Proceedings of ICADSS 2026 Symposium: FEEDback Genova – Urban Form as Catalyst for Transdisciplinary Creativity
DOI: 10.54254/2753-7064/2026.GE39452

Generating and Optimizing Urban Planning and Design Proposals with Generative Artificial Intelligence

Ziqi Yuan

Nanning University, Nanning, China
1519283170@qq.com

Abstract. Generative artificial intelligence (AI) is expanding its applications in urban planning and design from data analysis to proposal generation and optimization. This review examines twenty core publications across four areas: master planning and regulatory detailed planning, urban spatial form, urban digital twins, and urban design. Three representative cases are used to compare algorithmic modules, generative tasks, constraints, and evaluation approaches. The comparison indicates that deep reinforcement learning, generative adversarial networks, and diffusion models can support the exploration of spatial proposals at different scales, while digital twins can support evaluation and feedback. However, adaptation to local contexts and implementation feasibility require further validation. Organized around analysis, generation, evaluation, feedback, and optimization, the review emphasizes combining AI's ability to process data and explore alternative proposals with planners' professional knowledge and local experience to support human–AI collaborative design through goal setting, proposal revision, and value-based decision-making.

Keywords: generative artificial intelligence, urban planning and design, proposal optimization, human–AI collaboration

1. Introduction

As cities expand and urban systems become more complex, planning and design must reconcile land use, transportation, environmental considerations, and social needs. Traditional approaches that rely on planners' experience and manual exploration are limited in how efficiently they can explore proposals and require costly iterations when handling large datasets, coordinating multiple objectives and constraints, and exploring alternative proposals. Generative artificial intelligence (AI) offers new computational approaches to expand the range of candidate proposals and support design decisions [1].

Existing research has extended AI applications from urban data analysis to the generation and optimization of land use, building layouts, and urban form. Deep reinforcement learning, site-embedded generative adversarial networks (GANs), and multimodal diffusion models have demonstrated potential for sequential planning, conditional control, and human–computer interaction, respectively [2-4]. However, these models remain limited in their ability to adapt to local

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7064/2026.GE39452

social needs, historical and cultural contexts, policies and regulations, and practical construction conditions.

Through a literature review and case comparison, this study examines the sequence of analysis, generation, evaluation, feedback, and optimization across four areas: master planning and regulatory detailed planning, urban spatial form, urban digital twins, and urban design. It explores how AI's computational capabilities can be combined with planners' professional knowledge and local experience to support human–AI collaborative design that integrates goal setting, proposal revision, and value-based decision-making.

2. Literature review

2.1. Literature scope and analytical framework

This review selects 20 core publications based on thematic relevance and the availability of technical details, giving priority to original technical studies. It draws on the transparent reporting principles of PRISMA 2020 to describe the scope of the literature and the selection criteria [5]. The analysis considers two dimensions—design stages and modes of AI involvement—across master planning and regulatory detailed planning, urban spatial form, urban digital twins, and urban design. It compares generated outputs, spatial representations, algorithmic modules, constraints, evaluation, and stages at which people intervene to trace the connections among analysis, generation, evaluation, feedback, and optimization.

2.2. Master planning and regulatory detailed planning

Research on land-use generation first faces the challenge of translating objectives into model inputs. LUCGAN and its extended framework organize urban context and spatial connections into graph representations, use adversarial learning to generate land-use configurations, and incorporate quantification and evaluation stages [6]. IHPlanner adopts a hierarchical generation structure, using functional modules and attention mechanisms to translate human instructions into spatial configurations in an attempt to connect overall intentions with detailed layouts [7]. The former focuses on the relationship between contextual conditions and the distribution of generated outputs, whereas the latter uses higher-level objectives to guide lower-level layouts.

Reinforcement learning divides the planning process into a sequence of actions that can be evaluated. A study of land use in regional rail networks uses a graph neural network (GNN) to encode spatial relationships, excludes infeasible choices through hierarchical action masks, and incorporates coverage, access to opportunities, and equity into its objectives [8]. Another type of AI advisory system encodes land use, floor area ratios, and building coverage ratios in different color channels, allows users to specify road conditions, and then generates proposals for comparison [9]. These methods can help identify planning configurations, but both reward weights and color encodings simplify complex institutional conditions. Property rights, implementation timing, and policy trade-offs still require professional judgment.

2.3. Urban spatial form

Street networks and building layouts are major areas of generative research. A multistage street-network design framework connects data preparation, deep learning, and an interactive interface, allowing planners to provide local road guidance before the model completes the network [10].

DOI: 10.54254/2753-7064/2026.GE39452

GANmapper learns the relationships between streets and building footprints, while InstantCITY further addresses the synthesis of morphologically consistent vector building data and its transfer to new contexts [11, 12]. These studies further automate morphological representation, but their use in planning analysis still depends on the correct conversion of coordinates, topology, and attributes.

Research on conditional diffusion uses information such as land use, terrain, and intersections to condition road generation, first producing road images and then converting them into networks through thinning and vectorization [13]. A study using Stable Diffusion and ControlNet combines conditions such as textual land-use descriptions, roads, and water bodies to synthesize satellite-view urban images, evaluating the outputs using Fréchet inception distance (FID), kernel inception distance (KID), and human assessment [14]. These metrics reflect visual or distributional characteristics and are not yet sufficient to assess travel efficiency or planning compliance. By comparison, ESGAN improves control over building layouts through site conditions and an attribute discriminator module, indicating that generation quality depends on both how conditions are represented and how results are checked [3].

2.4. Urban digital twins

Digital twins provide generative systems with an environment for analyzing existing conditions and testing scenarios. The Herrenberg study integrates a 3D city model with traffic and wind simulations, alongside public participation, to support planning discussions [15]. The Dublin Docklands study uses an interactive 3D model to gather residents' feedback on urban spaces [16]. These efforts primarily connect data on existing conditions, simulation, and participation and cannot be directly equated with generative models, but they can support the comparison and evaluation of candidate proposals.

Research on generative spatial artificial intelligence proposes the concept of a large flow model that would use autoencoders to connect spatial structures with urban flow data, its framework remains in the design and development stages [17]. Another study of urban layouts uses latent representations of 3D urban blocks to generate layouts conditioned on roads and other factors [18]. Both suggest opportunities to integrate generated outputs with urban data platforms. However, evidence from practical applications remains insufficient to establish whether continuous updating, simulation calibration, and generative optimization can be reliably connected.

2.5. Urban design

Urban design concerns building massing, street frontages, public spaces, and the experience of place, its evaluation includes aspects that are difficult to capture through morphological similarity alone. Urban-GAN combines convolutional networks, case-based reasoning, and generative models to explore diverse urban design proposals [19]. Research on multimodal diffusion separates roads and land use, building layouts, and detailed visualization into distinct stages, uses text and control images to convey intent, and involves professionals in review and revision between stages [4]. The role of planners thus expands from selecting final images to controlling the process.

Performance-oriented research attempts to couple generative networks with simulators, as exemplified by the joint use of CP-GAN and the Urban Weather Generator to evaluate morphology and thermal-environment responses [20]. This allows proposals to be compared in terms of environmental performance, although simulation boundaries and parameters also affect the evaluation. Place-specific design must also account for who uses the space, which activities should

DOI: 10.54254/2753-7064/2026.GE39452

be retained, and how different groups benefit, social value cannot be inferred directly from visual quality.

3. Case analysis

This section examines three cases: community planning with deep reinforcement learning, building layout generation with ESGAN, and urban form design with multimodal diffusion. Focusing on reward feedback, attribute discrimination, and practitioner input during generation, it analyzes how planning constraints are incorporated into models, how evaluation leads to proposal revision, and how planners participate in generation and optimization.

3.1. Community spatial planning with deep reinforcement learning

Zheng et al. formulate the development of spatial proposals for community planning as a sequential decision-making process. Their approach represents site relationships using an urban topological graph and extracts features through a graph neural network to support an agent in allocating land uses and roads [2]. Separate policy networks handle land use and roads: the system first assigns functions and then selects roads from the boundaries of spatial units. Earlier decisions therefore shape the subsequent stage and may limit the scope of later adjustments. The key is to connect spatial analysis results with action policies so that the generation process can use evaluation signals linked to planning objectives.

Within this workflow, graph representations support analysis, policy networks perform generation, reward functions evaluate the resulting configurations, and policy training provides feedback and optimization. Planners participate in setting objectives and reviewing proposals, combining algorithmic search with professional judgment. The case demonstrates that planning evaluation can inform the generation process, rather than merely be used to rank completed proposals.

The reward function also defines the aspects of planning quality that the model can recognize. Local culture, the needs of disadvantaged groups, and implementation conditions may be overlooked during search if they are not incorporated into rewards or constraints. Accordingly, this case offers insights into decision modeling and objective decomposition for master planning and regulatory detailed planning, its experiments at the community scale cannot be directly extrapolated to the automation of urban planning in its entirety. Reporting results separately for different objectives can help reveal the trade-offs between competing interests that underlie an aggregate score.

3.2. Site-embedded building layout generation with ESGAN

ESGAN, proposed by Jiang et al., comprises three modules: image processing, layout generation, and post-processing [3]. Input processing organizes information about the site and its surroundings into a format the model can process. During generation, conditions are introduced through the conditional block, while the attribute discriminator constrains generated outputs to maintain attribute consistency during training. Post-processing converts the outputs into a form suitable for subsequent spatial representation and use.

This design integrates site analysis and generation control within a single workflow. The conditional block supplies information about "where to generate" and "which attribute requirements the output should satisfy," the attribute discriminator provides corresponding feedback, and post-

DOI: 10.54254/2753-7064/2026.GE39452

processing connects generated images to professional workflows. Its strength lies in translating site conditions and design indicators into specific modules, reducing the limitations of relying solely on visual similarity to control layouts.

However, attribute matching does not mean that all planning and engineering requirements have been met. Checks of building spacing, fire safety arrangements, and sunlight access require corresponding geometric and computational models. This review therefore treats ESGAN as a representative approach to controllable layout generation and does not equate its post-processing with a complete compliance review. Subsequent human–AI feedback should be able to identify where conflicts occur and allow planners to preserve satisfactory areas while modifying conditions locally.

3.3. Multimodal diffusion and stage-based human–AI collaboration

Building on Stable Diffusion and ControlNet, He et al. combine textual design intent with structural diagram conditioning to address roads and land use, building layouts, detailed design, and visualization in stages [4]. Applications in New York and Chicago, together with the participation of professional plannerswith input from professional planners, provide case-based evidence for understanding this mode of interaction.

The collaborative mechanism in this study centers on review at each stage. Users can inspect the results of the preceding step and adjust the text or control images before proceeding to the next stage of generation. Evaluation can thus be translated into new conditions through human revisions. The role of evaluation extends beyond screening candidates to limiting the propagation of outputs that depart from the design intent. The study assesses outputs in terms of fidelity, adherence to conditions, and diversity, among other dimensions, but these evaluations still need to be distinguished from assessments of statutory compliance and spatial performance.

Reinforcement learning is well suited to representing actions and their long-term effects, ESGAN emphasizes layouts under site constraints, and multimodal diffusion facilitates communication and refinement at successive stages. Differences in their indicators, data, and tasks preclude a universal ranking of their relative merits. Comparisons should instead consider whether constraints can be effectively incorporated into models, whether evaluation can lead to revisions, and whether planners can understand and control those revisions.

4. Discussion

4.1. From automated generation to closed-loop evaluation and feedback

The four areas can be connected through a shared workflow. During analysis, geographic information systems (GIS), imagery, and survey data are used to identify existing conditions and needs. Generation produces candidate land-use configurations and spatial forms. Evaluation examines geometric validity, planning indicators, and spatial performance. Feedback translates identified problems into revised conditions, rewards, or manual edits, after which optimization searches for feasible alternatives. Human–AI collaboration spans confirming goals, intermediate review, and final decision-making, it is not a separate step added at the end of the workflow.

The effectiveness of this closed loop depends on whether feedback is actionable. An aggregate score can indicate differences between candidates, but identified locations of road discontinuities, capacity exceedances, or sunlight conflicts can directly guide revisions. Existing GAN, diffusion, and reinforcement learning approaches each provide some interfaces for such feedback [2-4,13].

DOI: 10.54254/2753-7064/2026.GE39452

Consistent spatial identifiers and data representations are still needed to trace the relationship between local adjustments and overall performance.

This integration also concerns the scale of evaluation. Street networks can be compared in terms of connectivity and accessibility, while building layouts can be assessed for density, spacing, and open space. Design proposals require an assessment of how people experience a place in relation to actual activities. Only when indicators are linked to the corresponding spatial objects can the objectives affected by a particular revision be identified. Results for visual representation, topological relationships, and environmental performance should be reported separately to support planners' comparisons and trade-offs, preventing improvements in one indicator from compensating for unacceptable constraint violations.

4.2. Local context and professional judgment

Urban form is shaped by historical development, ways of life, and governance systems. Densities or block dimensions learned from existing examples may not be appropriate in another city. In historic districts and neighborhood regeneration projects, existing activities, affordable spaces, and relationships among residents may also need to be preserved. Such requirements can be partly translated into protected areas, functional requirements, and areas that must remain unchanged, but they cannot be fully reduced to model labels or a single reward.

Planners therefore need to interpret the task before generation, identify conflicts with local conditions during generation, and assess benefits and losses for different groups during evaluation. Public feedback collected through digital twins provides additional input to these judgments [15, 16]. Whether a collaborative system improves planning depends on whether that input actually changes goals or proposals, rather than merely on whether the interface supports interaction. Recording proposal versions, changes to conditions, and the rationale for human decisions helps make the design process open to review.

4.3. Evaluation boundaries and limitations of the field

Evaluation should distinguish among visual representation, rule compliance, simulated performance, and outcomes in actual use. Image-generation metrics cannot substitute for planning performance, and improvements in simulation do not necessarily translate into improvements after construction [14, 20]. One research direction is to combine hard-constraint checking with multi-objective evaluation and examine generalizability and the effects of collaboration through cross-city testing, module ablation studies, and controlled user studies. Results for conflicting performance objectives should be reported separately so that aggregate scores do not obscure equity or implementation issues.

5. Conclusion

This review has examined the technical approaches through which generative AI contributes to urban planning and design across four application areas. The methods in existing studies can generate land-use configurations, street networks, and building layouts for specific tasks and revise the results through conditioning, rewards, or human intervention. The three cases show that the planning value of generative capabilities depends on constraint representation, evaluative feedback, and human control. Digital twins can support updates to existing conditions and scenario

DOI: 10.54254/2753-7064/2026.GE39452

comparisons, although their continuous integration with generative optimization is still being explored.

Using a framework organized by design stage and mode of AI involvement, this review brings analysis, generation, evaluation, feedback, and optimization together as a human–AI collaborative process, providing a basis for comparing technical approaches across tasks. The core literature was selected for thematic relevance and consists predominantly of English-language studies, no exhaustive database-based bibliometric analysis was conducted. The studies differ in spatial scale, training conditions, and evaluation baselines. Accordingly, the synthesis primarily explains technical mechanisms and cannot establish industry-wide adoption rates or a unified performance ranking. The evidence also comes mainly from experiments and case-based validation, limiting conclusions about long-term implementation, local social impacts, and institutional compatibility. Future research could combine local rules, simulations that can be calibrated, and public participation to examine the practical feasibility of generated proposals and support planners in documenting their professional judgments and making value-based decisions.

References

[1] Jiang, F., Ma, J., Webster, C. J., et al. (2024). Generative urban design: A systematic review on problem formulation, design generation, and decision-making. *Progress in Planning, 180*: 100795. https://doi.org/10.1016/j.progress.2023.100795.

[2] Zheng, Y., Lin, Y., Zhao, L., et al. (2023). Spatial planning of urban communities via deep reinforcement learning. *Nature Computational Science, 3*: 748–762. https://doi.org/10.1038/s43588-023-00503-5.

[3] Jiang, F., Ma, J., Webster, C. J., et al. (2023). Building layout generation using site-embedded GAN model. *Automation in Construction, 151*: 104888. https://doi.org/10.1016/j.autcon.2023.104888.

[4] He, M., Liang, Y., Wang, S., et al. (2026). Human-guided urban form generation using multimodal diffusion models. *Building and Environment, 287*(Part B): 113892. https://doi.org/10.1016/j.buildenv.2025.113892.

[5] Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*: n71. https://doi.org/10.1136/bmj.n71.

[6] Wang, D., Fu, Y., Liu, K., et al. (2023). Automated urban planning for reimagining city configuration via adversarial learning: Quantification, generation, and evaluation. *ACM Transactions on Spatial Algorithms and Systems, 9*(1), 2:1–2:24. https://doi.org/10.1145/3524302.

[7] Wang, D., Wu, L., Zhang, D., et al. (2023). Human-instructed deep hierarchical generative learning for automated urban planning. *Proceedings of the AAAI Conference on Artificial Intelligence, 37*(4), 4660–4667. https://doi.org/10.1609/aaai.v37i4.25589.

[8] Shi, Z., Liang, Y., Wang, C., Zheng, Y., & Zhao, J. (2026). Transit-oriented land use planning for regional metro networks via deep reinforcement learning. *Transportation Research Part A: Policy and Practice, 211*, 105110. https://doi.org/10.1016/j.tra.2026.105110.

[9] Park, C., No, W., Choi, J., & Kim, Y. (2023). Development of an AI advisor for conceptual land use planning. *Cities, 138*, 104371. https://doi.org/10.1016/j.cities.2023.104371.

[10] Fang, Z., Qi, J., Fan, L., et al. (2022). A framework for human-computer interactive street network design based on a multi-stage deep learning approach. *Computers, Environment and Urban Systems, 96*, 101853. https://doi.org/10.1016/j.compenvurbsys.2022.101853.

[11] Wu, A. N., & Biljecki, F. (2022). GANmapper: Geographical data translation. *International Journal of Geographical Information Science, 36*(7), 1394–1422. https://doi.org/10.1080/13658816.2022.2041643.

[12] Wu, A. N., & Biljecki, F. (2023). InstantCITY: Synthesising morphologically accurate geospatial data for urban form analysis, transfer, and quality control. *ISPRS Journal of Photogrammetry and Remote Sensing, 195*, 90–104. https://doi.org/10.1016/j.isprsjprs.2022.11.005.

[13] Gu, X., Zhang, M., Lyu, J., & Ge, Q. (2024). Generating urban road networks with conditional diffusion models. *ISPRS International Journal of Geo-Information, 13*(6), 203. https://doi.org/10.3390/ijgi13060203.

[14] Wang, Q., Liang, Y., Zheng, Y., et al. (2025). Generative AI for urban planning: Synthesizing satellite imagery via diffusion models. *Computers, Environment and Urban Systems, 122*, 102339. https://doi.org/10.1016/j.compenvurbsys.2025.102339.

DOI: 10.54254/2753-7064/2026.GE39452

[15] Dembski, F., Wössner, U., Letzgus, M., et al. (2020). Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany. *Sustainability, 12*(6), 2307. https://doi.org/10.3390/su12062307.
[16] White, G., Zink, A., Codecá, L., & Clarke, S. (2021). A digital twin smart city for citizen feedback. *Cities, 110*, 103064. https://doi.org/10.1016/j.cities.2020.103064.
[17] Huang, J., Bibri, S. E., & Keel, P. (2025). Generative spatial artificial intelligence for sustainable smart cities: A pioneering large flow model for urban digital twin. *Environmental Science and Ecotechnology, 24*, 100526. https://doi.org/10.1016/j.ese.2025.100526.
[18] He, L., Kamath, H. G., Fei, S., Niyogi, D., & Aliaga, D. (2026). AI-based urban layout generation model. *npj Urban Sustainability, 6*, 87. https://doi.org/10.1038/s42949-026-00369-2.
[19] Quan, S. J. (2022). Urban-GAN: An artificial intelligence-aided computation system for plural urban design. *Environment and Planning B: Urban Analytics and City Science, 49*(9), 2500–2515. https://doi.org/10.1177/23998083221100550.
[20] Xiong, Z., Wang, M., Zhao, J., et al. (2026). Generative urban climate modeling using coupled GAN–UWG frameworks: Morphological strategies for transformation and adaptive heat mitigation. *Urban Climate, 65*, 102786. https://doi.org/10.1016/j.uclim.2026.102786.