Urban Vibrancy Embedding and Application on Traffic Prediction

Sumin Han
School of Computing
KAIST
Daejeon 34141, South Korea
suminkaist@gmail.com
ORCID: 0000-0002-4071-8469

Yeji Lee
Matrixity Corp.
Suwon 16226, South Korea
yeji@matrixity.io
ORCID: 0009-0009-4524-2517

Jisun An
Luddy School of Informatics
Indiana Univ. Bloomington
Bloomington, Indiana, USA
jisunan@iu.edu
ORCID: 0000-0002-4353-8009

Dongman Lee
School of Computing
KAIST
Daejeon 34141, South Korea
dlee@kaist.ac.kr
ORCID: 0000-0001-5923-6227

Abstract—**Urban vibrancy reflects the dynamic human activity within urban spaces and is often measured using mobile data that captures floating population trends. This study proposes a novel approach to derive Urban Vibrancy embeddings from real-time floating population data to enhance traffic prediction models. Specifically, we utilize variational autoencoders (VAE) to compress this data into actionable embeddings, which are then integrated with long short-term memory (LSTM) networks to predict future embeddings. These are subsequently applied in a sequence-to-sequence framework for traffic forecasting. Our contributions are threefold: (1) We use principal component analysis (PCA) to interpret the embeddings, revealing temporal patterns such as weekday versus weekend distinctions and seasonal patterns; (2) We propose a method that combines VAE and LSTM, enabling forecasting dynamic urban knowledge embedding; and (3) Our approach improves accuracy and responsiveness in traffic prediction models, including RNN, DCRNN, GTS, and GMAN. This study demonstrates the potential of Urban Vibrancy embeddings to advance traffic prediction and offer a more nuanced analysis of urban mobility.**

Index Terms—**Urban Vibrancy, Floating Population, Traffic Prediction, Knowledge Adaption**

I. INTRODUCTION

Urban vibrancy reflects the dynamic human activities that occur within urban spaces, and it is increasingly measured through mobile data capturing floating population[1] trends [1], [2]. This concept embodies not just the visible energy of a city but also the potential for interactions and the fundamental social conditions that shape urban life [3]. In terms of human activity, urban vibrancy can be expressed as the intensity and diversity of people's interactions within a city, providing a lens into how individuals engage with their environment on both physical and social levels [4].

---

[1]The floating population refers to individuals who temporarily reside in a specific area without being counted in the official census, often for work or educational purposes.

Understanding urban vibrancy poses significant challenges, primarily in terms of **Interpretation** and **Application**.

1. **Interpretation**: A major challenge lies in effectively clustering and interpreting patterns of urban vibrancy as observed through floating population data, requiring a systematic approach to categorize patterns and establish criteria for similarity. Conventional methods, such as spatiotemporal analysis [2], [5], kernel density estimation [6], and point-of-interest (POI) association [7], have been developed to capture and characterize urban vibrancy. However, these approaches often struggle with the demands of real-time data streams and lack the granularity needed to interpret complex, high-dimensional, or image-based datasets.

2. **Application**: Another challenge is determining how urban vibrancy derived from floating population data can be applied in various domains. Recent works have explored applications in urban vibrancy similar to multi-task learning. For example, research on citywide crowd flow prediction has used spatial-temporal neural networks to model long-range dependencies [8], while other studies have employed deep multi-view networks for tasks like taxi demand prediction [9]. These applications underscore the potential of urban vibrancy data, yet they have not fully harnessed the insights that real-time floating population trends can offer for traffic prediction and similar dynamic applications.

This study introduces an innovative method for deriving Urban Vibrancy embeddings from real-time floating population data to enhance traffic prediction models. Specifically, we employ variational autoencoders (VAE) to compress this data into actionable embeddings. These embeddings are then integrated with long short-term memory (LSTM) networks to predict future states and applied within a sequence-to-sequence framework for traffic forecasting.

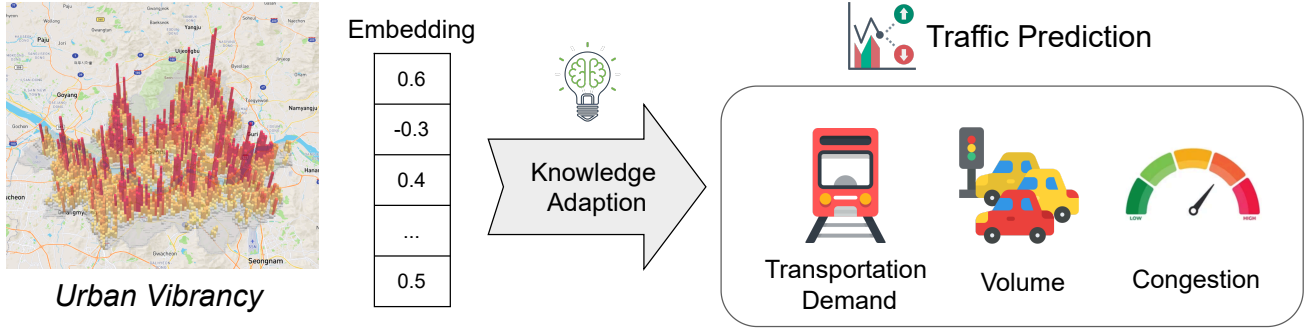Our contributions are as follows:

Fig. 1: Embedding Urban Vibrancy and Knowledge Adaption on Traffic Prediction.

1) Interpretation of Real-Time Embeddings: Through detailed principal component analysis (PCA), we visualize the generated embeddings, which reveal distinct temporal patterns such as weekday versus weekend trends, hourly variations, and seasonal shifts. This analysis provides intuitive insights into the underlying dynamics of urban vibrancy.
2) Dynamic Urban Vibrancy Embedding: By integrating VAEs and LSTMs, our approach embeds real-time urban data dynamically, enhancing the model's adaptability to changing urban conditions and improving its responsiveness to real-time data fluctuations.
3) Enhanced Model Responsiveness and Accuracy: Our method demonstrates substantial improvements in traffic prediction models, including RNN, DCRNN, GTS, and GMAN, by incorporating real-time urban insights. This makes it particularly suitable for the dynamic demands of smart city environments.
4) Comprehensive Experimental Validation: We validate our approach on real-world data from Seoul, Korea, showcasing the method's effectiveness. We also provide our code and dataset to the community to support reproducibility and further research [2].

This paper contributes to the field by offering a novel perspective on using urban vibrancy to drive advanced traffic prediction, facilitating a more nuanced and responsive approach to urban mobility analysis.

II. RELATED WORK

A. Urban Vibrancy using Floating Population

Urban vibrancy derived from floating population has become an increasingly important aspect of urban studies and planning. This concept refers to the dynamic energy and vitality that temporary or non-permanent residents bring to urban areas [3]. Floating populations, which include tourists, business travelers, and short-term migrants, contribute significantly to the diversity and intensity of human activities in cities [10]. These transient groups often engage in various social, economic, and cultural activities, thereby enhancing the overall

---

[2]https://github.com/suminhan/wka-net

vibrancy of urban spaces [11]. Researchers have utilized multisource urban big data, including mobile phone location data and social media check-ins, to analyze the spatiotemporal patterns of floating populations and their impact on urban vibrancy [12], [13]. Such studies have revealed that floating populations can significantly influence the attractiveness of neighborhoods, the use of public spaces, and the overall economic vitality of urban areas6. Understanding the relationship between floating populations and urban vibrancy is crucial for urban planners and policymakers, as it can inform decisions related to infrastructure development, public service provision, and strategies for enhancing the overall quality of urban life [3].

B. Knoweldge Adaption on Traffic Prediction

Conventional traffic prediction models like DCRNN [14], ASTGCN [15], and GraphWaveNet [16] incorporate external information as an additional input channel alongside traffic data but lack mechanisms to process this information into interpretable embeddings. While [17] introduces a neural network for cross-mode knowledge adaptation, it does not extend to graph-based spatio-temporal models. The MultiView Deep LSTM framework, designed for ride-hailing demand forecasting, captures features from order, speed, and weather views but lacks sufficient interpretability and does not fully exploit complex graph structures in spatio-temporal data [18].

III. PRELIMINARIES

The traffic prediction problem addressed in this study is defined as follows: The traffic value is represented as $X_t \in \mathbb{R}^{n_s \times n_f}$, where n_s denotes the number of sensors, and n_f represents multiple channels, such as "on" and "off" in the case of demand. Let $C_t \in \mathbb{R}^{n_w \times n_h \times n_c}$ represent the real-time floating population at timestamp t, where n_w and n_h are the width and height of the area, respectively, and n_c indicates the number of channels. Consequently, the traffic prediction task is defined as a p-to-q sequence prediction finding an optimal function h:

$$h(X_{p-t+1}, ..., X_t, C_{p-t+1}, ..., C_t) \rightarrow \hat{X}_{t+1}, ..., \hat{X}_{t+q}$$

IV. METHODS

In this study, we employ a Variational Autoencoder to extract Urban Vibrancy Embeddings, which are then used for traffic prediction. Specifically, we apply an LSTM-based sequence-to-sequence model to forecast future Urban Vibrancy Knowledge. The predicted embeddings are fed into the Decoder component of the pretrained Variational Autoencoder, where the error between the predicted and actual Floating Population values is calculated for training. Finally, we incorporate an Urban Knowledge-Informed Spatio-Temporal Embedding into existing models such as RNN, DCRNN, GTS, and GMAN, which we hypothesize will improve overall model performance.

Our Urban Vibrancy Knowledge extraction approach is inspired by the Vision-Memory-Control (VMC) framework proposed in [19], where components are divided into vision, memory, and control stages. In our method, embeddings of the current floating population scene serve as the "vision" component, and these embeddings are used to predict future scenes, replacing the control component with a prediction mechanism in a structure similar to VMC.
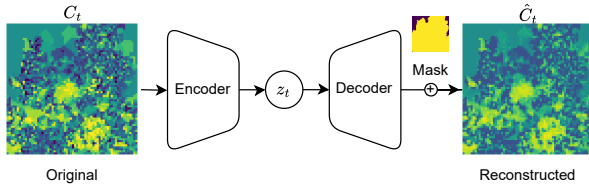
A. Variational Autoencoder (VAE)

Fig. 2: Variational Autoencoder

The Floating Population data is normalized to values between 0 and 1, similar to image pixel values, and represented in a grid format. Details regarding the preprocessing of the C values are outlined in Section V-A.

As illustrated in Figure 2, the Variational Autoencoder (VAE) model is applied to the refined Floating Population image at time τ, denoted as C_τ, to extract the embedding z_τ (i.e., $z_\tau \leftarrow \text{Encoder}(C_\tau) \in \mathbb{R}^d$). The Decoder then reconstructs $\hat{C}_\tau$ (i.e., $\hat{C}_\tau \leftarrow \text{Decoder}(z_\tau + \textit{Noise})$), and the model is trained using Mean Squared Error (MSE) loss between C_τ and $\hat{C}_\tau$. The VAE process involves the following steps:

$$\mathcal{L}_{\text{vae}} = \text{MSE}(C, \hat{C}) + D_{\text{KL}}(q_\phi(z|C) \parallel p(z))$$

In detail, an image mask with values of either 0 or 1 is applied to the final layer of the decoder, where these values indicate active or inactive floating population signals. This mask is concatenated into the final layer of the Decoder, followed by an additional dense layer to produce the final output. The mask helps the model avoid misinterpreting the data distribution and prevents training on null patterns from inactive cells.

B. LSTM for Forecasting Urban Vibrancy Embedding

The UVE-Seq2Seq model, depicted in Figure 3, leverages the VAE's Encoder (VE) and Decoder (VD) components to predict future Urban Vibrancy Embeddings ($\hat{z}_{t+1}, ..., \hat{z}_{t+q}$). Trained in Section IV-A, the model uses embeddings derived from the current Floating Population data to forecast future embeddings. During this process, the parameters of VE and VD remain frozen and are not updated, with only the Seq2Seq LSTM model undergoing training.

To predict these embeddings, the model first extracts an embedding z for each C_τ in the range $[t - p + 1, ..., t]$:

$$z_\tau \leftarrow \textbf{VE}(C_\tau)$$

Using these embeddings, the UV-ENCDEC model generates future embeddings ($\hat{z}_{t+1}, ..., \hat{z}_{t+q}$):

$$(\hat{z}_{t+1}, ..., \hat{z}_{t+q}) \leftarrow \textbf{UV-ENCDEC}(z_{t-p+1}, ..., z_t)$$

Each predicted embedding $\hat{z}_\tau$ is then transformed into an anticipated scene $\hat{C}_\tau$ using the VAE Decoder (VD):

$$\hat{C}_\tau \leftarrow \textbf{VD}(\hat{z}_\tau) \quad \text{for} \quad \tau \in [t+1, ..., t+q]$$

The model optimizes its predictions using Mean Squared Error (MSE) loss:

$$\mathcal{L}_{\text{UV-ENCDEC}} = \text{MSE}(C, \hat{C})$$

This sequence-to-sequence LSTM model is designed to capture and predict the future distributions of Floating Population, enhancing traffic prediction by embedding urban vibrancy data into the forecasting process.

C. Urban Vibrancy Knowledge Adaptation for Traffic Prediction

While some models incorporate external information by simply concatenating it with the existing input data, we found that this approach did not lead to significant performance improvements. Therefore, we developed the Spatio-Temporal Vibrancy Embedding (STVE) as a more effective way to adapt external knowledge. This embedding is then integrated into traffic prediction models to enhance their ability to utilize urban vibrancy information.

1) Spatio-Temporal Vibrancy Embedding (STVE): The extracted Urban Vibrancy Embedding (UVE) is derived from past Floating Population scenes for the previous p timesteps and forecasted for future q timesteps using the UV-ENCDEC model. This UVE is incorporated into traffic prediction models as supplementary information through Spatio-Temporal Embedding.

The concept of Spatio-Temporal Embedding, which plays a role similar to Positional Encoding in Transformers, is also utilized in works such as [20], [21]. It introduces markers that account for spatial and temporal differences. Here, we integrate Urban Vibrancy Embedding with temporal information to enhance traffic prediction.

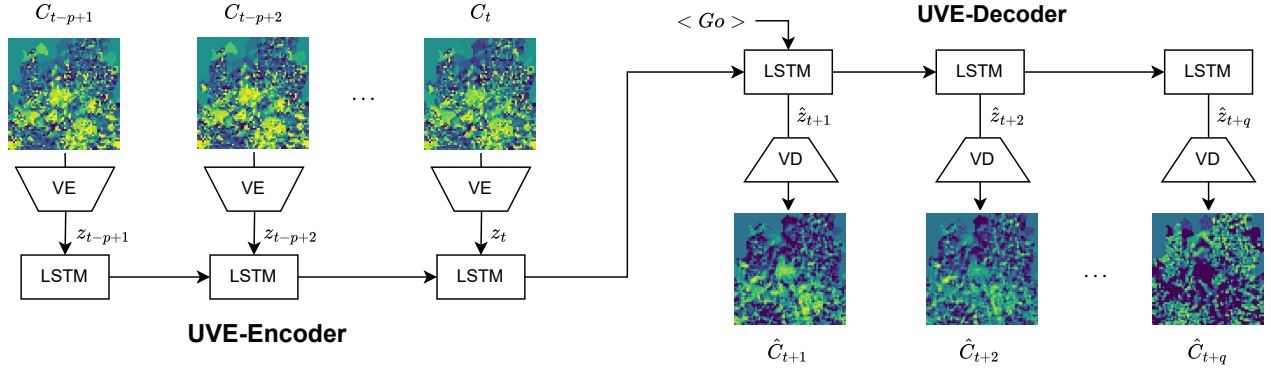The components of Spatio-Temporal Vibrancy Embedding are as follows:

Fig. 3: UVE-Seq2Seq: Forecasting Future Urban Vibrancy Embedding.

- $\mathbf{Z}_S \in \mathbb{R}^{n_s \times n_s}$: Sensor encoding, an identity matrix of size $n_s \times n_s$, which is a unique one-hot embeddings for each sensor.
- $\mathbf{Z}_T \in \mathbb{R}^{(p+q) \times (7+24)}$: Timestep embedding, where the day-of-week and time-of-day (24 hours) for each timestep are encoded using one-hot encoding into vectors of sizes $\mathbb{R}^7$ and $\mathbb{R}^{24}$, respectively.
- $\mathbf{Z}_V = (z_{t-p+1}, \ldots, z_t, \hat{z}_{t+1}, \ldots, \hat{z}_{t+q}) \in \mathbb{R}^{(p+q) \times d}$: Urban Vibrancy Embedding, consisting of embeddings from both observed and forecasted data.

The final Spatio-Temporal Urban Vibrancy Embedding ($\mathbf{Z}_{STV}$) combines these components as follows:

$$\mathbf{Z}_{STV} = f_1(\mathbf{Z}_S) + f_2(\mathbf{Z}_T || \mathbf{Z}_V) \in \mathbb{R}^{n_s \times (p+q) \times d}$$

where each f_1, f_2 is a 2-stacked fully connected layers with batch normalization, and $||$ denotes concatenation.

2) Application to Traffic Prediction Models: The Urban Knowledge-Informed Spatio-Temporal Embedding is integrated into various traffic prediction models as follows:

1) **GRU**: $\mathbf{Z}_{TV} = f(\mathbf{Z}_T || \mathbf{Z}_V) \in \mathbb{R}^{(p+q) \times d}$ is concatenated with the input at each GRU step to enhance temporal context.
2) **DCRNN, GTS**: In the Encoder, each input timestep X is expanded to d-dimensions using fully connected layers. These transformed inputs are then combined with urban vibrancy embedding as $f(X) + \mathbf{Z}_{STV,1:p} \in \mathbb{R}^{n_s \times p \times d}$. In the Decoder, $\mathbf{Z}_{STV,p+1:p+q}$ is utilized in place of the GO token.
3) **GMAN**: $\mathbf{Z}_{STV}$ replaces the traditional Spatio-Temporal Encoding (STE) to enrich the model with urban vibrancy information.

By incorporating STVE into these models, we enhance their ability to predict traffic patterns with an additional layer of urban context, leveraging real-time vibrancy information to improve accuracy and robustness. Detailed implementations can be referred from our public repository.

TABLE I: Data Statistics

	Floating Population	Subway Demand	Traffic Volume	Traffic Speed
# Nodes	64×64	275	73	263
Channel	1	2	1	1
Mean	570.99	725.47	1866.93	28.25
Std	1092.88	1129.04	1480.88	13.77
Timespan	Jan 1, 2017 $\sim$ Dec 31, 2018 (hourly, 17,520 steps)			

V. EXPERIMENTAL SETTINGS

A. Dataset

This study utilizes a comprehensive dataset from Seoul, which captures floating population data over a two-year period (2017-2018) across a central 22 km by 22 km area that includes major urban districts such as Gangnam and Gangbuk.[3] The land use of our research area is also depicted in Figure 5. The dataset records hourly variations in population density, amounting to $2 \times 365 \times 24$ observations, and is supplemented with subway ridership, traffic volume, and traffic speed data to evaluate improvements in predictive accuracy. Rigorous preprocessing was performed, with missing values imputed through temporal averaging of adjacent weeks to maintain data continuity and integrity. Table I presents key statistics for each data source, highlighting variability and providing essential context for model development. This high-resolution dataset lays a robust foundation for embedding real-time urban vibrancy insights into traffic prediction, enabling an in-depth exploration of how dynamic urban activity patterns impact traffic flows in one of the world's most vibrant metropolitan areas.

1) Preprocessing Floating Population: To preprocess the Floating Population variable C, we first apply z-score normalization for each cell individually using its respective mean (μ_{ij}) and standard deviation (σ_{ij}) from original values $C^o \in \mathbb{R}^{8760 \times 64 \times 64 \times 1}$: $C'_{ij} = \frac{C^o_{ij} - \mu_{ij}}{\sigma_{ij}}$. The normalized result is then capped within ± 2 standard deviations: $C'' =$

---

[3]The region is defined by $[126.8721, 37.4723, 127.1219, 37.6709]$.

(a) Weekday   (b) Weekend   (c) PCA Component 1 (x-axis): central Seoul areas.   (d) PCA Component 2 (y-axis): outer Seoul areas.
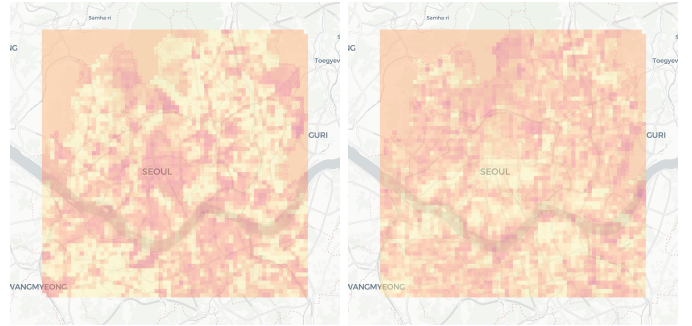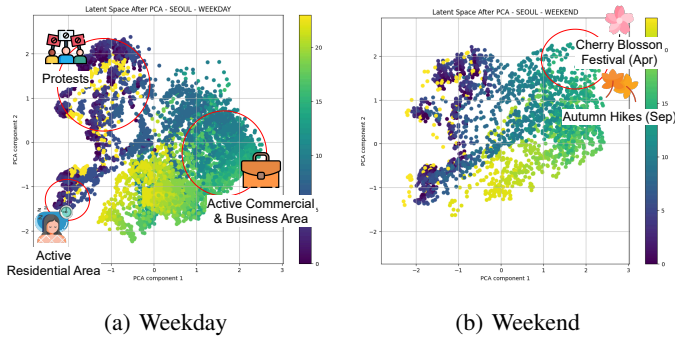
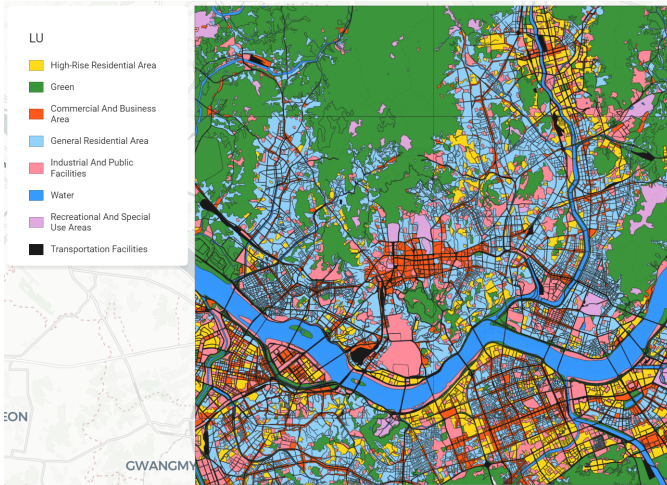Fig. 4: Comparison of Urban Vibrancy Embedding on Weekdays and Weekends

Fig. 5: Land Use of our Research Area

$\min(\max(C', -2), 2)$. Finally, the data is scaled to the [0, 1] range using $C = \frac{C'' + 2}{4}$.

2) Graph Construction: The graphs for Subway Demand, Traffic Volume, and Traffic Speed are constructed as follows: Subway Demand graphs connect adjacent stations and transfer stations, while Traffic Volume graphs link traffic sensors based on their proximity. For Traffic Speed, road segments are connected according to their physical distances.

B. Settings

We used the LibCity library[4] for our experiments. The model was trained with parameters $p = 6$ and $q = 6$, with each unit representing one hour. For each model, we applied the default learning rate and other hyperparameters provided by LibCity, ensuring that weekday and time-of-day information were consistently utilized.

The baseline models included RNN, DCRNN [14], GTS [22], and GMAN [21]. The dataset was divided into training, validation, and test sets with a ratio of 0.4, 0.1, and 0.5, respectively. The training and validation sets consisted of data

[4]https://libcity.ai/

from 2017, while the test set used only 2018 data. This setup ensured that the model was tested on a distinct, future year to evaluate its predictive performance.

Unlike traditional traffic prediction approaches, where data is split sequentially, we shuffled the training and validation sets within the 2017 data. This shuffling ensured that the validation set encompassed data from all seasons, rather than being biased toward winter at the end of the year. By including seasonal diversity, we aimed for a model that learns a balanced representation across various patterns and conditions throughout the year. The evaluation metric used for this experiment was Mean Absolute Error (MAE).

VI. RESULT

A. Interpretation of Urban Vibrancy Embedding

In this paper, real-time LTE floating population data, which reflects urban vibrancy, is spatio-temporal data. To facilitate a deeper understanding of this data and to enable multi-faceted interpretation, we employed data visualization techniques. Principal Component Analysis (PCA) was used to interpret and analyze the data through visualization.

1) Embedding Analysis of Temporal Urban Vibrancy Patterns: Figure 4 reveals distinct embedding patterns for weekdays and weekends, displayed in subplots (a) and (b), respectively. This facilitates a clear comparison between these temporal contexts. Figures 4c and 4d showcase the decoded results of PCA1 and PCA2. Plus, we find 5 PCA components can represent the most influential components in the dataset based on reconstruction error with a 99.7% threshold.

In South Korea, urban areas are categorized into residential, commercial, industrial, and green zones, while non-urban areas include managed, agricultural/forestry, and natural conservation zones. PCA1 primarily reflects population activity in Seoul's commercial and industrial zones, whereas PCA2 captures activity in residential, green, and non-urban zones. Thus, PCA1 and PCA2 serve as proxies for population movement within these distinct urban and non-urban sectors, highlighting temporal urban vibrancy patterns that vary by zone type.
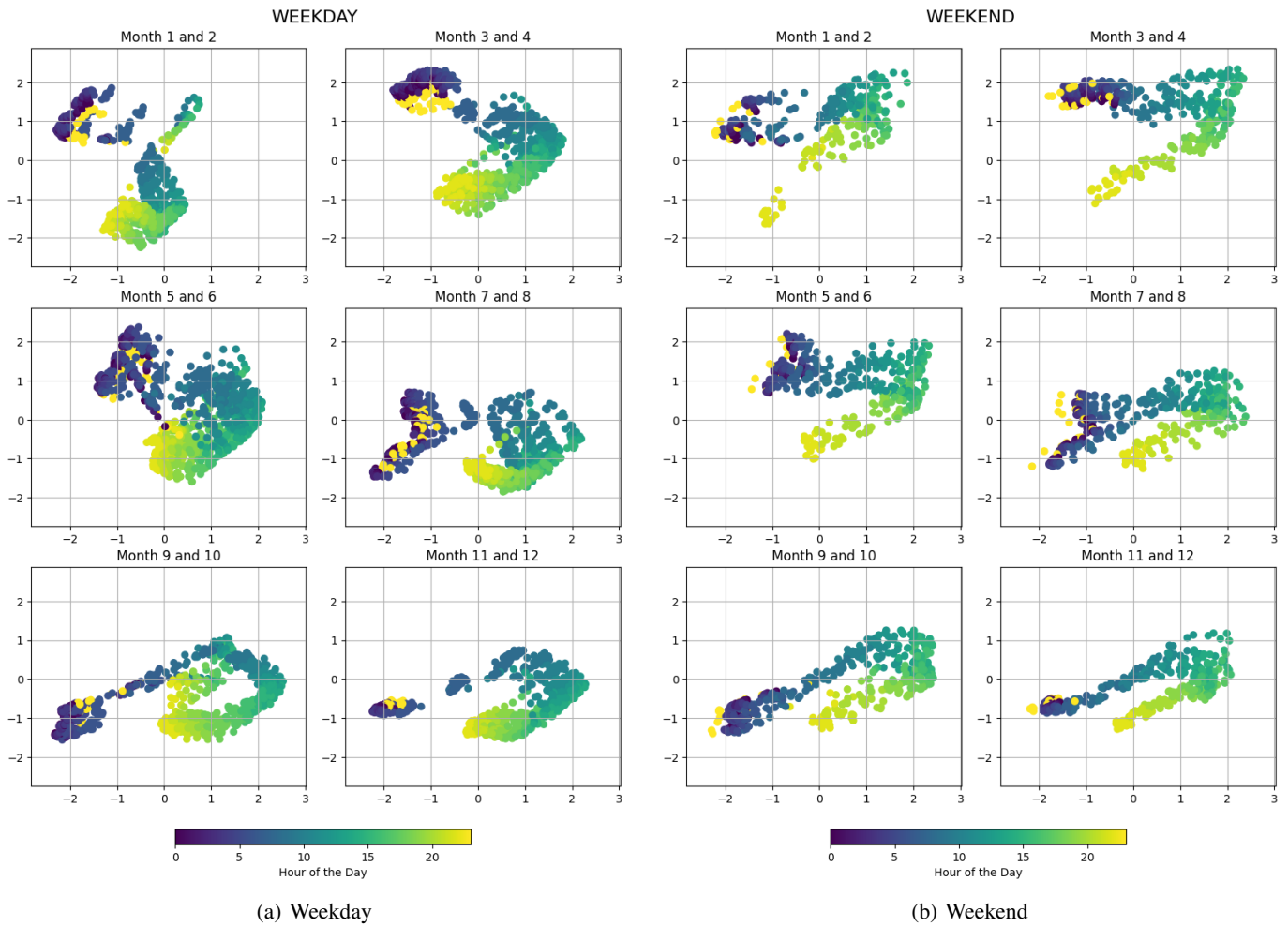
(a) Weekday

(b) Weekend

Fig. 6: Comparison of Urban Vibrancy Embedding on Weekdays and Weekends of year 2017.

2) Temporal Evolution and Seasonal Variations: Figure 6 shows the temporal evolution of urban vibrancy embeddings for weekdays and weekends, visualized at bimonthly intervals. Weekday embeddings tend to form concentrated clusters, indicating population convergence in specific areas due to work and commuting routines. Conversely, weekend embeddings are more dispersed, suggesting a broader array of activities with less structured movement as people engage in leisure and recreational pursuits.

A color gradient from purple (00:00) to yellow (23:00) illustrates the daily temporal flow, revealing a rotation in embedding values over time. Peak activity is observed from 10:00 to 15:00, while reduced activity occurs from 00:00 to 05:00, indicating population migration outside Seoul. Notably, from 20:00 to 23:00, a sharp increase in activity reflects significant movement in and out of Seoul. Weekday patterns show counterclockwise rotations, while weekend patterns display clockwise rotations with a noticeable upward trend from January to June in areas like parks and the Han River, due to outdoor activities.

Seasonal variations are also evident; winter months (January and February) show more concentrated clusters, while summer months (July and August) exhibit broader distributions, reflecting increased outdoor activities during warmer periods. These seasonal shifts correspond to changes in mobility, with summer promoting more outdoor movement and dispersed activities, while colder winter months limit outdoor engagements.

3) Urban Planning Implications and Impact of Unique Events: The observed temporal patterns provide valuable insights for urban planning and management. Weekday clustering suggests that urban transportation and infrastructure can be optimized to meet predictable commuting demands. In contrast, the dispersed weekend patterns call for a more flexible approach to accommodate varied leisure activities. Additionally, businesses might consider adjusting their hours or staffing levels to align with these urban vibrancy trends, particularly during weekends with heightened activity.

Unique events can also significantly affect urban vibrancy. For instance, from January to March 2017, large-scale protests of impeachment of ex-president in Seoul drew approximately 800,000 people to Jongno-gu. The embeddings from this period show unique rotational patterns compared to 2018 data,

reflecting the impact of this rare collective event on urban dynamics. These insights emphasize the need for responsive planning strategies that account for both routine and exceptional events, supporting a more adaptive and resilient urban environment.

In summary, this study offers a detailed view of urban vibrancy, capturing how movement patterns fluctuate across weekdays, weekends, and seasons. The findings provide valuable guidance for urban planning, infrastructure management, and public policy, contributing to a more adaptable urban environment that meets the diverse needs of city inhabitants.

B. Traffic Prediction Enhancement

Table II compares the performance of various models (RNN, DCRNN, GTS, GMAN) with different embeddings (None, TE, VE, TVE) for predicting SUBWAY, VOLUME, and SPEED metrics over horizons of 1, 2, 3, and 6.

The RNN model shows high errors without embeddings, with metrics like SUBWAY errors at 352.47 (horizon 1) and SPEED errors at 6.133. Adding TE, VE, or TVE improves performance, with TVE generally yielding the best results, such as reducing SUBWAY error to 73.34. DCRNN performs better than RNN, especially with VE, achieving a SUBWAY error of 91.26 (horizon 1) and SPEED error of 4.38. GTS generally has lower errors and benefits significantly from VE and TVE, with SUBWAY errors dropping to 57.77 and VOLUME errors to 138.21 (horizon 6) with TVE. GMAN shows good performance for VOLUME (169.41 with VE) and SPEED predictions but has higher SUBWAY errors compared to GTS and DCRNN.

Among embeddings, TE alone reduces errors but is outperformed by VE and TVE. VE improves predictions across all models, especially with DCRNN and GTS. TVE achieves the lowest errors in most cases, like a 57.77 SUBWAY error for GTS, combining temporal and vibrancy data effectively.

Errors generally increase with horizon length, but embeddings like VE and TVE help control this, particularly for VOLUME and SPEED. For specific metrics, GTS with VE or TVE performs best for SUBWAY, while DCRNN with VE or TVE excels in SPEED predictions.

In summary, embedding strategies, especially VE and TVE, significantly enhance predictive accuracy, particularly for GTS and DCRNN models. VE alone is effective, but TVE provides the lowest errors by adding a comprehensive context.

VII. Conclusion

In conclusion, this research presents a novel framework for enhancing traffic prediction through the use of Urban Vibrancy Embeddings derived from real-time floating population data. By employing Variational Autoencoders (VAE) to compress high-dimensional urban data into meaningful embeddings and integrating these with Long Short-Term Memory (LSTM) networks within a sequence-to-sequence forecasting framework, we capture and leverage the dynamic patterns of urban vibrancy effectively.

Our findings highlight the value of this approach in several key areas. First, the use of principal component analysis (PCA) allows for an intuitive interpretation of the temporal patterns embedded in urban vibrancy data, such as differences between weekdays and weekends, as well as seasonal and hourly shifts. Second, the dynamic embedding of real-time urban data improves model responsiveness and adaptability to fluctuating urban conditions, which is crucial for applications in smart city environments. Third, our experiments demonstrate notable improvements in the predictive accuracy of models like RNN, DCRNN, GTS, and GMAN, showing the versatility and effectiveness of our approach across different modeling architectures.

Through rigorous validation on real-world data from Seoul, we confirm the practical applicability and robustness of our method. By releasing our code and dataset, we aim to support reproducibility and encourage further exploration in this area. This study advances the field by providing a more nuanced and responsive approach to traffic prediction, integrating urban vibrancy insights to enhance the adaptability of models to the dynamic realities of urban mobility.

References

[1] C. Bergroth, O. Järv, H. Tenkanen, M. Manninen, and T. Toivonen, "A 24-hour population distribution dataset based on mobile phone data from helsinki metropolitan area, finland," *Scientific data*, vol. 9, no. 1, p. 39, 2022.

[2] W. K. Lee, S. Y. Sohn, and J. Heo, "Utilizing mobile phone-based floating population data to measure the spatial accessibility to public transit," *Applied geography*, vol. 92, pp. 123–130, 2018.

[3] A. Barreca, R. Curto, and D. Rolando, "Urban vibrancy: An emerging factor that spatially influences the real estate market," *Sustainability*, vol. 12, no. 1, p. 346, 2020.

[4] R. Fu, X. Zhang, D. Yang, T. Cai, and Y. Zhang, "The relationship between urban vibrancy and built environment: an empirical study from an emerging city in an arid region," *International journal of environmental research and public health*, vol. 18, no. 2, p. 525, 2021.

[5] K.-S. Lee, S. Y. You, J. K. Eom, J. Song, and J. H. Min, "Urban spatiotemporal analysis using mobile phone data: Case study of medium- and large-sized korean cities," *Habitat International*, vol. 73, pp. 6–15, 2018.

[6] Y. Zhang, W. Zhong, D. Wang, and F.-T. Lin, "Understanding the spatiotemporal patterns of nighttime urban vibrancy in central shanghai inferred from mobile phone data," *Regional Sustainability*, vol. 2, no. 4, pp. 297–307, 2021.

[7] C. Gao, S. Li, M. Sun, X. Zhao, and D. Liu, "Exploring the relationship between urban vibrancy and built environment using multi-source data: Case study in munich," *Remote Sensing*, vol. 16, no. 6, p. 1107, 2024.

[8] J. Feng, Y. Li, Z. Lin, C. Rong, F. Sun, D. Guo, and D. Jin, "Context-aware spatial-temporal neural network for citywide crowd flow prediction via modeling long-range spatial dependency," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 16, no. 3, pp. 1–21, 2021.

[9] H. Yao, F. Wu, J. Ke, X. Tang, Y. Jia, S. Lu, P. Gong, J. Ye, and Z. Li, "Deep multi-view spatial-temporal network for taxi demand prediction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, 2018.

TABLE II: Performance Comparison

		SUBWAY				VOLUME				SPEED			
	Horizon	1	2	3	6	1	2	3	6	1	2	3	6
RNN	NONE	352.47	352.96	351.77	351.71	196.39	212.95	220.95	231.52	6.1336	6.6175	6.9996	7.2999
	TE	77.33	80.59	**82.31**	**84.11**	190.47	197.21	201.07	209.95	**5.7842**	**5.9992**	**6.1433**	**6.3660**
	VE	78.76	85.26	92.26	98.22	189.87	197.06	200.50	204.83	6.1837	6.9754	7.5057	7.7835
	TVE	**73.34**	**77.44**	83.29	90.54	**187.40**	**193.68**	**196.81**	**199.58**	5.7868	6.0699	6.4011	6.8643
DCRNN	NONE	142.20	207.63	260.95	359.25	152.84	187.81	206.16	239.85	5.3059	6.7597	7.4098	8.4015
	STE	131.19	190.28	234.26	288.31	**145.08**	**164.11**	**174.69**	**191.94**	4.3802	5.1023	5.4427	5.9109
	SVE	**91.26**	**102.45**	**111.70**	**120.22**	163.80	183.82	192.08	202.13	4.7570	5.5517	5.8617	6.2579
	STVE	192.80	234.67	258.78	284.82	162.63	176.39	183.17	192.94	4.4395	5.2042	5.5707	5.9860
GTS	NONE	69.68	79.80	85.14	91.10	161.11	210.87	235.88	276.99	4.7113	5.6962	6.1930	6.7567
	STE	65.98	70.54	73.76	**79.12**	142.23	166.38	175.12	190.28	**4.2430**	**4.9347**	**5.2724**	**5.7986**
	SVE	62.73	72.29	77.92	83.53	155.11	179.22	188.87	200.35	4.4471	5.2555	5.6333	6.0569
	STVE	**57.77**	**67.32**	**72.87**	80.38	**138.21**	**163.32**	**174.82**	**187.72**	4.3116	5.0549	5.4554	5.9742
GMAN	NONE	523.22	515.00	512.95	533.61	426.75	341.96	311.69	549.61	10.0531	9.3465	9.2520	13.0090
	STE	312.69	**307.75**	**300.90**	**277.29**	169.41	184.02	192.87	205.50	5.1138	5.6817	5.9832	6.4349
	SVE	566.58	567.34	565.11	564.72	**164.71**	**177.08**	**185.95**	**200.01**	**5.0487**	5.6810	6.0083	6.6369
	STVE	**307.84**	312.39	309.23	311.89	165.76	178.50	186.91	199.11	5.0951	**5.6375**	**5.9554**	**6.3519**

[10] L. Chen, L. Zhao, Y. Xiao, and Y. Lu, "Investigating the spatiotemporal pattern between the built environment and urban vibrancy using big data in shenzhen, china," *Computers, Environment and Urban Systems*, vol. 95, p. 101827, 2022.

[11] W. Tu, T. Zhu, J. Xia, Y. Zhou, Y. Lai, J. Jiang, and Q. Li, "Portraying the spatial dynamics of urban vibrancy using multisource urban big data," *Computers, Environment and Urban Systems*, vol. 80, p. 101428, 2020.

[12] W. S. Lee, "How does air pollution affect floating population in metropolitan city: embedding-based approach," *Clean Technologies and Environmental Policy*, pp. 1–11, 2024.

[13] K. Jang, H. Suh, F. Haddad, *et al.*, "Urban street clusters: unraveling the associations of street characteristics on urban vibrancy dynamics in age, time, and day," *Urban Info*, vol. 3, p. 27, 2024.

[14] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*, 2018.

[15] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, pp. 922–929, 2019.

[16] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, 2019.

[17] C. Li, L. Bai, W. Liu, L. Yao, and S. T. Waller, "Knowledge adaption for demand prediction based on multi-task memory neural network," in *Proceedings of the 29th ACM international conference on information & knowledge management*, pp. 715–724, 2020.

[18] Y. Wu, H. Zhang, C. Li, S. Tao, and F. Yang, "Mvdlstm: Multiview deep lstm framework for online ride-hailing order prediction," *The Journal of Supercomputing*, vol. 78, no. 6, pp. 8531–8559, 2022.

[19] D. Ha and J. Schmidhuber, "Recurrent world models facilitate policy evolution," in *Advances in Neural Information Processing Systems 31*, pp. 2451–2463, Curran Associates, Inc., 2018.

[20] S. Han, Y. Park, M. Lee, J. An, and D. Lee, "Enhancing spatio-temporal traffic prediction through urban human activity analysis," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pp. 689–698, 2023.

[21] C. Zheng, X. Fan, C. Wang, and J. Qi, "Gman: A graph multi-attention network for traffic prediction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, pp. 1234–1241, 2020.

[22] C. Shang, J. Chen, and J. Bi, "Discrete graph structure learning for forecasting multiple time series," in *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*, OpenReview.net, 2021.