engineering proceedings

MDPI

Proceeding Paper

Random Seed Generation for Convergence of Large-Scale People Flow Prediction Using Generative Adversarial Networks and Rationality of Output †

Yu-Hsuan Lin [1,*], Yi-Chung Chen [1], Tzu-Yin Chang [2] and Rong-Kang Shang [2]

[1] Department of Computer Science and Engineering, National Chung Hsing University, Taichung 40227, Taiwan; chenyich@nchu.edu.tw
[2] National Science and Technology Center for Disaster Reduction, New Taipei 231007, Taiwan; geoct@ncdr.nat.gov.tw (T.-Y.C.); rkshang@ncdr.nat.gov.tw (R.-K.S.)
* Correspondence: kevin22362009@gmail.com
† Presented at 8th International Conference on Knowledge Innovation and Invention 2025 (ICKII 2025), Fukuoka, Japan, 22–24 August 2025.

Abstract

Since the emergence of big data analytics, forecasting future population distributions in specific regions has become a popular prominent research topic. Accurate predictions offer benefits for urban planners, traffic management authorities, and commercial stakeholders. However, most studies have concentrated on population dynamics within isolated locations, often overlooking broader spatial fluctuations across larger geographic areas. This narrow scope limits the practical utility of such predictions. Therefore, generative adversarial networks (GANs) have been employed to estimate population counts across multiple locations within expansive regions. Despite their potential, many GAN-based models encounter significant challenges when tasked with predicting numerous locations simultaneously, resulting in prolonged training times or failure to achieve convergence. To address these limitations, we developed a novel random number generation method to improve the training efficiency and convergence stability of GANs. We also set a new identification criterion to ensure that the large-scale population distributions generated by GAN closely reflect real-world conditions. The developed model in this study was validated using actual telecommunications-based pedestrian flow data from Taiwan, demonstrating its effectiveness and practical feasibility.

Keywords: pedestrian flow; GAN; population distribution

check for updates

Academic Editors: Teen-Hang Meen, Chun-Yen Chang and Cheng-Fu Yang

Published: 24 February 2026

Citation: Lin, Y.-H.; Chen, Y.-C.; Chang, T.-Y.; Shang, R.-K. Random Seed Generation for Convergence of Large-Scale People Flow Prediction Using Generative Adversarial Networks and Rationality of Output †. *Eng. Proc.* **2025**, *120*, 69. http://doi.org/10.3390/engproc2025120069

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Since the advent of big data analytics, forecasting future population distributions in regions has been important. Accurate predictions offer substantial value to urban planners, traffic management authorities, and commercial stakeholders by informing infrastructure development, resource allocation, and strategic planning. However, previous studies have focused on population dynamics within isolated locations, often neglecting broader spatial fluctuations across larger geographic areas. This limited scope reduces the practical applicability of such predictive models.

To address this gap, generative adversarial networks (GANs) are used to estimate population counts across multiple locations within expansive regions (Figure 1). While

GANs offer promising capabilities, significant challenges exist with simultaneously predicting data for numerous locations. These challenges include prolonged training times and difficulties in achieving stable convergence.

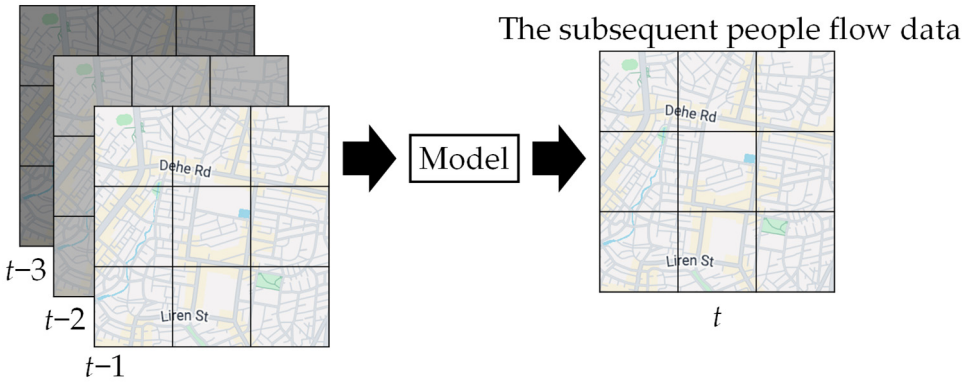


Figure 1. Pedestrian flow prediction process.

To address the challenges, we developed a novel random number generation model that enhances the training efficiency and convergence reliability of GANs. The model introduces a new identification criterion to ensure that the large-scale population distributions generated by GAN closely resemble actual conditions. The developed model was validated using real-world telecommunications-based pedestrian flow data from Taiwan, demonstrating its effectiveness and practical feasibility.

Various supervised machine learning models are used to forecast pedestrian flow. For instance, Vermeulen et al. [1] employed the K-nearest neighbor algorithm to predict passenger volumes during specific events. Li et al. [2] developed a pedestrian flow prediction model based on long short-term memory (LSTM), while Zhang et al. [3] utilized a spatio-temporal residual network (ST-ResNet) to forecast citywide passenger movement. However, Wang et al. [4] and Wang et al. [5] argued that these approaches depend exclusively on historical data and lack adaptability to dynamic fluctuations in pedestrian flow. Consequently, they advocated for the use of generative AI techniques, which are better suited to learning realistic variations in crowd behavior and are not constrained by past states. Their experimental results support the efficacy of this approach.

While generative models improve prediction accuracy for variable pedestrian flow, they introduce challenges. Specifically, GAN models struggle with convergence and may produce plausible results for individual grid cells that fail to align with the overall spatial distribution. These limitations underscore the need for improved strategies. To address these issues, we incorporated historical pedestrian flow values into the random variable inputs of GANs. By embedding authentic past data into the generative process, the developed model facilitates convergence and ensures that the generated population distributions remain closely aligned with real-world conditions. Finally, the model is evaluated using actual data to validate its effectiveness and reliability.

2. Method

We extended the architecture of GAN [5] to validate the efficacy of the proposed method. Figure 2 presents a workflow of the architecture. Initially, a series of preprocessing techniques is employed to systematically organize the original telecommunications pedestrian flow data into grid data for the entire city, hereinafter referred to as pedestrian flow slice data. Subsequently, the slice of data is input into the designated GAN for training purposes, thereby completing the training of the overall model.

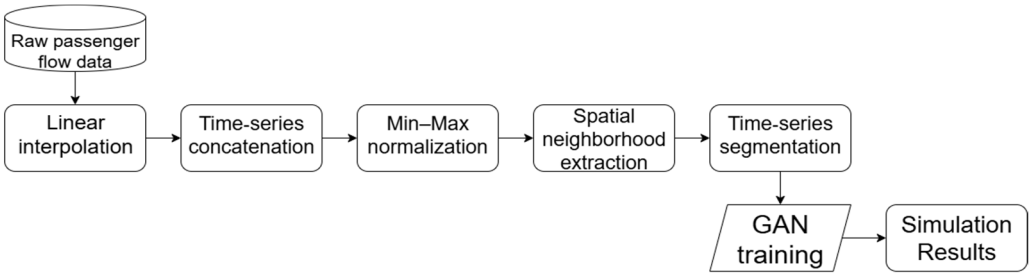


Figure 2. Workflow of the proposed method.

2.1. Data Preprocessing

For comparative analysis with the model proposed in Ref. [5], linear interpolation was used to impute missing values in the crowd flow dataset. Time series construction and normalization, where crowd flow data from the same grid across different time points are concatenated into a time series and normalized such that the maximum value is 1 and the minimum is -1. Spatial neighborhood extraction, in which a central grid is selected and the crowd flow time series of its $m \times m$ spatial neighborhood is chosen, is retrieved. The choice of m must reflect the characteristics of the local business district. A larger value of m introduces noise and distorts modeling outcomes, while a smaller value may fail to capture the spatial dynamics of the district. Temporal segmentation is conducted, where the $m \times m$ time series data are divided based on their timestamps. Given N total time points, the dataset is split into N slices, each representing the spatial crowd flow distribution at a specific time.

2.2. Model Structure

The GAN architecture used in this study resembles that of Ref. [5] (Figure 3). The primary distinction pertains to the input and layer structure of the model. The subsequent explanation delineates the various components of our model. First, the input to the generator G consists of an $m \times m$ matrix with two channels; one channel contains the actual crowd flow data from the previous time point $(t-1)$, while the other channel comprises an $m \times m$ random noise matrix intended to introduce diversity in generation. The output of generator G is also an $m \times m$ crowd flow distribution matrix, representing the predicted crowd flow at the subsequent time point. Regarding the discriminator D, it receives an $m \times m$ matrix with two channels as input. The first channel corresponds to the crowd flow at time $(t-1)$, and the second channel contains either the real or generated crowd flow at time t. It outputs a binary classification probability to ascertain whether the input sample is real or generated. In the training process, the parameters of both the generator and discriminator are updated alternately: D is held constant during the update of G to encourage the generation of more realistic samples; conversely, G remains fixed while D is updated to enhance the discriminator's ability to identify real samples. This training cycle is repeated until the generator attains the capability to produce high-quality short-term crowd flow predictions under the specified historical conditions.

The details of the developed model in this study are different from those of Ref. [5]. Both the generator and discriminator in Ref. [5] utilize convolutional neural networks, comprising input layers, convolutional layers, flattening layers, and fully connected layers. However, the use of convolutional and flattening layers is inappropriate for modeling pedestrian flow distribution. The significance of convolutional layers involves identifying similar patterns across different regions, yet map regions often contain distinct stores or facilities, rendering the appearance of such patterns unlikely. Consequently, except for employing convolutional layers at the preliminary stages of the generator and discriminator

to incorporate inputs from two channels, we utilized fully connected layers in the remaining sections to analyze variations in pedestrian flow across different grids.

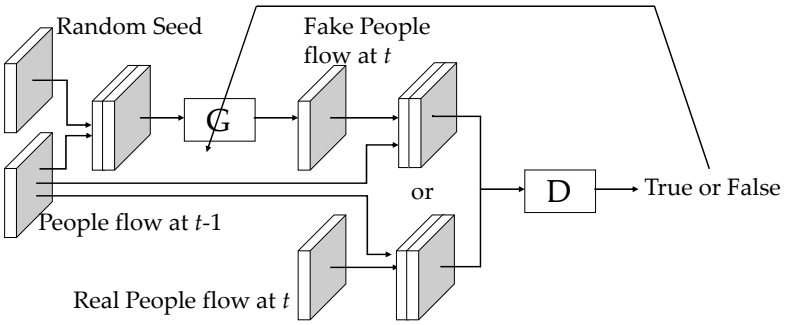


Figure 3. GAN architecture used in this study.

2.3. Model Validation Method

To accurately assess the contribution of conditional input to the performance of the model, we used a control group in which the generator *G* of the control group receives an $m \times m$ matrix of single-channel random noise, and the output is an $m \times m$ crowd flow prediction matrix. The discriminator D within the control group accepts an $m \times m$ single-channel matrix as input, which is either the real crowd flow values or the fake crowd flow generated by the generator, and produces a binary classification probability to determine whether the input sample is real or generated. Aside from the conditional channel, the network structure, layer configuration, and all training hyperparameters of the control group are consistent in the main model. This configuration ensures that the evaluation isolates the effect of incorporating historical conditional information on convergence speed, model stability, and prediction accuracy. By comparing performance metrics against a control group, the advantages of conditional GANs in pedestrian flow prediction tasks are demonstrated. To enhance the reliability of the experimental results, ten models were trained for each region using different initial parameter settings. This approach mitigates the uncertainty introduced by the random initialization of GAN weights. The root mean square error (RMSE) from the models was averaged, and the resulting curve was plotted to suppress fluctuations arising from individual runs, thereby ensuring more robust and consistent outcomes.

3. Simulations

3.1. Dataset

We utilized authentic telecommunications data obtained from Taipei, Taiwan, for spatiotemporal prediction analysis. The dataset encompasses a longitudinal range from 120.990° E to 122.005° E and a latitudinal range from 24.940° N to 25.165° N, spanning the period from 1 January 2018 to 31 December 2019. Each cell within the generated image grid represents an area of 500 by 500 m on the ground. We used four high-traffic commercial zones: the Shilin Business District, the Xinyi Business District, the Zhongxiao Business District, and the Banqiao Houzhan Business District. These locations were chosen due to their diversity of customer groups present, which effectively reflects the varied behavior patterns of people within different scenarios in the Taipei metropolitan area. Hence, these areas serve as ideal sites for validating the people flow prediction model.

3.2. Model Hyperparameters

This generator accepts a tensor of shape $3 \times 3 \times 2$ (comprising the actual flow of individuals at the previous moment and random noise of identical dimensions) as input. Initially, a 1×1 convolutional layer with 8 channels and padding set to the same is

employed to facilitate cross-channel feature fusion without altering the spatial dimensions. This is followed by Batch Normalization and a tanh activation function to constrain the output within the range $[-1, 1]$. Subsequently, the features are flattened into vectors, and two fully connected layers with 32 and 16 hidden units, respectively, each accompanied by Batch Normalization and tanh activation, are utilized to enable interaction and compression of global information. The final fully connected layer outputs a 9-dimensional vector, which is reshaped into a $3 \times 3 \times 1$ prediction matrix.

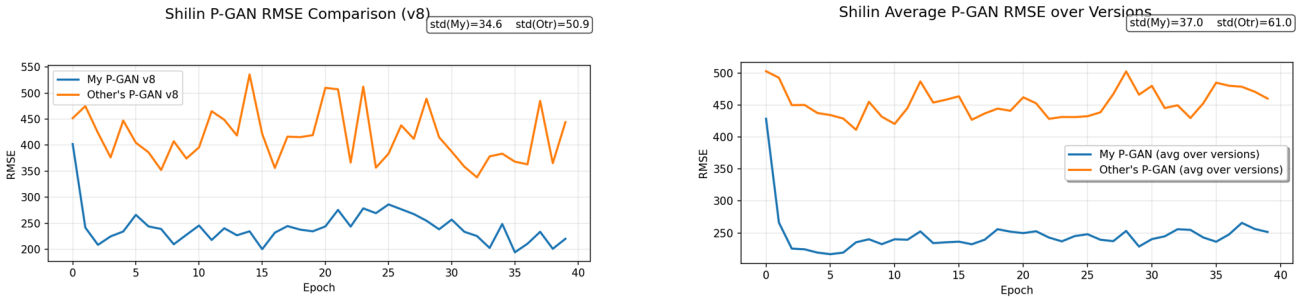
In the discriminator, we determined the number of convolution channels, the number of fully connected layer units, the activation function, and the normalization method to achieve the best distinction between real and generated samples. The discriminator takes a tensor of shape $3 \times 3 \times 2$ (including conditional channels and sample channels) as input. It passes through a layer of 1×1 convolution (output channel number 8, padding = 'same'), which completes information fusion only in the channel dimension without changing the spatial size, so that the model quickly learns the local correspondence between conditions and samples; then batch normalization is added to stabilize training and accelerate convergence. Then, the convolutional feature map is flattened into a one-dimensional vector and sequentially input into two fully connected layers containing 32 and 16 hidden units. Each layer uses the Leaky rectified linear unit (ReLU) ($\alpha = 0.2$) activation function, which retains negative information, prevents neurons from dying, and effectively alleviates internal covariate shift with the assistance of BatchNormalization. Finally, through a classification layer containing 2 output units and softmax activation, the output corresponds to the probability of "real" and "generated" samples.

To ensure a fair evaluation of the effectiveness of the conditional input, the control group model is entirely consistent with this model concerning the overall architecture and hyperparameter configurations of each layer. The sole distinction lies in the fact that its input comprises only a single-channel 3×3 tensor and does not include the historical flow condition channel.

3.3. Simulation Results

In this study, we used RMSE as an indicator to evaluate the model. Specifically, for each training batch, the generator uses the historical passenger flow data at the previous time $(t - 1)$ as a conditional input and generates a 3×3 grid passenger flow prediction matrix corresponding to time t. Then, this prediction matrix is flattened with the true observation matrix after inverse normalization, and the RMSE of each batch is calculated, and finally all batch results are averaged to obtain the average prediction error of the model on the entire data set. This approach not only quantifies the prediction deviation of individual batches but also suppresses occasional fluctuations through the averaging process between batches, ensuring that the evaluation results are more robust and comparable. The control group also uses the same RMSE calculation process, only comparing the 3×3 passenger flow matrix generated by random noise with the true passenger flow matrix, without the need for historical conditional input.

Figure 4a shows the RMSE curves of the GAN and control group models in the Shilin area at different training epochs. The blue curve represents the model developed in this study, and the orange curve represents the control group model. The vertical axis shows RMSE, and the standard deviation (SD) of RMSE of each model is marked in the upper right corner of the chart. The RMSE of the developed model is maintained in the range of about 200–300, while the control group is mostly between 350 and 500, which means that the model developed has a significantly smaller average error and higher prediction accuracy in each short-term passenger flow prediction.



(**a**) (**b**)

Figure 4. Comparison of RMSE changes between our model and the control model in Shilin area: (**a**) RMSE curve of a single version; (**b**) RMSE curve after averaging ten versions.

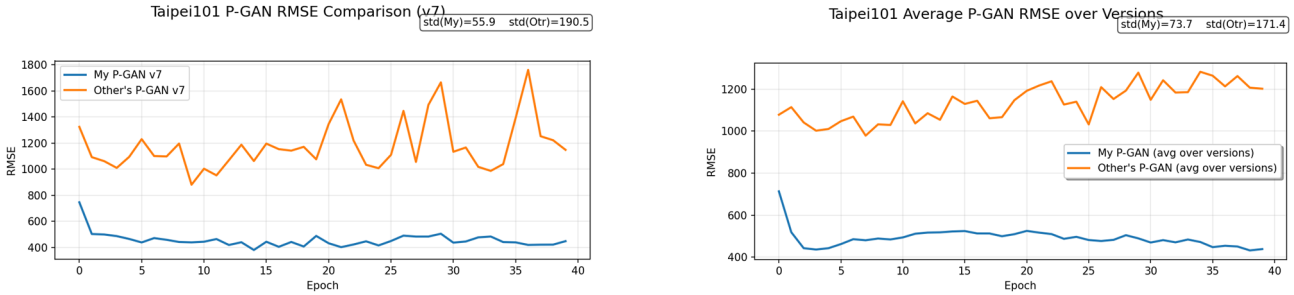
Along the epoch axis, RMSE drops rapidly and stabilizes in the first 5–10 epochs. The control group does not decrease significantly until more than 30 epochs. In other words, the model has achieved optimal performance after fewer training iterations, with lower training costs and shorter development cycles. The SD of the RMSE of the model is 34.6 (vs. 50.9), indicating that its performance fluctuates less between different epochs. A lower SD means that the model adapts more evenly to the training data and is less likely to experience large error fluctuations due to batch differences or noise, which improves the repeatability and reliability of the prediction results.

The lower average error, faster convergence speed, and smaller performance fluctuations all indicate that our GAN is not only more accurate in short-term crowd flow prediction tasks, but also more efficient and robust.

Figure 4b shows the average RMSE curves of the developed GAN model and the control model for 10 versions at different training epochs in the Shilin area. The average RMSE falls between about 210–270, while the control model is mostly concentrated between 420 and 500, indicating that the GAN proposed in this study can maintain significantly lower prediction errors under various settings. From the convergence behavior, the developed model dropped and stabilized in the first 5 epochs. In contrast, the control model requires a longer training time to see a significant performance improvement, indicating that the architecture of the developed model has a faster convergence speed. In addition, the RMSE of the developed model fluctuated less between the ten versions (an SD of 37.0), which is much smaller than 61.0 of the control group. This result shows that even if the average of the ten versions is taken, this method shows a consistently excellent trend, verifying its effectiveness.

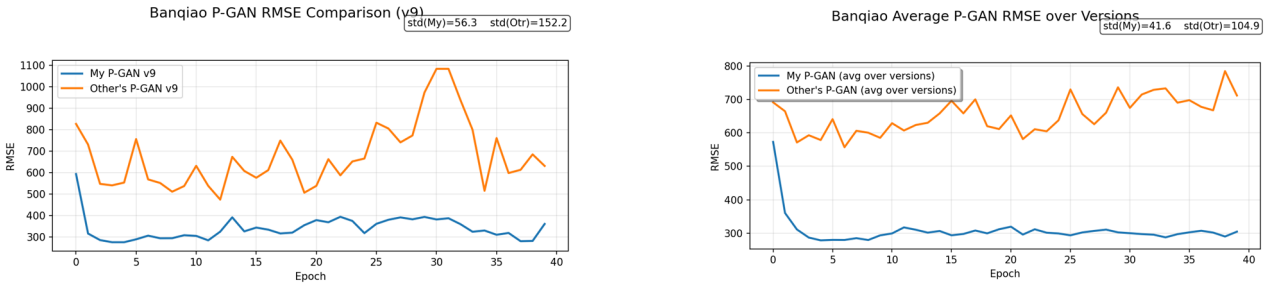
Figures 5 and 6 illustrate the training processes for Xinyi Business District and Banqiao Houzhan Business District using our method and the control method, respectively. The results are consistent with those of Figure 4 for the Shilin Business District. RMSE of the developed either with a single model or the average of 10 models, converges faster than the control group and achieves a lower value. This further confirms the effectiveness of the developed model.

In the final experimental region, the Zhongxiao Business District, results diverged from those observed in other areas, warranting specific discussion. Between the 10th and 25th training epochs, the model's RMSE temporarily increased to nearly 1000, reducing the performance gap with the control group. This fluctuation reflects the highly variable pedestrian flow in Zhongxiao, characterized by distinct peak periods during lunch hours and substantial evening traffic driven by dining and entertainment activities. Moreover, heterogeneous sources of foot traffic—such as MRT stations, shopping centers, and office buildings—contribute to significant spatial variability in flow levels across different blocks at the same time.



(**a**) (**b**)

Figure 5. Comparison of RMSE changes between our model and the control model in Xinyi Business District: (**a**) RMSE curve of a single version; (**b**) RMSE curve after averaging ten versions.

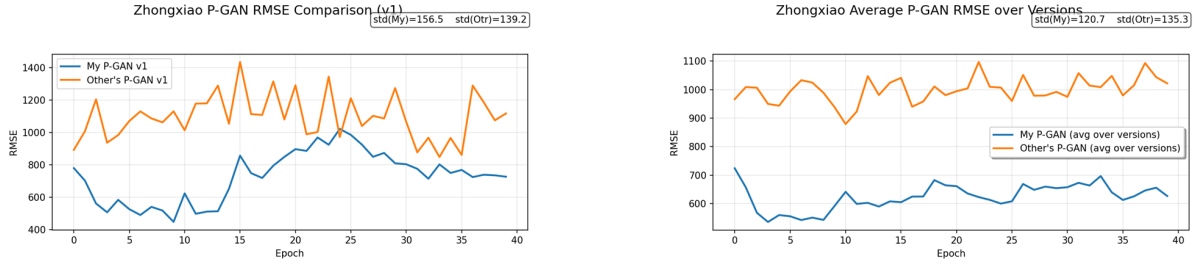


(**a**) (**b**)

Figure 6. Comparison of RMSE changes between our model and the control model in Banqiao Houzhan Business District: (**a**) RMSE curve of a single version; (**b**) RMSE curve after averaging ten versions.

During training, the generator captures these complex patterns, occasionally modeling secondary peaks and atypical scenarios. This exploratory behavior leads to transient overfitting and RMSE instability. However, after the 25th epoch, the model reorients its learning trajectory, and the RMSE declines to below 700, demonstrating its capacity for self-correction. Despite mid-training fluctuations, the model consistently achieves a lower average error than the control group, underscoring its effectiveness in handling dynamic urban environments.

Figure 7b presents the average RMSE across ten independently initialized variants of the GAN and its control group counterpart within the Zhongxiao Business District, plotted against the number of training iterations. The results corroborate the findings in Figure 4b: the proposed model converges more rapidly to a stable performance level and maintains a consistently low error margin across repeated trials, indicating superior generalization capability.



(**a**) (**b**)

Figure 7. Comparison of RMSE changes between our model and the control model in Zhongxiao Business District: (**a**) RMSE curve of a single version; (**b**) RMSE curve after averaging ten versions.

4. Conclusions

Recent research has increasingly focused on forecasting population dynamics in specific urban locations. Historically, scholars have employed a range of techniques, including supervised learning models and generative AI. Generative AI often yields superior predictive performance compared to traditional supervised approaches. However, these advanced models are not without limitations, frequently encountering challenges such as slow training speeds and convergence instability. To address these issues, we introduced a novel GAN-based random number generation method designed to enhance the efficiency and reliability of population flow predictions. The model was evaluated using real-world telecommunications-based population flow data from Taiwan, demonstrating its effectiveness and practical feasibility.

Author Contributions: Conceptualization, Y.-H.L. and Y.-C.C.; methodology, Y.-H.L. and Y.-C.C.; software, Y.-H.L.; validation, Y.-H.L. and Y.-C.C.; formal analysis, Y.-H.L., and Y.-C.C.; investigation, Y.-H.L. and Y.-C.C.; resources, T.-Y.C. and R.-K.S.; data curation, T.-Y.C. and R.-K.S.; writing—original draft preparation, Y.-H.L.; writing—review and editing, Y.-C.C.; visualization, Y.-H.L.; supervision, Y.-C.C.; project administration, Y.-C.C. and T.-Y.C.; funding acquisition, Y.-C.C. and T.-Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science and Technology Council, Taiwan, grant numbers 113-2121-M-005-003 and 114-2121-M-005-006.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The people flow data used in this paper cannot be made public because it is regulated by the privacy law of the Taiwan government.

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

References

1. Vermeulen, J.L.; Hillebrand, A.; Geraerts, R. A comparative study of k-nearest neighbour techniques in crowd simulation. *Comput. Animat. Virtual Worlds* **2017**, *28*, e1775. [CrossRef]
2. Li, Y.; Cao, H. Prediction for Tourism Flow based on LSTM Neural Network. In Proceedings of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things, Qufu, China, 19–21 October 2017.
3. Zhang, J.; Zheng, Y.; Qi, D. Deep spatio-temporal residual networks for citywide crowd flows prediction. In Proceedings of the 31th AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
4. Wang, S.; Cao, J.; Chen, H.; Peng, H.; Huang, Z. SeqST-GAN: Seq2Seq generative adversarial nets for multi-step urban crowd flow prediction. *ACM Trans. Spat. Algorithms Syst. (TSAS)* **2020**, *6*, 1–24. [CrossRef]
5. Wang, Y.T.; Lee, C.L.; Chen, Y.C. Anomaly Detection of Crowd Flow by Using Ensemble Generative Adversarial Networks. In Proceedings of the 31st International Conference on Geoinformatics, Toronto, ON, Canada, 14–16 August 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.