International Journal of Geo-Information

MDPI

Article

Mining Spatiotemporal Mobility Patterns Using Improved Deep Time Series Clustering

Ziyi Zhang [1,*], Diya Li [2], Zhe Zhang [1,2] and Nick Duffield [1,3]

[1] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA; zhezhang@tamu.edu (Z.Z.); duffieldng@tamu.edu (N.D.)
[2] Department of Geography, Texas A&M University, College Station, TX 77843, USA; diya.li@tamu.edu
[3] Texas A&M Institute of Data Science, Texas A&M University, College Station, TX 77843, USA
* Correspondence: zyzhang@tamu.edu

Abstract: Mining spatiotemporal mobility patterns is crucial for optimizing urban planning, enhancing transportation systems, and improving public safety by providing useful insights into human movement and behavior over space and time. As an unsupervised learning technique, time series clustering has gained considerable attention due to its efficiency. However, the existing literature has often overlooked the inherent characteristics of mobility data, including high-dimensionality, noise, outliers, and time distortions. This oversight can lead to potentially large computational costs and inaccurate patterns. To address these challenges, this paper proposes a novel neural network-based method integrating temporal autoencoder and dynamic time warping-based K-means clustering algorithm to mutually promote each other for mining spatiotemporal mobility patterns. Comparative results showed that our proposed method outperformed several time series clustering techniques in accurately identifying mobility patterns on both synthetic and real-world data, which provides a reliable foundation for data-driven decision-making. Furthermore, we applied the method to monthly county-level mobility data during the COVID-19 pandemic in the U.S., revealing significant differences in mobility changes between rural and urban areas, as well as the impact of public response and health considerations on mobility patterns.

Keywords: spatiotemporal data mining; mobility patterns; time series clustering; deep learning

Citation: Zhang, Z.; Li, D.; Zhang, Z.; Duffield, N. Mining Spatiotemporal Mobility Patterns Using Improved Deep Time Series Clustering. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 374. https://doi.org/10.3390/ijgi13110374

Academic Editors: Wolfgang Kainz, Lan Mu and Jue Yang

Received: 10 September 2024
Revised: 11 October 2024
Accepted: 22 October 2024
Published: 24 October 2024

1. Introduction

Spatiotemporal mobility pattern mining is a crucial area of data mining that focuses specifically on identifying and analyzing movement patterns within mobility data as they shift over time and across different locations. These patterns have provided useful insights in a wide range of applications, such as urban planing [1–5], traffic management [6–8], and disaster reduction [9–12]. With advancements in data storage and processing capabilities, real-world applications now have the capacity to store and retain data over extended periods [13]. Consequently, mobility data across various domains are increasingly being preserved in the form of time series data, facilitating more comprehensive analysis and long-term insights.

Researchers have developed various methods leveraging time series data to perform related analysis such as classification [14,15], clustering [10,16–19], forecasting [20–24], and causal inference [25–27]. Among these tasks, time series clustering is particularly challenging but important. Time series clustering has gained significant attention for its effectiveness across diverse fields, including climate [28–30], energy [31,32], and the environment [33,34]. Despite the advances in time series clustering, several challenges persist, including high dimensionality, noise, outliers, and time distortions. These issues often lead to increased computational costs and reduced accuracy in clustering outcomes.

An appropriate representation can help mitigate high dimensionality, noise, and outliers, and reduce computational costs, leading to more efficient and accurate time

series clustering. Many studies have been carried out with a focus on dimensionality reduction (DR) techniques. For example, principal component analysis (PCA)-based K-means algorithm [16,35,36] has been proposed to achieve faster and more accurate time series clustering. This method constructs common projection axes as the prototype for each cluster, enabling more efficient clustering by reducing dimensionality while preserving essential data characteristics. Non-negative matrix factorization (NMF) is also employed as a DR technique prior to applying the K-Means clustering algorithm [37,38].

With the advancement of neural networks (NNs), deep representation learning [39] has focused on constructing meaningful and compact data representations. This approach aims to capture essential features from raw data, enabling more efficient processing and improving the performance of downstream tasks such as classification, clustering, and forecasting. Deep time series clustering [40] exemplifies how deep representation learning can be used to analyze time series data. In this approach, temporal representations are first learned through advanced deep representation learning techniques, and then clustering algorithms are applied to temporal representations to extract meaningful clusters. Researchers have devoted considerable effort to improving deep time series clustering. For example, the authors in [40–45] employed the convolutional autoencoder (CAE) for extracting representations and K-means as the clustering algorithm. To better capture temporal information, some studies developed a temporal autoencoder (TAE) by integrating a recurrent neural network (RNN) into the convolutional autoencoder, which resulted in improved performance [46–49].

However, we found that existing methods often overlook several critical characteristics of mobility datasets, making it challenging to apply these methods directly to mobility data. These challenges include determining whether the extracted temporal representations are suitable for downstream clustering algorithms, as well as addressing issues related to high dimensionality, noise, outliers, and time distortions in their methods. In an effort to overcome these obstacles within a comprehensive framework, we introduce a novel method that combines a neural network-based temporal autoencoder with a dynamic time warping-based K-means clustering algorithm. Additionally, we propose an optimization strategy to improve the quality of generated temporal representations, making them more suitable for the downstream K-means clustering algorithm. This approach not only mitigates the impact of high dimensionality, noise, outliers, and time distortions but also ensures that the features extracted are well-suited for specific applications. The efficiency of our proposed algorithm is demonstrated through extensive experiments on synthetic and real-world spatiotemporal mobility data, proving its exceptional ability to precisely identify spatiotemporal mobility patterns. The significant contributions of this work are outlined as follows:

- We formalize the task of mining spatiotemporal mobility patterns using a deep time series clustering approach and introduce a comprehensive framework that integrates a neural network-based temporal autoencoder with a dynamic time warping-based K-means clustering algorithm.
- We propose an optimization strategy to enhance the generated temporal representations, making them more suitable for the downstream K-means clustering algorithm.
- Extensive experiments on both synthetic and real-world mobility datasets demonstrate that our proposed approach outperforms existing time series clustering techniques, especially when dealing with data characterized by high dimensionality, noise, outliers, and time distortions.
- We reveal significant differences in mobility changes between rural and urban areas, as well as the influence of public response and health considerations on mobility in the U.S. by applying our method to mobility data during the COVID-19 pandemic.

We have organized our paper as follows: Section 2 presents the data used in this study and our proposed methodology; Section 3 presents the comparative performance evaluation, module contribution analysis, and a case study on mining spatiotemporal

mobility patterns during the COVID-19 pandemic in the U.S.; Section 4 discusses the limitations and provides directions for future research, while Section 5 concludes the study.

2. Materials and Methods

2.1. Data

We used three distinct real-world time series datasets to thoroughly evaluate the effectiveness of our method in identifying and analyzing spatiotemporal mobility patterns. The first, Melbourne Pedestrian Time Series Data, consists of detailed hourly pedestrian counts collected from 10 distinct locations across Melbourne, Australia, throughout the entirety of 2017. These data originate from an automated system (Figure 1) designed to monitor pedestrian activity in the city, with the goal of providing valuable insights into pedestrian movement patterns. By analyzing these data, urban planners can better understand the flow of people through different parts of the city, enabling more informed decisions about urban infrastructure and planning. The placement of sensors in various locations allows for the analysis of spatial variations in pedestrian traffic, while the continuous data collection across a full year enables a comprehensive temporal analysis, taking into account seasonal changes, holidays, and other factors that influence pedestrian behavior.

Similarly, the second, Chinatown Pedestrian Time Series Data, offers monthly pedestrian counts for Chinatown-Swanston Street (North) in Melbourne, Australia, also for the year 2017. This dataset is categorized based on whether the counts were collected on regular weekdays or weekends, providing a granular view of temporal variations in pedestrian activity. By examining these data, we were able to evaluate our method's capability to recognize not only spatial patterns but also temporal variations, such as differences in pedestrian traffic between weekdays and weekends. This level of detail is crucial for understanding how different times of the week influence foot traffic in specific areas, which is valuable for businesses, city planners, and local authorities.

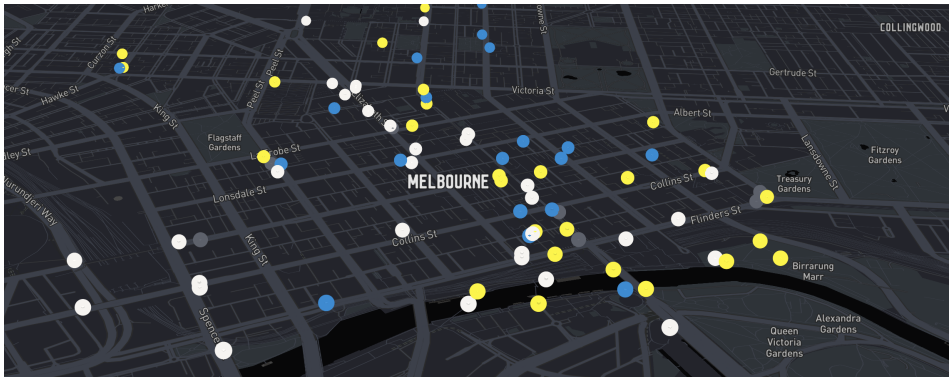

Figure 1. The pedestrian counting system in the city of Melbourne is a network of sensors, represented in various colors, located throughout key areas of the city to count pedestrian movements in real time. This system helps monitor foot traffic trends, providing valuable insights into urban mobility, crowd behavior, and the impact of major events such as festivals and public holidays, as well as changes in mobility patterns.

The third, COVID-19 Mobility Time Series Data, is derived from the Trips by Distance datasets published by the Bureau of Transportation Statistics in the United States. This dataset provides county-level mobility information for the COVID-19 pandemic, tracking the average distance people traveled from their homes across 3142 counties between March 2020 and May 2020. This period is particularly significant as it coincides with the early stages of the pandemic when policies such as lockdowns, social distancing measures, and travel restrictions were being rapidly implemented and modified across the country. The dataset offers a unique opportunity to examine how these dynamic policies influenced human mobility patterns on a large scale. By using these data as a case study, we aimed to identify spatiotemporal mobility patterns under rapidly changing conditions, providing

insights into how public health interventions and shifting regulations impacted movement across different regions. This analysis helps to illustrate the broader applicability of our method to real-world phenomena influenced by both spatial and temporal factors.

In addition to real-world data, we used a synthetic dataset to evaluate the performance of our proposed method in handling noise, outliers, and time distortions, as visualized in Figure 2. We followed the model presented by [50] and then generated these time series with different trends as follows:

$$\begin{aligned} f_1(t) &= (5+\alpha)\cdot g_{[a,b]}(t) + \epsilon, \\ f_2(t) &= (5+\alpha)\cdot g_{[a,b]}(t)\cdot \frac{t-a}{b-a} + \epsilon, \\ f_3(t) &= (5+\alpha)\cdot g_{[a,b]}(t)\cdot \frac{b-t}{b-a} + \epsilon. \end{aligned} \tag{1}$$

In this context, t ranges from 1 to 128, and a is defined as an integer-valued uniform random variable within the range from 16 to 32. The difference $b-a$ follows an integer-valued uniform distribution with bounds from 32 to 96. Both α and ϵ are variables drawn from a standard normal distribution $\mathcal{N}(0,1)$. Additionally, $g_{[a,b]}$ represents the indicator function defined over the interval $[a,b]$, which can be expressed as follows:

$$g_{[a,b]}(t) = \begin{cases} 1 & a \le t \le b \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $f_1(t), f_2(t)$, and $f_3(t)$ represent the generation functions for the time series with different trends.

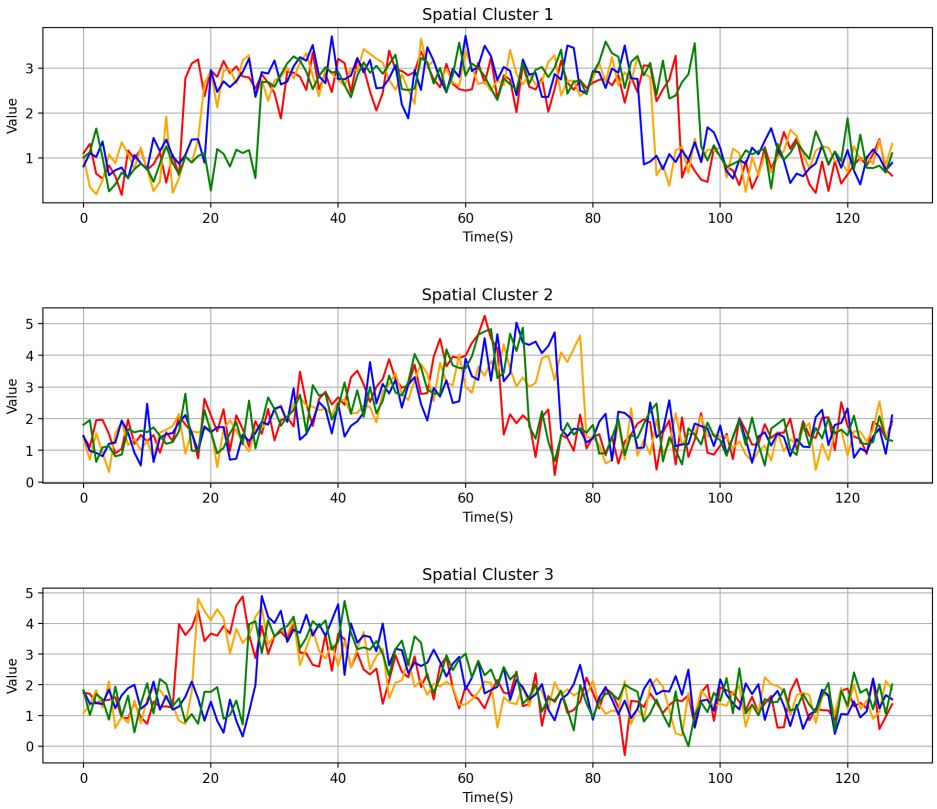

Figure 2. This synthetic time series data includes noise, outliers, and time distortions. We assume the presence of three distinct spatial clusters. Locations within the same spatial cluster, represented by different colors, typically exhibit similar patterns or trends in time series data.

2.2. Methods

This section details our two-stage, improved deep time series clustering (I-DTSC) method for uncovering spatiotemporal mobility patterns. The first stage is a neural network-based temporal autoencoder for learning the temporal representations. The second stage is to utilize a dynamic time warping-based K-means clustering algorithm on learned temporal representations to generate k spatial clusters. After generating k spatial clusters, we visualize them on a map and explore the spatiotemporal patterns. The framework is illustrated in Figure 3, and we discuss our proposed method in the following parts.

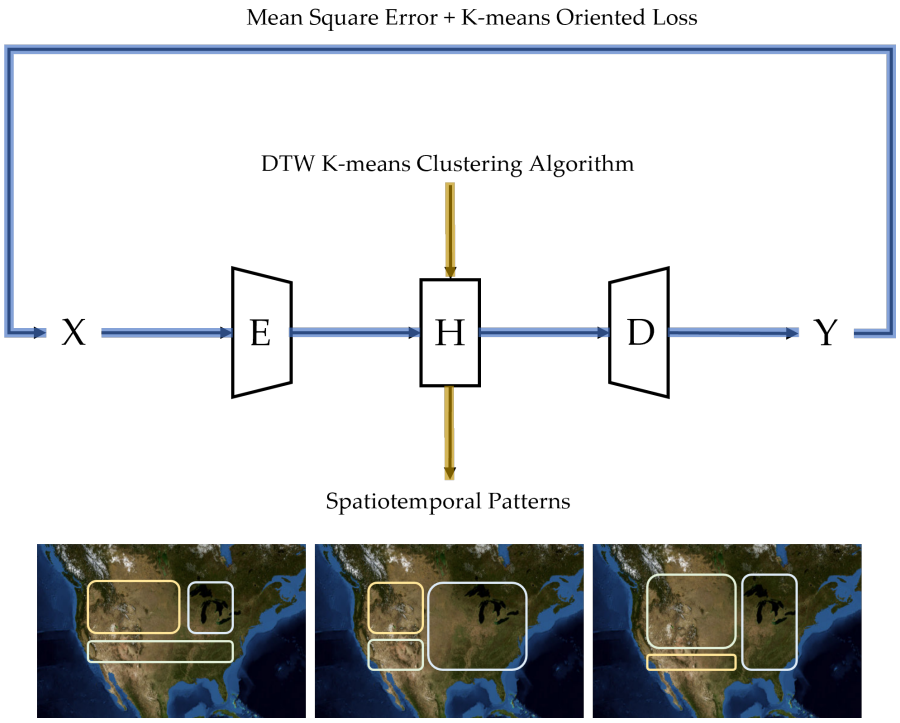

Figure 3. The workflow of our proposed method. X represents the input time series data with location coordinates, E denotes the encoding process, and H is the latent space where temporal representations are learned. D is the decoding process, which maps the temporal representations to Y. After generating the temporal representations in H, apply the dynamic time warping-based K-means (DTW K-means) clustering algorithm to produce spatiotemporal patterns and visualize them on the map using location coordinates. Exploring dynamic spatiotemporal mobility patterns, represented by squares in different colors, enables better decision-making.

Given a collection of time series data from n locations, denoted by $\mathcal{X} = (x_1, x_2, ..., x_n)$, where each time series $x_i \in \mathbb{R}^T$. We introduce a non-linear mapping function $E(\cdot) : x_i \rightarrow h_i$ for encoding, and another non-linear mapping function $D(\cdot) : h_i \rightarrow y_i$ for decoding. Specifically, we design the non-linear mapping $E(\cdot)$ by employing the initial two layers as a one-dimensional convolutional neural network (1D-CNN) layer followed by a max-pooling layer, aimed at capturing short-term patterns within the data. A 1D-CNN is a specialized type of neural network designed for processing sequential data. Unlike their 2D counterparts, which are well-suited for image data, 1D-CNNs are ideal for analyzing time series data, audio signals, and any form of data that unfold over time or space. A latent representation of the n-th feature map of the existing layer is given by the following form:

$$h_i^n = \sigma(x_i \times W^n + b^n), \tag{3}$$

where W represents the filters, b is the corresponding bias for the n-th feature map, and σ denotes the activation function (e.g., sigmoid, ReLU), and $\times$ indicates the 1D convolution operation. The next two layers are Bi-directional Long Short-Term Memory (Bi-LSTM)

layers, which are for learning temporal changes in two directions of time. A Bi-LSTM network is a special type of recurrent neural network that learns from sequences of data, like sentences or time series. Unlike a regular recurrent neural network that only looks at data from the beginning to the end, Bi-directional LSTMs extend the standard LSTM models by utilizing two LSTMs applied to the input data. In the first pass, an LSTM processes the input sequence in its original order (i.e., the forward layer). In the second pass, the input sequence is reversed and fed into another LSTM (i.e., the backward layer). This bi-directional approach enhances the model's ability to learn long-term dependencies, thereby improving its overall accuracy. For each element in the input sequence, each layer of standard LSTMs performs the following computations:

$$\begin{aligned} i_t &= \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{ii}), \\ f_t &= \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}), \\ g_t &= \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}), \\ o_t &= \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}), \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t, \\ h_t &= o_t \odot \tanh(c_t). \end{aligned} \tag{4}$$

Here, h_t represents the hidden state at time t, c_t is the cell state at time t, and x_t denotes the input at time t. The term h_{t-1} refers to the hidden state from the previous time step or the initial hidden state at $t = 0$. The variables i_t, f_t, g_t, and o_t correspond to the input, forget, cell, and output gates, respectively. σ is the sigmoid function, tanh is the hyperbolic tangent function, and $\odot$ denotes the Hadamard (element-wise) product.

After the non-linear mapping $E(\cdot)$, the original temporal sequence $x_i \in \mathbb{R}^T$ is encoded into a considerably reduced representation $h_i \in \mathbb{R}^d$, where d represents the dimensionality of the temporal representation. The selection of d may vary, being either larger or smaller than T. In our experiments, we opt for $d < T$, aiming to preserve the most significant information in a lower-dimensional space, while enhancing computational efficiency when processing high-dimensional datasets. The decoding process $D(\cdot)$ is facilitated by a fully-connected neural network, which reconstructs the temporal representations back into a new time series $y_i \in \mathbb{R}^T$. In the optimization process, the mean square error (MSE) is employed as the metric to quantify reconstruction loss, as demonstrated by the subsequent formula:

$$\mathcal{L}_{MSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n} ||y_i - x_i||_2}. \tag{5}$$

Beyond merely preserving the most informative representation through the reconstruction loss expressed in Equation (5), we incorporate an additional K-means oriented loss. This facilitates the formation of cluster structures within the learned temporal representations, yielding representations that are particularly amenable to the K-means clustering algorithm. Specifically, by considering a static matrix $H \in \mathbb{R}^{n \times d}$, the challenge of finding the optimal solution in K-means clustering is reinterpreted as a problem of maximizing the trace in the context of the Gram matrix. By employing spectral relaxation, the K-means objective is transformed into a formulation that sidesteps local minima, thereby promising optimal solutions:

$$\mathcal{L}_{K-means} = Tr(H^T H) - Tr(F^T H^T H F), \tag{6}$$

where Tr denotes the matrix trace. The close-form solution for F is derived by assembling the first k singular vectors of H, following the principles outlined in the Ky Fan theorem [17]. Unlike static matrices, in our approach, the latent temporal representation H is dynamically learned within the network. This dynamic learning process effectively serves as a regularization term for H; the overall objective is to minimize the below equation:

$$\min_{H} \mathcal{L}_{MSE} + \lambda \mathcal{L}_{K-means}, \tag{7}$$

where λ is a balancing factor that regulates the trade-off between the mean squared error and the K-means oriented loss.

After generating temporal representations, we apply a dynamic time warping-based K-means clustering algorithm to these representations. Dynamic time warping can be seen as an extension of Euclidean distance, which offers local (non-linear) alignment. The calculating procedure of dynamic time warping can be illustrated as follows: given two temporal sequences $P \in \mathbb{R}^n$ and $Q \in \mathbb{R}^n$, formulate an $n \times n$ matrix, with the element at intersection i, j quantifying the Euclidean between p^i and q^j; the goal of dynamic time warping distance is to identify the path of minimum cumulative distance across this matrix. Define this path by $M = \{m_1, m_2, ..., m_k\}$, where the dynamic time warping distance is determined as the smallest sum of Euclidean distance along the path M:

$$M^* = \arg\min_M \left(\sqrt{\sum_{i=1}^{k} m_i}\right). \tag{8}$$

This whole process can be summarized in Algorithm 1.

Algorithm 1 Dynamic Time Warping

procedure INPUT(P, Q) ▷ Time Series P with length n; Time Series Q with length n.
  for i = 1 to n **do**
    for j = 1 to n **do**
      $Matrix(i, j) \leftarrow \sqrt{(p^i - q^j)^2}$
  Define a path through the matrix:$M = m_1, m_2, ..., m_k, k \geq n$
  $DTW \leftarrow argmin_M(\sqrt{\sum_{m=1}^{K} m_k})$
  return DTW

The detailed procedure of applying a dynamic time warping-based K-means clustering algorithm to these representations is outlined in Algorithm 2, which operates on temporal representations $H \in \mathbb{R}^{n \times d}$, where n is the number of locations and d is the dimension of temporal latent representations. The algorithm starts by selecting k central temporal representations from H at random to serve as initial centroids. Subsequently, each temporal representation is assigned to the nearest centroid, determined by the smallest dynamic time warping distance. The next step involves updating the centroids by recalculating the mean position of the temporal representations assigned to each centroid, followed by a re-evaluation of the dynamic time warping distances between each temporal representation and the new centroids. This iterative process continues until the assignments of temporal representation to centroids stabilize and no further reassignments occur.

Algorithm 2 Dynamic Time Warping-based K-means on Temporal Representations

procedure INPUT(H, k) ▷ n-size temporal representations: $H \in \mathbb{R}^{n \times d}$; number of clusters: $k < n$.
  Randomly initialize k centroids $(c_1, c_2, ..., c_k) \in H$.
  Calculate dynamic time warping distance between each temporal representation in $H = (h_1, h_2, ..., h_n)$ and each centroid in $(c_1, c_2, ..., c_k) \in H$.
  Assign the temporal representation in H to the centroid c_i by selecting the one with the smallest dynamic time warping distance from it among all k centroids, and finally form k spatial clusters: $C = (c_1, c_2, ..., c_k)$.
  Recalculate the new k central temporal representations in $C = (c_1, c_2, ..., c_k)$.
  Stop until C will not change; otherwise, repeat from Step 3.
  return C

The final step involves visualizing and interpreting the spatial clusters. The generated k spatial clusters are represented as $C = (c_1, c_2, ..., c_k)$, where each c_i consists of a set of geographic locations that exhibit similar trends or patterns. These locations are then projected onto a map using their coordinates, allowing us to explore spatiotemporal patterns. To quantify the spatial clusters, we introduce a metric called the average travel distance of a cluster, defined as follows:

$$\textit{Average Travel Distance (ATD)} = \frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T} dist_{n,t}, \tag{9}$$

where N represents the number of locations within a cluster, T denotes the number of time steps, and $dist_{n,t}$ is the travel distance in the n-th location at time step t.

The *ATD* reflects the average distance traveled within a cluster over a specified number of time steps. It quantifies how much movement occurs, on average, within the geographic locations of a cluster over time. This can provide insights into the mobility patterns of the regions in the cluster, indicating whether certain areas have high or low travel activity. By comparing *ATD* across clusters, we can assess the relative mobility within different geographic areas, potentially linking these patterns to underlying socioeconomic or environmental factors.

3. Results

3.1. Comparative Performance Evaluation

To evaluate the effectiveness of our proposed method in identifying mobility patterns, we first formulate the problem as follows: consider a set of n geographic locations, each represented by a time series $\mathcal{X} = \{x_1, x_2, ..., x_n\}$, where each time series $x_i \in \mathbb{R}^T$, where x_i^t denotes the observation at time t for location i, and T is the total number of time steps. The goal is to group these time series into k spatial clusters, denoted by $\{c_1, c_2, \ldots, c_k\}$, such that locations within the same cluster exhibit similar temporal patterns. The objective can be formulated as finding a partition $\{c_1, c_2, \ldots, c_k\}$ of the set of time series $\{x_1, x_2, ..., x_n\}$ that minimizes within-cluster variance while maximizing between-cluster variance. This can be expressed as follows:

$$\min_{\{c_1, c_2, ..., c_k\}} \sum_{j=1}^{k} \sum_{x_j \in c_j} dist(x_j, z_j), \tag{10}$$

where z_j is the centroid of cluster c_j, and $\text{dist}(x_j, z_j)$ is a distance measure between the time series x_j and the cluster centroid z_j.

To verify the clustering, time series classification can be applied to the clusters. Given a set of labels $\{l_1, l_2, \ldots, l_k\}$ corresponding to known categories, the classification task assigns each time series x_i to one of the labels based on its cluster assignment. The accuracy of this classification provides a measure of how well the clustering captures meaningful, discriminative patterns in the data. We have access to ground truth information, allowing us to evaluate the cluster results against these ground truth locations. Therefore, we choose to employ the Rand Index, which can be expressed as follows:

$$\textit{Rand Index} = \frac{2 \cdot (TP + TN)}{n(n-1)}. \tag{11}$$

In this context, true positives (TP) represent the count of time series pairs accurately grouped within the same spatial cluster. Conversely, true negatives (TN) refer to the count of time series pairs accurately assigned to distinct spatial clusters, while n denotes the size of the spatiotemporal dataset.

We compared our proposed method with some widely used methods and present the results in Table 1. Our proposed method demonstrates superior performance by achieving the highest average rand index. The synthetic dataset, characterized by noise, outliers, and

time distortions, presents significant challenges for traditional K-means clustering methods. These traditional methods fail primarily due to their reliance on Euclidean distance, which lacks robustness to time distortions.

Although our proposed temporal autoencoder and dynamic time warping-based K-means clustering algorithm without K-means oriented loss (I-DTSC w/o KL) shows improvement over most clustering algorithms, its performance is close to that of the DTW-based K-means (DTW K-Means) approach. We acknowledge that the efficient computation is due to the advantage brought by the dimensionality reduction of the temporal autoencoder, but the potential of the temporal autoencoder has not been fully realized, leading to the neglect of some important information in the original time series.

The complete framework of I-DTSC exhibited exceptional robustness to noise, outliers, and time distortions. This approach guides the generation of more reliable and informative representations, thereby facilitating superior performance. In the case of the Melbourne dataset, our analysis indicates an absence of time distortions in this dataset, allowing traditional K-means to perform adequately; thus, the advantages of DTW K-means clustering algorithms are less pronounced in this context. Despite this, our proposed method outperforms other time series clustering techniques in identifying spatial clusters from time series collected in different locations. For the Chinatown dataset, which aims to differentiate between weekday and weekend time series collected at the same location, most methods fail to discern this distinction. This failure is often due to the subtlety of the differences; for instance, weekend mobility may still exceed typical levels during the evening and overnight periods. Conversely, our proposed method effectively identifies these critical differences, distinguishing between weekday and weekend patterns. Consequently, our method exhibits the highest performance in this context.

Table 1. Comparative results on synthetic and real-world mobility data. (Metric: Rand Index).

Category	Methods	Synthetic Mobility Data	Melbourne Mobility Data	Chinatown Mobility Data
Variants of K-means	K-means	0.6824 ± 0.0094	0.8767 ± 0.0064	0.5414 ± 0.0008
	DTW K-means	0.8836 ± 0.0124	0.8584 ± 0.0054	0.5714 ± 0.0094
DR-based K-means	NMF + K-means	0.5898 ± 0.0068	0.8514 ± 0.0075	0.5036 ± 0.0008
	PCA + K-means	0.7259 ± 0.0003	0.8699 ± 0.0046	0.5376 ± 0.0047
NN-based K-means	CAE + K-means	0.6575 ± 0.0012	0.8612 ± 0.0049	0.6712 ± 0.0075
	TAE + K-means	0.6891 ± 0.0021	0.8718 ± 0.0163	0.7088 ± 0.0035
Ours	I-DTSC w/o KL	0.8804 ± 0.0071	0.8601 ± 0.0049	0.6932 ± 0.0154
	I-DTSC	0.9117 ± 0.0036	0.8834 ± 0.0013	0.9055 ± 0.0099

3.2. Module Contribution Analysis

To better understand our proposed method, we first examined the role of the temporal autoencoder, as depicted in Figure 4. By employing our designed architecture and focusing on minimizing the mean square error between the original and reconstructed time series, we ensured that the temporal representations preserved essential characteristics, such as trends, while simultaneously reducing data dimensionality and mitigating noise and outliers. However, we found that only using the temporal autoencdoer does not address time distortions. This observation is also supported by performance comparisons, demonstrating that traditional K-means clustering methods struggle with datasets characterized by significant time distortions.

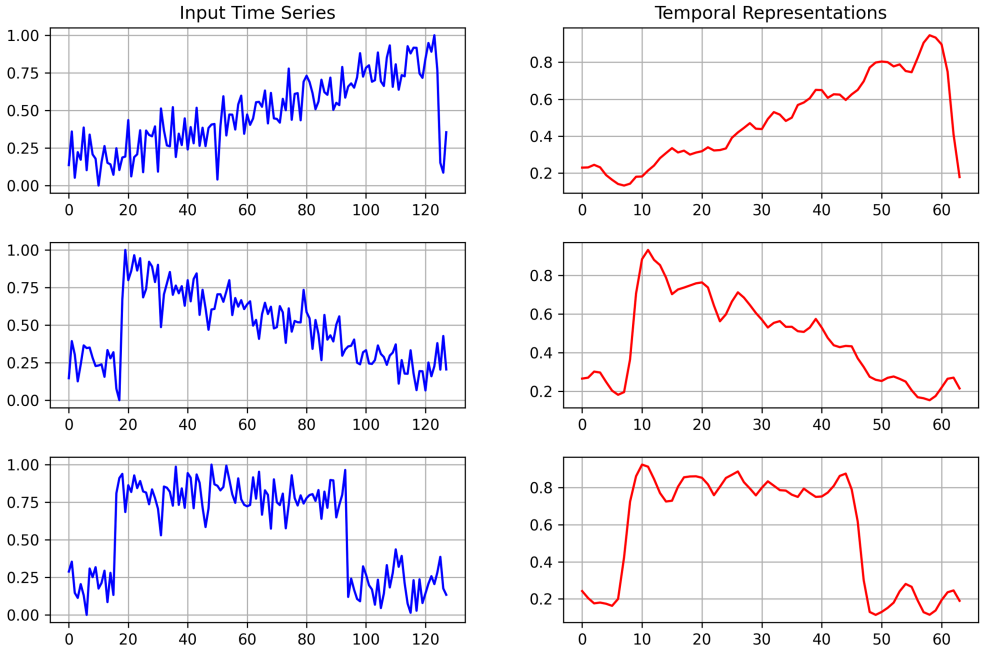

Figure 4. The temporal representations (**right**; red), with a reduced dimensionality of 64 compared to the original input time series dimension of 128 (**left**; blue), successfully preserved essential characteristics, such as trends. This reduction in dimensionality not only retained critical information but also helped mitigate noise and outliers.

Time distortions are a prevalent issue, particularly in real-world time series datasets. The traditional K-means clustering algorithm, which employs Euclidean distance as the similarity metric, is efficient and requires low computational resources, with a complexity of $O(n)$. However, this method often fails when confronted with time distortions, as the utilization of Euclidean distance cannot adequately align time points affected by distortions. A more suitable solution is to employ a DTW K-means clustering algorithm, which can yield more accurate results. Nonetheless, this strategy significantly increases computational demands, with a time complexity of $O(n^2)$. In our proposed method, we aim to achieve accurate results while also reducing time complexity. By compressing the original time series of length n to temporal representation with a shorter length m, where $m < n$, we apply the DTW K-means clustering algorithm not on the original series but on the temporal representations. This adjustment lowers the time complexity to $O(m^2)$. Our experimental findings suggest that the temporal autoencoder and DTW K-means clustering algorithm can mutually enhance each other's performance, achieving a balance between accuracy and computational efficiency.

In our experiments, we also analyzed the role of K-means oriented loss illustrated in Equation (6) for optimization. We conducted experiments on the Chinatown time series dataset to compare learning temporal representations with and without the K-means oriented loss. Figure 5 illustrates the results: without the K-means oriented loss, the representations appear scattered and chaotic; however, integrating the K-means oriented loss guides the process of learning temporal representations, resulting in structured temporal representations.

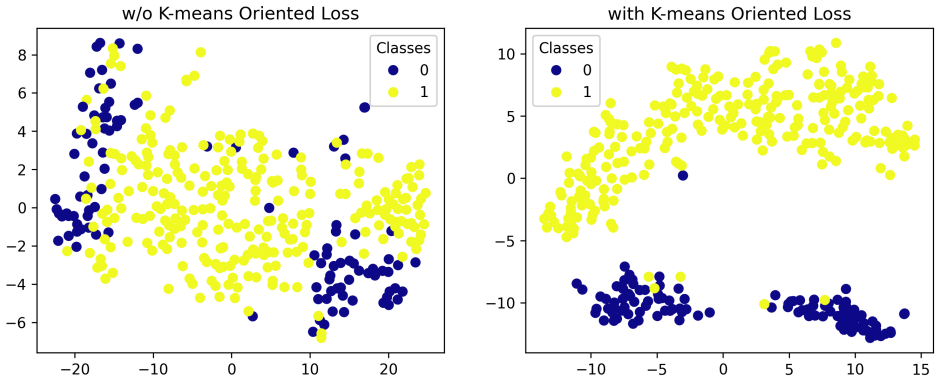

Figure 5. The visualization using t-SNE on the Chinatown time series dataset compares temporal representations without K-means oriented loss (**left**) and temporal representations with K-means oriented loss (**right**).

3.3. Case Study: Mining Spatiotemporal Mobility Patterns During COVID-19 Pandemic

The COVID-19 pandemic has caused a substantial negative impact on public health, infected more than 704 million people, and caused 7.01 million deaths worldwide as of September 2024. The United States government had issued social distancing measures to restrict gatherings, domestic and international travel, and business activities to reduce the COVID-19 cases and mortality rate. In response to the crisis, researchers developed various spatial models to analyze the distributions of the disease patterns and their dependency on factors such as mobility [11,51–53], social [54,55], economic [56,57], and demographic factors [31,58]. Furthermore, some existing methods were focused on spatiotemporal data modeling in COVID-19-related research to improve decision-making [59–64].

As detailed in this section, we applied our proposed method to uncover spatiotemporal patterns hidden in the COVID-19 mobility data. We adopted the elbow method to identify the optimal cluster number. Specifically, the elbow method is a widely used heuristic in mathematical optimization, which involves selecting a point-referred to as the "Elbow" or "Knee of the curve" beyond which the benefits of further improvements do not justify the additional cost [65]. In clustering, this approach suggests choosing a number of clusters such that adding more clusters does not significantly enhance the model's accuracy. We conducted several experiments with different values of k across three months. Based on the results, we selected $k = 4$ as the optimal number of clusters. In other words, we categorized the 3142 counties into 4 groups, indicating that counties within the same group exhibit similar mobility patterns.

Figure 6 presents the spatiotemporal mobility patterns derived from our proposed method using four distinct colors to represent varying levels of mobility, where lighter shades indicate higher mobility. The results offer granular insights into mobility patterns during the three key phases of the COVID-19 pandemic: before the implementation of stay-at-home orders, during the stay-at-home period, and throughout the reopening phase. In the following sections, we delve into detailed observations of spatiotemporal mobility patterns, along with potential factors that contribute to these patterns.

Before the national pandemic response fully ramped up, the Midwest and Southern regions of the U.S. exhibited relatively high mobility levels, as indicated by the green and yellow shades across counties in states like Iowa, Kansas, Missouri, and Texas. These areas, being more rural in nature, likely saw higher movement due to daily activities such as agricultural work and commuting over larger distances. Additionally, many of these regions did not implement early restrictions, contributing to sustained levels of mobility during the initial phase of the pandemic. In contrast, coastal and urban areas, such as those on the East Coast (New York, New Jersey, Massachusetts) and the West Coast (California, Washington), displayed more moderate mobility levels. While formal restrictions had yet to be implemented in these areas, voluntary reductions in movement may have already begun

as media coverage of the pandemic intensified. Urban centers, which typically rely more on public transportation and have higher population densities, experienced a shift in mobility patterns as non-essential travel was curtailed earlier compared to more rural areas.

During the stay-at-home period, the map reveals a significant nationwide reduction in mobility, with most counties shaded in deep purple, indicating low movement across both urban and rural areas. These widespread declines reflect the effectiveness of stay-at-home orders, which kept the majority of the population indoors except for essential activities like grocery shopping and medical appointments. However, there were notable exceptions in states such as Wyoming, Nebraska, and parts of Texas, where mobility remained relatively higher, as indicated by green and yellow shades. This can be attributed to essential industries, particularly in rural and agricultural regions, that required continued operation, looser restrictions in states with lower population densities, and localized outbreak responses where fewer COVID-19 cases were reported, resulting in less urgency to limit movement [66].

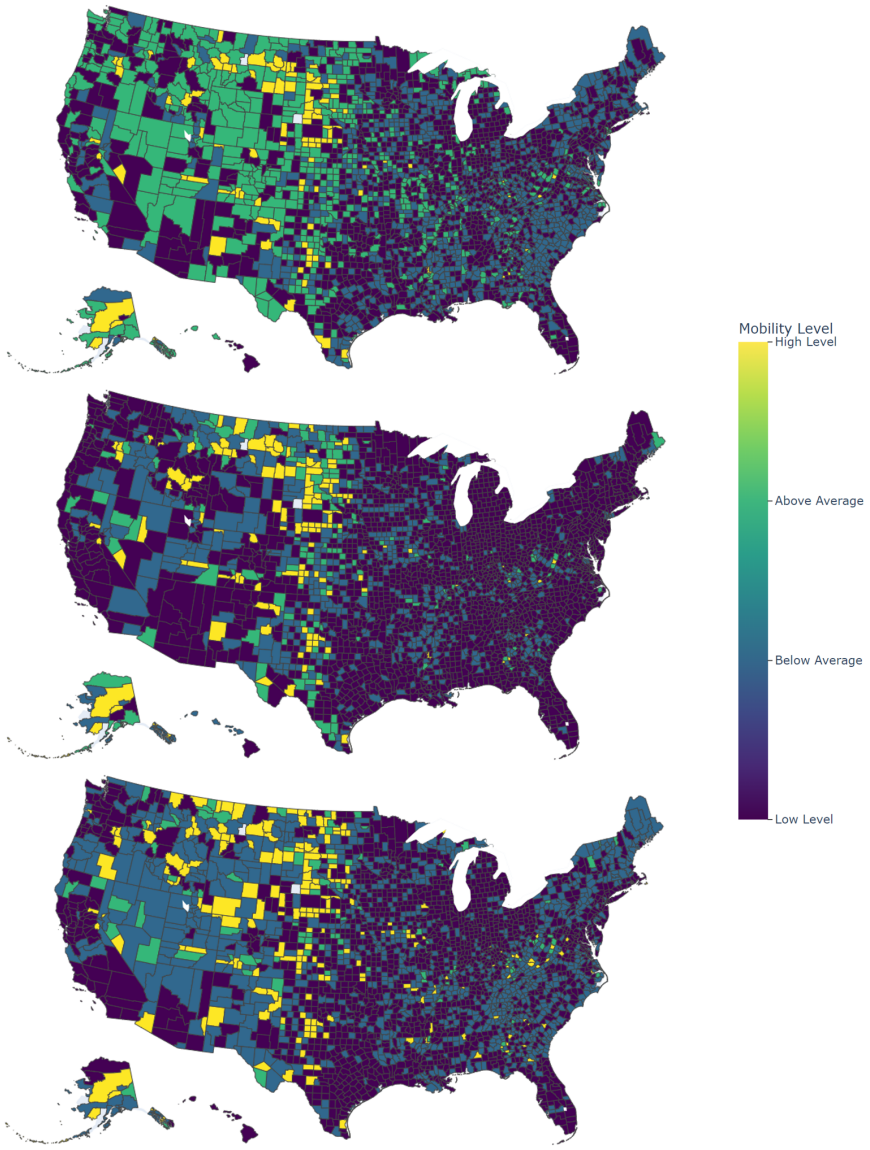

Figure 6. Spatiotemporal mobility patterns during the COVID-19 pandemic in three phases: before stay-at-home, during stay-at-home, and reopening (**top** to **bottom**). Mobility is categorized into four levels: high, above average, below average, and low, reflecting different movement trends across regions.

The bottom map reflects the reopening period when states began to relax restrictions and reopen their economies, leading to a partial recovery in mobility. In regions like the Midwest (Nebraska, Iowa, Kansas) and the South (Texas, Florida, Georgia), mobility levels noticeably increased, as indicated by the green and yellow shades. These states were among the first to prioritize economic recovery, allowing residents to return to work, shopping, and other activities, even amid rising COVID-19 cases. In contrast, coastal and urban areas, particularly on the East Coast (New York, New Jersey, Massachusetts) and the West Coast (California, Oregon, Washington), experienced slower mobility recovery, as shown by the persistent purple areas [52]. This slower rebound can be attributed to several factors, including ongoing caution and phased reopening strategies, especially in regions hit hardest early in the pandemic, where sectors operated under strict guidelines to avoid a second wave. Additionally, many large companies in cities like New York and San Francisco continued to enforce work-from-home policies, reducing the need for commuting and contributing to lower mobility. In states like Florida and Texas, the increased mobility may also reflect the reopening of tourist destinations and leisure activities, alongside higher traffic to commercial centers as restrictions on dining, retail, and entertainment venues were eased.

Across all three maps, clear differences emerge between rural and urban regions in terms of mobility patterns. Rural areas consistently show higher mobility both before and during the pandemic, likely due to the essential nature of agricultural and logistical work, which required continued movement. In contrast, urban areas experienced more significant reductions in mobility during the stay-at-home orders, likely due to stricter enforcement and the widespread ability to transition to remote work [67]. Mobility trends also closely align with state and local policy decisions regarding stay-at-home orders and reopening strategies. States in the Midwest and South, which reopened earlier, saw a quicker return to higher mobility, while states that adopted more cautious approaches experienced more gradual increases. Additionally, public response and health considerations influenced mobility. In areas with higher COVID-19 case rates or where perceived risk was greater, residents may have voluntarily reduced movement even during the reopening phase, as seen in lower mobility levels in states like New York and California.

Figure 7 presents the *ATD* across the stages: before the stay-at-home policies, during the stay-at-home orders, and during the reopening phase, categorized by high, above average, below average, and low mobility levels. The analysis reveals a significant reduction in mobility patterns immediately following the implementation of social distancing policies in March, with a gradual increase observed in May as reopening orders were enacted. This trend underscores the direct relationship between mobility shifts and the stringency of social distancing measures enforced by states. The results also highlight that some states were slower to curtail mobility in response to the pandemic, likely due to a combination of less severe outbreaks, more lenient local regulations, or lower levels of public concern and awareness. Overall, the result demonstrates that mobility across most states declined sharply in response to the onset of COVID-19 and the corresponding implementation of public health guidelines [68].

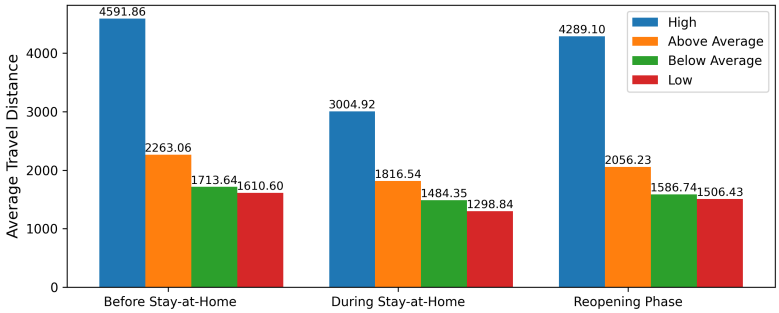

Figure 7. Average travel distance (*ATD*) across different stages from March to May 2020 (in meters).

4. Discussion

4.1. Limitations

It is important to point out some limitations of our work. First, although we used data from several locations, including Melbourne, Australia, and the United States, they are not enough to represent global mobility patterns, as our analysis might have miss important variations, particularly in underrepresented regions like developing countries. A wider geographical area needs to be considered to better understand a broader range of spatiotemporal patterns.

Second, through extensive experiments, we found that the choice of method should be tailored to the specific dataset. For datasets with minimal noise, outliers, or time distortions, or where these factors do not significantly impact the results, applying K-means with Euclidean distance directly can be both efficient and effective. However, in our study, we operated under the assumption of uncertain data quality and consider the worst-case scenario, where high-dimensionality, noise, outliers, and time distortions are present. Under these conditions, we developed a method that balances both accuracy and efficiency to handle these challenges effectively. Additionally, we used the elbow method to determine the optimal number of spatial clusters. However, in practical situations, where ground truth about spatial clusters is unavailable, the number of clusters may need to be adjusted based on specific applications, with the primary goal being to support better decision-making.

Third, we acknowledge that mobility is influenced by various factors, and understanding how mobility patterns evolve on their own may not be sufficient for informed decision-making. Thus, a deeper exploration into how specific socioeconomic conditions, such as economic status and population density, affect mobility patterns would enhance the analysis. Additionally, this study primarily focused on short-term mobility changes using only three months of mobility data from the COVID-19 pandemic in the U.S. Long-term analysis is necessary to gain a more comprehensive understanding of spatiotemporal mobility patterns.

4.2. Future Work

Here, we propose two future directions that should be explored. First, using more advanced neural network architectures, such as hybrid models combining convolutional, recurrent, or attention-based networks, along with improved optimization strategies, could better capture the relationships between multiple inputs and support long-term analysis. These enhanced architectures would provide a deeper understanding of the intricate dynamics within mobility data. Additionally, employing more suitable clustering algorithms, such as density-based or hierarchical methods, could enhance both the efficiency and accuracy of clustering the generated representations, leading to more nuanced insights into mobility patterns.

Second, we need to address more challenging situations, such as handling time series data with missing values, irregular sampling, and incomplete location information. In many real-world mobility datasets, data collection may be incomplete due to sensor failures, communication errors, or irregular intervals in reporting. Developing robust methods to impute missing values or work effectively with irregular time intervals is essential. Techniques such as interpolation, time series imputation models, and incorporating probabilistic methods could help mitigate the effects of missing data. Additionally, extending the models to handle irregularly spaced data points could improve their applicability and reliability in real-world scenarios, ensuring more accurate analysis and pattern detection.

5. Conclusions

This study proposes an improved deep time series clustering method for effectively mining spatiotemporal mobility patterns. By integrating a neural network-based temporal autoencoder with a dynamic time warping-based K-means clustering algorithm, we successfully address challenges such as high dimensionality, noise, outliers, and time distortions in mobility data, resulting in a more robust clustering approach. Extensive experiments

on both synthetic and real-world datasets demonstrated the effectiveness of our method. Applying our approach to mobility data of the COVID-19 pandemic in the U.S. revealed significant differences in mobility patterns between rural and urban areas. Rural areas consistently exhibited higher mobility, both before and during the pandemic, likely due to the essential nature of agricultural and logistical work that required continued movement. In contrast, urban areas experienced more substantial reductions in mobility during stay-at-home orders, likely due to stricter enforcement and the widespread adoption of remote work. Additionally, public response and health considerations influenced mobility patterns. In regions with higher COVID-19 case rates or greater perceived risk, residents voluntarily reduced movement, even during reopening phases, as observed in lower mobility levels in states like New York and California. This application highlights the broader implications of our proposed method in potential fields such as public health, epidemiology, transportation analysis, urban planning, smart cities, and environmental monitoring.

Author Contributions: Conceptualization, Ziyi Zhang, Zhe Zhang, and Nick Duffield; methodology, Ziyi Zhang; software, Ziyi Zhang and Diya Li; validation, Ziyi Zhang; formal analysis, Ziyi Zhang; investigation, Diya Li; resources, Diya Li; data curation, Diya Li; writing—original draft preparation, Ziyi Zhang; writing—review and editing, Zhe Zhang; visualization, Ziyi Zhang; supervision, Zhe Zhang and Nick Duffield; project administration, Zhe Zhang and Nick Duffield; funding acquisition, Zhe Zhang and Nick Duffield. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by U.S. National Science Foundation (NSF) Grant #2339174 CAREER: A Cyberinfrastructure Enabled Hybrid Spatial Decision Support System for Improving Coastal Resilience to Flood Risks and #2321069 Collaborative Research: CyberTraining: Implementation: Small: Broadening Adoption of Cyberinfrastructure and Research Workforce Development for Disaster Management.

Data Availability Statement: The Melbourne Pedestrian Time Series Data and the Chinatown Pedestrian Time Series Data (https://www.cs.ucr.edu/~eamonn/time_series_data/, accessed on 15 January 2024), as well as the COVID-19 Mobility Time Series Data (https://data.bts.gov/Research-and-Statistics/Trips-by-Distance/w96p-f2qv, accessed on 10 September 2021), were utilized in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ATD	Average Travel Distance
BiLSTM	Bi-directional Long Short-Term Memory
CAE	Convolutional Autoencoder
DR	Dimensionality Reduction
DTW	Dynamic Time Warping
I-DTSC	Improved Deep Time Series Clustering
KL	K-means oriented Loss
MSE	Mean Square Error
NN	Neural Network
NMF	Non-negative Matrix Factorization
PCA	Principal Component Analysis
RI	Rand Index
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
TAE	Temporal Autoencoder
TN	True Negative
TP	True Positive
1D-CNN	One-Dimensional Convolutional Neural Network

References

1. Hasan, S.; Schneider, C.M.; Ukkusuri, S.V.; González, M.C. Spatiotemporal patterns of urban human mobility. *J. Stat. Phys.* **2013**, *151*, 304–318. [CrossRef]
2. Alessandretti, L.; Sapiezynski, P.; Lehmann, S.; Baronchelli, A. Multi-scale spatio-temporal analysis of human mobility. *PLoS ONE* **2017**, *12*, e0171686. [CrossRef] [PubMed]
3. Caldeira, A.M.; Kastenholz, E. Spatiotemporal tourist behaviour in urban destinations: A framework of analysis. *Tour. Geogr.* **2020**, *22*, 22–50. [CrossRef]
4. Zhai, H.; Gu, B.; Wang, Y. Evaluation of policies and actions for nature-based solutions in nationally determined contributions. *Land Use Policy* **2023**, *131*, 106710. [CrossRef]
5. Zhai, H.; Gu, B.; Zhu, K.; Huang, C. Feasibility analysis of achieving net-zero emissions in China's power sector before 2050 based on ideal available pathways. *Environ. Impact Assess. Rev.* **2023**, *98*, 106948. [CrossRef]
6. Mao, F.; Ji, M.; Liu, T. Mining spatiotemporal patterns of urban dwellers from taxi trajectory data. *Front. Earth Sci.* **2016**, *10*, 205–221. [CrossRef]
7. Neves, F.; Finamore, A.C.; Madeira, S.C.; Henriques, R. Mining actionable patterns of road mobility from heterogeneous traffic data using biclustering. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 6430–6445. [CrossRef]
8. Xiao, G.; Wang, Y.; Wu, R.; Li, J.; Cai, Z. Sustainable maritime transport: A review of intelligent shipping technology and green port construction applications. *J. Mar. Sci. Eng.* **2024**, *12*, 1728. [CrossRef]
9. Song, X.; Zhang, Q.; Sekimoto, Y.; Shibasaki, R. Prediction of human emergency behavior and their mobility following large-scale disaster. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 5–14.
10. Zhang, Z.; Li, D.; Zhang, Z.; Duffield, N. A time-series clustering algorithm for analyzing the changes of mobility pattern caused by COVID-19. In Proceedings of the 1st ACM SIGSPATIAL International Workshop on Animal Movement Ecology And Human Mobility, Beijing, China, 2 November 2021; pp. 13–17.
11. Hu, N.; Zhang, Z.; Duffield, N.; Li, X.; Dadashova, B.; Wu, D.; Yu, S.; Ye, X.; Han, D.; Zhang, Z. Geographical and temporal weighted regression: Examining spatial variations of COVID-19 mortality pattern using mobility and multi-source data. *Comput. Urban Sci.* **2024**, *4*, 6. [CrossRef]
12. Li, D.; Zhang, Z.; Alizadeh, B.; Zhang, Z.; Duffield, N.; Meyer, M.A.; Thompson, C.M.; Gao, H.; Behzadan, A.H. A reinforcement learning-based routing algorithm for large street networks. *Int. J. Geogr. Inf. Sci.* **2024**, *38*, 183–215. [CrossRef]
13. Aghabozorgi, S.; Shirkhorshidi, A.S.; Wah, T.Y. Time-series clustering—A decade review. *Inf. Syst.* **2015**, *53*, 16–38. [CrossRef]
14. Zhao, B.; Lu, H.; Chen, S.; Liu, J.; Wu, D. Convolutional neural networks for time series classification. *J. Syst. Eng. Electron.* **2017**, *28*, 162–169. [CrossRef]
15. Zhang, Z.; Li, D.; Song, Z.; Duffield, N.; Zhang, Z. Location-Aware social network recommendation via temporal graph networks. In Proceedings of the 7th ACM SIGSPATIAL Workshop on Location-Based Recommendations, Geosocial Networks and Geoadvertising, Hamburg, Germany, 13 November 2023; pp. 58–61.
16. Li, L.; Prakash, B.A. Time series clustering: Complex is simpler! In Proceedings of the 28th International Conference on Machine Learning (ICML-11), Bellevue, WA, USA, 28 June–2 July 2011; pp. 185–192.
17. Ma, Q.; Zheng, J.; Li, S.; Cottrell, G.W. Learning representations for time series clustering. In Proceedings of the Advances in Neural Information Processing Systems 32, Vancouver, BC, Canada, 8–14 December 2019.
18. Xu, K.; Chen, L.; Wang, S. Data-driven kernel subspace clustering with local manifold preservation. In Proceedings of the 2022 IEEE International Conference on Data Mining Workshops (ICDMW), Orlando, FL, USA, 28 November–1 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 876–884.
19. Xu, K.; Chen, L.; Wang, S. A multi-view kernel clustering framework for categorical sequences. *Expert Syst. Appl.* **2022**, *197*, 116637. [CrossRef]
20. Rasul, K.; Seward, C.; Schuster, I.; Vollgraf, R. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; PMLR; pp. 8857–8868.
21. Zhou, T.; Ma, Z.; Wen, Q.; Wang, X.; Sun, L.; Jin, R. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. In Proceedings of the International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022; PMLR; pp. 27268–27286.
22. Zhang, Z.; Ren, S.; Qian, X.; Duffield, N. Towards invariant time series forecasting in smart cities. In Proceedings of the Companion Proceedings of the ACM on Web Conference 2024, Singapore, 13–17 May 2024; pp. 1344–1350.
23. Xu, K.; Chen, L.; Patenaude, J.M.; Wang, S. Rhine: A regime-switching model with nonlinear representation for discovering and forecasting regimes in financial markets. In Proceedings of the 2024 SIAM International Conference on Data Mining (SDM), Houston, TX, USA, 18–20 April 2024; SIAM: Philadelphia, PA, USA, 2024; pp. 526–534.
24. Xu, K.; Chen, L.; Patenaude, J.M.; Wang, S. Kernel representation learning with dynamic regime discovery for time series forecasting. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Taiwan, China, 7–10 May 2024; Springer: Berlin/Heidelberg, Germany, 2024; pp. 251–263.
25. Pamfil, R.; Sriwattanaworachai, N.; Desai, S.; Pilgerstorfer, P.; Georgatzis, K.; Beaumont, P.; Aragam, B. Dynotears: Structure learning from time-series data. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Online, 26–28 August 2020; PMLR; pp. 1595–1605.

26. Tank, A.; Covert, I.; Foti, N.; Shojaie, A.; Fox, E.B. Neural granger causality. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 4267–4279. [CrossRef]
27. Zhang, Z.; Ren, S.; Qian, X.; Duffield, N. Learning Flexible Time-windowed Granger Causality Integrating Heterogeneous Interventional Time Series Data. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Barcelona, Spain, 25–29 August 2024; pp. 4408–4418.
28. Steinbach, M.; Tan, P.N.; Kumar, V.; Klooster, S.; Potter, C. Discovery of climate indices using clustering. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24–27 August 2003; pp. 446–455.
29. Ji, M.; Xie, F.; Ping, Y. A dynamic fuzzy cluster algorithm for time series. In *Abstract and Applied Analysis*; Wiley Online Library: Hoboken, NJ, USA, 2013; Volume 2013, p. 183410.
30. Elangasinghe, M.; Singhal, N.; Dirks, K.; Salmond, J.; Samarasinghe, S. Complex time series analysis of PM10 and PM2. 5 for a coastal site using artificial neural network modelling and k-means clustering. *Atmos. Environ.* **2014**, *94*, 106–116. [CrossRef]
31. Kontis, V.; Bennett, J.E.; Rashid, T.; Parks, R.M.; Pearson-Stuttard, J.; Guillot, M.; Asaria, P.; Zhou, B.; Battaglini, M.; Corsetti, G.; et al. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. *Nat. Med.* **2020**, *26*, 1919–1928. [CrossRef] [PubMed]
32. Iglesias, F.; Kastner, W. Analysis of similarity measures in times series clustering for the discovery of building energy patterns. *Energies* **2013**, *6*, 579–597. [CrossRef]
33. Shumway, R.H. Time-frequency clustering and discriminant analysis. *Stat. Probab. Lett.* **2003**, *63*, 307–314. [CrossRef]
34. Huang, X.; Li, Z.; Lu, J.; Wang, S.; Wei, H.; Chen, B. Time-series clustering for home dwell time during COVID-19: What can we learn from it? *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 675. [CrossRef]
35. Yang, K.; Shahabi, C. A PCA-based similarity measure for multivariate time series. In Proceedings of the 2nd ACM International Workshop on Multimedia Databases, Washington, DC, USA, 13 November 2004; pp. 65–74.
36. Li, H. Multivariate time series clustering based on common principal component analysis. *Neurocomputing* **2019**, *349*, 239–247. [CrossRef]
37. Zhou, L.; Du, G.; Tao, D.; Chen, H.; Cheng, J.; Gong, L. Clustering multivariate time series data via multi-nonnegative matrix factorization in multi-relational networks. *IEEE Access* **2018**, *6*, 74747–74761. [CrossRef]
38. He, G.; Wang, H.; Liu, S.; Zhang, B. CSMVC: A multiview method for multivariate time-series clustering. *IEEE Trans. Cybern.* **2021**, *52*, 13425–13437. [CrossRef]
39. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [CrossRef]
40. Alqahtani, A.; Ali, M.; Xie, X.; Jones, M.W. Deep time-series clustering: A review. *Electronics* **2021**, *10*, 3001. [CrossRef]
41. Tavakoli, N.; Siami-Namini, S.; Adl Khanghah, M.; Mirza Soltani, F.; Siami Namin, A. An autoencoder-based deep learning approach for clustering time series data. *SN Appl. Sci.* **2020**, *2*, 1–25. [CrossRef]
42. Özgül, O.F.; Bardak, B.; Tan, M. A convolutional deep clustering framework for gene expression time series. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**, *18*, 2198–2207. [CrossRef]
43. Mousavi, S.M.; Zhu, W.; Ellsworth, W.; Beroza, G. Unsupervised clustering of seismic signals using deep convolutional autoencoders. *IEEE Geosci. Remote. Sens. Lett.* **2019**, *16*, 1693–1697. [CrossRef]
44. Karadayı, Y.; Aydin, M.N.; Öğrenci, A.S. A hybrid deep learning framework for unsupervised anomaly detection in multivariate spatio-temporal data. *Appl. Sci.* **2020**, *10*, 5191. [CrossRef]
45. Chadha, G.S.; Islam, I.; Schwung, A.; Ding, S.X. Deep convolutional clustering-based time series anomaly detection. *Sensors* **2021**, *21*, 5488. [CrossRef] [PubMed]
46. Madiraju, N.S. Deep Temporal Clustering: Fully Unsupervised Learning of Time-Domain Features. Master's Thesis, Arizona State University, Tempe, AZ, USA, 2018.
47. Trosten, D.J.; Strauman, A.S.; Kampffmeyer, M.; Jenssen, R. Recurrent deep divergence-based clustering for simultaneous feature learning and clustering of variable length time series. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3257–3261.
48. Ienco, D.; Interdonato, R. Deep multivariate time series embedding clustering via attentive-gated autoencoder. In Proceedings of the Advances in Knowledge Discovery and Data Mining: 24th Pacific-Asia Conference, PAKDD 2020, Singapore, 11–14 May 2020; Proceedings, Part I 24; Springer: Berlin/Heidelberg, Germany, 2020; pp. 318–329.
49. Lafabregue, B.; Weber, J.; Gançarski, P.; Forestier, G. End-to-end deep representation learning for time series clustering: A comparative study. *Data Min. Knowl. Discov.* **2022**, *36*, 29–81. [CrossRef]
50. Saito, N. Local Feature Extraction and Its Applications Using a Library of Bases. Ph.D.Thesis, Yale University, New Haven, CT, USA, 1994.
51. Zhou, Y.; Liu, X.C.; Grubesic, T. Unravel the impact of COVID-19 on the spatio-temporal mobility patterns of microtransit. *J. Transp. Geogr.* **2021**, *97*, 103226. [CrossRef]
52. Gao, S.; Rao, J.; Kang, Y.; Liang, Y.; Kruse, J. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSpatial Spec.* **2020**, *12*, 16–26. [CrossRef]

53. Nouvellet, P.; Bhatia, S.; Cori, A.; Ainslie, K.E.; Baguelin, M.; Bhatt, S.; Boonyasiri, A.; Brazeau, N.F.; Cattarino, L.; Cooper, L.V.; et al. Reduction in mobility and COVID-19 transmission. *Nat. Commun.* **2021**, *12*, 1–9. [CrossRef] [PubMed]
54. Bavel, J.J.V.; Baicker, K.; Boggio, P.S.; Capraro, V.; Cichocka, A.; Cikara, M.; Crockett, M.J.; Crum, A.J.; Douglas, K.M.; Druckman, J.N.; et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **2020**, *4*, 460–471. [CrossRef]
55. Clemente-Suárez, V.J.; Navarro-Jiménez, E.; Moreno-Luna, L.; Saavedra-Serrano, M.C.; Jimenez, M.; Simón, J.A.; Tornero-Aguilera, J.F. The impact of the COVID-19 pandemic on social, health, and economy. *Sustainability* **2021**, *13*, 6314. [CrossRef]
56. Brodeur, A.; Gray, D.; Islam, A.; Bhuiyan, S. A literature review of the economics of COVID-19. *J. Econ. Surv.* **2021**, *35*, 1007–1044. [CrossRef]
57. Padhan, R.; Prabheesh, K. The economics of COVID-19 pandemic: A survey. *Econ. Anal. Policy* **2021**, *70*, 220–237. [CrossRef]
58. Shah, P.; Owens, J.; Franklin, J.; Mehta, A.; Heymann, W.; Sewell, W.; Hill, J.; Barfield, K.; Doshi, R. Demographics, comorbidities and outcomes in hospitalized Covid-19 patients in rural southwest Georgia. *Ann. Med.* **2020**, *52*, 354–360. [CrossRef]
59. Li, D.; Chaudhary, H.; Zhang, Z. Modeling spatiotemporal pattern of depressive symptoms caused by COVID-19 using social media data mining. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4988. [CrossRef]
60. Zhu, S.; Bukharin, A.; Xie, L.; Santillana, M.; Yang, S.; Xie, Y. High-resolution spatio-temporal model for county-level COVID-19 activity in the US. *ACM Trans. Manag. Inf. Syst. (TMIS)* **2021**, *12*, 1–20. [CrossRef]
61. Liu, L.; Hu, T.; Bao, S.; Wu, H.; Peng, Z.; Wang, R. The spatiotemporal interaction effect of COVID-19 transmission in the United States. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 387. [CrossRef]
62. Wang, Y.; Liu, Y.; Struthers, J.; Lian, M. Spatiotemporal characteristics of the COVID-19 epidemic in the United States. *Clin. Infect. Dis.* **2021**, *72*, 643–651. [CrossRef] [PubMed]
63. Maiti, A.; Zhang, Q.; Sannigrahi, S.; Pramanik, S.; Chakraborti, S.; Cerda, A.; Pilla, F. Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustain. Cities Soc.* **2021**, *68*, 102784. [CrossRef] [PubMed]
64. Song, Z.; Zhang, Z.; Lyu, F.; Bishop, M.; Liu, J.; Chi, Z. From Individual Motivation to Geospatial Epidemiology: A Novel Approach Using Fuzzy Cognitive Maps and Agent-Based Modeling for Large-Scale Disease Spread. *Sustainability* **2024**, *16*, 5036. [CrossRef]
65. Liu, F.; Deng, Y. Determine the number of unknown targets in open world based on elbow method. *IEEE Trans. Fuzzy Syst.* **2020**, *29*, 986–995. [CrossRef]
66. Cuadros, D.F.; Branscum, A.J.; Mukandavire, Z.; Miller, F.D.; MacKinnon, N. Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. *Ann. Epidemiol.* **2021**, *59*, 16–20. [CrossRef]
67. Huang, Y.; Li, R. The lockdown, mobility, and spatial health disparities in COVID-19 pandemic: A case study of New York City. *Cities* **2022**, *122*, 103549. [CrossRef]
68. Badr, H.S.; Du, H.; Marshall, M.; Dong, E.; Squire, M.M.; Gardner, L.M. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 1247–1254. [CrossRef]