The consolidation of weekly and daily data further improves the prediction performance.

The CNN1D gets poor performance in two disease datasets. A possible reason is that the CNN would filter small values, which may ignore the temporal dynamics within the incubation period.

The number of hidden neurons has little effect on the prediction performance. A possible reason is that the disease transmission may have not strong periodic events. Hence, the model cannot be well trained for disease datasets.

From the data aspect, the benchmark methods have better performance on daily data than weekly data. For two disease datasets, the RMSE value is decreased by 15.1% and 3.2% when fed

with daily data and weekly data, respectively. This reveals that fine-grained datasets provide sufficient the temporal dynamics. However, for the HB datasets, MLR has a poorer performance on daily data than weekly data. A possible reason is that the daily HB time series greatly fluctuates, which is not belong to linear distributions.

The experiments on the combination of daily data and weekly data have better performance than only one kind of data. For the experiments of DGR on HFMD datasets, the RMSE value is decreased by 26.3% at most. For the experiments of DGR on HB datasets, the RMSE value is decreased by 12.6% at most. This demonstrates that the temporal dynamics from the different granularities have overlaps but are different. This reveals that the dual-grained data are essential.

The proposed DGDR outperforms the benchmark methods, and the improvements are significant. For the experiments on weekly data, the RMSE value is decreased by 12.2% and 22.78% at most on HFMD datasets and HB datasets, respectively. For the experiments on daily data, comparing the DGDR with other methods, the RMSE value is decreased by 18.82% and 26.51% at most, respectively. It demonstrates that the proposed DR component is efficient to capture temporal dynamics from both equal-grained time series and fine-grained time series. For two disease datasets, DGDR decreases the RMSE value by 8.0% and 3.3% when compared with DGR. The DGR is designed for dual-grained time series, but it only has the ability in representing the time step dimension. This shows the impact of the DR component in capturing and consolidating the temporal dynamics.

In summary, the main reasons for the improvement are as follows: (1) dual-grained data can provide with more sufficient

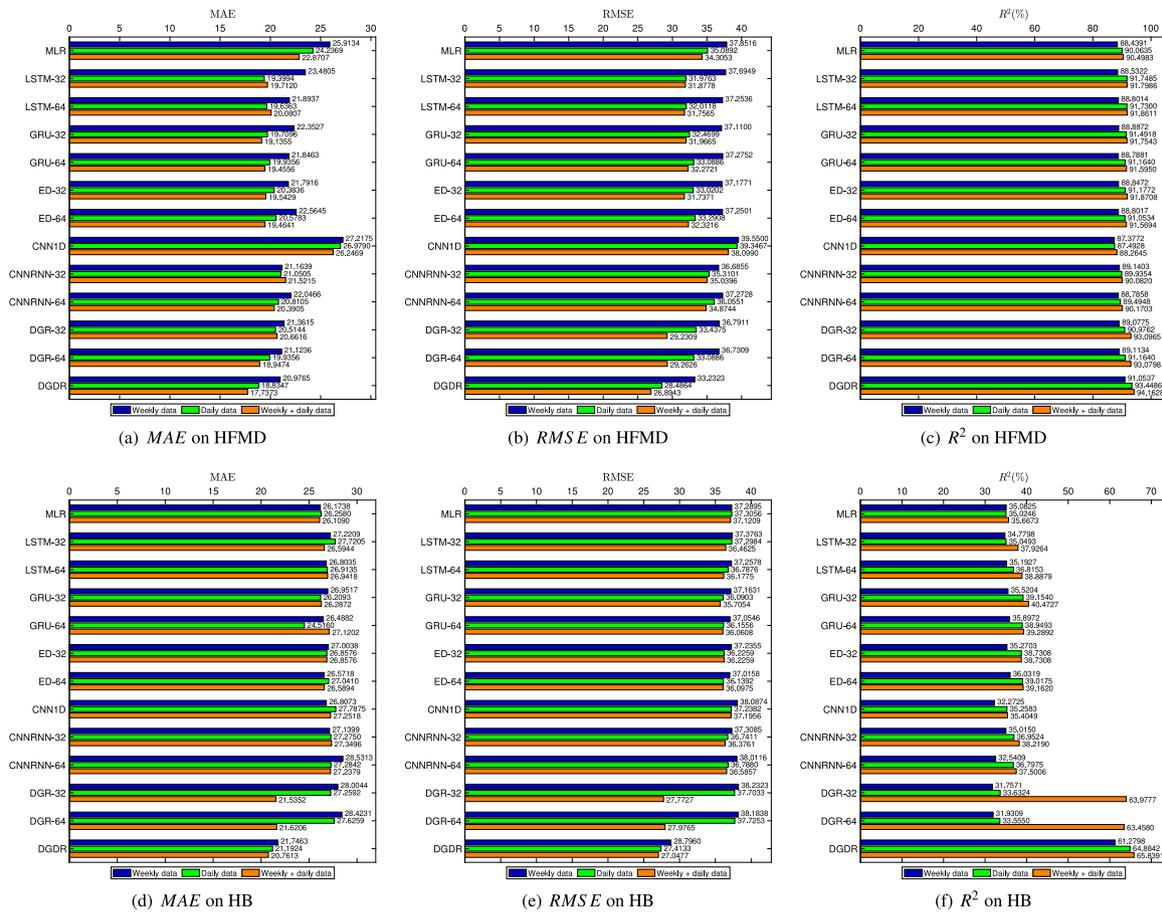


Fig. 6. The performance comparisons of all methods. For both HFMD and HB datasets, d_1 is both fixed at 4, and d_2 is fixed at 11 and 2, respectively.

temporal dynamic; (2) the DR component can efficiently capture the temporal dynamic from the dual-grained time series. They both efficiently help DGDR to learn the various kinds of temporal patterns.

7. Conclusion

This paper proposed a common infectious disease prediction method, which is named Dual-grained Directional Representation (DGDR), to predict the outpatient number in the upcoming week. The core directional representation (DR) component is designed to extract the temporal dynamics of a time series from three directions with high fusion ability. The proposed DGDR learns those representations from both equal-time series data and fine-grained time series data.

Extensive experiments on the hand, foot, and mouth disease (HFMD) dataset and the hepatitis beta (HB) dataset show the effectiveness of the DGDR. The experimental analyses show the HFMD high-incidence period is around 7 weeks. Moreover, the DGDR has been discussed and its predictions have been adopted by the Xiamen City Center for Disease Control and Prevention (XMCCDC).

In the future, we will further study the associations between different infectious diseases, which would provide better control of disease transmission.

CRediT authorship contribution statement

Peisong Zhang: Validation, Writing – original draft. **Zhijin Wang:** Conceptualization, Methodology, Software, Supervision. **Yaohui Huang:** Visualization, Investigation. **Mingzhai Wang:** Data curation, Formal analysis, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgments

This work was partly supported by the Natural Science Foundation of Fujian Province (CN) (nos. 2021J01859 and 2022J01335). Thanks to the Xiamen City Center for Disease Control and Prevention (XMCCDC) for sharing the data. We gratefully acknowledge the editor and anonymous reviewers for their efforts and time in improving the paper.

References

- [1] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, *Lancet* 395 (10224) (2020) 565–574, [http://dx.doi.org/10.1016/S0140-6736\(20\)30251-8](http://dx.doi.org/10.1016/S0140-6736(20)30251-8).
- [2] X. Li, H.M. Krumholz, W. Yip, K.K. Cheng, J. De Maeseneer, Q. Meng, E. Mossialos, C. Li, J. Lu, M. Su, et al., Quality of primary health care in China: challenges and recommendations, *Lancet* 395 (10239) (2020) 1802–1812, [http://dx.doi.org/10.1016/S0140-6736\(20\)30122-7](http://dx.doi.org/10.1016/S0140-6736(20)30122-7).

- [3] G. Pullano, L. Di Domenico, C.E. Sabbatini, E. Valdano, C. Turbelin, M. Debin, C. Guerri, C. Kengne-Kuetche, C. Souty, T. Hanslik, et al., Underdetection of cases of COVID-19 in France threatens epidemic control, *Nature* 590 (7844) (2021) 134–139, <http://dx.doi.org/10.1038/s41586-020-03095-6>.
- [4] J.Y. Ko, M.L. Danielson, M. Town, G. Derado, K.J. Greenlund, P.D. Kirley, N.B. Alden, K. Yousey-Hindes, E.J. Anderson, P.A. Ryan, S. Kim, R. Lynfield, S.M. Torres, G.R. Barney, N.M. Bennett, M. Sutton, H.K. Talbot, M. Hill, A.J. Hall, A.M. Fry, S. Garg, L. Kim, C.-N.S. Team, Risk factors for coronavirus disease 2019 (COVID-19)-associated hospitalization: COVID-19-associated hospitalization surveillance network and behavioral risk factor surveillance system, *Clin. Infect. Dis.* 72 (11) (2020) e695–e703, <http://dx.doi.org/10.1093/cid/ciaa1419>.
- [5] Z. Wang, B. Cai, COVID-19 cases prediction in multiple areas via shapelet learning, *Appl. Intell.* 52 (1) (2022) 595–606, <http://dx.doi.org/10.1007/s10489-021-02391-6>.
- [6] Y. Huang, P. Zhang, Z. Wang, Z. Lu, Z. Wang, HFMD cases prediction using transfer one-step-ahead learning, *Neural Process. Lett. Early View* (2022) <http://dx.doi.org/10.1007/s11063-022-10795-9>.
- [7] Y. Xiang, Y. Jia, L. Chen, L. Guo, B. Shu, E. Long, COVID-19 epidemic prediction and the impact of public health interventions: A review of COVID-19 epidemic models, *Infect. Dis. Model.* 6 (2021) 324–342, <http://dx.doi.org/10.1016/j.idm.2021.01.001>.
- [8] Z. Wang, Y. Huang, B. He, Dual-grained representation for hand, foot, and mouth disease prediction within public health cyber-physical systems, *Softw. Pract. Exp.* 51 (11) (2021) 2290–2305, <http://dx.doi.org/10.1002/spe.2940>.
- [9] L. Wang, A. Adiga, J. Chen, A. Sadilek, S. Venkatramanan, M.V. Marathe, CausalGNN: Causal-based graph neural networks for spatio-temporal epidemic forecasting, in: *Proceedings of 36th AAAI Conference on Artificial Intelligence, AAAI Press, 2022*, pp. 12191–12199, *Virtual*.
- [10] R. Rosenfeld, R.J. Tibshirani, Epidemic tracking and forecasting: Lessons learned from a tumultuous year, *Proc. Natl. Acad. Sci.* 118 (51) (2021) e2111456118, <http://dx.doi.org/10.1073/pnas.2111456118>.
- [11] Z. Wang, Q. Su, G. Chao, B. Cai, Y. Huang, Y. Fu, A multi-view time series model for share turnover prediction, *Appl. Intell. Early View* (2022) <http://dx.doi.org/10.1007/s10489-021-02979-y>.
- [12] S. Bai, J.Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, 2018, *CoRR*, [arXiv:1803.01271](https://arxiv.org/abs/1803.01271).
- [13] E. Hosenzade, S. Haratzadeh, CNNpred: CNN-based stock market prediction using a diverse set of variables, *Expert Syst. Appl.* 129 (2019) 273–285, <http://dx.doi.org/10.1016/j.eswa.2019.03.029>.
- [14] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, G.W. Cottrell, A dual-stage attention-based recurrent neural network for time series prediction, in: *International Joint Conference on Artificial Intelligence, Melbourne, Australia, 2017*, pp. 2627–2633, <http://dx.doi.org/10.24963/ijcai.2017/366>, ijcai.org.
- [15] B. Zhou, G. Yang, Z. Shi, S. Ma, Interpretable temporal attention network for COVID-19 forecasting, *Appl. Soft Comput.* 120 (2022) 108691, <http://dx.doi.org/10.1016/j.asoc.2022.108691>.
- [16] P. Zhang, Z. Wang, G. Chao, Y. Huang, J. Yan, An oriented attention model for infectious disease cases prediction, in: *Proceedings of the 34th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Vol. 13343, Springer, Kitakyushu, Japan, 2022*, http://dx.doi.org/10.1007/978-3-031-08530-7_11.
- [17] L. Li, Y. Jiang, B. Huang, Long-term prediction for temporal propagation of seasonal influenza using transformer-based model, *J. Biomed. Inform.* 122 (2021) 103894, <http://dx.doi.org/10.1016/j.jbi.2021.103894>.
- [18] Y.-C. Chen, P.-E. Lu, C.-S. Chang, T.-H. Liu, A time-dependent SIR model for COVID-19 with undetectable infected persons, *IEEE Trans. Netw. Sci. Eng.* 7 (4) (2020) 3279–3294, <http://dx.doi.org/10.1109/TNSE.2020.3024723>.
- [19] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai, et al., Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions, *J. Thorac. Dis.* 12 (3) (2020) 165–174, <http://dx.doi.org/10.21037/jtd.2020.02.64>.
- [20] H. Kamarthi, L. Kong, A. Rodriguez, C. Zhang, B.A. Prakash, When in doubt: Neural non-parametric uncertainty quantification for epidemic forecasting, in: *Proceedings of 35th Advances in Neural Information Processing Systems, 2021*, pp. 19796–19807, *Virtual*.
- [21] A. Hernandez-Matamoros, H. Fujita, T. Hayashi, H. Perez-Meana, Forecasting of COVID19 per regions using ARIMA models and polynomial functions, *Appl. Soft Comput.* 96 (2020) 106610, <http://dx.doi.org/10.1016/j.asoc.2020.106610>.
- [22] B. Adhikari, X. Xu, N. Ramakrishnan, B.A. Prakash, EpiDeep: Exploiting embeddings for epidemic forecasting, in: A. Teredesai, V. Kumar, Y. Li, R. Rosales, E. Terzi, G. Karypis (Eds.), *Proceedings of the 25th ACM Conference on Knowledge Discovery and Data Mining, ACM, Anchorage, USA, 2019*, pp. 577–586, <http://dx.doi.org/10.1145/3292500.3330917>.
- [23] G. Panagopoulos, G. Nikolentzos, M. Vazirgiannis, Transfer graph neural networks for pandemic forecasting, in: *Proceedings of 35th AAAI Conference on Artificial Intelligence, AAAI Press, 2021*, pp. 4838–4845, *Virtual*.
- [24] G. Lai, W. Chang, Y. Yang, H. Liu, Modeling long- and short-term temporal patterns with deep neural networks, in: *International Conference on Research and Development in Information Retrieval, ACM, Ann Arbor, MI, USA, 2018*, pp. 95–104, <http://dx.doi.org/10.1145/3209978.3210006>.
- [25] S. Yang, J. Wu, C. Ding, Y. Cui, Y. Zhou, Y. Li, M. Deng, C. Wang, K. Xu, J. Ren, B. Ruan, L. Li, Epidemiological features of and changes in incidence of infectious diseases in China in the first decade after the SARS outbreak: an observational trend study, *Lancet Infect. Dis.* 17 (7) (2017) 716–725, [http://dx.doi.org/10.1016/S1473-3099\(17\)30227-X](http://dx.doi.org/10.1016/S1473-3099(17)30227-X).
- [26] China Center for Disease Control and Prevention, A research report on the cognitive status of Chinese hepatitis B Patients1, 2005, http://http://www.chinacdc.cn/jkzt/crb/zl/bdxgy/jyszL_2211/201103/t20110314_28179.html. (Accessed 2005).
- [27] R. Kumar, F. Al-Turjman, L. Srinivas, M. Braveen, J. Ramakrishnan, ANFIS for prediction of epidemic peak and infected cases for COVID-19 in India, *Neural Comput. Appl.* (2021) 1–14, <http://dx.doi.org/10.1007/s00521-021-06412-w>.
- [28] A. Kara, Multi-step influenza outbreak forecasting using deep LSTM network and genetic algorithm, *Expert Syst. Appl.* 180 (2021) 115153, <http://dx.doi.org/10.1016/j.eswa.2021.115153>.
- [29] D. Radojicic, S. Kredatus, The impact of stock market price Fourier transform analysis on the gated recurrent unit classifier model, *Expert Syst. Appl.* 159 (2020) 113565, <http://dx.doi.org/10.1016/j.eswa.2020.113565>.
- [30] Y. Cui, C. Zhu, G. Ye, Z. Wang, K. Zheng, Into the unobservables: A multi-range encoder-decoder framework for COVID-19 prediction, in: *Proceedings of the 30th ACM International Conference on Information & Knowledge Management, Association for Computing Machinery, 2021*, pp. 292–301, <http://dx.doi.org/10.1145/3459637.3482356>.
- [31] N. Ben Yahia, M. Dhiaeddine Kandara, N. Bellamine BenSaoud, Integrating models and fusing data in a deep ensemble learning method for predicting epidemic diseases outbreak, *Big Data Res.* 27 (2022) 100286, <http://dx.doi.org/10.1016/j.bdr.2021.100286>.
- [32] I.E. Livieris, E. Pintelas, P. Pintelas, A CNN-LSTM model for gold price time-series forecasting, *Neural Comput. Appl.* 32 (23) (2020) 17351–17360, <http://dx.doi.org/10.1007/s00521-020-04867-x>.
- [33] X. Yin, H. Yi, J. Shu, X. Wang, X. Wu, L. Yu, Clinical and epidemiological characteristics of adult hand, foot, and mouth disease in northern Zhejiang, China, May 2008–November 2013, *BMC Infect. Dis.* 14 (1) (2014) 1–8, <http://dx.doi.org/10.1186/1471-2334-14-251>.
- [34] J. Xiao, Q. Zhu, F. Yang, S. Zeng, Z. Zhu, D. Gong, Y. Li, L. Zhang, B. Li, W. Zeng, X. Li, Z. Rong, J. Hu, G. He, J. Sun, J. Lu, T. Liu, W. Ma, L. Sun, The impact of enterovirus A71 vaccination program on hand, foot, and mouth disease in Guangdong, China: A longitudinal surveillance study, *J. Infect. Early View* (2022) <http://dx.doi.org/10.1016/j.jinf.2022.06.020>.
- [35] China Center for Disease Control and Prevention, Health tips for may 1st holiday in 2021, 2021, http://www.chinacdc.cn/yw_9324/202104/t20210425_230152.html. (Accessed 2021).