

RESEARCH

Open Access



A novel graph neural network based approach for influenza-like illness nowcasting: exploring the interplay of temporal, geographical, and functional spatial features

Jiajia Luo^{1†}, Xuan Wang^{1†}, Xiaomao Fan^{2†}, Yuxin He³, Xiangjun Du¹, Yao-Qing Chen¹ and Yang Zhao^{1*}

Abstract

Background Accurate and timely monitoring of influenza prevalence is essential for effective healthcare interventions. This study proposes a graph neural network (GNN)-based method to address the issue of cross-regional connectivity in predicting influenza outbreaks, aiming to achieve real-time and accurate influenza prediction.

Methods We proposed a GNN-based approach with dual topology processing, capturing both geographical and socio-economic associations among counties/cities. The model inputs consist of weekly matrices of influenza-like illness (ILI) rates at city level, along with geographical topology and functional topology. The model construction involves temporal feature extraction through 1-dimensional gated causal convolution, spatial feature embedding through graph convolution, and additional adjustments to enhance spatiotemporal interaction exploration. Evaluation metrics include four commonly used measures: root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and Pearson correlation (Corr).

Results Our approach for predicting influenza outbreaks achieves competitive performance on real-world datasets (Corr = 0.8202; RMSE = 0.0017; MAE = 0.0013; MAPE = 0.0966), surpassing established baselines. Notably, our approach exhibits excellent capability in accurately and timely capturing short-term influenza outbreaks during the flu season, outperforming competitors across all evaluation metrics.

Conclusion The incorporation of dual topology processing and the subsequent fusion mechanism allows the model to explore in-depth spatiotemporal feature interactions. Demonstrating superior performance, our approach shows great potential in early detection of flu trends for facilitating public health decisions and resource optimization.

Keywords Influenza prediction, Spatiotemporal interaction, Graph neural network (GNN), Functional topology, Socio-economic association

[†]Jiajia Luo, Xuan Wang, and Xiaomao Fan contributed equally to this work.

*Correspondence:

Yang Zhao

zhaoy393@mail.sysu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Influenza poses a serious and enduring global public health concern. Each year, a substantial number of individuals suffer from severe cases of influenza, ranging from 3 to 5 million. The associated mortality is estimated to reach from 2.9 to 6.5 million deaths [1]. The influenza viruses have the capacity to instigate both epidemics and pandemics, leading to the explosion of infections in susceptible communities and eventually culminating in significant socio-economic burdens for families and society [2–4]. Ensuring accurate and timely influenza prediction is of urgent necessity, as it empowers health departments to formulate effective strategies for influenza-related endeavors.

In recent years, machine learning and statistical methods have been extensively applied in influenza prediction due to their simplicity and efficiency [5, 6]. External variables, such as internet search data, weather records, and holiday information, have been validated as correlated factors with influenza prevalence [7]. Least absolute shrinkage and selection operator (LASSO), support vector machine (SVM) and tree-based methods enable integrating these external variables to enhance the accuracy of influenza predictions [6, 8, 9]. Nevertheless, procuring these external factors can pose challenges, and there may exist the potential influence of “data hubris” [10]. Additionally, autoregressive integrated moving average (ARIMA), along with other sequence-based approaches, offer a flexible framework capable of capturing temporal patterns [11, 12]. However, unforeseen patterns may occasionally emerge in low-latitude regions characterized by volatile influenza activity [6]. Relying solely on temporal patterns can not fully address the complex task of accurately predicting influenza outbreaks.

Deep learning techniques, along with novel methods derived from real-world data, have garnered significant attention [13–15]. The advancement of these technologies has paved the way for their application in predicting influenza outbreaks. By leveraging deep learning algorithms such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph neural networks (GNNs), it becomes possible to extract temporal and spatial patterns that facilitate the prediction of cross-regional influenza dynamics [16–22]. Among them, GNNs excel in capturing and propagating information across graph connections, enabling efficient handling of irregular or dynamic links between entities [23]. While CNNs are typically suited for grid-like data and RNNs for sequential data [24, 25], they might not match the performance of GNNs in effectively capturing spatial and temporal dependencies. GNNs have been widely applied across diverse

domains. In social networks, they support community detection and behavior prediction [26, 27]. In biology and chemistry, GNNs aid molecular property prediction, drug discovery, and protein design, as exemplified by ProteinMPNN, which encodes protein structures as graphs [28]. Additionally, GNNs enhance recommendation systems, traffic forecasting, and tasks in natural language processing involving structural representations [29–31]. Moreover, current research also incorporates the self-attention mechanism to investigate the spatiotemporal correlations in influenza [32, 33]. However, self-attention might not sufficiently encompass the relationships among diverse regions, potentially restricting its effectiveness in handling a broad spectrum of cross-regional interactions.

Spatial patterns play a vital role in shaping the distribution of influenza. For instance, the interplay of population density and regional connectivity frequently serves as a critical framework for elucidating the dynamics of disease transmission [34]. High degrees of correlations in influenza activity are observed among geographically adjacent areas, aligning with Tobler’s first law of geography (TFL): “Everything is related to everything else, but near things are more related than distant things” [35]. In general, as the distance between infected and susceptible cities increases, influenza transmission weakens sharply. However, the modern highly developed transportation networks have mitigated the dominant role of this geographical distance. As a result, there still exists the persistence of remote transmission events because of factors like work commuting and air traffic [36]. Existing influenza spatiotemporal prediction approaches predominantly deal with the spatial features in geographical aspect, neglecting the functional spatial factors [16, 21]. Figure 1 illustrates the scenario where central cities A, B, C, and D have close connections with geographically surrounding cities (indicated by blue dashed lines). Additionally, there are active functional interactions among these central cities in terms of transportation, economy, trade, tourism and so on (indicated by red dashed lines). Therefore, capturing the functional interconnections between regions is crucial for accurately modeling the transmission dynamics of influenza, as it can rapidly spread to regions that may appear geographically distant. Moreover, integrating geographical spatial dynamics with functional socio-economic factors enhances the predictive performance by incorporating diverse aspects of transmission dynamics and regional interactions.

To address the aforementioned limitations of the existing models for influenza prediction, we propose a GNN based approach to examine the innovative temporal-geospatial-functional spatial correlations. This study presents the following key contributions:

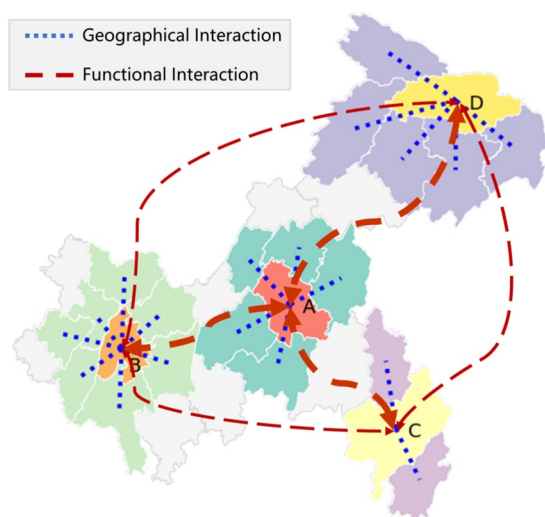


Fig. 1 Spatial Interactions and Influences Between Cities. A, B, C, and D are the central cities in each of the respective regions. Blue dashed lines represent geographical adjacency interactions, while red dashed lines represent cross-regional long-distance functional interactions

- Alongside the geographical topology, we construct the additional functional topology by calculating the threshold correlation coefficient matrix among node sequences. This functional topology serves as a concentrated reflection of both functional socio-economic effects on influenza and co-occurrence patterns of influenza prevalence, expanding spatial feature dimensionality and providing valuable information.
- By incorporating dual-channel parallel spatio-temporal graph convolutional modules, the developed model facilitates the simultaneous learning of both temporal-geographical and temporal-functional spatial feature interactions. The comprehensive fusion of these interactions information enhances the predictive performance of our model.
- We validate our model on real-world influenza surveillance datasets, and the results emphasize the critical importance of introducing functional topology. Particularly, during the initial phases of rapid influenza transmission escalation in the flu season, our model demonstrates the exceptional capability of capturing and responding to data fluctuations and trends.
- Our model's simple and flexible structure, along with its proven robustness, highlights its strong application potential.

The remaining sections of this paper are organized as follows: [Related works](#) section presents a comprehensive

review of related research on the influenza prediction approach. In [Methods](#) section, we outline the implementation of the proposed model. The results of our evaluation scheme and comparison with classical influenza prediction approaches are presented in [Experiments](#) and [Results](#) sections. [External validation](#) section presents the model's prediction results on another dataset to validate its generalization ability. [Discussion](#) section offers a detailed discussion. Finally, [Conclusion](#) section concludes the paper with key remarks.

Related works

Previous research works have developed various modeling methods for influenza forecasting. ARIMA and seasonal ARIMA models are commonly employed when the influenza data have a clear trend and seasonality. For instance, He et al. utilized ARIMA to predict the positive rates of various types of influenza viruses in children [8]. Influenza follows cyclical and seasonal patterns, with its peak in cases consistently occurring during the winter season, typically from December to March in temperate or subtropical regions of the northern hemisphere [37, 38]. ARIMA is adept at capturing the trends, seasonality, and cyclic variations in influenza activity, thus facilitating accurate and timely predictions of influenza outbreaks [11]. However, its linear assumptions constrain adaptability to non-linear and abrupt variations, posing challenges in handling sudden changes and incomplete data due to its strong dependence on historical information.

Exogenous data, such as internet search query volumes, have been combined with machine learning techniques to improve influenza prediction accuracy. For instance, Google Flu Trends (GFT), introduced in 2008, utilized Google search queries related to influenza to monitor ILI [39]. However, the reliability of GFT's predictions was compromised by changes in Google's search algorithms and user behavior [10]. More recently, researchers have explored diverse data sources, including Baidu Index for influenza search patterns in China [40], and integrated Google search trends, weather information, electronic health records (EHR), and social media data for short-term influenza forecasting in France [41]. Despite these advancements, the majority of these studies have primarily operated at a broader scale, and there may exist limitations in their applicability for meticulous influenza monitoring at the city/county level.

The fusion of spatial interdependencies and temporal patterns has demonstrated improved efficacy in cross-regional influenza prediction, surpassing traditional sequence-based prediction approaches [16]. Rapid advancements in deep learning algorithms have allowed for innovative investigations on spatiotemporal interactions, expanding prediction capabilities. For instance,

He et al. [16] employed the spatiotemporal residual network (ST-ResNet) algorithm for influenza nowcasting. This approach transformed the ILI data for each region into a format that resembles a picture, using information about the region's shape. By combining time-series data with geographical information, the ST-ResNet algorithm improved the accuracy for cross-regional influenza forecasts. Similarly, Wang et al. [42] utilized convolutional operations to extract spatial neighborhood information. These spatial features were then combined to form a new time series, and RNNs were employed to capture temporal patterns. However, these spatiotemporal prediction approaches primarily focus on geographical spatial information, neglecting the valuable functional spatial information that arise from various socio-economic factors.

Exploring the correlations among regions, which share similar attributes but geographically distant, is crucial for understanding the spread of influenza [43]. We term the connectivity driven by socio-economic factors, extending beyond geographical links, as “functional topology”. Zhou et al. [44] incorporated international air transportation data and diverse internet-sourced information into a multi-modal Hidden Markov Model. This approach aimed to predict influenza outbreaks across 106 countries and regions worldwide, considering global travel patterns and diverse data sources. Similarly, Guo et al. [45] examined spatial correlations among ten U.S. regions through historical influenza data's correlation coefficients, employing multiple regression analysis to explore connections between ILI data in different regions. This analysis helped discover dependencies and patterns that contribute to a better understanding of influenza transmission. However, the study of functional topology in influenza prediction, as well as the combination of functional and geographical topologies, remains a relatively new and unexplored area. There is a pressing need for more comprehensive and meticulous research in integrating these two topologies.

Additionally, recent research has also brought up regional influenza prediction models using self-attention mechanisms. Jung et al. [32] and Moon et al. [33] employed self-attention mechanisms to identify similarities among influenza patterns across regions. Although the self-attention mechanism is proficient at capturing similarity and correlation, its reliance on local information restricts its potential to explore spatiotemporal associations across regions. Moreover, interpretability is crucial for reliable and convincing influenza predictions, but the self-attention mechanism, while effective in capturing statistical patterns, lacks interpretability. In contrast, GNNs excel with explicit graph structure (spatial information) and interpretability in spatiotemporal convolutional interactions.

Methods

ILI data reported by CDC is usually at the county/city level, limiting the availability of meticulous spatial distribution information at finer granularity. This inherent limitation is a primary factor restricting the predictive performance of many spatiotemporal prediction models. However, from a macro perspective, various counties/cities are interconnected nodes with specific influence and correlation in space. Therefore, it is more appropriate to utilize graph neural networks at this coarser level of spatial information granularity. In this section, we elaborate on the process of data preparation and model construction.

Notations

Indicator of influenza prevalence intensity: The influenza-like illness rate (ILI_Rate) [46] refers to the proportion of outpatient and inpatient visits attributed to flu-like symptoms. It serves as an ideal indicator for assessing the intensity of influenza prevalence within a specific temporal and spatial scope. For ILI surveillance data reported at the county/city level with weekly intervals, we calculate the ILI_Rate to evaluate the influenza intensity. This calculation is depicted in (1).

$$\text{ILI_Rate}_{t,k} = \frac{N_{t,k}^{\text{ILI}}}{N_{t,k}^{\text{Visits}}} \quad (1)$$

where t denotes the week number, k denotes the county/city; $N_{t,k}^{\text{ILI}}$ and $N_{t,k}^{\text{Visits}}$ represent the total number of patients diagnosed with ILI, the total number of outpatient and inpatient visits within county/city k during week t , respectively.

Problem definition

ILI_Rate nowcasting task: The ILI_Rate nowcasting task involves learning a mapping function $f(\cdot)$ to predict the county/city-level ILI_Rate for a specific week. This prediction is based on the historical ILI_Rate sequences spanning M weeks, as well as consideration of the geographical relationships (G_{geo}) and functional connectivity (G_{func}) among counties/cities. This process is illustrated in (2).

$$Y = f(\mathbf{X}, G_{\text{geo}}, G_{\text{func}}) \quad (2)$$

where Y represents the predicted ILI_Rate for specific week, $\mathbf{X} = \{X^1, X^2, \dots, X^M\}$ denotes the historical ILI_Rate sequences for the past 1 to M weeks, and G_{geo} and G_{func} refer to the geographical and functional topologies defined in [Model inputs](#) section.

Model inputs

The overall architecture of our proposed Dual-Topology Spatio-temporal Graph Convolutional Networks (Dual-Topo-STGCN) is illustrated in Fig. 2. For the task of influenza spatiotemporal nowcasting, the inputs include historical weekly ILI_Rate data for each node, a graph representing geographical relationships and another graph representing functional connectivity.

Input 1

ILI_Rate matrix $X^{T \times N}$.

The ILI_Rate matrix is denoted as $X^{T \times N}$, where T is the length of the historical ILI_Rate sequence and N is the number of nodes which correspond to counties or cities; X_i^t represents the ILI_Rate of node i during week t ; X_i represents the historical ILI_Rate sequence of node i ; X^t collectively characterizes the ILI_Rate for all nodes during week t across the entire country or region.

Input 2

Geographical topology G_{geo} .

It is crucial to understand how the interactions between geographically adjacent counties/cities impact on influenza prevalence. We describe the geographical adjacency of counties/cities as a graph $G_{\text{geo}} = (V, A_{\text{geo}}, E_{\text{geo}})$.

$V = \{v_1, v_2, \dots, v_N\}$ represents the set of nodes characterizing various counties/cities, where N denotes the number of nodes in the graph. E_{geo} represents the set of edges, where an edge e_{ij} exists between node v_i and node v_j if they are geographically adjacent. $A_{\text{geo}} \in \mathbb{R}^{N \times N}$ denotes the adjacency matrix of the counties/cities network topology graph. In the adjacency matrix A_{geo} , elements with edges are represented as 1, while non-adjacent elements are represented as 0.

Input 3

Functional topology G_{func} .

Another much more crucial aspect of spatial relationships involves close intercity mobility, which is driven by a

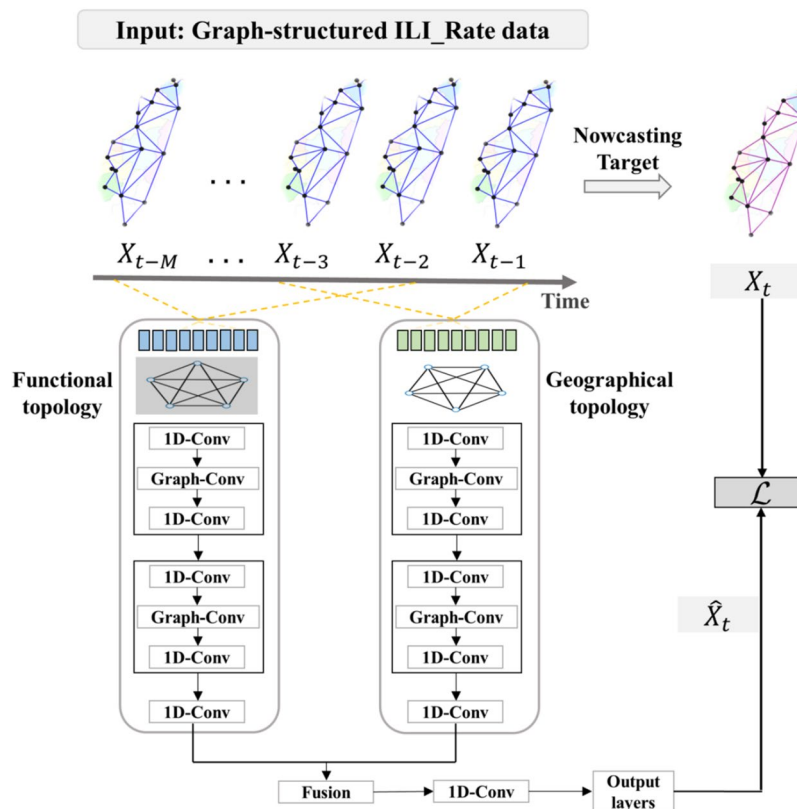


Fig. 2 Illustration of the architecture of the Dual-Topo-STGCN model. $X_{t-M}, \dots, X_{t-1}$ represent the input graph-structured and normalized ILI_rate feature matrices. Functional topology and Geographical topology refer to the normalized adjacency matrices representing functional spatial relationships and geographical connections, respectively. 1D-Conv denotes the one-dimensional causal convolution operation with a gating mechanism, and Graph-Conv represents the graph convolution operation. Fusion refers to the integration of dual-channel features, which can involve operations such as selecting the maximum or minimum value, averaging, weighted averaging, or concatenation, depending on the application scenario. Output layers are fully connected layers that perform step-by-step dimensionality reduction on the fused features

multitude of functional socio-economic factors. Similarly, to describe this functional connectivity, we develop a graph denoted as $G_{\text{func}} = (V, A_{\text{func}}, E_{\text{func}})$.

Same as G_{geo} , V denotes the set of N nodes in the graph. $A_{\text{func}} \in \mathbb{R}^{N \times N}$ represents the similarity matrix of these nodes. In the construction of A_{func} , we calculate the correlation coefficient matrix of individual nodes using the aforementioned ILI_Rate matrix. This correlation relationship serves as a summary of the influenza fluctuation patterns and co-occurrence characteristics among nodes. To maintain the sparsity of G_{func} , a threshold is established. Elements exceeding the threshold are considered significant and are preserved as edges, while those below are set to 0. This sparsity constraint helps reduce computational complexity and allows for more efficient analysis and modeling of the functional topology. Depending on the sparsity setting, this threshold changes accordingly. E_{func} denotes the set of edges in the functional topology G_{func} . In this context, an edge e_{ij} exists between nodes v_i and v_j if their similarity in terms of influenza prevalence exceeds the aforementioned threshold, indicating a potential strong connectivity between these nodes.

Feature extraction

To capture the spatiotemporal interactions of influenza data, we utilize a 1D causal convolution with gating for temporal feature extraction and a graph convolutional layer for incorporating spatial information. These two operations are represented as “1D-Conv” and “Graph-Conv” in the model architecture diagram (Fig. 2).

Temporal dependence modelling

To extract temporal features, we employ a 1-dimensional (1D) causal convolution with a gated mechanism. This method consists of three parallel convolutional layers, each utilizing a 1x3 kernel. For each node's sequence, the first layer performs a standard convolution operation, while the second layer applies a sigmoid activation function to its convolutional output. The output of the second layer then serves as weight to be combined with the result of the first layer, implementing gated linear units (GLU). Subsequently, this combined output is added to the result of the third layer before applying a Rectified Linear Unit (ReLU) activation function. This entire process results in the final output of the temporal convolution.

The definition of temporal gated convolution is illustrated in (3).

$$\begin{aligned} X_{\mathcal{T}} &= \Gamma *_{\mathcal{T}} X_{\text{in}} \\ &= \text{ReLU}(P \odot \sigma(Q) + H) \in \mathbb{R}^{(M-K_t+1) \times C_o} \end{aligned} \quad (3)$$

where Γ denotes temporal convolutional kernel, $*_{\mathcal{T}}$ denotes temporal convolution, and X_{in} is the input time

series of nodes. P , Q and H are the outputs of the three stacked convolutional layers. P and Q serve as the inputs of gates in GLU, respectively, while H represents the residual connections; $\odot$ denotes the element-wise Hadamard product. The sigmoid gate $\sigma(Q)$ controls which input P of the current states plays a key role in time series. M is the length of time series, K_t is the size of convolutional kernel and C_o denotes the number of channels in the output features.

Spatial feature embedding

Following the extraction of temporal features, we employ a graph convolutional layer to capture the spatial information between nodes. Each node is influenced to varying degrees by its connected nodes, encompassing both geographical adjacency and cross-regional socio-economic interactions. For each individual node, we aggregate information from its neighborhood nodes, and then fuse it with its own intrinsic features. These synthesized features are then transformed by a learnable weight matrix, ultimately yielding updated features enriched with spatial information. The graph convolution is shown in (4).

$$\begin{aligned} X_{\mathcal{ST}} &= \Theta *_G X_{\mathcal{T}} \\ &= \text{ReLU}(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \cdot X_{\mathcal{T}} \cdot \Theta) \end{aligned} \quad (4)$$

where Θ denotes the graph convolutional kernel, a learnable parameter matrix used for convolving input features; $*_G$ denotes graph convolution; $X_{\mathcal{T}}$ is the input feature matrix for graph convolution, which is also the output feature matrix of temporal convolution. $D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ signifies the symmetric normalized adjacency matrix, which is multiplied with the input temporal feature matrix $X_{\mathcal{T}}$ to embed the spatial feature.

Model construction and loss

As shown in Fig. 2, for a given set of sequences, we have two distinct topologies: G_{geo} and G_{func} , each representing different spatial relationships. A single-pathway network encounters challenges in simultaneously capturing the interactions between temporal features and multiple types of spatial features. To address this issue, we adopt a strategy of duplicating the data processing pathways and connecting them in parallel. Each pathway is responsible for handling its own spatiotemporal convolutions and feature extraction.

Within each pathway, a sequence of operations forms a spatiotemporal feature extraction block, including gated temporal convolution, graph convolution, and another gated temporal convolution. After concatenating several of these spatiotemporal convolution blocks, the extracted features are fused through a specific mechanism.

Subsequently, these fused features pass through a final gated temporal convolution layer, then being input to the output layer. This output layer is composed of a series of fully connected layers with dimension reduction, mapping the extracted features to ILI_Rate predictions for a future week.

In this study, we evaluate our model's performance using the L2 loss. The loss function for ILI_Rate prediction is defined as shown in (5).

$$\mathcal{L} = \|\hat{X}_t - X_t\|^2 \quad (5)$$

where $\hat{X}_t$ denotes the predicted ILI_Rate and X_t corresponds to the ground truth.

Experiments

Data source

In this study, we employed real-world ILI datasets to validate our developed prediction model. The ILI surveillance data [47], including ILI cases at county/city level from Centers for Disease Control, Taiwan province, China (Taiwan CDC) has been published on a weekly basis since 2008, providing researchers with a wealth of continuous, comprehensive, and reliable data for studies. Additionally, Taiwan's geographical characteristics make it an interesting case for this research. The main island of Taiwan is surrounded by the sea, which minimizes the influence of cross-border land transportation on the spread of diseases. The island's scattered counties/cities can each be regarded as a node, and the hubs of maritime and aerial transportation are primarily some key nodal cities. This distribution characteristic is a typical represent of the clustering of developed-underdeveloped urban areas in most regions. Such generality assures the generalization capability of functional topological information in influenza transmission.

The dataset covers a 12-year period, from week 14, 2008, to week 10, 2020, spanning from the initiation of ILI reporting in Taiwan (19 counties/cities of the main island) to the commencement of control measures against COVID-19.

Data preprocessing and model implementation

Data collection & normalization

Weekly county/city-level ILI data from the Taiwan CDC website were downloaded. ILI case counts for different age groups and total emergency department visits were aggregated. The ILI_Rate was then calculated using Equation (1).

The ILI_Rate dataset was normalized for each county/city, and the two input adjacency matrices were also normalized.

Sample generation

The dataset includes 624 weeks (12 years) of continuous ILI_Rate monitoring data. The normalized data were split along the time axis into training, validation, and test sets with a ratio of 8:2:2. Specifically, the first 8 years (416 weeks) were used as the training set, the next 2 years (104 weeks) as the validation set, and the last 2 years (104 weeks) as the test set. There was no overlap between the datasets to prevent information leakage.

In each dataset, input sequences and corresponding labels were generated using a sliding window approach. The input sequence length (window size) was set to 52 weeks, and the corresponding label was for 1 week (predicting the next week). The window slid forward by 1 week each time (step size = 1), and the extracted "sequence + label" pair formed one sample. This process resulted in 364 samples for the training set, 52 for the validation set, and 52 for the test set.

Model implementation and parameter setting

Learning Rate: The initial learning rate (LR) was set to 0.003. During training, a LR scheduler dynamically adjusted the LR based on the model's performance. The main parameters of the scheduler were: mode='min' (to minimize the validation loss), factor=0.2 (LR adjustment ratio, i.e., the new LR = previous LR $\times$ 0.2), and patience=5 (the LR would be reduced if there was no improvement in validation performance for 5 consecutive iterations).

Early Stopping Mechanism: To prevent overfitting, an early stopping scheme [48] was employed. During each iteration, the validation loss was monitored to determine when to stop trainings. Specifically, the patience threshold was set to 20. After each iteration, the current validation loss was compared with the previous iteration's loss. If the validation loss decreased, it indicated that the model was improving, and the current model parameters were saved as the best model with counter=0. If the validation loss increased, indicating overfitting risk, the previous best model parameters were retained, and the counter was incremented. When the counter exceeded the patience threshold after several consecutive iterations with increased validation loss, training stopped, and the previously saved best model was used for testing.

Other Hyperparameters: The number of spatiotemporal convolution blocks was set to 2, which provided the best performance. The batch size was 26, with a maximum of 1000 iterations. For the graph convolution, the number of spatial feature channels was set to 16, and the output feature channels were set to 64.

Ablation study

To investigate the impact of geographical and functional topologies, we conducted a series of ablation experiments introducing variations of our model. These variations include a model utilizing only geographical topology, referred to as Geo-STGCN, and a model focusing on functional connectivity, denoted as Func-STGCN. Each variant aims to assess the impact of specific topological information on the predictive performance. In these experiments, separate models were trained using either geographical or functional topology, and their performance was compared to our comprehensive Dual-Topo-STGCN model, which integrates features from both topologies.

Models for comparison

We conducted comparative analyses by comparing the proposed model to various well-established prediction approaches. To ensure the fairness of the comparison, consistent data preprocessing procedures were employed for all considered models. These procedures included utilizing identical dataset partitioning methods and ensuring uniform sliding window sizes.

Autoregressive integrated moving average (ARIMA)

The most commonly used ARIMA [49] model with parameters p , d , and q was considered as the first baseline. In practice, the differencing order d was determined through the Augmented Dickey-Fuller (ADF) [50] test, and the optimal p and q values were selected by minimizing the corrected Akaike Information Criterion (AIC) [51] in a stepwise method. Once the model was fitted with these parameters, we applied it for nowcasting.

Long short-term memory (LSTM)

Influenza surveillance data commonly exhibit seasonal and periodic trends [52]. Recurrent Neural Networks (RNNs) establish iterative relationships between hidden layers at different time steps, effectively capturing influenza's temporal dynamics. LSTM networks [53], as a variant of RNNs, excel in influenza prediction by preserving short-term and long-term memory chains through interactive updates.

Convolutional long short-term memory (Conv-LSTM)

Conv-LSTM [54] is a hybrid architecture that combines the strengths of Convolutional Neural Networks (CNNs) and LSTM networks. In this framework, each LSTM cell is fed with input represented as images or two-dimensional matrices encompassing spatial features. The integration of spatial information derived through

convolutional operations with temporal sequence data enables the model to capture both spatial and temporal dependencies.

Graph attention network (GAT)

Graph Attention Networks (GAT) [55] leverage the self-attention mechanism to assign different importance weights to each neighboring node, enabling the model to capture complex, heterogeneous relationships in graph-structured data. Unlike traditional Graph Neural Networks, GAT dynamically computes attention coefficients, allowing the model to focus more on relevant neighbors and effectively handle irregular graph structures. For influenza prediction task, GAT can model spatial dependencies between regions or cities.

Evaluation metrics

After applying the aforementioned models, we obtained the county/city level predictions for the last 52 weeks (from week 11 in 2019 to week 10 in 2020) in the test set. To evaluate the predictive performance of each model, four metrics were employed: the root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE), and Pearson correlation coefficient (Corr).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\%$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

$$Corr = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

where $\hat{y}_i$ and y_i indicate the value of predicted ILI_Rate and actual ILI_Rate respectively.

Results

We evaluate the predictive performance of each model using previously mentioned various metrics. For the purpose of fair comparison, we calculate the average values of these metrics for each model across the 19 counties/cities. Additionally, we analyze the prediction results in more granular time periods to compare their performance in these meaningful and distinct intervals. The results are summarized in Table 1.

Table 1 Aggregated nowcasting results for 19 counties/cities on average in different time periods

Time Period / Model	Corr	RMSE	MAE	MAPE
Entire Time Period (week 1 - 52)				
ARIMA	0.7281	0.0019	0.0013	0.1017
LSTM	0.7932	0.0018	0.0013	0.1019
Conv-LSTM	0.7678	0.0019	0.0013	0.1004
GAT	0.5994	0.0026	0.0020	0.1565
Dual-Topo-STGCN	0.8202	0.0017	0.0013	0.0966
Non-Flu Season (week 1 - 35)				
ARIMA	0.5302	0.0013	0.0010	0.0785
LSTM	0.6793	0.0012	0.0010	0.0752
Conv-LSTM	0.6756	0.0011	0.0009	0.0696
GAT	0.3809	0.0020	0.0016	0.1276
Dual-Topo-STGCN	0.6569	0.0012	0.0010	0.0793
Flu Season (week 36 - 50)				
ARIMA	0.6336	0.0028	0.0022	0.1423
LSTM	0.7086	0.0027	0.0021	0.1436
Conv-LSTM	0.7324	0.0029	0.0023	0.1467
GAT	0.6181	0.0035	0.0029	0.1968
Dual-Topo-STGCN	0.8024	0.0024	0.0019	0.1294

Model performances during the entire time period

The “Entire Time Period” section in Table 1 presents the predicted results of our model (Dual-Topo-STGCN) and baselines over the continuous 52 weeks. It can be easily observed that our model exhibits outstanding performance across all evaluation metrics, with an average Corr of 0.8202, RMSE of 0.0017, MAE of 0.0013, and MAPE of 0.0966. It is worth mentioning that the attention mechanism is considered effective in capturing spatial dependencies between nodes, so the prediction performance of GAT is promising. However, the results show that its performance even falls short of traditional models, and it is outperformed by ARIMA. The classical deep learning algorithms LSTM and Conv-LSTM demonstrate relatively similar predictive performance. This suggests that gridded spatial information, primarily geographical aspect, has a limited effect on influenza transmission. It appears that LSTM and Conv-LSTM are not significantly worse than our Dual-Topo-STGCN in average performance; however, in situations where influenza spreads rapidly, their performances become inadequate.

Model Performances during the Flu Season

According to the introduction of Taiwan CDC [56], influenza prevalence in Taiwan typically shows a noticeable temporal pattern, which begins in late November, reaching its peak at the end of the year or the start of the following year, and stabilizing in February and March.

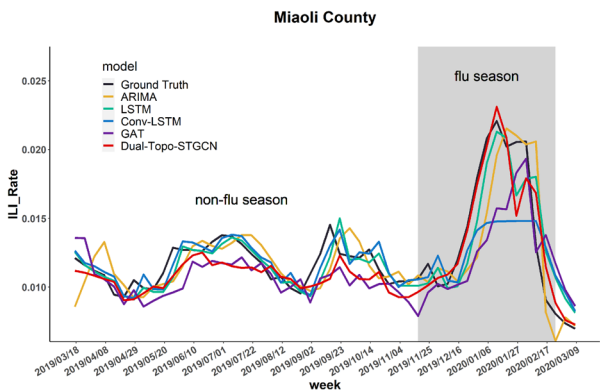


Fig. 3 Comparison of prediction by different models in Miaoli County (Shaded area indicates flu season, our model captures and responds to data fluctuations most rapidly and accurately during the surge of influenza cases)

This period is referred to as the flu season each year. We extract the test results of this period (week 36–50) to calculate the aforementioned metrics, and the results are presented in the “Flu Season” section of Table 1. It can be observed that, in comparison to the baseline models, our model exhibits the best predictive performance during the flu season by reaching the highest Corr and the lowest RMSE, MAE and MAPE on average. This superior performance across all evaluation metrics demonstrates the effectiveness of our model in accurately forecasting flu trends.

Similarly, from the visualization of different models’ prediction curves (example of Miaoli County as shown in Fig. 3), we can distinctly observe that as the flu season approaches, the number of cases rapidly increases. Conventional models heavily rely on historical data from previous weeks to make predictions, which leads to an apparent delay in responding to the rapid fluctuations of data. In contrast, our model demonstrates a remarkable ability to capture the early-stage upward trends of influenza prevalence and promptly respond to such rising trends, surpassing all baseline models by a significant margin. This highlights its relevance in addressing emerging phenomena. All 19 cities’ visualization in flu season is shown in Fig. 4.

Model Performances during the Non-Flu Season

In contrast to the flu season, the non-seasonal influenza season is characterized by a concealed state of influenza prevalence with a small number of cases. From the comparison of prediction curves in “non-flu season” period in Fig. 3, the curves are intertwined, making it difficult to determine which model performs better from visual judgement. This intertwined situation is also evident in other cities. We calculated the evaluation metrics during

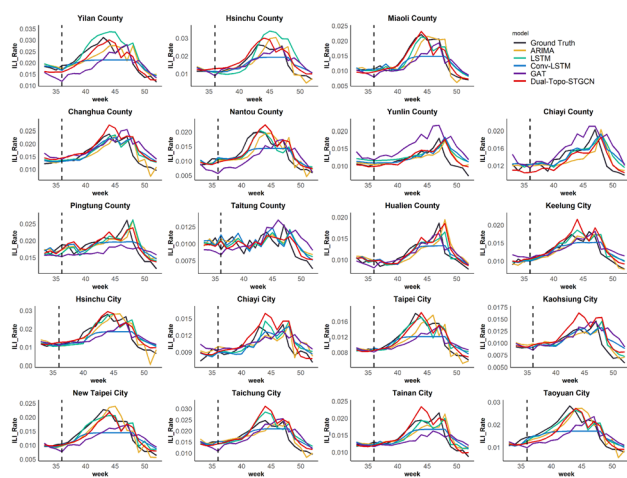


Fig. 4 Visualization of the prediction results of 19 counties/cities in Flu Season. The dashed line represents the boundary between the non-flu season and the flu season

the non-flu season (weeks 1–35), and the results are presented in the “Non-Flu Season” section of Table 1. Contrary to the flu season, many conventional models, such as the commonly used ARIMA, demonstrate relatively good predictive performance during the non-flu season. This may be because the data fluctuations are smaller during the non-flu season, benefiting models that rely on historical data averages for predictions. However, such smoothing is ineffective in responding to urgent situations of rapid influenza transmission. Cities with a larger short-term increase in ILI_Rate, as shown in Fig. 4, are more likely to experience excessively smooth predictions from Conv-LSTM, demonstrating this point.

Ablation results

Table 2 provides a comparative analysis of predictive outcomes for several cities in the ablation experiments during the entire time period. It can be noted that for nearly all cities, the performance of single-topology model utilizing functional topology is consistently better than that of model using geographical topology. This underscores the significant influence of functional connectivity on the transmission of influenza.

However, this better-performed model still exhibits relatively high levels of errors and a low Corr overall, with no significant advantage over the baseline models (average results for Func-STGCN: Corr=0.7345, RMSE=0.0021, MAE=0.0016, MAPE=0.1310). When combining both topologies, the performance of the combined model (Dual-Topo-STGCN) was significantly improved. Through the integration of these two types of spatial features and the implementation of thorough information fusion, the Dual-Topo-STGCN model

achieves a noteworthy reduction of errors and enhancement of Corr.

External validation

To verify the generalization ability of the model in other regions, we applied our model to the ILI_Rate dataset [57] from the United States and compared it with other models. The CDC’s influenza surveillance network divides the 48 contiguous states of the U.S. into 10 regions based on geographic areas. After obtaining and organizing the regional-level ILI_Rate data, we selected data from Week 40 of 2004 to Week 11 of 2020, covering a total of 15.5 years (806 weeks). The data was divided along the timeline into training, validation, and test sets in a ratio of 11.5:2:2. All other operations were consistent with the data from Taiwan province. The average results of various metrics in the 10 regions are shown in Table 3.

From Table 3, we can observe that our model achieves optimal results across all evaluation metrics on the test set, whether for the entire year or during the flu season and non-flu season. This suggests that its predictive performance outperforms the other baseline models.

To visually observe the performance differences among models, we selected the prediction results of Region 2, which has the highest population density, for visualization. As shown in Fig. 5, it can be observed that, during the non-flu season, the performance differences among the models are not very noticeable. However, when the flu season arrives, ARIMA, LSTM, and our Dual-Topo-STGCN are all able to sensitively capture the upward trend in the data and respond accordingly. LSTM loses its ability to accurately predict after following the first small peak, while ARIMA’s ability to identify peaks is somewhat weaker, often overestimating the peak value. In contrast, our model not only stays closely aligned with the ground truth throughout, but also demonstrates stronger responsiveness and adaptability to complex data fluctuations. This robustness in continuous prediction highlights the model’s ability to learn sufficiently rich information from historical data and spatial patterns.

Discussion

Functional topology and its geographical relevance

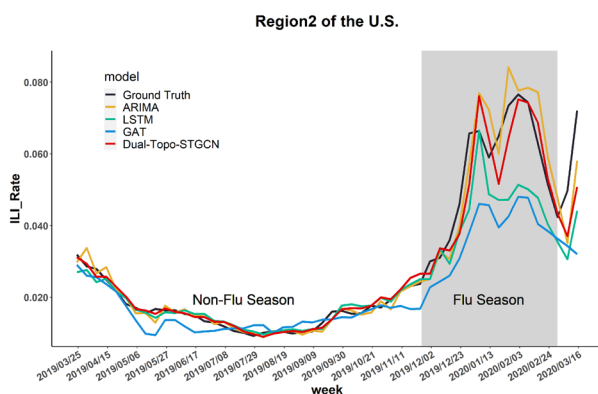
In this study, it is validated that the influence of functional socio-economic factors on influenza prevalence is rooted in and extends beyond geographical features. By visualizing the functional topology structure, we can observe information consistent with the geographical topology, as depicted in Fig. 6. Specifically, a high degree of connectivity among adjacent counties/cities exists in the geographical topology (Fig. 6a), while this close interrelationship among local regions is similarly evident in the functional topology (Fig. 6b). This indicates that

Table 2 Comparison of model performances across multiple cities considering various topological information

Metric	Model	Keelung City	Hsinchu City	Taoyuan City	Taipei City	New Taipei City	Yilan County	Miaoli County	Nantou County	Pingtung County
Corr	Geo-STGCN	0.7811	0.8050	0.8343	0.8819	0.8414	0.7903	0.8325	0.6234	0.6730
	Func-STGCN	0.7550	0.8137	0.8499	0.8559	0.8742	0.8419	0.8577	0.8376	0.7513
	Dual-Topo-STGCN	0.8492	0.9043	0.8842	0.9147	0.9180	0.8969	0.9057	0.8744	0.7976
RMSE	Geo-STGCN	0.0021	0.0027	0.0023	0.0012	0.0019	0.0034	0.0019	0.0027	0.0024
	Func-STGCN	0.0020	0.0027	0.0022	0.0014	0.0017	0.0026	0.0017	0.0022	0.0017
	Dual-Topo-STGCN	0.0015	0.0019	0.0020	0.0010	0.0014	0.0022	0.0015	0.0016	0.0016
MAE	Geo-STGCN	0.0015	0.0018	0.0016	0.0009	0.0014	0.0031	0.0015	0.0018	0.0020
	Func-STGCN	0.0014	0.0019	0.0016	0.0011	0.0013	0.0023	0.0012	0.0017	0.0013
	Dual-Topo-STGCN	0.0011	0.0014	0.0015	0.0008	0.0010	0.0019	0.0011	0.0012	0.0012
MAPE	Geo-STGCN	0.1344	0.1427	0.1205	0.0931	0.1205	0.1590	0.1352	0.1550	0.1091
	Func-STGCN	0.1255	0.1516	0.1213	0.1185	0.1111	0.1143	0.1063	0.1724	0.0717
	Dual-Topo-STGCN	0.0923	0.1044	0.0977	0.0758	0.0777	0.0965	0.0896	0.1108	0.0666

Table 3 Aggregated Nowcasting Results for the 10 Regions of the U.S. in Different Time Periods

Time Period / Model	Corr	RMSE	MAE	MAPE
Entire Time Period (week 1 - 52)				
ARIMA	0.9497	0.0068	0.0037	0.1184
LSTM	0.9606	0.0073	0.0045	0.1579
GAT	0.9436	0.0115	0.0080	0.3036
Dual-Topo-STGCN	0.9688	0.0053	0.0032	0.1028
Non-Flu Season (week 1 - 35)				
ARIMA	0.9657	0.0023	0.0015	0.1023
LSTM	0.9715	0.0022	0.0017	0.1407
GAT	0.8792	0.0052	0.0040	0.3034
Dual-Topo-STGCN	0.9796	0.0018	0.0013	0.0889
Flu Season (week 36 - 50)				
ARIMA	0.7686	0.0116	0.0085	0.1569
LSTM	0.7974	0.0120	0.0099	0.1860
GAT	0.8173	0.0175	0.0153	0.2906
Dual-Topo-STGCN	0.8332	0.0085	0.0067	0.1271

**Fig. 5** Comparison of prediction by different models in Region2 of the U.S. (Shaded area indicates the flu season). Compared to the baseline models, our model consistently stays close to the Ground Truth

the functional topology encapsulates key geographical information.

On the other hand, the interconnections between counties/cities within the functional topology correspond well to real geographical characteristics, population, economy, and other factors. For example, population density is one of the most significant factors influencing the transmission of influenza. Regarding population distribution, Taiwan has a predominantly concentrated population on its western side and sparser population in the central and eastern regions due to the presence of numerous mountains (Fig. 6c). In the functional topology, denser connections and co-occurrences

are observed on the western side of the Taiwan island. Conversely, the four eastern counties - Yilan, Hualien, Taitung, and Pingtung - with sparse population and geographical isolation, exhibit very limited interactions with each other and with other counties/cities. This intercity connectivity precisely corresponds to the population distribution pattern. These findings indicate that the impact of functional socio-economic factors is consistent with geographical characteristics.

Furthermore, from the results in Table 2, we can observe that the spatial patterns depicted by the functional topology play a crucial role in improving the performance of the influenza prediction model. In fact, this effect is not limited to our proposed Dual-Topo-STGCN model. In the GAT model used for comparison, we also attempted to replace the input geographical topology with the functional topology. The prediction results showed a significant improvement in GAT's performance, which further supports the view that "functional topology can provide more useful information".

Synergizing geographical and functional spatial features

The individual interactions of temporal-geographical or temporal-functional spatial features were not able to effectively enhance the predictive performance of the model. Spatial patterns revealed by a single topology tend to be one-sided, resulting in an inadequate characterization of real-world spatial features and an insufficient clarification of spatial relationships between regions. Furthermore, during the extraction of spatiotemporal interaction features, the presence of non-patterned and irrelevant information in the single topology can contaminate the extracted features, thereby weakening the model's predictive capability. The challenges faced by the single-topology model lie in its struggle to distinguish between true spatiotemporal patterns and other irrelevant information, leading to suboptimal predictive performance.

In fact, considering only a single spatial pattern is undeniably narrow, especially in the geographical aspect. This limitation is why the suboptimal outcomes were observed in many studies that solely focus on geographical spatial features and apply convolutional operations [16, 21]. The improvements in their model's performance are primarily driven by innovations in exploring temporal patterns, with spatial features playing a secondary role. Moreover, the highly anticipated combination of the (multi-head) attention mechanism and GNNs-specifically the GAT model-has failed to show competitive results across multiple datasets. One reason for this is that it does not explicitly handle the temporal dependencies within the features. More importantly, it only considers a single type

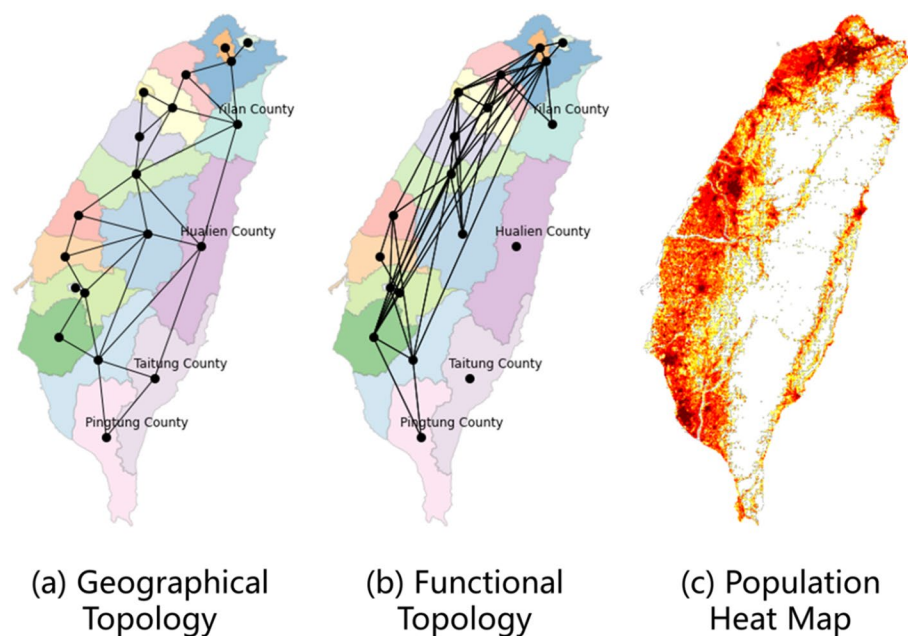


Fig. 6 Illustration of the two types of topologies. The geographical topology focuses more on depicting the adjacency relationships between counties/cities (a), while the functional topology illustrates the functional connections between counties/cities (b), which more closely align with the actual population distribution (c)

of spatial relationship between nodes, without accounting for multiple spatial patterns simultaneously.

However, in our proposed dual-topo model, the fusion of two types of spatiotemporal interaction features enables us to leverage the synergistic or complementary nature of these features. This integration helps mitigate the impact of unrelated variations by capturing more robust and meaningful patterns, ultimately enhancing the model's predictive accuracy. The results of the ablation experiments shown in Table 2 perfectly validate this statement.

Additionally, through our exploration of geographical topology, we discovered that introducing the surrounding background area, such as the ocean or open land, as a new node to delineate the positional relationships of the target nodes, led to a significant improvement in model performance. The same performance improvement was also evident in Geo-STGCN and even Dual-Topo-STGCN. Before introducing the background node, the average performance metrics of Dual-Topo-STGCN across the 19 counties/cities were: Corr = 0.6849, RMSE = 0.0022, MAE = 0.0017, MAPE = 0.1362. After the introduction, they improved to: Corr = 0.8202, RMSE = 0.0017, MAE = 0.0013, MAPE = 0.0966, showing a significant performance boost. This exciting discovery suggests that exploring spatial features introduces a new perspective, namely, seeking background information beyond the

target nodes to assist and enhance the characterization of spatial relationships within the target area.

Critical timeliness in flu season response

By forecasting influenza outbreaks with high accuracy and precision, the model enables timely and effective intervention strategies, optimizing resource allocation and mitigating cross-regional transmission risks. During the non-flu season, certain conventional models exhibit considerable predictive performance with minimal variation. Compared to the flu season, they consistently exhibited significantly lower errors. The reason is that conventional models can conveniently generate stable and balanced predictions based on relatively unchanged historical observations when the flu tends to remain latent or spreads at a low level.

However, we should be cautious of the inertia caused by excessive stability and balance. It is more important to focus on early warnings for the onset of influenza during the flu season. With the flu season approaching, there will be a rapid surge in cases within a short period, thereby imposing a quick and substantial impact and threat to public health. Therefore, if we can make accurate and reliable predictions, we will be better equipped to handle the epidemic calmly. Additionally, having a precise determination of the peak timing will enable us to regulate the allocation of healthcare resources effectively, thereby averting unnecessary waste.

Achieving the objective of accurate prediction is not easily accomplished by conventional models, as they exhibit limited sensitivity to the fluctuating variations in data. Conversely, our model performed exceptionally well for this task. It demonstrated rapid responsiveness during the initial stages of the influenza outbreak, maintaining close conformity to the ground truth in its predictions. Moreover, it effectively captured the subsequent declining trend following the peak.

Model complexity, robustness, and application potential

For a predictive model, it is crucial to consider both the model's complexity and robustness. While keeping the size of the validation and test sets constant, the U.S. dataset has more training samples and fewer nodes compared to the Taiwan dataset. Under the same preprocessing and model architecture, the prediction results of the GAT model on these two datasets are strikingly different: it performs relatively well on the U.S. dataset but performs poorly on the Taiwan dataset (as shown in Tables 1, 3 and Figs. 3, 5). This is because GAT uses multi-head attention, which not only increases model complexity but also introduces greater randomness due to multiple randomly initialized weight matrices. As a result, it requires a sufficiently large training sample size, and the increase in the number of nodes further amplifies the need for more samples. In contrast, our Dual-Topo-STGCN model has a simpler and more flexible structure, requiring fewer samples while being capable of simultaneously capturing temporal dependencies and various complex spatial patterns. It achieves very stable prediction results on both datasets.

This highlights its potential for applications: in regions where infectious disease surveillance systems are established later and available historical data is limited, leading to fewer training samples, our model can provide reliable predictions. Moreover, its computational overhead is very low (training typically completes within a few minutes on an A6000 GPU). The lightweight nature of the model and its low computational resource requirements also facilitate deployment in underdeveloped regions.

Conclusion

In this study, we propose a novel approach based on graph neural networks with dual topology for spatiotemporal ILI Rate prediction. Our model is independent of external variables, features a simple architecture with high computational efficiency, and demonstrates strong performance. We comprehensively explore the various factors influencing influenza transmission among cities, including temporal effects, geographical relationships, and functional topological connectivity. By employing two parallel spatiotemporal interaction pathways, our

model extracts valuable information from historical data and spatial features. The subsequent feature fusion yields integrated high-level features, leading to a significant improvement in predictive performance.

Our approach shows advantage in promptly detecting upward trends during the critical initial stage of flu season, outperforming all the compared methods. The results are of significant importance for health decisions, individual protection, and maintaining public health. By accurately forecasting influenza trends, our model can assist in timely deployment of healthcare resources, thereby mitigating the impact of flu outbreaks. Furthermore, accurately assessing the post-peak decline can minimize healthcare resource cost, avoid excessive preventive measures, and expedite the return to normalcy for both society and individuals. Moreover, the robustness and flexibility of our model make it especially valuable in regions with limited historical data, ensuring reliable predictions even in challenging scenarios.

Although our model has shown promising results, it is important to acknowledge its limitations. Currently, the model relies solely on ILI data and two topologies, without considering external variables such as transportation data, passenger flow, and population dynamics, which could enhance its predictive power. In the future, we plan to refine the model by incorporating additional data sources, exploring more complex topologies, and enabling real-time updates to enhance its utility in public health decision-making. To further improve scalability and accuracy, we will apply advanced optimization techniques such as adaptive learning and reinforcement learning. Additionally, we aim to enhance the model's interpretability through visualization tools and feature importance analysis, thereby increasing trust and usability in real-world healthcare applications. By addressing these limitations, we aim to further improve the model's robustness and applicability to various infectious diseases in dynamic settings.

Acknowledgements

Supported by High-performance Computing Public Platform (Shenzhen Campus) of SUN YAT-SEN UNIVERSITY.

Authors' contributions

JL, XW, XF, YH, XD, YC and YZ co-conceived the study. YZ secured funding. The investigation was performed by JL, XW, XF, YH, XD, YC. YZ supervised study conduct. Analysis was performed by JL, XW, XF, and supervised by all other authors. JL and XF drafted the manuscript, and all other authors assisted in writing. The final manuscript was approved by all authors.

Funding

This work was supported in part by the National Key Research and Development Program of China under grant number 2023YFC2307305, and the Shen Zhen-Hong Kong-Macao Science and Technology Project Fund under grant number SGDX20210823104003030 and SGDX20210823103403028.

Data availability

The datasets generated and/or analysed during the current study are available in the Taiwan CDC Open Data Portal, <https://data.cdc.gov.tw/en/dataset/>

hi-outpatient-emergency-visit-influenza-like-illness, and the U.S. CDC Open Data Portal, <https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html>.

Declarations

Ethics approval and consent to participate

This study did not involve human participants, data, or tissue. It was conducted using only aggregated and anonymized data. Institutional review board approval was not required.

Consent for publication

Not applicable.

Competing interest

The authors declare no competing interests.

Author details

¹School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen 518107, Guangdong, China. ²College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518118, Guangdong, China. ³College of Urban Transportation and Logistics, Shenzhen Technology University, Shenzhen 518118, Guangdong, China.

Received: 25 March 2024 Accepted: 24 January 2025

Published online: 01 February 2025

References

- World Health Organization. Influenza (Seasonal). 2023 [cited 2024 Nov 27]. [https://www.who.int/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)).
- de Francisco N, Donadel M, Jit M, Hutubessy R. A systematic review of the social and economic burden of influenza in low-and middle-income countries. *Vaccine*. 2015;33(48):6537–44.
- Putri WC, Muscatello DJ, Stockwell MS, Newall AT. Economic burden of seasonal influenza in the United States. *Vaccine*. 2018;36(27):3960–6.
- Macias AE, McElhaney JE, Chaves SS, Nealon J, Nunes MC, Samson SI, et al. The disease burden of influenza beyond respiratory illness. *Vaccine*. 2021;39:A6–14.
- Hung SK, Wu CC, Singh A, Li JH, Lee C, Chou EH, et al. Developing and validating clinical features-based machine learning algorithms to predict influenza infection in influenza-like illness patients. *Biomed J*. 2023;46(5):100561.
- Cheng HY, Wu YC, Lin MH, Liu YL, Tsai YY, Wu JH, et al. Applying machine learning models with an ensemble approach for accurate real-time influenza forecasting in Taiwan: Development and validation study. *J Med Internet Res*. 2020;22(8):e15394.
- Woo H, Cho Y, Shim E, Lee JK, Lee CG, Kim SH. Estimating influenza outbreaks using both search engine query data and social media data in South Korea. *J Med Internet Res*. 2016;18(7):e177.
- He Z, Tao H. Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study. *Int J Infect Dis*. 2018;74:61–70.
- Rutland BE, Weese JS, Bolin C, Au J, Malani AN. Human-to-dog transmission of methicillin-resistant *Staphylococcus aureus*. *Emerg Infect Dis*. 2009;15(8):1328.
- Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. *Science*. 2014;343(6176):1203–5.
- Newbold P. ARIMA model building and the time series analysis approach to forecasting. *J Forecast*. 1983;2(1):23–35.
- Wang Z, Chakraborty P, Mekaru SR, Brownstein JS, Ye J, Ramakrishnan N. Dynamic poisson autoregression for influenza-like-illness case count prediction. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM (Association for Computing Machinery); 2015. p. 1285–94.
- Fatima U, Hina S, Wasif M. A novel global clustering coefficient-dependent degree centrality (GCCDC) metric for large network analysis using real-world datasets. *J Comput Sci*. 2023;70:102008.
- Fatima U, Hina S. Efficient algorithm for maximal clique size evaluation. *Int J Adv Comput Sci Appl*. 2019;10(7):444–52.
- Yang L, Yang J, He Y, Zhang M, Han X, Hu X, et al. Enhancing infectious diseases early warning: A deep learning approach for influenza surveillance in China. *Prev Med Rep*. 2024;43:102761.
- He Y, Zhao Y, Chen Y, Yuan HY, Tsui KL. Nowcasting influenza-like illness (ILI) via a deep learning approach using google search data: An empirical study on Taiwan ILI. *Int J Intell Syst*. 2022;37(3):2648–74.
- Zhu X, Fu B, Yang Y, Ma Y, Hao J, Chen S, et al. Attention-based recurrent neural network for influenza epidemic prediction. *BMC Bioinformatics*. 2019;20(18):1–10.
- Kara A. Multi-step influenza outbreak forecasting using deep LSTM network and genetic algorithm. *Expert Syst Appl*. 2021;180:115153.
- Zhu H, Chen S, Lu W, Chen K, Feng Y, Xie Z, et al. Study on the influence of meteorological factors on influenza in different regions and predictions based on an LSTM algorithm. *BMC Public Health*. 2022;22(1):1–17.
- Jiang P. Nowcasting influenza using Google flu trend and deep learning model. In: Proceedings of the 2020 2nd International Conference on Economic Management and Cultural Industry (ICEMCI 2020). Paris: Atlantis Press; 2020. p. 407–16.
- Xi G, Yin L, Li Y, Mei S. A deep residual network integrating spatial-temporal properties to predict influenza trends at an intra-urban scale. In: Proceedings of the 2nd ACM SIGSPATIAL international workshop on AI for geographic knowledge discovery. New York: ACM (Association for Computing Machinery); 2018. p. 19–28.
- Athanasios M, Fragkizidis G, Zarkogianni K, Nikita KS. Long Short-term Memory-Based Prediction of the Spread of Influenza-Like Illness Leveraging Surveillance, Weather, and Twitter Data: Model Development and Validation. *J Med Internet Res*. 2023;25:e42519.
- Wu Z, Pan S, Chen F, Long G, Zhang C, Philip SY. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst*. 2020;32(1):4–24.
- Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput*. 2019;31(7):1235–70.
- Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Trans Neural Netw Learn Syst*. 2021;33(12):6999–7019.
- Wu Y, Fu Y, Xu J, Yin H, Zhou Q, Liu D. Heterogeneous question answering community detection based on graph neural network. *Inf Sci*. 2023;621:652–71.
- Casas S, Gulino C, Liao R, Urtasun R. Spagann: Spatially-aware graph neural networks for relational behavior forecasting from sensor data. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE; 2020. p. 9491–7.
- Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, et al. Robust deep learning-based protein sequence design using ProteinMPNN. *Science*. 2022;378(6615):49–56.
- Gao C, Wang X, He X, Li Y. Graph neural networks for recommender system. In: Proceedings of the fifteenth ACM international conference on web search and data mining. New York: ACM (Association for Computing Machinery); 2022. p. 1623–5.
- Bui KHN, Cho J, Yi H. Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues. *Appl Intell*. 2022;52(3):2763–74.
- Wu L, Chen Y, Shen K, Guo X, Gao H, Li S, et al. Graph neural networks for natural language processing: A survey. *Found Trends Mach Learn*. 2023;16(2):119–328.
- Jung S, Moon J, Park S, Hwang E. Self-attention-based deep learning network for regional influenza forecasting. *IEEE J Biomed Health Inform*. 2021;26(2):922–33.
- Moon J, Jung S, Park S, Hwang E. RESEAT: Recurrent Self-Attention Network for Multi-Regional Influenza Forecasting. *IEEE J Biomed Health Inform*. 2023;27(5):2585–96.
- Eggo RM, Cauchemez S, Ferguson NM. Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. *J R Soc Interface*. 2011;8(55):233–43.
- Miller HJ. Tobler's first law and spatial analysis. *Ann Assoc Am Geogr*. 2004;94(2):284–9.
- Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L, et al. Human mobility and the spatial transmission of influenza in the United States. *PLoS Comput Biol*. 2017;13(2):e1005382.

37. Yang JR, Hsu SZ, Kuo CY, Huang HY, Huang TY, Wang HC, et al. An epidemic surge of influenza A (H3N2) virus at the end of the 2016–2017 season in Taiwan with an increased viral genetic heterogeneity. *J Clin Virol*. 2018;99:15–21.
38. Shu YL, Fang LQ, de Vlas SJ, Gao Y, Richardus JH, Cao WC. Dual seasonal patterns for influenza, China. *Emerg Infect Dis*. 2010;16(4):725.
39. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457(7232):1012–4.
40. Liang F, Guan P, Wu W, Huang D. Forecasting influenza epidemics by integrating internet search queries and traditional surveillance data with the support vector machine regression model in Liaoning, from 2011 to 2015. *PeerJ*. 2018;6:e5134.
41. Poirier C, Hsuen Y, Bouzillé G, Cuggia M, Lavenu A, Brownstein JS, et al. Influenza forecasting for French regions combining EHR, web and climatic data sources with a machine learning ensemble approach. *PLoS ONE*. 2021;16(5):e0250890.
42. Wang R, Wu H, Wu Y, Zheng J, Li Y. Improving influenza surveillance based on multi-granularity deep spatiotemporal neural network. *Comput Biol Med*. 2021;134:104482.
43. Kamran H, Aleman DM, Carter MW, Moore KM. Spatio-Temporal Clustering of Multi-Location Time Series to Model Seasonal Influenza Spread. *IEEE J Biomed Health Inform*. 2023;27(4):2138–48.
44. Zhou X, Yang F, Feng Y, Li Q, Tang F, Hu S, et al. A spatial-temporal method to detect global influenza epidemics using heterogeneous data collected from the Internet. *IEEE/ACM Trans Comput Biol Bioinforma*. 2017;15(3):802–12.
45. Guo X, Xiong NN, Wang H, Ren J. Design and analysis of a prediction system about influenza-like illness from the latent temporal and spatial information. *IEEE Trans Syst Man Cybern Syst*. 2021;52(1):66–77.
46. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health*. 2020;2(2):e85–93.
47. Taiwan Centers for Disease Control. Health insurance outpatient and emergency visits - influenza. Taiwan CDC. [cited 2024 Nov 27]. Available from: <https://data.cdc.gov.tw/en/dataset/hi-outpatient-emergency-visit-influenza-like-illness>.
48. Rice L, Wong E, Kolter Z. Overfitting in adversarially robust deep learning. In: *Proceedings of the International Conference on Machine Learning*. Cheltenham: PMLR; 2020. p. 8093–104.
49. Box GE, Jenkins GM, Reinsel GC, Ljung GM. *Time series analysis: forecasting and control*. Hoboken: Wiley; 2015.
50. Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc*. 1979;74(366a):427–31.
51. Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control*. 1974;19(6):716–23.
52. Lai G, Chang WC, Yang Y, Liu H. Modeling long-and short-term temporal patterns with deep neural networks. In: *The 41st international ACM SIGIR conference on research & development in information retrieval*. 2018. pp. 95–104.
53. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735–80.
54. Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: *Advances in Neural Information Processing Systems 28 (NeurIPS 2015)*. Red Hook: Curran Associates, Inc.; 2015. p. 802–10.
55. Velickovic P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y, et al. Graph attention networks. *Stat*. 2017;1050(20):10–48550.
56. Centers for Disease Control and Prevention. Vaccine (influenza). Updated 2024 Jun 3 [cited 2024 Nov 27]. <https://www.cdc.gov.tw/Category/Page/LnrbufSY9uTylprhOZ15NA>.
57. Centers for Disease Control and Prevention (CDC). National, regional, and state level outpatient illness and viral surveillance. CDC. [cited 2024 Nov 27]. Available from: <https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.