Exploration of virtual city construction and optimization based on deep learning

Zihao Jiang

School of Computer and Communication Engineering, Northeastern University, 110819, China

jzh20030424@outlook.com

Abstract. With continuous artificial intelligence and computer graphics technology, virtual cities are receiving widespread attention as an essential digital twin technology, . The core issue of this study is how to choose appropriate neural networks and algorithms to build models to construct virtual cities. The research methods include literature search, research and improvement of deep learning algorithms, and exploration of multi-model combinations. The research conclusion shows that choosing appropriate neural networks and algorithms is the key to building high-quality virtual cities, and targeted improvement and optimization of deep learning algorithms can further improve the accuracy and efficiency of virtual city construction. The strategy of multi-model combination also shows its unique advantages. By integrating different neural networks and algorithms, people can fully utilize their advantages and compensate for each other's deficiencies. With the advancement of technology, more innovative methods and technologies will be applied to this, which will help to build a more realistic virtual world and promote the development and application of virtual cities.

Keyword: Virtual city, deep learning, neural network, multimodal model.

1. Introduction

As an outstanding representative of digital twin technology, the virtual city uses technologies such as 5G/6G, the Internet of Things, blockchain, urban models, spatiotemporal big data, data computation, and artificial intelligence to build a smart city based on digital twins, gradually becoming a hot spot in scientific research and application [1, 2]. In the construction of a virtual city, selecting appropriate neural networks and algorithms to construct accurate and efficient models is the core issue of this study. Currently, we live in an era of data explosion. The physical city, as the foundation of the virtual city, contains a large amount of multi-format data, including texts, images, and videos [3]. The rise of deep learning technology allows us to capture and extract subtle features and patterns of change in cities from these data. This study aims to construct a more realistic and detailed virtual city through a literature review, research and improvement of deep learning algorithms, and exploration of multi-model integration strategies. The research results show that selecting suitable neural networks and algorithms is crucial for building high-quality virtual cities. We can further enhance the accuracy and efficiency of virtual city construction by targeted improvements and optimizations of deep learning algorithms. In addition, the strategy of multi-model integration demonstrates its unique advantages. By integrating different neural networks and algorithms, we can fully utilize their strengths and compensate for their weaknesses, achieving a more comprehensive and precise virtual city construction. With the continuous advancement of technology, more innovative methods and techniques will emerge, injecting new vitality into the construction and development of virtual cities and pushing their applications in simulating actual urban development, disaster simulations, game design, and other fields to new heights.

2. The Application of Deep Learning in Virtual City Construction

As a branch of artificial intelligence, deep learning simulates the working mode of human brain neural networks, and experiences often correspond to data stored in the form of features, possessing robust feature learning and representation capabilities. In constructing virtual cities, deep learning can extract valuable features through learning from a large amount of urban data, providing an accurate model for generating virtual cities.

2.1. Feature learning

The essence of deep learning is to build machine learning models with many hidden layers and massive training data to learn more useful features, ultimately improving the accuracy of classification or prediction. So, "deep modelling" is the means, and "feature learning" is the goal [4]. In the construction of virtual cities, the primary application of deep learning technology is to learn and simulate various city features. This includes but is not limited to various aspects such as urban images, geographical layout, road planning, traffic patterns, population distribution, and environmental data. By constructing deep neural network models, deep learning can automatically learn and recognize these features from massive urban data and generate realistic virtual urban landscapes.

2.2. Multimodal data fusion

Through multimodal data fusion technology, deep learning can effectively integrate data from different sources and forms. This fusion enriches virtual cities' information dimension and makes learning of urban features more comprehensive and accurate.

More realistic virtual urban landscapes and buildings can be generated through deep learning techniques. The model can learn details such as textures, lighting, and shadows from the real world, making virtual cities visually closer to the real world.

Deep learning enables virtual cities to have more advanced interactive functions. By building an intelligent interaction system, users can interact in real cities, such as exploring the interior of buildings and participating in urban activities. This interactivity enhances user immersion and provides more urban planning and management possibilities.

3. Selection of neural networks

Selecting the appropriate neural network is crucial in constructing virtual cities based on deep learning.Different neural networks have different advantages in feature extraction, model representation, and computational efficiency.

3.1. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a particular type of deep feedforward network, and the CNN model mainly includes the input layer, convolutional layer, pooling layer, fully connected layer, and output layer [5].

Convolutional neural networks perform excellently in processing image data, utilizing basic structures such as convolutional layers, pooling layers, and fully connected layers to learn and extract relevant features on their own and utilize them [6]. In the construction of virtual cities, it can handle the texture, color, and other features of urban landscapes.

DOI: 10.54254/2755-2721/69/20241478

3.2. Recurrent Neural Network (RNN) and Long Short Term Memory Network (LSTM)

Recurrent neural networks refer to deep cognitive patterns used to process sequence information. Including loop units: The loop units in RNN allow signals to be freely transmitted at various time steps. This makes RNN suitable for managing time. RNN performs well in processing time series data, making it suitable for simulating urban dynamic changes such as traffic flow, population migration, and other patterns. However, when dealing with long sequences, RNN may face the problem of vanishing or exploding gradients.

To address this issue, Long Short Term Memory (LSTM) networks have been proposed. Long Short Term Memory Network (LSTM) is a variant of RNN that effectively solves the gradient problem when processing long sequences by introducing gating mechanisms and memory units, thus possessing good long-term dependency modeling ability. The proposal of LSTM has further promoted the development of RNN, and now it is more common to refer to LSTM structured neural networks as RNN [7]. In the construction of virtual cities, LSTM can be used to simulate and predict the long-term development trends of cities, such as urban planning, population growth, etc.

3.3. Generative Adversarial Network (GAN)

Generative Adversarial Network (GAN) consists of two parts: a generator and a discriminator. Research has shown that GAN enables models to explain the accurate distribution patterns of data through a game between the two, thereby generating new image data commonly used in image processing in computer science and medicine [9, 10]. Generative adversarial networks can learn unfamiliar and familiar data suitable for tasks to adapt to diverse design conditions and have good application prospects in urban design [11].

In virtual city construction, GAN can generate highly realistic elements such as urban buildings and landscapes. By training the generator to generate fake city images and then judging their authenticity through a discriminator, the generator can produce increasingly realistic city images after multiple iterations of training.

3.4. Deep Reinforcement Learning Network (DRL)

Deep reinforcement learning combines the advantages of deep learning and reinforcement learning, enabling agents to learn complex decision-making processes. Research has shown that DRL has also succeeded in applications such as robot control, computer vision, natural language processing, healthcare, and game theory [12]. In virtual cities, DRL can be used to train intelligent transportation systems, optimize traffic signal control strategies, or train intelligent agents in virtual cities for autonomous navigation and task completion.

3.5. Transformer network

Transformer is a neural network structure based on a self-attention mechanism, which has a strong learning ability for long-distance dependencies, strong multimodal fusion ability, and more interpretable models. It can play a significant role in image classification, object detection, image segmentation, recognition tasks, image enhancement, image generation, and video processing [13]. Transformer networks can play various roles in virtual cities, mainly due to their powerful sequence modelling ability and self-attention mechanism, which can be applied to time series prediction in virtual cities, such as predicting traffic flow, crowd flow, urban event detection, energy consumption, etc.

3.6. Autoencoder

Autoencoder is an unsupervised neural network model that can automatically learn practical abstract features from unlabeled samples, commonly used for data dimensionality reduction or feature learning. In virtual city construction, autoencoders can help extract key features of city images or data for subsequent modeling and analysis [14, 15].

3.7. Capsule Networks

In October 2017, the proposal of capsule networks caused a considerable sensation in the academic community and was once considered a theoretical alternative to convolutional neural networks. Firstly, it replaces the max pooling method in convolutional neural networks with dynamic routing algorithms, which can avoid the loss of more important information. Secondly, its output is in the form of vectors, representing not only the size of the input response but also its direction or attitude information. Moreover, capsule networks require much less data than convolutional neural networks to achieve good results [16].

4. Algorithm optimization and model improvement

In order to further improve the accuracy and efficiency of virtual city construction, we need to optimize and improve the selected neural networks and algorithms. For example, the model's generalization ability and convergence speed can be improved by adjusting the network structure, optimizing the loss function, and introducing regularization terms. In addition, advanced technologies such as ensemble learning and transfer learning can enhance the model's stability and adaptability.

4.1. Algorithm optimization and improvement using natural language algorithm models as an example

Taking the analysis of natural language algorithm models as an example, the selection and optimization of neural networks and algorithms is a crucial process that directly affects the performance and accuracy of the model. In natural language processing, commonly used neural network models include Recurrent Neural Network (RNN), Long Short Term Memory Network (LSTM), and Transformer. These models perform well in processing sequential data and can capture contextual information in text. After selecting the appropriate neural network model, it is also necessary to optimize the algorithm to improve the performance of the model. By conducting thorough study and implementing proper parameter initialisation, the training process of neural networks can be expedited. Implementing a learning rate decay approach to dynamically modify the learning rate in order to accommodate varying training stages; Utilising methods such as L1 regularisation, L2 regularisation, and dropout to improve the model's capacity to generalise; Optimise the parameter update technique to minimise the model's loss function and enhance the speed and accuracy of convergence. Terminate the training process once the validation error starts to rise in order to avoid overfitting the model.By comprehensively applying optimization and improvement measures, we can significantly improve the performance and accuracy of natural language processing algorithm models, providing stronger support for practical applications [17-19].

4.2. The specific application of method optimization and model improvement in virtual city construction

In the construction of virtual cities, we can also draw inspiration from algorithm optimization and model improvement strategies in natural language processing. For example, using models such as RNN, LSTM, or Transformer, we can better simulate urban dynamic changes and capture dependencies in time series data. By using parameter initialization, learning rate adjustment, and regularization techniques, the training efficiency and model's generalization ability can be improved, ensuring the stability and authenticity of the virtual city. In addition, introducing integrated learning and transfer learning strategies can accelerate the model's adaptation process in new scenarios and improve the flexibility and scalability of virtual city construction.

4.3. Building virtual urban life based on the combination of multiple models

In February 2024, OpenAI released its first AI video model Sora, while Google also released an updated version of Gemini 1.5, from Pika and Gemini 1.0 to Sora and Gemini 1.5. The continuous breakthroughs in AI multimodal models have brought new development opportunities for the technology of building virtual cities. The key to combining multiple models is integrating different data models and algorithms to build a comprehensive virtual city framework, creating a more realistic, rich, and interactive urban experience.

The key to combining multiple models lies in integrating different data models and algorithms. This includes multiple aspects such as geographic information models, building models, population behavior models, and transportation models. Geographic information models describe the spatial layout and geographical features of cities; Building models create various types of buildings and facilities in virtual cities; The population behavior model simulates the daily activities and interactive behavior of virtual residents; The traffic model simulates the traffic flow and management of cities. By integrating these models, we can build a comprehensive virtual city framework that provides users a richer and more authentic urban experience.

5. Conclusion

This study explores multiple levels of constructing virtual cities based on deep learning. It provides a detailed analysis of the specific applications of technologies such as deep learning and neural networks in virtual city construction. At the same time, it explores the practical utility of algorithm optimization and model improvement. Through in-depth research, it has been found that deep learning has demonstrated outstanding capabilities in capturing urban features and optimizing the generation and rendering of virtual cities, providing innovative solutions to overcome the challenges in virtual city construction. Looking ahead to the future, with the deep integration and application of cutting-edge technologies such as virtual reality, augmented reality, artificial intelligence, cloud computing and big data processing, quantum computing, etc., we are confident in injecting more substantial technological power into the development of virtual cities, in order to create more realistic, intelligent, and sustainable virtual cities, and contribute more to the development of human society.

References

[1] Zhuang Cunbo, Liu Jianhua, Xiong Hui, Ding Xiaoyu, Liu Shaoli&Weng Gang (2017). The connotation, architecture, and development trend of product digital twins Computer Integrated Manufacturing System (04), 753-768 Doi: 10.13196/j.cims.2017.04.010

[2] Li Deren (2020). New Heights in the Construction of Smart Cities in Digital Twin Cities. China Survey and Design (10), 13-14

[3] Niu Qiang, Lu Xiangyi&Wei Wei (2018). Preliminary Discussion on Virtual Smart Cities - Exploration of Smart City Construction Methods Based on Virtual City Space. Urban Architecture (15), 26-30. doi: 10.19892/j.cnki. csjz. 2018. 15.006

[4] Yu Kai, Jia Lei, Chen Yuqiang&Xu Wei (2013). Yesterday, Today, and Tomorrow of Deep Learning. Computer Research and Development (09), 1799-1804

[5] Zheng Yuanpan, Li Guangyang&Li Ye (2019). Application Research Review of Deep Learning in Image Recognition. Computer Engineering and Applications (12), 20-36

[6] Lu Hongtao&Zhang Qinchuan (2016). Application Research Review of Deep Convolutional Neural Networks in Computer Vision. Data Acquisition and Processing (01), 1-17. doi: 10.16337/j.1004-9037.2016.01.001

[7] Peng Huajian&Min Yuan (2023). Application of Deep Learning in Image Analysis of Construction Projects Proceedings of the 17th National Conference on Signal and Intelligent Information Processing and Application (pp. 551-555). China Communications Wuhan Zhixing International Engineering Consulting Co., Ltd; China Communications Second Navigation Engineering Bureau Co., Ltd; Doi: 10.26914/c.cnkihy.2023.054831

[8] Yangliu (2022). Overview of the Application of Deep Learning in Transportation. Information Technology and Informatization (02), 192-195

[9] LEDIG C, THEIS L, HUSZ F, et al. PhotoRealistic Single Image Super Resolution Using a Generative Adversarial Network [J] Ar Xiv, 2016: 105-114

[10] Wang Kunfeng, Gou Chao, Duan Yanjie, et al. Research progress and prospects of generative adversarial network GAN [J]. Journal of Automation. 2017, 03:321-332

DOI: 10.54254/2755-2721/69/20241478

[11] Liu Yuezhong, Stokes Ludi&Yang Yang (2018). Research on Urban Design Based on Conditional Generative Adversarial Networks. Journal of Architecture (09), 108-113
[12] Liu Quan, Zhai Jianwei, Zhang Zongchang, Zhong Shan, Zhou Qian, Zhang Peng&Xu Jin (2018). Overview of Deep Reinforcement Learning. Journal of Computer Science (01), 1-27
[13] Liu Wenting&Lu Xinming (2022). Research progress on Transformer based on computer vision. Computer Engineering and Applications (06), 1-16
[14] Zheng Yin, Chen Quanqi&Zhang Yujin (2014). Deep learning and its new progress in target and behavior recognition. Chinese Journal of Image and Graphics (02), 175-184
[15] Laijie, Wang Xiaodan, move forward, Song Yafei&Quanwen (2021). Overview of Autoencoders and Their Applications. Journal of Communications (09), 218-230
[16] Lu Chunyan (2019). Improvement of Capsule Networks and Their Applications in Image Generation. Master's Thesis, Southwest University https://kns.cnki.net/kcms2/article/abstract?v=YoFA4grnCX5hvY5hZSsa67mm90KLlljN1d1nHkOoOdCeYZCIAFBo9JP -73mypM0ugu jiq4eU2cEA4BDHicXseDqkUjUjXz_lsEJ3b-DStL92NRQccnxal9oQol2ueBcY5l9UjgGttUNBgnUKQJYDw==uniplatform=NZKPTlanguage=CHS
[17] He Kai (2020). Master's degree in research and application of text classification based o n natural language processing (thesis, Nanjing University of Posts and Telecommunica tions). Master's degree https://link.cnki.net/doi/10.27251/d.cnki.gnjdc.2020.000274doi:10.27251/d.cnki.gnjdc.2020.000274.
[18] Che Wanxiang, Dou Zhicheng, Feng Yansong, Gui Tao, Han Xianpei, Hu Baotian& Zhao Yanyan (2023). Natural Language Processing in the Era of Big Models: Challenges, Opportunities, and Development. Chinese Science: Information Science (09), 1645-1687
[19] Zhang Kaizhen, Li Yanwu, Liu Bo, Li Jie, Xie Hui&Zhang Zhongyi (2021). Research on hyperparameter adjustment strategy based on VGG16 network. Technology and Innovation (22), 10-13. doi: 10.15913/j.cnki. kjycx.2021.22.05