Manuscript Number: CEUS-D-21-00877

Simulating large-scale urban land-use patterns and dynamics using the U-Net deep learning architecture

Dear Mr. Wang,

Thank you for submitting your manuscript to Computers, Environment and Urban Systems.

I have completed my evaluation of your manuscript. The reviewers recommend reconsideration of your manuscript following major revision. I invite you to resubmit your manuscript after addressing the comments below. Please resubmit your revised manuscript by Jul 06, 2022.

When revising your manuscript, please consider all issues mentioned in the reviewers' comments carefully: please outline every change made in response to their comments and provide suitable rebuttals for any comments not addressed. Please note that your revised submission may need to be re-reviewed.

To submit your revised manuscript, please log in as an author at https://www.editorialmanager.com/ceus/, and navigate to the "Submissions Needing Revision" folder.

Computers, Environment and Urban Systems values your contribution and I look forward to receiving your revised manuscript.

Kind regards,      
Alison Heppenstall     
Associate Editor

Computers, Environment and Urban Systems

**Reviewer #1:**

The paper reports the result of a study using deep learning to model urban land development by applying the U-Net architecture to a study area of the North China Plain. The subject addressed in this article is worthy of investigation and appropriate to the journal.

Thanks for confirming the value of this study.

However, the methodology is not clearly descried w.r.t the data used for the development of the two U-Net models and the novel methodological contributions.

In our first submission, we noted the the different historical periods with different names (1994–2006 and 2006–2018 are referred to as “base,” 2018 and 2030 are called “target,” and the true 2018 urban map is called “reference”) and described the U-Net as “U-Net trained on base/target years.” Such notation could confuse the readers because they need to roll back to the method section to discern the differences between “U-Nets trained on base/target years.”

To emphasize the differences of the two U-Net develop these two models, we added a table in the introduction section and append a suffix to the two U-Net models as below. We deleted the description on training phases (i.e., 1994–2006 and 2006–2018 are referred to as “base,” etc.), and used the U-Net-A/B rather than “U-Net trained on base/target years” to identify the deep learning model through the manuscript.

Table 1. The U-Net developed in this study. (This table is inserted above the workflow chart)

|  |  |  |
| --- | --- | --- |
| Model name | Training years | Project year |
| U-Net-A | 1994, 2006 | 2018 |
| U-Net-B | 2006, 2018 | 2030 |

To clarify the data used to train the U-Net models, we added a table in the “Data processing” section.

Table 3. The data used to train U-Net models. (This table is inserted in the “2.4. Data preprocessing”)

|  |  |  |  |
| --- | --- | --- | --- |
| Data | Time | Source | Resolution |
| Urban land use | 1994, 2006, 2018 | Wang et al. (2021) | 30 m |
| Elevation | 2000\* | Shuttle Radar Topography Mission | 30 m |
| Slope | 2000\* | Derived from elevation | 30 m |

\* Note, the elevation and slope data were used in both U-Net-A and U-Net-B despite their being acquired in 2000.

The novel methodological contributions, comparing to CA models, of deep learning architecture (U-Net) in simulating urban development mainly lies on 1) reducing subjectivity upon tweaking model weights/parameters, and 2) introducing computer vision technologies to “see” urban growth (i.e., sensitive to spatial features) rather than regression in a neighborhood scop to compose driving factors into transition rules.

That been said, the U-Net itself is not an innovative architecture (cutting-edge perspective) in deep learning field. So, we focus the innovation of U-Net on how it could add up to the current CA-mainstreamed urban simulation study: 1) automated to extract transition rules and 2) extract large scale (256 \* 256 pixels window comparing to ~ 10\*10 window of CA models) spatial features into simulation. However, it is not appropriate to discussion novel methodological contributions in the methods section. Instead, we discussed how the U-Net could contribute to urban simulation studies in the discussion section (4.1. Mimicking real-world urban patterns and dynamics, and 4.2. Reducing subjectivity by automatically constructing transition rules).

Somehow, the conclusions are not well supported by the study results, e.g., "The U-Net model successfully captured neighborhood effects, gravity effects, and linear expansion along transportation routes in the urban dynamics", where "linear expansion along transportation routes" seems to be concluded based on visual inspections of 11 smaller areas.

We agree that the “neighborhood effects, gravity effects, and linear expansion effects” could be abstracted by more appropriate metrics in theory, but current landscape indexes have limited power to reflect these characteristics (Frazier and Kedron 2017; Pontius et al. 2008; Uuemaa et al. 2009). Indeed, assess the quality of output features remains a significant challenge even in the deep learning community because of the versatile traits in spatial signals that beyond the mathematical description of simple shape indexes (Wang et al. 2020; Ghosh et al. 2020): it is common to evaluate the performances of deep learning models by observe the resulted images (Gonog and Zhou 2019; Xu et al. 2018). However, we still computed the landscape shape index for readers so that they comprehend the performance of U-Net simulation with a basic estimation.

In this study, we deliberately illustrated 11 small-scale (1km – 10km) regions to present that U-Net has learned some inherent patterns (i.e., the neighborhood, gravity, and linear effects) of urban development. The reason for choosing small-scale regions is that urban expansion simulation usually performs less accurate in small areas than in larger regions (Pontius et al. 2008). If small area looks sensible, readers are more likely to be convinced that the deep learning architecture could simulate the urban development well in the overall scale.

To explain to readers of our visual inspection-based pattern evaluating, we added the description of “It is difficult to evaluate the learned with landscape metrics (Frazier & Kedron, 2017; Pontius et al., 2008; Uuemaa, Antrop, Roosaare, Marja, & Mander, 2009). To assess the learned patterns, we randomly selected 11 evenly distributed regions in the research area and visually inspect them” in the “2.7 validation and accuracy assessment” section. Meanwhile, we also discussed the cause of choosing visual inspection over shape metrics to evaluate learned patterns (sophistication in spatial pattern abstraction), the potential bias (cherry picking), and the measures in this study to reduce such biases (11 small-areas that is random and evenly distributed) in the “4. discussion” section.

The highlight states "U-Net achieves similar accuracies with CA models with lengthen calibration". What do you mean about "lengthen calibration"? How does your study result support "similar accuracies"? What makes the U-Net models better than CA models? Perhaps, the highlight needs to be revised and corresponding quantitative comparison and discussion are needed in the discussion section.

One of my main concerns is the description of the data used in the study, which are not clear and consistent considering the following listed example lines. You actually used land use maps of 1994, 2006 and 2018 only, plus elevation/slope data (not sure if also of these three years), but the way the data are described is confusing.

Line 97-98: "used high-accuracy maps of urban land-use between 1993 and 2012 to train the model" and tested the model by the projected spatial distribution of urbanization for 2018

Line 100-101: THIS trained model was then used to project patterns in 2030

Line 108-109: "We used Landsat data to map urban development in the study area for the years 1994, 2006, and 2018" - I think you are talking about "map land use"

Line 111-113: here you start talk about using different years of data for two different U-net models.

Line 196-199: what are the spatial resolutions of land use, elevation, and slope data, assuming they are rasters?

Line 213-215: what's urban map here? Here, what data were used to train which model?

Line 226: what are the input images here? Land use maps?

Line 285-286: "to historical urban land-use maps for the period 1994-2006" - the expression implies maps of more than two years

The terms used throughout the manuscript need to be clearly-defined or understood and consistent, e.g., urban dynamic map, urban map, urban land-use map, urban land use map, urban pixel number, urban pixels, urban images

Line 204-205: any previous studies that prove "little additional useful information" provided by distance factors?

What message does Fig. 5 convey? How does it show dynamics of urban land use? What do you mean about "urban dynamic map"? Do you really need this large map while showing the details as small snapshots?

Line 310-312: Is SuiXian a large town? What are those smaller villages which have the same phenomena - high transition potential?

Figure 10 - do you really need this large map while showing the details as small snapshots?

Figure 11 - do you really need this large map while showing the details as small snapshots?

Fig. 10, 11 and 13 show 11 smaller urban areas, no large urban centers. How will you generalize the observations from the patterns around these areas to more general urban development?

What's the purpose of including projection of 2030 urban land use? Is it just to demonstrate the use of the trained U-Net models?

**Reviewer #2:**

I have reviewed this article carefully; I agree that it is a promising new idea to apply U-Net to urban growth simulation and prediction.

However, this method focuses on the characteristics of images (i.e., urban land use landscape), rather than understanding the process of urban growth from the driving factors (i.e., the combined effects of nature and society, etc.) that drive landscape evolution. Therefore, this method cannot understand the mechanism of urban evolution.

On the other hand, the author does not strongly prove that the simulation accuracy of this method is better than the results of traditional CA-based result.

I noticed that the author mentioned in the article that U-NET achieved higher accuracy than CA in some cities. But this does not mean that U-NET performs better in a large area, such as the study area of this paper.

Publication bibliography

Frazier, Amy E.; Kedron, Peter (2017): Landscape Metrics: Past Progress and Future Directions. In *Curr Landscape Ecol Rep* 2 (3), pp. 63–72. DOI: 10.1007/s40823-017-0026-0.

Ghosh, Bhaskar; Dutta, Indira Kalyan; Totaro, Michael; Bayoumi, Magdy (2020): A Survey on the Progression and Performance of Generative Adversarial Networks, pp. 1–8. DOI: 10.1109/ICCCNT49239.2020.9225510.

Gonog, Liang; Zhou, Yimin (2019): A Review: Generative Adversarial Networks, pp. 505–510. DOI: 10.1109/ICIEA.2019.8833686.

Pontius, Robert Gilmore; Boersma, Wideke; Castella, Jean-Christophe; Clarke, Keith; Nijs, Ton de; Dietzel, Charles et al. (2008): Comparing the input, output, and validation maps for several models of land change. In *Ann Reg Sci* 42 (1), pp. 11–37. DOI: 10.1007/s00168-007-0138-2.

Uuemaa, Evelin; Antrop, Marc; Roosaare, Jüri; Marja, Riho; Mander, Ülo (2009): Landscape metrics and indices : an overview of their use in landscape research. In *LIVING REVIEWS IN LANDSCAPE RESEARCH* 3 (1), pp. 1–28. Available online at http://biblio.ugent.be/publication/695518.

Wang, Jinzhu; Hadjikakou, Michalis; Bryan, Brett A. (2021): Consistent, accurate, high resolution, long time-series mapping of built-up land in the North China Plain. In *GIScience & Remote Sensing*, pp. 1–17. DOI: 10.1080/15481603.2021.1948275.

Wang, Zhengwei; Healy, Graham; Smeaton, Alan F.; Ward, Tomás E. (2020): Use of Neural Signals to Evaluate the Quality of Generative Adversarial Network Performance in Facial Image Generation. In *Cogn Comput* 12 (1), pp. 13–24. DOI: 10.1007/s12559-019-09670-y.

Xu, Qiantong; Huang, Gao; Yuan, Yang; Guo, Chuan; Sun, Yu; Wu, Felix; Weinberger, Kilian (2018): An empirical study on evaluation metrics of generative adversarial networks. DOI: 10.48550/arXiv.1806.07755.