

Spatial planning of urban communities via deep reinforcement learning

Received: 15 March 2023

Accepted: 27 July 2023

Published online: 11 September 2023

 Check for updates

Yu Zheng^{1,2}, Yuming Lin^{1,2}, Liang Zhao³, Tinghai Wu³, Depeng Jin^{1,2} & Yong Li^{1,2}✉

Effective spatial planning of urban communities plays a critical role in the sustainable development of cities. Despite the convenience brought by geographic information systems and computer-aided design, determining the layout of land use and roads still heavily relies on human experts. Here we propose an artificial intelligence urban-planning model to generate spatial plans for urban communities. To overcome the difficulty of diverse and irregular urban geography, we construct a graph to describe the topology of cities in arbitrary forms and formulate urban planning as a sequential decision-making problem on the graph. To tackle the challenge of the vast solution space, we develop a reinforcement learning model based on graph neural networks. Experiments on both synthetic and real-world communities demonstrate that our computational model outperforms plans designed by human experts in objective metrics and that it can generate spatial plans responding to different circumstances and needs. We also propose a human–artificial intelligence collaborative workflow of urban planning, in which human designers can substantially benefit from our model to be more productive, generating more efficient spatial plans with much less time. Our method demonstrates the great potential of computational urban planning and paves the way for more explorations in leveraging computational methodologies to solve challenging real-world problems in urban science.

Urban communities have become centers of innovation, creativity and opportunity, attracting people from all walks of life who seek access to urban amenities, such as entertainment, education, healthcare and employment. Effective spatial planning of urban communities is crucial for economic activities and sustainable development of cities, playing an important role in shaping the way people interact with their environment and with each other, creating unique urban cultures and identities. Modern urban planning tends to be vehicle oriented, which favors centralized functionalities and automobile-dependent transportation, bringing many challenges to the urban life. Despite the severe traffic congestion problem, vehicular transportation accounts

for about 72% of greenhouse gas emissions and over 70% of NO₂ emissions in urban areas, which increases the risk of climate change and harms residents' health. In addition, the high car-dependency leads to inequality in accessing urban services^{1,2}, such as education and medical care. Meanwhile, the COVID-19 pandemic since early 2020 also reveals the vulnerability of cities during lockdowns, which calls for more attention to local production and consumption within communities compared with traditional long-distance commuting³. Therefore, a transformation is anticipated in urban planning, shifting the focus from vehicle-oriented approaches to people-oriented ones. This change will emphasize local neighborhoods and promote a decentralized and

¹Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing, P. R. China. ²Department of Electronic Engineering, Tsinghua University, Beijing, P. R. China. ³Department of Urban Planning and Design, Tsinghua University, Beijing, P. R. China.

✉e-mail: liyong07@tsinghua.edu.cn

efficient spatial layout. Notably, the concept of a 15-minute city has gained increasing popularity in both planning new urban communities and renovating existing ones^{4–8}, depicting the future of urban planning in which residents can reach essential services within 15 min by walking or cycling^{4,9,10}, which can substantially improve the accessibility, sustainability and quality of urban life.

The 15-minute city represents one of the current urban planning trends that call for layouts with high spatial efficiency within urban communities. As a long-standing engineering task in urbanization, urban spatial planning has no general solution in practice, and it is often obtained by human designers through multiple rounds of analysis, discussion and iteration. To help reduce the heavy burden of human designers, spatial planning tools with large-scale data involved have been developed, from traditional spreadsheets to the introduction of geographic information system in the 1970s and to the recent advances in adopting artificial intelligence (AI) to perform urban analysis and site design. For example, AI helps in the generation of building layouts¹¹ and architectural design¹². In fact, researchers have been thinking of the automation of urban planning for decades^{13–17}, and much effort has been made to release urban planners from tedious tasks, by designing mathematical and computational models for cities^{13–16} and developing planning support tools¹⁷. With these methods addressing the computational part of urban planning, human designers can concentrate on the more challenging tasks that require creativity and hard work, such as public engagement, land ownership and public rights of way. However, existing methods do not address the fundamental and challenging sub-task, the specific layout of land use and roads, which is more of a computational task, but still depends on the experiences and intuition of human experts, far from automation^{18,19}. Considering current issues and future needs in urban planning, harnessing the powerful computing capabilities of AI for land use and road layout is attractive. First, because of insufficient experience, human intuition can be unreliable and suboptimal when optimizing multiple objective metrics in a huge solution space, for example, minimizing the distance between various services and residential areas in a 15-minute city. On the contrary, these quantifiable metrics can be well captured by computational AI approaches, especially deep reinforcement learning (DRL) algorithms, which automatically learn from millions of trials interacting with an environment, gradually approach the optimal solution and obtain strategies beyond human ability. Second, generating spatial plans solely by human designers is time consuming, which takes several weeks of iterations for one single planning task, and consequently makes it impossible to meet the demands of rapid urbanization. Different from human designers, AI models are more generic and efficient. Specifically, well-trained models in one planning scenario can generalize to various other scenarios. Meanwhile, spatial plans can be obtained through fast model inference, which takes only a few milliseconds. Moreover, besides achieving notable performance in instant feedback-based decision tasks such as video games²⁰, DRL^{21–24} have achieved great success in similar episodic planning tasks that previously relied on human experiences and intuitions, such as the game of Go²⁵, chemical synthesis²⁶, solving math puzzles²⁷ and chip design²⁸. Therefore, utilizing computational AI models to generate spatial layouts can play a critical role in urban planning.

In this work, we propose a DRL-based urban-planning model capable of generating land use and road layouts for urban communities. Compared with chip design²⁸ and the game of Go²⁵, which have regular (grid) conditions, spatial planning is more challenging because of the existence of different forms of geometries. To tackle the diversity and irregularity of urban blocks, we propose an urban contiguity graph to describe the topology of urban geometries, with urban geographic elements as nodes and spatial contiguity as edges. The graph formulation allows us to capture the essential spatial relationship in any form of community. Consequently, we formulate spatial planning as a sequential decision-making problem on the graph and conduct planning at

the topology level instead of geometry. Another major challenge of spatial planning is the huge solution space and the even larger accompanying action space. The action space for a moderate-size community can easily surpass 4,000¹⁰⁰ (4,000 possible actions per step with 100 steps in total for a community spatial plan), making exhaustive search infeasible. To reduce the action space, we train an AI agent consisting of one value network and two policy networks, which searches for decent planning strategies by efficient exploration and exploitation in the huge action space. Specifically, the value network predicts the quality of spatial plans based on the completion of the 15-minute-city concept, including the efficiency of transportation, service and ecology. The other two policy networks are developed for the AI agent to select the locations for land use and roads, respectively. In this way, the action space is substantially narrowed down by sampling actions from the policy networks and estimating rewards with the value network. To obtain effective representations of urban geographic elements, we further develop a graph neural network (GNN)-based state encoder²⁹ that utilizes message propagation and neighbor aggregation on the urban contiguity graph, capturing the spatial relationship between land, road segments and junctions. This GNN state encoder is shared across the value network and policy networks to facilitate reward prediction and location selection. Finally, after massive trials and errors, the AI agent is trained to generate spatial plans that are much more efficient than those generated by human experts.

We demonstrate the effectiveness of our proposed DRL model through extensive experiments on spatial planning. Given the same initial conditions and planning constraints, our method substantially outperforms state-of-the-art algorithms and human experts, improving objective metrics on spatial efficiency by over 48.6%. Particularly, with existing real-world communities as initial conditions, our model can generate land-use renovation plans that remarkably improve residents' accessibility to various facilities by over 18.5% compared with those generated by human experts. Considering the maturity and complexity of the current urban-planning approaches, it is critical to demonstrate the possibility that the proposed framework can collaborate well with human designers. Therefore, based on the presented DRL model, we propose a human–AI collaborative workflow in which human designers focus on conceptual prototyping and utilize our model to accomplish the heavy and time-consuming planning work. We demonstrate that human designers can benefit from the human–AI collaborative workflow, which outperforms the fully human-labor workflow in both objective planning metrics and subjective blind tests consisting of 100 professional human designers, with over 3,000 times speedup on the time cost. Furthermore, we show that our model can learn universal planning skills from simple scenarios and leverage them to design in different styles for large-scale complicated planning tasks, such as the green community and the service community. Our model's transferability and flexibility make it possible to automate the urban-planning process, which can greatly assist human designers and make them more productive.

Results

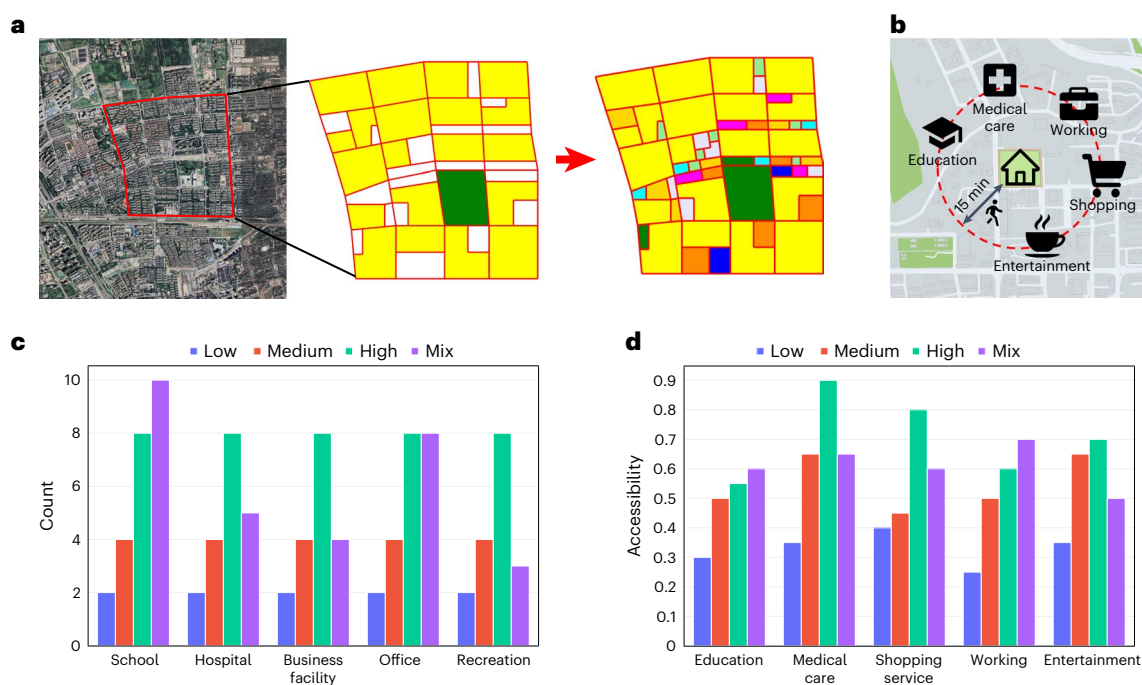
General spatial planning performance

We first demonstrate the fundamental spatial planning capability of the proposed DRL framework by comparing it with existing baseline methods, including rule-based heuristics and genetic algorithm (see Supplementary Sections 2.1 and 2.2 for details of the experimental setup and baseline methods, respectively), as shown in Table 1. We repeated the experiments five times and calculated the *P* value with *t* test, which illustrates the statistical significance of the improvements brought by our proposed model (Table 1). We experiment on both synthetic and real-world communities, with quantitative planning requirements, such as the area and the number of different land-use types (Methods). The synthetic scenario is a square community with a side length of 2.4 km under a 4 × 4 grid secondary road initial conditions

Table 1 | Comparison with baseline methods

Method	Synthetic grid			Real-world HLG		Real-world DHM	
	Service	Ecology	Traffic	Service	Ecology	Service	Ecology
Centralized	0.4514	0.6090	0.5199	0.5833	0.5178	0.5533	0.5562
Decentralized	0.4867	0.6624	0.5008	0.6000	0.5549	0.6067	0.5248
GSCA	–	–	–	0.6100	0.4310	0.5355	0.4578
GA	0.5000	0.5008	0.6019	0.5700	0.6312	0.4333	0.5905
DRL with MLP	0.5625	0.7571	0.6857	0.6300	0.7061	0.6621	0.6104
Our method	0.6833**	0.9171**	0.9384**	0.7100*	0.7303*	0.7484**	0.8976**
Improvement over GA (%)	36.66	83.13	55.91	24.56	15.70	72.72	52.01
Improvement over DRL with MLP (%)	21.48	21.13	36.85	12.70	3.43	13.03	47.05

Bold values indicate the best performance. We conducted a one-sided t test to evaluate the statistical significance of the performance improvement. * $P < 0.1$; ** $P < 0.05$

**Fig. 1 | Demonstration of community renovation and a 15-minute city.**

a, Community renovation. We replicate the roads, residential blocks and parks from a real-world community and leave other areas as vacant lands for renovation. The agent places different types of facilities to maximize the accessibility of service for residents in the community. Please refer to Supplementary Table 1 for the meanings of different colors. **b**, A 15-minute city, which means that five basic residential needs can be satisfied within 15 min by walking or cycling. **c**, Facility needs. We vary the needs for five different facilities (school, hospital, business, office and recreation) that correspond to the five basic services

(education, medical care, shopping, working and entertainment). We investigate low needs (2 counts per facility), medium needs (4 counts per facility), high needs (8 counts per facility) and mixed needs (10, 5, 4, 8 and 3 counts for the five facilities, respectively). **d**, Service accessibility performance under different needs. We show the 15-min-circle index for the five basic services of the generated community plan under different facility needs. The radical value means the proportion of residential blocks that can access the corresponding service within 15 min.

(Extended Data Fig. 1). The real-world scenarios are two existing communities, Huilongguan CP-02 (HLG) and Dahongmen (DHM) communities in Beijing, collected and processed from OpenStreetMap using the Python package OSMnx³⁰ and geopandas. Both of the two real-world communities have an area of about 4 km² (Fig. 1a and Supplementary Fig. 11a). To inspect the quality of the community plan, we evaluate the efficiency of the spatial layout according to the existing literature on service³¹, ecology³² and traffic^{33–35}, which provide a comprehensive evaluation regarding the accessibility to basic urban services, the coverage of greenness and the efficiency and rationality of the road network. Specifically, ‘Service’ measures the layout efficiency of facilities that

satisfy the different needs of residents (for example, offices, schools and hospitals satisfy the needs for work, education and medical care, respectively), and we calculate the proportion of facility categories that can be accessed within 15 min by walking or cycling. ‘Ecology’ focuses on the layout efficiency of parks and green open space, which reflects the residents’ physical and psychological needs for greenness, and the ecological coverage proportion is calculated as the metric. ‘Traffic’ measures the efficiency of road transportation, which is a combination of road density and connectivity (see details of the metrics in ‘Methods’). All three metrics are in the range of [0, 1] with larger values indicating better planning performance, although 1 may not be

attainable because of the limitations of the number of functionalities. For the synthetic grid community, our model goes through a full cycle of two stages, land-use planning and road planning, to accomplish the layout of both aspects, whereas for the real-world communities, we investigate the task of community renovation, in which the roads are already built and our model only needs to accomplish the first stage of land-use planning, which involves the layout of multiple facilities such as schools and hospitals. This setting is reasonable because rebuilding the existing road network is also rare in actual urban renovation projects. Because the road network has not been modified, we do not report the traffic metric for real-world communities.

We first compare our model with two rule-based heuristics, centralized and decentralized approaches. From the results in Table 1, we can observe that the centralized heuristic is outperformed by our model in all three metrics: service from 0.6833 to 0.4514 (−33.94%), ecology from 0.9171 to 0.6090 (−33.60%) and traffic from 0.9384 to 0.5199 (−44.60%) in the synthetic case; service from 0.7100 to 0.5833 (−17.85%) and ecology from 0.7303 to 0.5178 (−29.10%) in the real-world HLG case; and service from 0.7484 to 0.5533 (−26.07%) and ecology from 0.8976 to 0.5562 (−38.03%) in the real-world DHM case, reflecting the drawbacks of the current urban-planning approaches that highly depend on long-distance transportation. Meanwhile, although the decentralized heuristic outperforms the traditional centralized method with respect to most metrics: service from 0.4514 to 0.4867 (+7.82%) and ecology from 0.6090 to 0.6624 (+8.77%) in the synthetic case; service from 0.5833 to 0.6000 (+2.86%) and ecology from 0.5178 to 0.5549 (+7.16%) in the real-world HLG case; and service from 0.5533 to 0.6067 (+9.65%) and ecology from 0.5562 to 0.5248 (−5.65%) in the real-world DHM case, it is still much worse than our DRL approach: service from 0.6833 to 0.4867 (−28.77%), ecology from 0.9171 to 0.6624 (−27.77%) and traffic from 0.9384 to 0.5008 (−46.63%) in the synthetic case; service from 0.7100 to 0.6000 (−15.49%) and ecology from 0.7303 to 0.5549 (−24.02%) in the real-world HLG case; and service from 0.7484 to 0.6067 (−18.93%) and ecology from 0.8976 to 0.5248 (−41.53%) in the real-world DHM case, indicating that handcrafted rules are inadequate to achieve an effective layout with optimal efficiency.

The above heuristics are based on fixed and non-optimizable rules, whereas results show that neither centralized nor decentralized rules may be optimal. Thus, we further include an optimizable baseline method based on the widely used genetic algorithm (GA). Results in Table 1 show that GA indeed outperforms the two rule-based heuristics on traffic efficiency for the synthetic community with substantial relative improvements from 0.5008 to 0.6019 (+20.19%), showing the advantage of optimizable parameters over fixed density-first rules. However, the spatial efficiency of land use attained by GA is roughly the same as or even worse than the two heuristics, much worse than our DRL approach: service from 0.6833 to 0.5000 (−26.83%), ecology from 0.9171 to 0.5008 (−45.39%) and traffic from 0.9384 to 0.6019 (−35.86%) for the synthetic community; service from 0.7100 to 0.5700 (−19.72%) and ecology from 0.7303 to 0.6312 (−13.57%) for the real-world HLG community; and service from 0.7484 to 0.4333 (−42.10%) and ecology from 0.8976 to 0.5905 (−34.21%) for the real-world DHM community. It is worthwhile to notice that GA is tested as a sub-module in the whole framework because it is difficult for GA to express the problem of block division. We also include a geometric set-coverage model with adaptations for single-step planning (GSCA). However, as the planning conditions are dynamically changing with newly planned land use, the geometric set-coverage approximation cannot capture well the long-term spatial efficiency, with its performance similar to the GA and rule-based heuristics, substantially outperformed by DRL approaches. The results illustrate the intrinsic complexity of the planning task, as well as the necessity of parametric value estimation and effective exploration in a reduced action space, which is accomplished in the proposed DRL framework.

To verify the effectiveness of our designed graph model and GNN, we introduce another DRL baseline that replaces the GNN with a multi-layer perceptron (MLP) model. Results in Table 1 show that GNN plays an essential role, improving the efficiency of spatial layout substantially: service from 0.5625 to 0.6833 (+21.5%), ecology from 0.7571 to 0.9171 (+21.1%) and traffic from 0.6857 to 0.9384 (+36.9%) for the synthetic community; service from 0.6300 to 0.7100 (+12.7%) and ecology from 0.7061 to 0.7303 (+3.4%) for the real-world HLG community; and service from 0.6621 to 0.7484 (+13.0%) and ecology from 0.6104 to 0.8976 (+47.1%) for the real-world DHM community. The spatial relationship between different geographic elements (lands, roads and junctions) directly reflects the actual planning performance, such as the contiguity between residential areas and facilities. In our framework, we use GNN to capture such relationships, and the message-passing mechanisms in GNN effectively extract the neighborhood information for each node. Therefore, the obtained embedded representations from GNN summarize well the current planning conditions and efficiency, which benefits the training of better value prediction and policy optimization. Please refer to Supplementary Section 2.3 for details on the comparison with baseline methods. Meanwhile, the number of GNN layers and the dimension of node embedding also play important roles in the planning performance: for example, too many GNN layers can lead to over-smoothing but too few may exhibit limited ability in capturing topological information (see the detailed hyper-parameter study in Supplementary Section 2.8).

Demonstration of planning a 15-minute city

The 15-minute city means that most residential needs can be satisfied within 15 min by walking or cycling (Fig. 1b), which will reduce the dependency on vehicular transportation and achieve a more accessible, equitable, low-carbon and resilient city. As shown in Fig. 1a, we renovate the real-world HLG community to make it fit the 15-minute-city concept. In other words, with existing residential blocks fixed, our agent can generate spatial plans of various facilities to improve the layout efficiency of the community. In our framework, we set up multiple land-use types corresponding to residential needs, for example, office, business, school, hospital and recreation for working, shopping, education, medical care and entertainment. In 15-minute-city planning, our model lays out these functional land-use types in a more decentralized way, such that various facilities are within walking distance from existing residential blocks. To evaluate the service accessibility, we calculate the number of residential blocks that are less than 500 m from a corresponding facility for each of the five service types, as an indicator for satisfying the service need within 15 min.

The effect of the 15-minute city is closely related to the needs of facilities, that is, requirements on the amount of different land-use types. We set the planning needs as predefined configurations of the environment in our framework because it is usually determined by a broader class of elements such as demographic, economic, social and public demand. The DRL agent will generate spatial plans that maximize the spatial efficiency under all these requirements. We investigate community renovation with different facility needs, including low, medium, high and a mixed needs of five land-use types, as shown in Fig. 1c. Specifically, we utilize a pretrained model of our framework and generate community renovation plans through direct model inference under new planning requirements without fine-tuning. Figure 1d shows the proportion of residential blocks that can satisfy five basic services within 15 min. With a low need of facilities, the accessibility of service is limited, with only about 30% of the residential blocks achieving the 15-minute-city goal. After increasing the needs of facilities from 2 to 4 (medium), service accessibility improves effectively, and an average of 55% of communities reaches the goal of the 15-minute city. In addition, the gain from further increasing the needs (high) tends to diminish, for example, the proportion of communities that can reach recreation sites for entertainment only increases from 65% to 70%. Meanwhile,

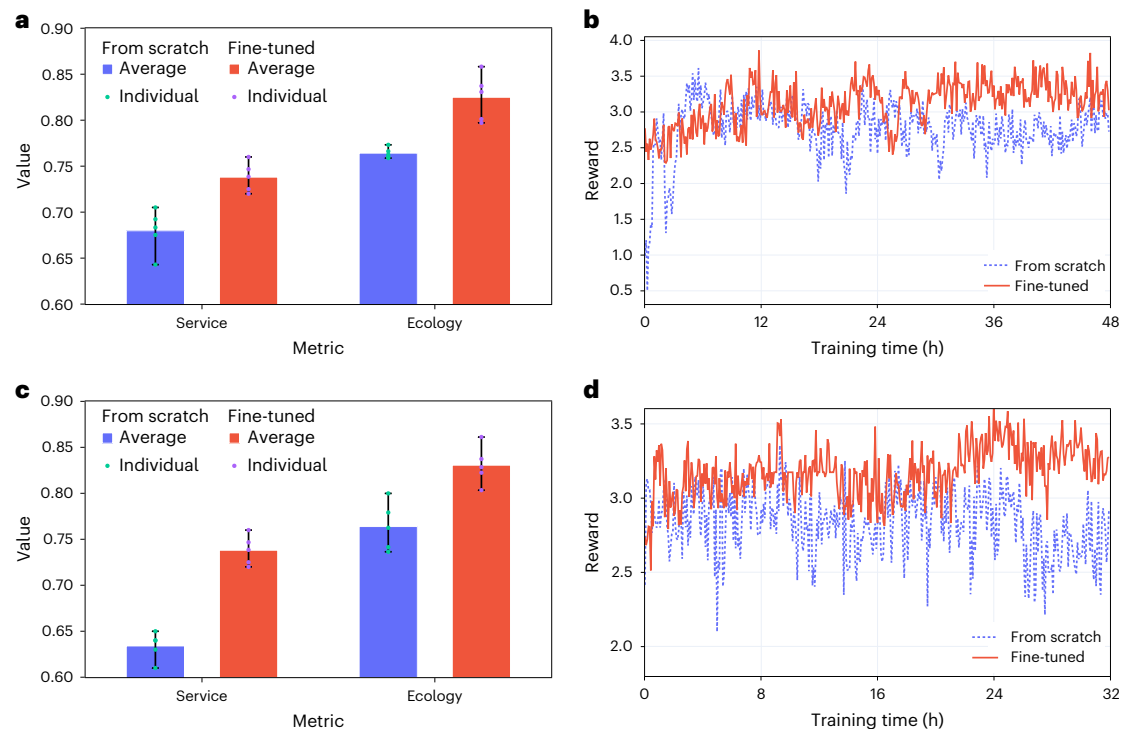


Fig. 2 | Demonstration of transferring pretrained models to different spatial planning tasks. a, b, Transfer between different community scales, from a small grid community with 3×3 blocks to a large grid community with 4×4 blocks. **a**, Service and ecology performance for training from the pretrained model versus training from scratch. The bar chart shows the average performance over five repeated experiments and the dots indicate individual performance of each experiment. The error bars show the maximum and minimum metric values of the five repeated experiments. **b**, The convergence plot of the two models.

c, d, Transfer between different community forms, from simple synthetic grid community with 3×3 blocks to the complicated real-world community in Fig. 1a. **c**, Service and ecology performance for training from the pretrained model versus training from scratch. The bar chart shows the average performance over five repeated experiments and the dots indicate individual performance of each experiment. The error bars show the maximum and minimum metric values of the five repeated experiments. **d**, The convergence plot of the two models.

the renovation planning requirements for different communities vary, such as higher demand for schools and offices in communities with more young residents (mix in Fig. 1c). The corresponding results in Fig. 1d demonstrate our framework's ability to design a community according to customized requirements, in which education and work are more accessible. In real-world scenarios, the needs of different services vary depending on economic constraints or political realities, and our framework is flexible in generating spatial plans of different service requirements (see results on the other DHM community in Supplementary Section 2.5).

Demonstration of model transferability

With varying objectives, requirements and initial conditions, community spatial planning tends to have quite different goals, which requires the planning model to be able to transfer between different planning tasks. Specifically, the transferability makes it possible to first obtain a decent pretrained model under moderate scale and simplified conditions, and then adapt the pretrained model to much more complicated planning scenarios, which can substantially reduce the time cost of model training. In previous experiments of the 15-minute city, we show that our DRL model can achieve successful transfer between different planning needs of the same community (Fig. 1c,d), by conducting direct model inference with different service-facility supplies. We now further investigate two more challenging settings to study the transferability of our model, which are transfer from small to large scale and transfer from simple to complex scenario (Supplementary Section 2.6).

As we focus on community-level spatial planning, which usually concerns an area of a few square kilometers and a side length of less

than 3 km (ref. 36), we investigate how our model transfers between different scales of synthetic grid communities, from a small community of 3×3 blocks (about 3 km^2) to a large community of 4×4 blocks (about 6 km^2), with a doubled decision-sequence length. Figure 2a shows the performance of the above transfer process. Training from a pretrained model achieves much better spatial efficiency than training a new model from scratch, making a relative progress of about 7.95% in the service metric and improving the ecology metric by about 4.57%. The improvements are substantial and stable, as shown by the small error bars in Fig. 2a. We also illustrate the convergence curves of the two models in Fig. 2b, in which the pretrained model starts from a much higher reward and adapts to the larger-scale community with a better final performance. It is worth noting that transferring to even larger scales is technically possible; however, the expanded area will go beyond the focus of a community-level planning task and involves much more complicated factors.

The results of transferring between different forms of communities are similar. Figure 2c demonstrates the performance of transferring from the simple grid scenario to the complicated real-world HLG community renovation scenario. Training from a pretrained model achieves a substantially higher service metric with relative progress of about 4.69%, and the ecology metric is also better than training from scratch by about 4.63%. The convergence plot in Fig. 2d also shows that the pretrained model starts from a higher position (+12.5% episodic reward) and ends with much better performance (+13.8% episodic reward) than the randomly initialized model. The above experiments verify that our model can learn universal skills for improving spatial layout efficiency from small-scale and simple spatial planning tasks,

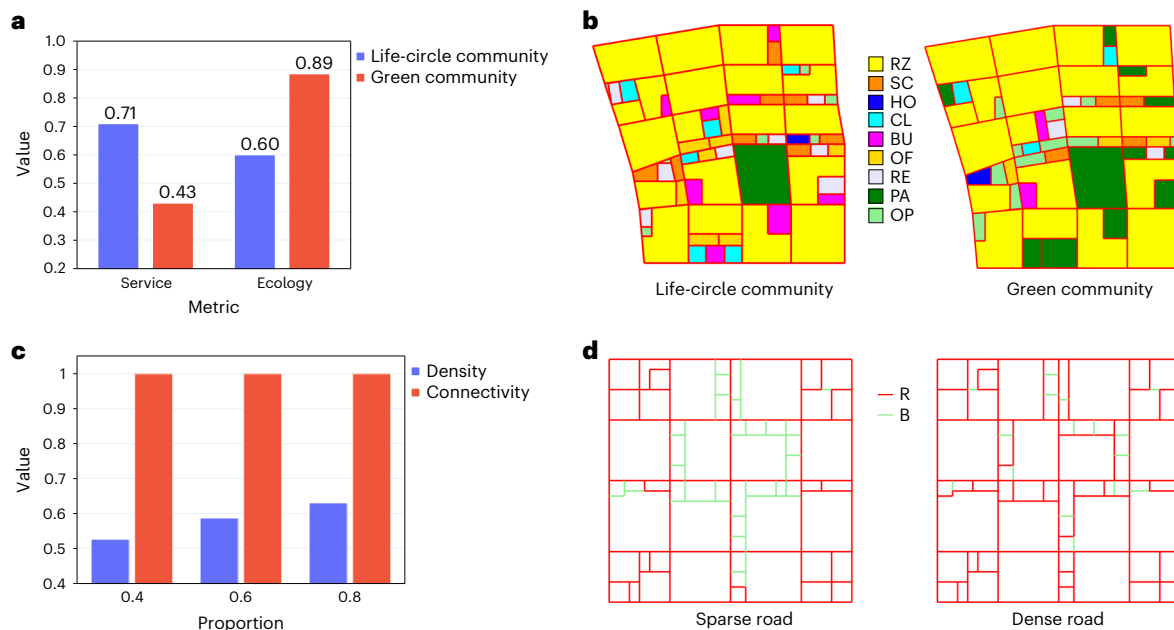


Fig. 3 | Demonstration of community plans of different styles. a,b, Spatial planning for different land-use attributes. **a,** Service and ecology performance for life-circle community and green community. **b,** The generated spatial plans for the life-circle community and the green community. Please refer to Supplementary Table 1 for the color legend. **c,d,** Spatial planning for different

road densities. **c,** The density and connectivity of planned roads with different proportions of land-use boundaries being selected. **d,** The generated plans for different road densities. Please refer to Supplementary Table 2 for the color and acronym legend.

and such skills can be successfully transferred to different spatial planning tasks that are much larger and more complicated.

Generating plans of different styles

There is no one-size-fits-all design in urban planning, and different planning styles essentially influence urban activities in the planned city³⁷. Thus, it is of vital importance for urban-planning models to be able to generate multiple candidate plans in different styles for selection. Particularly, spatial planning for communities can be oriented differently for the sustainable development of cities^{38–40}. For example, service-oriented planning emphasizes the accessibility of urban services and results in a life-circle community. Ecology-oriented planning emphasizes the coverage of parks and green space and leads to a green community. Meanwhile, communities can have sparse or dense road networks and show different accessibility. We now demonstrate the capability of our framework in generating community plans of different styles.

We first investigate community renovation with different land-use settings. We adjust the planning requirements, increasing the number of facilities in life-circle-community planning and the number of parks in green-community planning. In addition, we also adapt the reward weight, α in equation (1), to emphasize different aspects, with higher service weight ($\alpha = 2.0$) for life-circle-community planning and lower service weight ($\alpha = 0.5$) for green-community planning. Figure 3a shows the performance with respect to service and ecology for the two generated spatial plans. The planned life-circle community achieves 0.71 in service, 65.1% higher than that achieved by the planned green community that conversely achieves much better (+48.3%) ecology value. Figure 3b illustrates the two generated spatial plans, in which our framework lay outs facilities dispersively for the life-circle community and places parks evenly for the green community, demonstrating the ability of our framework to design for different land-use settings (see Supplementary Section 2.8 for more results on reward weight).

We then study different forms of road planning in the synthetic grid community. Our framework can easily generate sparser or denser road networks by varying the total number of steps in the road planning task, that is, changing the termination step to decide how many land-use boundaries are selected and built into roads by the agent. We set the total number of steps as 40%, 60% and 80% of the feasible candidates (land-use boundaries), respectively. Figure 3c shows the density and connectivity of the generated road plans under different proportions. Our framework generates road plans of different densities according to planning needs, that is, denser roads with a larger number of planning steps, and ensures optimal connectivity (1.0) in all cases. As shown in Fig. 3d, in both sparse (40%) and dense (80%) road plans, the generated road networks contain only one connected component and no dead-end road. All these results demonstrate that our framework can generate spatial plans of different styles in both land-use attributes and road density.

Comparison and collaboration with professional human designers

We have demonstrated the advantage of our DRL framework in optimizing spatial efficiency over rule-based heuristics and GAs. One natural question is how the framework performs compared with human designers, because most of the urban plans are accomplished by human designers currently. Therefore, we invited eight professional and experienced urban planners (see Supplementary Table 4 for information on invited planners) to conduct the real-world HLG and DHM community spatial planning tasks and evaluated their performance. These experts were given the same initial conditions and planning requirements as our DRL model, and we evaluated the same two efficiency metrics, service and ecology. Figure 4a,b shows the planning performance on two communities, and we can observe that compared with human planners with many years of experience, our DRL model still

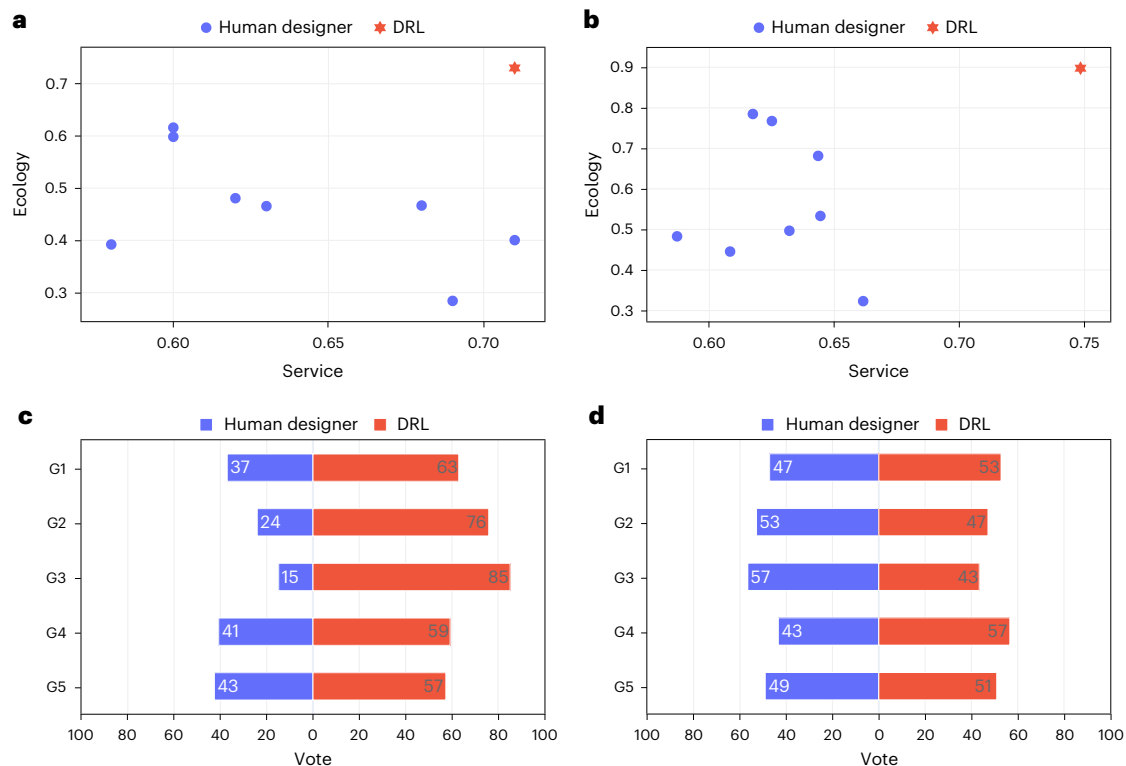


Fig. 4 | Comparison and collaboration with professional human designers.

a,b, Comparison of service and ecology metrics between our DRL model and eight invited professional human designers. **a,** The performance on HLG. **b,** The performance on DHM communities. The red hexagram indicates the performance of our DRL model and the blue dots indicate the performance of

human designers. **c,d,** Votes received by human designers and our DRL model in blind tests. The red and blue bars indicate the number of votes received by our DRL model and human designers, respectively, for each group from 100 professional designers as participants. **c,** Five groups for the HLG community. **d,** Five groups for the DHM community.

has a substantial advantage in searching for higher spatial efficiency, achieving Pareto optimal against all eight human designers in both communities. Specifically, the DRL model improves service and ecology efficiency by 13.6% and 15.4%, respectively, against the best human designer in the DHM community, and improvements against the human average are larger, surpassing 15.38% and 59.65% in service and ecology efficiency, respectively. Notably, our model achieves more balanced performance on the efficiency of different facilities; for example, the DRL model attains the best accessibility in three out of the five different basic residential needs for the DHM community (see Supplementary Figs. 5 and 6 and Supplementary Section 2.4 for details on comparison with professional human designers).

With the better performance of optimizing spatial efficiency, our AI model can help liberate human designers from heavy planning work and improve their productivity. Because human designers are good at conceptual prototyping and AI performs well in searching a large solution space, we propose a human–AI collaborative workflow (Supplementary Fig. 7a). In this workflow, human designers generate conceptual plans through centers and axes (Supplementary Fig. 7b) that are fed to AI models and make adjustments to the spatial plans generated by AI (see ‘Methods’ for details on the human–AI workflow). We compare this human–AI collaborative workflow with a fully human-labor workflow in objective efficiency metrics, subjective blind tests and time cost. Results show that the human–AI collaboration can substantially reduce the time consumption of planning without compromising but even improving the objective efficiency metrics (Supplementary Table 5). For example, in the DHM community, the human–AI collaborative workflow achieves Pareto optimal in service and ecology, improving their efficiency by 12.3% and 5.0%, respectively. Meanwhile, the time cost for AI (<1s) to

generate spatial plans is 3,000 times less than that for human designers (50–100 min). Nevertheless, it only takes about 2 days to obtain a well-trained AI model, which is much less than years for training a human designer.

Conversely, subjective judgments of spatial plans that cannot be easily quantified exist, for example, creativity, systemicity or targetedness^{41,42}. Thus, we invited 100 post-graduate-level human designers to participate in a blind test. In this test, they were asked to choose from two spatial plans based on personal experience and preference while unaware of whether the spatial plans are generated by human designers or AI (ten groups of two spatial plans, see Supplementary Fig. 8 for the blind test). Figure 4c,d shows the votes received by human designers and our DRL model; participants do not show clear preference in most cases, but AI wins substantially more votes in several spatial plans (G2 and G3 for HLG community). Through objective metric evaluation and subjective blind test, we demonstrate the feasibility of improving the productivity of human designers by collaboration with our proposed AI framework (see Supplementary Section 2.9 for details on the results of human–AI workflow).

It is worth noting that the proposed AI framework is not a replacement for human planners but a supporting tool to increase their productivity. In the above human–AI collaborative workflow, we take out the tedious computational work and leave it to the AI, allowing human planners to focus more on conceptual work, such as generating suitable axes and centers at the beginning. Furthermore, updating planning principles and ideas along with the continuous advancement of urban science requires the substantial participation of human planners. We regard cities as the masterpiece of human intelligence, and with the development of AI technology, humans will play a more critical role in creating a sustainable urban life.

Discussion

With urban planning being a long-standing problem, researchers have devoted decades of efforts to developing computational models and supporting tools for it in order to automate its process^{13–17}. Human urban planners are now several orders of magnitude more productive with geographic information system tools than they were decades ago, yet tedious layout tasks still need to be done manually. In our study, we demonstrate the potential of data-driven methods in urban planning, a task that previously relied heavily on human experiences and intuitions. In particular, our model can achieve much better planning performance as the training data increases. Because of the limitation of computational resources, we only collect training samples (interactions with the environment) in parallel using multiple threads on a single server and train the neural networks on a single graphics processing unit (GPU). In our experiments, we collect a million-scale dataset for each community, and the training lasts for about 48 h (Supplementary Section 2.7). Even though the number of generated spatial plans exceeds 1 million, it is still not large enough compared with the datasets leveraged in similar DRL-based tasks such as Go playing²⁵ and chip design²⁸. Efforts could be made to further increase the data volume to train our model, for example, to collect planning episodic samples on distributed clusters and train the model with multiple GPUs. It is worth noting that the decomposed three sub-spaces (what to plan, where to plan and how to plan) may be jointly optimized by the agent; however, the problem will become much more complicated and require much more training samples. Moreover, the what to plan part can be extended to include other important elements regarding urban sustainability, such as public transport routes and stations. Furthermore, different functionalities can be considered and planned simultaneously to better capture their synergies.

Despite the objective metrics and subjective concepts that have been incorporated into our framework, spatial planning also involves other subjective evaluation metrics such as esthetic and artistic scores, which can be further included in our framework. For example, some recent works^{43,44} have investigated training DRL models to meet human preferences or to align with human values. Meanwhile, it is worthwhile to notice that our current framework is guided mainly by static metrics and can generate community plans with high spatial efficiency. Beyond static planning based on spatial efficiency indicators, our framework can have major impacts when combined with a city simulator or a digital twin city via hierarchical multi-agent reinforcement learning⁴⁵, enabling a complete cycle of city-wide planning, construction and development, through optimizing spatial layouts based on dynamic evolutionary results such as simulated pedestrian and vehicular traffic flow⁴⁶.

In addition, this study provides a reference model for data-driven urban planning, whose technical framework may be extended from the community level to the entire city. However, scaling the method to the city level requires a substantial increase in training samples collected from distributed clusters, as well as larger neural networks trained with multiple GPUs on multiple servers. Additionally, planning for an entire city is a much more complex task, with diverse goals that include economic growth and resident health. Evaluating the impact of a city-level plan through a few static metrics is nearly impossible. Therefore, a city simulator can be used to comprehensively compare different plans and provide effective feedback to the planning model. The proposed method can serve as a starting point for automated planning at a larger scale, paving the way for a bright future.

Our experimental results demonstrate that generating spatial plans through DRL is viable and effective and can be flexibly adjusted to various planning concepts and styles. Practical urban planning can be much more complicated because of the tendency for various constraints and rules. For example, business zones are better to be placed near subway stations to maximize their economic benefits and schools should not be too narrow to fit the playground. In most of our

experiments, we disregarded these hundreds of rules to investigate spatial planning in a relatively clean and concise setting. In addition, critical issues regarding land ownership, public rights of way, urban segregation and gentrification need to be considered in real-world urban planning. Fortunately, with necessary and reasonable adaptations, our proposed approach can well handle these rules and political realities of decision-making in practical planning. For instance, these experiential rules can be easily integrated into our framework by adding action masks. Moreover, our framework can generate spatial plans that are fully consistent with planning rules from practical realities without damaging spatial efficiency (Supplementary Section 2.10 for details).

A spatial layout is a combinatorial result of both human creativity and strategic planning, in which it is always indispensable for human designers to provide initial conceptual insights. The presented model we develop is by no means a replacement for human designers but rather serves as an assistant for them. Human designers and AI models have distinct specialties in synthesizing conceptual prototypes and optimizing spatial efficiency, and we demonstrate a possible simple way of human–AI collaboration. In our experiments, we show that the AI model completes the complex and tedious planning work on the initial conceptual blueprints provided by human designers, which greatly automates spatial planning and substantially increases the productivity of human designers. By showcasing the collaboration with professional human designers, the proposed framework can inspire advanced urban-planning approaches, such as an iterative workflow between human designers and AI models to generate spatial plans based on multiple rounds of feedback.

Urban planning is never a simple game of selecting locations for land use and roads but rather a complicated interaction among multiple stakeholders, such as the government, citizens, planners and developers. Besides assisting planners to accelerate the process of spatial layout, our proposed approach can bring broader benefits to other participants. For example, our developed planning environment can provide precise quantifiable indicators of the 15-minute city, helping citizens and the government compare different community plans. In addition, public platforms can be constructed by introducing customized options to the proposed model, which can facilitate the participation of residents and developers in the planning process. Our proposed framework demonstrates the possibility of higher engagement of all participants, serving as a small step towards a more transparent and inclusive future of urban planning.

Methods

Problem formulation

We formulate the problem of community spatial planning as a sequential Markov decision process (MDP), an interactive process between the planning agent and the environment, in which the agent observes the ‘state’ (the current conditions of the community), takes an ‘action’ (placement of an urban functionality) at each step and receives ‘reward’ (effect of the planned result) signaled by the environment, which undergoes a ‘transition’ (changes of the layout) according to the agent’s action. We utilize DRL to learn an effective policy that maps states to actions with a parameterized neural network. The neural network is optimized towards higher spatial efficiency through massive training under the MDP, with millions of training samples of the 4-tuple (state, action, reward and transition). As illustrated in Extended Data Fig. 1, our MDP is composed of two consecutive stages:

- (1) Land-use planning. Given the initial road conditions, the agent places functionality blocks one at a time, either near existing roads or near boundaries of previously placed land use. After all the functionalities and open spaces are allocated, a reward regarding the efficiency of land use is returned to the agent, which treats different land use as an integrated system. The final land-use plan becomes the initial condition of road planning.

- (2) Road planning. Boundaries of planned land use are viable locations for road construction. The agent builds roads iteratively, turning one boundary into a road segment at a single step. Stopped at a predefined termination step, a reward considering the transportation efficiency is returned to the agent.

The reward is only calculated at the last step of each stage to summarize the performance of land-use planning and road planning, respectively, and all intermediate steps receive a reward of 0. We define the reward for land use and road layout based on the 15-minute-city concept^{4,8}, which emphasizes the spatial efficiency to facilitate active transportation such as walking and cycling instead of automobiles. The two reward terms are calculated as follows:

$$r_L = \alpha \text{Service} + \text{Ecology}, \quad (1)$$

$$r_R = \text{Traffic}, \quad (2)$$

where Service measures the community life-circle index in the 15-minute city, Ecology measures the coverage of green space and parks, Traffic is a combination of road density and connectivity (Methods) and α serves as a hyper-parameter indicating the weight of service performance in the land-use reward. With the calculated reward values, we use proximal policy optimization to update the parameters of value and policy networks⁴⁷. We first train the agent for the land-use planning task until convergence and then train the agent to build roads with the optimal land-use plan obtained in the first stage. After two stages of training, the AI agent is able to design communities with an efficient spatial layout of both land use and roads.

Graph model. Different from previous Go²⁵ and chip design²⁸ tasks, urban planning is more challenging because of its much larger degrees of freedom in the problem form. Specifically, the conditions of previous tasks are regular, for example, placing stones on a 19×19 board or placing rectangular macros onto a grid chip canvas, which can be represented by pixels (raster). By contrast, the conditions for community spatial planning are diverse and irregular because the road corners and land blocks are usually not orthogonal. To accurately describe urban geographic elements, including land blocks (L), segments of roads and land-use boundaries (S) and junctions between roads and land-use boundaries (J), we use vector representations, which have been proven to have substantial advantages over raster representations in urban planning⁴⁸, and consist of the following three geometries:

- (1) ‘Polygon’ that describes a vacant land to be planned (for example, L1 in the right part of Extended Data Fig. 2a) or an already planned land block (for example, L2 in the right part of Extended Data Fig. 2a) with the coordinates of the land boundary;
- (2) ‘LineString’ that represents a road segment (for example, S3 in the right part of Extended Data Fig. 2a) or a boundary edge of a land block (for example, S9 in the right part of Extended Data Fig. 2a) with the coordinates of the start and end points; and
- (3) ‘Point’ that stands for the junctions between roads and land-block boundaries (for example, J2 and J7 in the right part of Extended Data Fig. 2a) with their coordinates.

We transform all geographic elements into the above three categories of geometries and then represent the whole community as a graph, in which nodes are the geometries and edges stand for the spatial contiguity relationship between these geometries, that is, two nodes are connected if the underlying two geometries touch each other. Each node stores its geographic information as the node features, including the type, coordinates, width, height, length and area of the geometry. In this way, spatial planning can be transformed as a problem of making choices on a dynamic graph (Extended Data Fig. 2), in which the graph evolves according to the agent’s actions.

In the land-use planning task, the agent selects one L–J edge that connects a vacant land and a junction, placing a given functionality at the location specified by the corresponding L and J (Extended Data Fig. 2a). In each step, the topology of the contiguity graph changes because the newly placed functionality generates new nodes and edges. New nodes include the new functionality itself, its boundaries, new junctions and split segments. New edges indicate the newly established spatial contiguity. Similarly, in the road planning task, the agent selects one S node that is currently a boundary and constructs it as a road segment (Extended Data Fig. 2b). Although the topology remains the same, the graph’s attributes alter because the selected node’s type changes from boundary to road. Through the problem reformulation with the graph model, we can now handle the irregular urban blocks and unify the two seemingly distinct stages of land use and road planning on one single graph.

Action space design. Another major challenge of urban planning is the huge action space, which is almost infinite in the original continuous space, and still too large in the reduced discrete graph space. The contiguity graph continues to grow as we place a functionality at each step, resulting in a large graph with thousands of nodes and edges. A typical spatial plan of a 2 km by 2 km community can take a total number of 100 planning steps in each stage, and the contiguity graph can have 4,000 edges and 1,000 nodes, which makes the action space $4,000^{100}$ and $1,000^{100}$ for the two stages, respectively. In addition, valid actions are extremely sparse in the space, and a substantial portion of actions is of low quality and will lead to unreasonable results, such as placing a facility in the center of a vacant land without connecting roads. Therefore, it is crucial to reduce the action space and avoid unreasonable actions.

To address this challenge, we propose a general DRL framework in which an intelligent agent perceives and makes decisions in a reduced graph space, and the environment handles urban elements in the original geographic space and generates graph states according to the geographic spatial layout. Meanwhile, we decompose the entire action space into a Cartesian product of three sub-spaces, including what to plan, where to plan and how to plan, and let the DRL agent focus on the core issue of where to plan. The first sub-space of what to plan can be eliminated by fixing the planning order of different land-use types through domain knowledge, allowing land-use types that are more dependent on the initial road network to be planned earlier (Methods). To avoid apparently improper actions in where to plan, we impose planning constraints on the agent’s actions, with an action mask that blocks out unreasonable options, that is, only L–J edges and S nodes are candidates for the two planning stages, respectively. After selecting one L–J edge for a given land-use type, the functionality is placed in the corresponding land block (L node) at the location of the corresponding junction (J node), whose shape and size are determined by predefined rules that maximize the reuse of existing roads and boundaries (Methods); thus, the last sub-space of how to plan is effectively eliminated. Through these designs, we narrow the action space to a solvable scale and filter out most unreasonable actions, enabling efficient optimization for DRL algorithms. In summary, the original problem of spatial planning is successfully transformed into a standard sequential decision-making process on a graph with moderate action space.

Framework. After the above problem reformulation and action space design, we propose a DRL framework in which an AI agent learns to lay out land use and roads by interacting with the spatial planning environment, as illustrated in Extended Data Fig. 3. The sequential MDP (Extended Data Fig. 3e,f) contains the following key components:

- (1) States summarize the current spatial plan with the previously introduced contiguity graph containing rich node features, and other information, such as statistics of different land use types.

- (2) Actions indicate the locations to place the current land use or construct a new road segment, which are transformed from the selected edges or nodes in the contiguity graph.
- (3) Rewards are 0 for all intermediate steps, except for the last step in each stage, in which it evaluates the spatial efficiency of land use and roads.
- (4) Transitions describe the changes of the layout given the selected location, and the transitions occur in both the original geographic space (new land use and road on the map) and the transformed graph space (new topology and attributes of the graph).

At each step, the agent represents the state by encoding the graph with a GNN⁴⁹. Via multiple message passing and non-linear activation layers, the GNN state encoder generates effective representations of edges, nodes and the whole graph (Extended Data Fig. 3a), which will be leveraged by the value and policy networks (Extended Data Fig. 3b–d). Specifically, because choosing locations for land use is equivalent to selecting edges on the graph, the land-use policy network takes the edge embeddings and scores each edge with an edge-ranking MLP, as shown in Extended Data Fig. 3b. The obtained score for each edge indicates the sampling probability of the corresponding edge, which is returned to the environment and becomes the probability of placing the land use at the location specified by that edge. Similarly, in road planning, the road policy network takes node embeddings and scores each node with a node-ranking MLP (Extended Data Fig. 3d), outputting the probability of choosing one land block boundary and building it as a road segment. Finally, the value network takes in the graph embedding that summarizes the whole community and predicts the planning rewards with a fully connected layer (Extended Data Fig. 3c). To master the skills of spatial planning, millions of spatial plans are accomplished by the proposed model to search the large solution space during the training process, which is utilized as real-time training data to update the parameters of the neural network.

Detailed methodology

Initial conditions. As introduced in the paper, we use vector geometries including Polygon, LineString and Point to describe urban geographic elements. Specifically, there are ten types of land blocks that are represented as Polygon, including the initial vacant land to be planned, and nine different functionality types, which are residential (RZ), school (SC), hospital (HO), clinic (CL), business (BU), office (OF), recreation (RE), park (PA) and open space (OP). In addition, there are two types of segments (roads and land-use boundaries) that are represented by LineString and one type of junction (intersections between roads and land-use boundaries) that is represented by Point. Therefore, a community is faithfully represented by a table of geometries, in which each row is a geographic element with three columns of ID, type and geometry. Initial conditions of a community consist of all the original land blocks, roads and intersections, whose accurate coordinates are recorded by their corresponding geometries in the table of geometries. In the synthetic grid community, we experiment on a basic community with a size of 2.4 km by 2.4 km, containing 16 rectangular vacant lands, 40 horizontal or vertical initial road segments and 25 road intersections, as shown in the first step of Extended Data Fig. 1. In the real-world community, we replicate the road network of HLG and DHM communities in Beijing from OpenStreetMap using OSMnx³⁰ and geopandas, reserve residential blocks and leave other areas as vacant land to be renovated. Finally, we obtain two communities of around 4 km² as shown in Fig. 1a and Supplementary Fig. 11a.

Planning needs and requirements. Before carrying out the actual spatial planning, we need to determine the planning needs and requirements, which serve as the configuration of the planning environment. The planning need describes the amount that each land-use type has

to achieve, either in area or in number, for example, residential blocks of 50% community area and three hospitals. Meanwhile, we also have requirements on the minimum area (in square meters) of each planned block; for example, the area of one school is at least 10,000 m². Supplementary Table 3 shows an example of the planning needs and requirements for a community, in which 15% of the community area needs to be planned as parks; thus, it serves as a green community. Only spatial plans that satisfy all the needs and requirements are considered as successful episodes and reserved as training samples, and those failed episodes are discarded. In our framework, the planning needs and requirements are configurations for the environment, making our model highly flexible in generating spatial plans. Specifically, once we obtain a well-trained model under one configuration, we can simply change the configuration and directly perform model inference without re-training to generate plans for different planning needs and requirements, such as the community plans of different service supplies in Fig. 1c,d.

Planning order of land-use types. As introduced in the paper, in order to reduce the huge action space, we fix the planning order of different land-use types based on domain knowledge and make the agent focus on the core task of selecting locations. Because feasible locations are next to existing junctions, land use that is planned earlier will be closer to initial roads with more convenient traffic. Therefore, we first plan those facilities that depend more on roads, including hospitals (clinics), schools and recreation. Meanwhile, at later steps of land-use planning, the shape of feasible vacant lands tends to be more irregular and fragmented, which is not suitable for residential blocks that usually occupy a whole plot of land; thus, we distribute residential blocks after planning the above road-dependent facilities. Finally, we arrange those land-use types that are not much demanding in land shapes. After all the planning needs are satisfied, the remaining vacant lands are assigned as open spaces. In summary, the planning order in our framework is fixed as follows: hospital, school, clinic, recreation, residential, park, office, business and open space. Letting the agent determine the order of land-use types may be an alternative approach. However, it will make the problem much more complicated, as the action space is increased drastically. In practice, our fixed order generates sound spatial plans.

Land cutting. In land-use planning, the environment receives the action from the agent, which is the selected L–J edge, and cuts a new land from the corresponding land block (L node) at the location of the corresponding junction (J node). We develop a rule-based system with expert knowledge incorporated to determine the shape and size of the new land. The rule-based system is roughly composed of three steps: (1) Determine the relationship between J and L, such as in the middle of a road or at the corner. (2) Determine the reference line along existing boundaries from junction J, which can be I-shape, L-shape and U-shape. (3) Determine the length of inward extension from the reference line into the block L, forming the final sliced new land. The three steps are conducted according to expert knowledge, in order to meet the planning requirements and fit the current plan as closely as possible.

State. Our state contains three parts: (1) urban contiguity graph, (2) current object to be placed and (3) community statistics. We construct a graph to represent the current community information as illustrated in Extended Data Fig. 2, in which nodes are urban geographic elements and edges indicate the spatial contiguity relationship. We compute rich geographic attributes as node features, including the type, coordinates, area, length, width and height of the underlying urban element. The edges are represented by a sparse adjacency matrix. As for the current object to be placed, its type is determined by the environment according to planning needs and planning order, that is, the environment will traverse the planning order and transit to the next type if the planning needs for the previous type have been satisfied. We treat the

current object as a virtual isolated node, with its type feature provided by the environment and other node features left as default values. Lastly, community statistics include the area and count of different land-use types in the current plan, as well as the planning needs, which summarize the current conditions and the progress of spatial planning.

GNN state encoder. In our framework, the policy is used to select edges or nodes in the graph and the value is used to evaluate the whole graph. Thus, it is crucial to obtain effective encodings of nodes, edges and the whole graph for efficient policy-making and value evaluation. As shown in Extended Data Fig. 3a, we propose a GNN-based model to encode the current state. The message-passing and neighbor-aggregation mechanisms in GNN make it possible for each node to access the information of its neighbor nodes, which reflects the nearby situation in the original geographic space and is critical to the layout efficiency. Formally, we first concatenate all the node features and obtain an initial embedding for each node with a linear layer as follows:

$$v_i^0 = \text{Linear}(T_i \parallel C_i \parallel A_i \parallel L_i \parallel W_i \parallel H_i), \quad (3)$$

where T_i is the one-hot encoding of the type feature and the last five features are the coordinates, area, length, width and height, respectively. Then we feed node embeddings into several graph convolutional layers. In the k th graph convolutional layer, nodes broadcast their embeddings to the connected edges to achieve message passing. Each edge takes the passed message and obtains its edge embedding via a fully connected layer and a non-linear tanh activation function as follows:

$$e_{ij}^k = \frac{\tanh(\text{fc}^k(v_i \parallel v_j)) + \tanh(\text{fc}^k(v_j \parallel v_i))}{2}, \quad (4)$$

and then each node absorbs the embeddings of its connected edges to achieve neighbor aggregation as follows:

$$v_i^k = v_i^{k-1} + \frac{1}{n_i} \sum_{j \in N_i} e_{ij}^k, \quad (5)$$

where N_i denotes the neighbor set of node v_i and n_i is the number of neighbors. By stacking multiple graph convolutional layers, the obtained embeddings for each urban element (node or edge) can capture the nearby geographic conditions. We use the output of the last graph convolutional layer, v_i^L and e_{ij}^L , as the representations for nodes and edges, where L is a hyper-parameter in our model. Meanwhile, we adopt mean pooling on node and edge embeddings as follows:

$$v_{\text{mean}} = \frac{1}{M} \sum_{i=1}^M v_i^L, \quad (6)$$

$$e_{\text{mean}} = \frac{1}{N} \sum_{i,j \in E} e_{ij}^L, \quad (7)$$

$$g^L = v_{\text{mean}} \parallel e_{\text{mean}}, \quad (8)$$

where g^L is the obtained graph-level representation that describes the conditions of the whole community.

Besides the urban contiguity graph, the other two parts of the state, namely the current object to be placed v_c and the community statistics h_s , are also encoded. As introduced previously, the current object is regarded as a virtual isolated node with the same set of node features; thus, we reuse the linear layer to encode the current object:

$$v_c = \text{Linear}(T_c \parallel C_c \parallel A_c \parallel L_c \parallel W_c \parallel H_c), \quad (9)$$

where T_c is the type of the current object to be placed determined by the environment, and other features are default values. As for community statistics, we encode them via a two-layer MLP:

$$h_s = \text{MLP}(\text{area}_1 \parallel \dots \parallel \text{area}_9 \parallel \text{count}_1 \parallel \dots \parallel \text{count}_9), \quad (10)$$

where the area and count of all nine different land-use types are concatenated first and then fed into an MLP. Now all three parts of the state have been encoded by vectorized representations. Particularly, we now have representations for nodes, edges and the whole graph, which are adopted separately by policy and value networks.

Action. As illustrated in Extended Data Fig. 2a, land-use planning is reformulated as a sequential MDP in which the agent selects an edge in a dynamic graph. Therefore, the action space for land-use planning is the probability distribution of choosing from N edges, and we sample from this distribution to obtain the action. Similarly, road planning is a sequential MDP of choosing nodes as shown in Extended Data Fig. 2b; thus, the action space for road planning is the probability distribution over M nodes, which is sampled to generate the node selection action. In addition, as introduced previously, we impose constraints on the action space; for example, the agent can only select L–J edges (between vacant lands and junctions) and S nodes (land-use boundaries) to avoid unreasonable spatial plans. Thus, we calculate a mask in each step that indicates feasible options, and the probability distribution will be multiplied by the mask, allowing only feasible edges or nodes to be sampled as actions.

Policy and value networks. As shown in Extended Data Fig. 3b–d, we develop separate policy networks to take actions in the policy and value stages, respectively, as well as a value network to predict the performance of spatial plans. The three networks share the same state encoder to obtain state representations, taking full advantage of the GNN. Policy networks generate the probability distribution by scoring the edges and nodes in the graph, and then sample from this distribution to take actions. Meanwhile, the value network evaluates the whole graph to predict spatial efficiency, providing feedback for the community plan. In this section, we introduce the detailed design of the three networks.

Land-use policy network. In land-use planning, the agent places the current object at the location specified by the selected edge. The effect of edge selection is related to both the edge and the current object; for example, placing a hospital next to an already planned hospital may lead to low service efficiency. Therefore, the land-use policy network considers both the edge and the current object as input. As shown in Extended Data Fig. 3b, a feed-forward network, which is an edge-ranking MLP, is developed to score each edge:

$$s(e_{ij}) = \text{FF}_{\text{land}}(e_{ij}^L \parallel v_c \parallel e_{ij}^L - v_c \parallel e_{ij}^L \cdot v_c), \quad (11)$$

where the difference and the inner product of e_{ij}^L and v_c are also concatenated to emphasize the relationship between the current object to be planned and those already planned land use. The scores are converted to a probability distribution over all edges using softmax:

$$\text{Prob}(e_{ij}) = \frac{e^{s(e_{ij})}}{\sum_{s,t \in E} e^{s(e_{st})}}, \quad (12)$$

which is sampled to select an edge.

Road policy network. In road planning, the agent selects one boundary node and plans a road at its location. Different from that of land-use planning, the topology of the graph is stable with no new nodes to be added. Thus, there is no need to include the current object. Meanwhile,

the road policy network takes node embeddings as input, which already contain neighbor geographic information through message passing of GNN. As in Extended Data Fig. 3d, another node-ranking feed-forward MLP is adopted to score each node:

$$s(v_i) = \text{FF}_{\text{road}}(v_i). \tag{13}$$

The score is also transformed to probability with a softmax operator:

$$\text{Prob}(v_i) = \frac{e^{s(v_i)}}{\sum_{j \in N} e^{s(v_j)}}, \tag{14}$$

and we sample from this probability distribution to select one node.

Value network. As shown in Extended Data Fig. 3c, we develop a value network to judge the current planning situations and predict the planning performance. Because it is an overall evaluation of the entire community, we take the graph-level representation as the input of the value network. Meanwhile, we also include community statistics. Specifically, we concatenate graph representations and the statistics embedding, and adopt a fully connected layer to predict the performance:

$$\hat{v} = \text{fc}(g^t \parallel h_s), \tag{15}$$

where \hat{v} is the estimated value of the current plan.

Reward. We train the policy networks to optimize the efficiency of spatial layout, with respect to service, ecology and traffic. As in equations (1) and (2) of this paper, we define reward functions that give a comprehensive evaluation of the above metrics. Meanwhile, the reward values can be quickly computed within tens of milliseconds given a spatial plan, making it possible to collect large-scale samples for training DRL models. In this section, we introduce how the three metrics are calculated. It is worth noting that our framework is flexible and can be extended to include more metrics in the reward.

Service. We adopt the concept of 15-minute life circle⁴, which requires that basic services of the community be reachable for residents within 15 min by walking or cycling. Specifically, as demonstrated in Fig. 1b,c, we consider five different basic services, each of which is related to one or two facilities, that is, education (school), medical care (hospital, clinic), working (office), shopping (business) and entertainment (recreation). Therefore, the 15-minute life circle means that the distances between facilities and residential zones need to be less than the walking distance of 15 min, which is set as 500 m in our experiments. We define the service metric as the proportion of accessible services within 500 m, and the metric is averaged for all residential zones³¹. Formally, given a community spatial plan p , the service metric is calculated as follows:

$$d(i,j) = \min\{\text{EucDis}(\text{RZ}_i, \text{FA}_1^j), \dots, \text{EucDis}(\text{RZ}_i, \text{FA}_{n_j}^j)\}, \tag{16}$$

$$\text{Service}_i = \frac{1}{5} \sum_{j=1}^5 \mathbb{1}[d(i,j) < 500], \tag{17}$$

$$\text{Service} = \frac{1}{n_{\text{RZ}}} \sum_{i=1}^{n_{\text{RZ}}} \text{Service}_i, \tag{18}$$

where EucDis is the Euclidean distance, $d(i,j)$ is the minimum distance for the i th residential zone RZ_i to access the j th service that is provided by facility FA^j and n_j is the total number of facilities FA^j . Service_i is the 15-min life-circle metric for the i th residential zone, and we average over all n_{RZ} residential zones to obtain the final service metric for the whole community. This service metric guides the agent to arrange facilities

in a more decentralized way and close to residential zones, which is critical for increasing the ability of community services.

Ecology. The ecology of a community is important to the physical and mental health of residents; thus, we include an ecology metric that measures the layout efficiency of parks and open spaces. In general, parks and open spaces serve the residents who live in the neighborhood, and we hope they can serve as many residential areas as possible. Formally, we define the ecological serving range as the region within 300 m from a park or an open space, and the ecology metric measures the proportion of residential areas that are covered by the ecological serving range³². The metric is calculated as follows:

$$\text{ESR} = \text{Union}\{\text{Buffer}(\text{PA}_1, 300), \dots, \text{Buffer}(\text{PA}_{n_{\text{PA}}}, 300), \text{Buffer}(\text{OS}_1, 300), \dots, \text{Buffer}(\text{OS}_{n_{\text{OS}}}, 300)\}, \tag{19}$$

$$A_{\text{RZ}} = \sum_{i=1}^{n_{\text{RZ}}} \text{Area}(\text{RZ}_i), \tag{20}$$

$$A_{\text{RZ}}^e = \sum_{i=1}^{n_{\text{RZ}}} \text{Area}(\text{Intersection}(\text{RZ}_i, \text{ESR})), \tag{21}$$

$$\text{Ecology} = \frac{A_{\text{RZ}}^e}{A_{\text{RZ}}}, \tag{22}$$

where $\text{Buffer}(\text{PA}_i, 300)$ and $\text{Buffer}(\text{OS}_i, 300)$ represent the regions that extend the park and open space 300 m outward, which is their serving range, and ESR is the ecological serving range that combines the serving range of all parks and open spaces. The ecology metric encourages the agent to maximize A_{RZ}^e ; thus, the greenness of the community plan is promoted.

Traffic. For the second stage of road planning, we evaluate the traffic efficiency from three perspectives, including density³³, connectivity³⁴ and spacing³⁵. Road density is the ratio of the total length of roads to the land area. Connectivity is a network characteristic that reflects the strength of how different parts of a network are linked with each other, and we choose the number of connected components and the number of dead-end roads. To achieve appropriate road spacing, we also include two terms to penalize too large (>600 m) and too small (<100 m) spacing. Formally, the traffic metric is calculated as follows given the road plan p_{R} and the converted graph g_{R} from the planned road network:

$$T_{\text{density}} = \frac{\text{Length}(p_{\text{R}})}{A_c}, \tag{23}$$

$$T_{\text{connectivity}} = \frac{1}{\text{NCC}(g_{\text{R}})} + \frac{1}{1 + \sum_{v \in g_{\text{R}}} \mathbb{1}[\text{Degree}(v) = 1]}, \tag{24}$$

$$T_{\text{spacing}} = \frac{1}{1 + \sum_{r \in p_{\text{R}}} \mathbb{1}[\text{Length}(r) > 600]} + \frac{1}{1 + \sum_{r \in p_{\text{R}}} \mathbb{1}[\text{Length}(r) < 100]}, \tag{25}$$

$$\text{Traffic} = \frac{1}{3}(T_{\text{density}} + T_{\text{connectivity}} + T_{\text{spacing}}), \tag{26}$$

where Length calculates the length of a road segment, A_c is the area of the community, NCC calculates the number of connected components in a network and Degree calculates the degree of a node in the graph. Combining the three perspectives, the traffic metric encourages the agent to plan denser roads and, at the same time, guarantees

connectivity and appropriate spacing, without creating dead-end roads or planning too-long or too-short road segments.

Model training. We train our model for hundreds of iterations to learn the skills of spatial planning. In each iteration, we collect training samples of a few thousand episodes and update the parameters of our model using proximal policy optimization⁴⁷. Specifically, the loss function is a combination of policy loss, policy entropy and value loss. Policy loss is a surrogate clipped objective to improve the policy with safe exploration, which is calculated as follows:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}, \tag{27}$$

$$L_{policy} = \min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}_t), \tag{28}$$

$$\hat{A}_t = Q(s_t, a_t) - V(s_t), \tag{29}$$

where θ is the parameters of our model, $r_t(\theta)$ is the ratio of the probability of the new policy to the old policy, \hat{A}_t is the advantage function and clip restricts the update to be not too large. The entropy loss controls the balance between exploitation and exploration, which is calculated as follows:

$$L_{entropy} = \text{Entropy}[\text{Prob}(a_1), \dots, \text{Prob}(a_{n_a})], \tag{30}$$

where n_a is the total number of actions that equals to M (edges) or N (nodes) in different planning stages, and Prob is obtained by policy networks according to equations (12) and (14). We use mean squared error loss to supervise the value prediction:

$$L_{value} = \text{MSE}(\hat{v}_t, R_t), \tag{31}$$

where R_t is the return value from ground-truth and \hat{v}_t is estimated by value network according to equation (15). The final loss function is a weighted sum of the above three terms:

$$L = L_{policy} + \beta L_{entropy} + \gamma L_{value}, \tag{32}$$

where β and γ are hyper-parameters in our model.

Model inference. After we obtain a well-trained model, we perform model inference to generate community plans. We use the policy networks and compute the probability distribution over different actions according to equations (12) and (14), that is, the probability of selecting different edges and nodes. Then the most likely action is chosen to place land use or road at the location specified by the action:

$$a = \text{argmax}\{\text{Prob}(a_1), \dots, \text{Prob}(a_{n_a})\}, \tag{33}$$

where n_a is M or N for land use and road planning, respectively. It is worth noting that we can directly perform model inference under a different setup without re-training, and the results are illustrated in Fig. 1c,d.

Integration with manually designed planning concepts. The DRL framework is not designed for replacing human designers but serves as an intelligent assistant to improve the productivity of human designers. Specifically, AI models are good at optimizing spatial efficiency in large solution spaces, whereas human designers are good at conceptual prototyping. Therefore, we design a new workflow, in which human and AI collaboratively accomplish urban-planning tasks and leverage their respective expertise. As shown in Supplementary Fig. 7a, we propose a workflow with four key steps of conceptualization,

planning, adjustment and evaluation, where AI takes responsibility for the planning step. In the workflow, human designers can leave the heavy and specific planning work to the AI, and they only need to provide relatively abstract conceptual planning and make adjustments to the spatial plans generated by the AI. We represent planning concepts as two major types, center and axis, and each concept is related to one or several land-use functions. For example, in the left part of Supplementary Fig. 7b, the RE center in the HLG community represents the concept that encourages recreation zones near the specified location. Similarly, in the right part of Supplementary Fig. 7b, the BU&OF axis in the DHM community represents the concept that expects a business and office core along the specified band region. We feed the initial conditions of the community and the planning concepts to the model, and then train our DRL model to realize the planning concepts while at the same time optimizing spatial efficiency.

As the concept of center and axis is essentially the spatial relationship between specific land-use functions and predefined locations, it can be easily integrated into our framework. Specifically, we utilize customized reward functions to implement planning concepts, that is, we add a reward to reflect the extent of consistency with the planning concept. For the center concept, we calculate the fraction of concept-related land-use functions in the region near the specified center location as follows:

$$r_c = \frac{1}{n_c} \sum_{j=1}^{n_c} \mathbb{1}[T_j \in T_c], \tag{34}$$

where n_c is the number of land-use blocks within 100 m from the predefined center and T_c is the land-use function related to the concept. This reward encourages the DRL agent to place concept-related land-use functions near the predefined center location. For the axis concept, we calculate the expansion in the parallel direction and the coverage fraction in the vertical direction for the concept-related land-use functions, respectively, as follows:

$$r_a^p = \frac{L_a^p}{L_a}, \tag{35}$$

$$r_a^v = \frac{1}{n_a} \sum_{j=1}^{n_a} \mathbb{1}[T_j \in T_a], \tag{36}$$

$$r_a = \frac{1}{2}(r_a^p + r_a^v), \tag{37}$$

where L_a is the length of the axis, L_a^p is the distance of projected points of concept-related land-use functions on the axis, n_a is the number of land use blocks within 100m from the predefined axis and T_a is the land-use function related to the concept. This reward encourages the DRL agent to place concept-related land-use functions as evenly as possible in the band area around the axis. The concept reward is combined with efficiency reward, including service and ecology, by weighted sum. Through jointly optimizing efficiency and concept rewards, the DRL agent learns to improve spatial efficiency on the basis of realizing predefined planning concepts.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The geographic data of communities used in this work is collected from OpenStreetMap (<https://www.openstreetmap.org>) using Python 3.8.0 with packages osmnx>=1.1.2 and geopandas>=0.11.1. We provide the data for the three adopted communities in our experiments.

The data supporting the results of this study is available on Zenodo (<https://doi.org/10.5281/zenodo.8175420>)⁵⁰ and GitHub (<https://github.com/tsinghua-fib-lab/DRL-urban-planning>). Source data are provided with this paper.

Code availability

The code used in this research can be found at Zenodo (<https://doi.org/10.5281/zenodo.8175420>)⁵⁰ and Github (<https://github.com/tsinghua-fib-lab/DRL-urban-planning>).

References

- Xu, Y., Olmos, L. E., Abbar, S. & González, M. C. Deconstructing laws of accessibility and facility distribution in cities. *Sci. Adv.* **6**, eabb4112 (2020).
- Weiss, D. J. et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
- Nouvellet, P. et al. Reduction in mobility and Covid-19 transmission. *Nat. Commun.* **12**, 1090 (2021).
- Moreno, C., Allam, Z., Chabaud, D., Gall, C. & Pralong, F. Introducing the “15-minute city”: sustainability, resilience and place identity in future post-pandemic cities. *Smart Cities* **4**, 93–111 (2021).
- Ferrer-Ortiz, C., Marquet, O., Mojica, L. & Vich, G. Barcelona under the 15-minute city lens: mapping the accessibility and proximity potential based on pedestrian travel times. *Smart Cities* **5**, 146–161 (2022).
- Allam, Z., Nieuwenhuijsen, M., Chabaud, D. & Moreno, C. The 15-minute city offers a new framework for sustainability, liveability, and health. *Lancet Planet. Health* **6**, e181–e183 (2022).
- Allam, Z., Bibri, S. E., Chabaud, D. & Moreno, C. The “15-minute city” concept can shape a net-zero urban future. *Humanit. Soc. Sci. Commun.* **9**, 126 (2022).
- Weng, M. et al. The 15-minute walkable neighborhoods: measurement, social inequalities and implications for building healthy communities in urban China. *J. Transp. Health* **13**, 259–273 (2019).
- Balletto, G., Ladu, M., Milesi, A. & Borruso, G. A methodological approach on disused public properties in the 15-minute city perspective. *Sustainability* **13**, 593 (2021).
- Noworól, A., Kopyciński, P., Hatat, P., Salamon, J. & Hotuj, A. The 15-minute city—the geographical proximity of services in Krakow. *Sustainability* **14**, 7103 (2022).
- Merrell, P., Schkufza, E. & Koltun, V. Computer-generated residential building layouts. *ACM Trans. Graph.* **29**, 181 (2010).
- Landes, J. In *Artificial Intelligence in Urban Planning and Design* (eds As, I. et al.) Ch. 14 (Elsevier, 2022).
- Forrester, J. W. *Urban Dynamics* Vol. 11 (MIT Press, 1970).
- Lowry, I. S. A model of metropolis. Santa Monica, CA. RAND Corporation https://www.rand.org/pubs/research_memoranda/RM4035.html (1964).
- Wilson, A. G. Models in urban planning: a synoptic review of recent literature. *Urban Stud.* **5**, 249–276 (1968).
- Batty, M. Big data, smart cities and city planning. *Dialogues Hum. Geogr.* **3**, 274–279 (2013).
- Harris, B. & Batty, M. Locational models, geographic information and planning support systems. *J. Plan. Educ. Res.* **12**, 184–198 (1993).
- Wang, D., Fu, Y., Wang, P., Huang, B. & Lu, C.-T. Reimagining city configuration: automated urban planning via adversarial learning. In *Proc. 28th International Conference on Advances in Geographic Information Systems* 497–506 (Association for Computing Machinery, 2020).
- Dahal, K. R. & Chow, T. E. A GIS toolset for automated partitioning of urban lands. *Environ. Model. Softw.* **55**, 222–234 (2014).
- Mnih, V. et al. Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015).
- Agostinelli, F., McAleer, S., Shmakov, A. & Baldi, P. Solving the Rubik’s cube with deep reinforcement learning and search. *Nat. Mach. Intell.* **1**, 356–363 (2019).
- Fan, C., Zeng, L., Sun, Y. & Liu, Y.-Y. Finding key players in complex networks through deep reinforcement learning. *Nat. Mach. Intell.* **2**, 317–324 (2020).
- Novati, G., de Laroussilhe, H. L. & Koumoutsakos, P. Automating turbulence modelling by multi-agent reinforcement learning. *Nat. Mach. Intell.* **3**, 87–96 (2021).
- Wang, Z., Xu, Y., Wang, D., Yang, J. & Bao, Z. Hierarchical deep reinforcement learning reveals a modular mechanism of cell movement. *Nat. Mach. Intell.* **4**, 73–83 (2022).
- Silver, D. et al. Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017).
- Segler, M. H., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604–610 (2018).
- Fawzi, A. et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **610**, 47–53 (2022).
- Mirhoseini, A. et al. A graph placement methodology for fast chip design. *Nature* **594**, 207–212 (2021).
- Xue, J. et al. Quantifying the spatial homogeneity of urban road networks via graph neural networks. *Nat. Mach. Intell.* **4**, 246–257 (2022).
- Boeing, G. OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **65**, 126–139 (2017).
- Vich, G., Gómez-Varo, I. & Marquet, O. in *Resilient and Sustainable Cities* (eds Allam, Z. et al.) Ch. 3 (Elsevier, 2023).
- Kabisch, N., Strohbach, M., Haase, D. & Kronenberg, J. Urban green space availability in European cities. *Ecol. Indic.* **70**, 586–596 (2016).
- Cui, Y., Yu, Y., Cai, Z. & Wang, D. Optimizing road network density considering automobile traffic efficiency: theoretical approach. *J. Urban Plan. Dev.* **148**, 04021062 (2022).
- Liao, S.-M., Cheng, C.-H. & Chen, L.-S. The planning and construction of a large underpass crossing urban expressway in Shanghai: an exemplary solution to the traffic congestions at dead end roads. *Tunn. Undergr. Space Technol.* **81**, 367–381 (2018).
- Goto, A. & Nakamura, H. A study on appropriate road spacing for the functionally hierarchical network planning. *Transp. Res. Procedia* **25**, 3817–3825 (2017).
- Thornton, L. E. et al. Operationalising the 20-minute neighbourhood. *Int. J. Behav. Nutr. Phys. Act.* <https://doi.org/10.1186/s12966-021-01243-3> (2022).
- Lynch, K. & Rodwin, L. A theory of urban form. *J. Am. Inst. Plann.* **24**, 201–214 (1958).
- Campbell, S. Green cities, growing cities, just cities?: urban planning and the contradictions of sustainable development. *J. Am. Plann. Assoc.* **62**, 296–312 (1996).
- Burton, E. The compact city: just or just compact? A preliminary analysis. *Urban Stud* **37**, 1969–2006 (2000).
- Jabareen, Y. R. Sustainable urban forms: their typologies, models, and concepts. *J. Plan. Educ. Res.* **26**, 38–52 (2006).
- Oliveira, V. & Pinho, P. Evaluation in urban planning: advances and prospects. *J. Plan. Lit.* **24**, 343–361 (2010).
- Lichfield, N., Kettle, P. & Whitbread, M. *Evaluation in the Planning Process: The Urban and Regional Planning Series* Vol. 10 (Elsevier, 2016).
- Christiano, P. F. et al. Deep reinforcement learning from human preferences. *Adv. Neural Inf. Process. Syst.* **30** (2017).

44. Koster, R. et al. Human-centred mechanism design with democratic AI. *Nat. Hum. Behav.* **6**, 1398–1407 (2022).
45. Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C. & Socher, R. The AI economist: taxation policy design via two-level deep multiagent reinforcement learning. *Sci. Adv.* **8**, eabk2607 (2022).
46. Mohammadi, N. & Taylor, J. E. Thinking fast and slow in disaster decision-making with smart city digital twins. *Nat. Comput. Sci.* **1**, 771–773 (2021).
47. Schulman, J., Wolski, F., Dhariwal, P., Radford, A. & Klimov, O. Proximal policy optimization algorithms. Preprint at <https://doi.org/10.48550/arXiv.1707.06347> (2017).
48. Yao, J., Murray, A. T., Wang, J. & Zhang, X. Evaluation and development of sustainable urban land use plans through spatial optimization. *Trans. GIS* **23**, 705–725 (2019).
49. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations* (2017).
50. Zheng, Y. *tsinghua-fib-lab/drl-urban-planning*: Release v0.1 *Zenodo* <https://doi.org/10.5281/zenodo.8175420> (2023).

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under grant 2020AAA0106000 to Y.L. and the National Natural Science Foundation of China under U1936217, U20B2060 and 61971267 to Y.L. This work was supported in part by Tsinghua-Toyota Joint Research Institute Inter-disciplinary Program to L.Z. and Tsinghua-Cambridge Joint Research Initiative Fund to L.Z. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript. We are also grateful for the insightful discussion with Y. Yuan and S. Zheng at Tsinghua University.

Author contributions

Y.Z., L.Z. and Y. Li jointly launched this research. Y.Z., Y. Lin, L.Z. and Y. Li contributed ideas. Y.Z., Y. Lin and Y. Li designed the research methods and provided the research outline. Y.Z. developed the DRL framework. Y.Z. and Y. Lin performed the experiments. Y. Lin, L.Z., T.W.,

D.J. and Y. Li provided critical revisions. D.J. and Y. Li managed the project. All authors jointly analyzed the results and participated in the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43588-023-00503-5>.

Correspondence and requests for materials should be addressed to Yong Li.

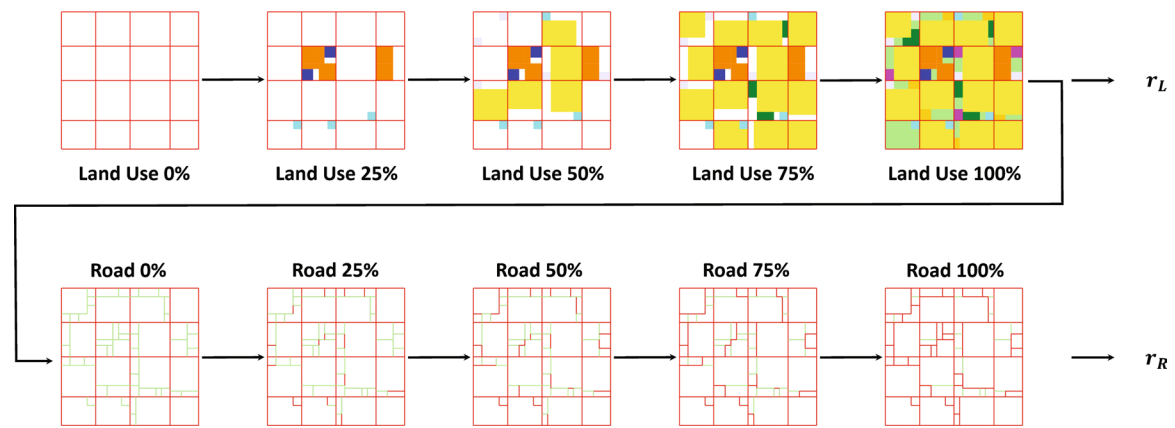
Peer review information *Nature Computational Science* thanks Paolo Santi, Weinan Zhang, Thomas W. Sanchez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Fernando Chirigati, in collaboration with the *Nature Computational Science* team. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

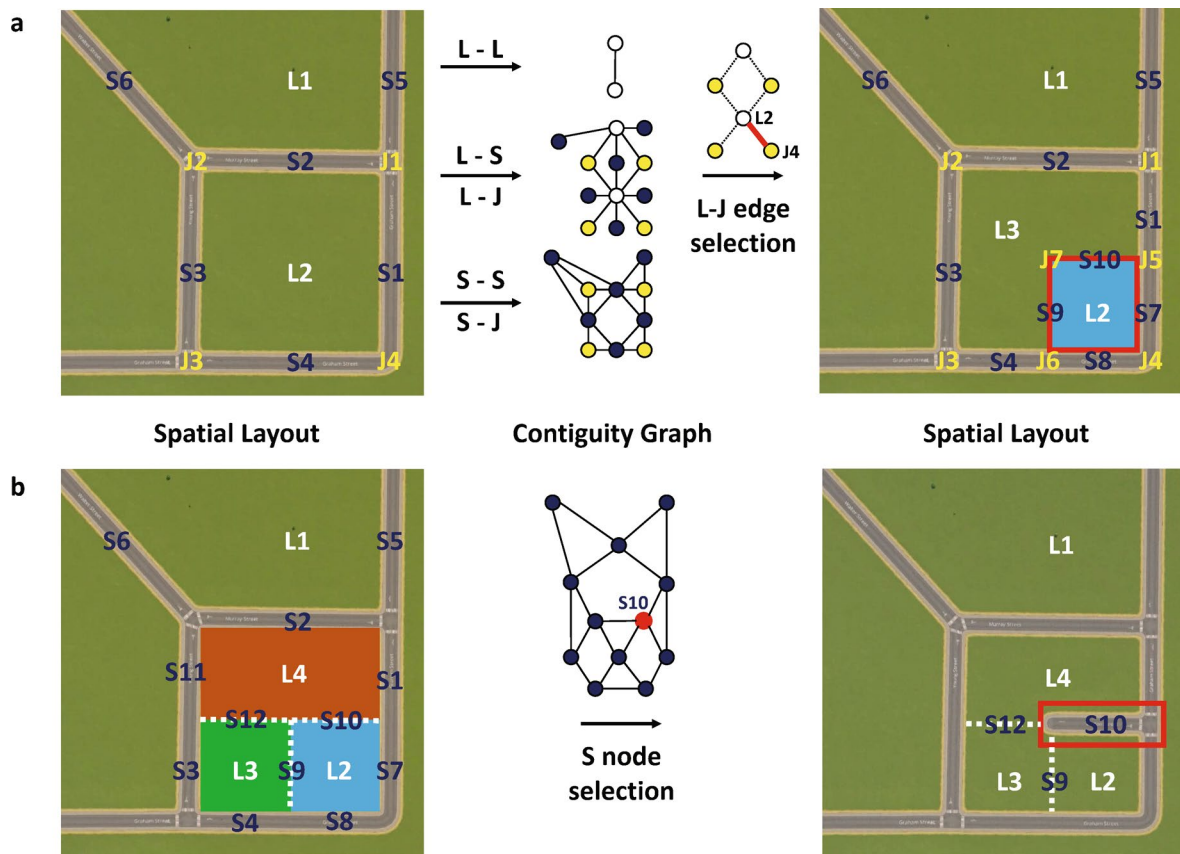
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023



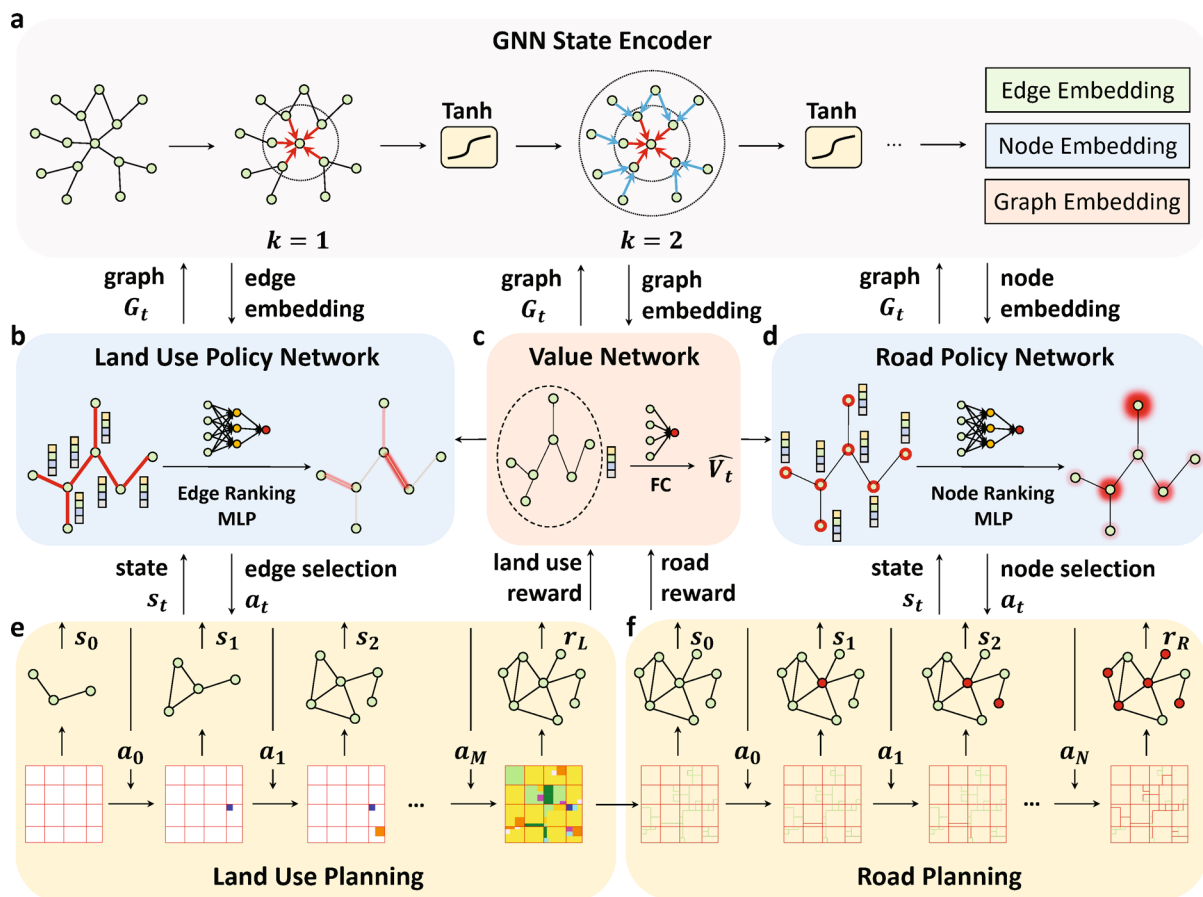
Extended Data Fig. 1 | Problem formulation of community-level spatial planning. The community-level spatial planning is formulated as a two-stage sequential decision-making problem. In the first stage of land use planning, the agent places different functionalities and open spaces until all the planning needs are satisfied. A reward r_L regarding the layout efficiency of all the land use as a whole is returned after the entire community is filled. The result of

land use planning serves as the initial conditions of the second stage of road planning, where one land use boundary is selected at each step and planned as a road segment. After a predefined termination step, a reward r_R considering the efficiency of road transportation is returned, and all intermediate steps have a reward of 0. Please refer to Supplementary Table 1 for the meaning of different colors.



Extended Data Fig. 2 | Reformulation of community spatial planning with a contiguity graph. a, Land use planning. A graph is constructed based on the contiguity relationship between urban geographical elements. Nodes consist of three different categories, including land (L), segment (S), and junction (J). Land can be vacant land to be planned or already planned land use. Segment can be a road or a land use boundary. Junction is the intersection between roads and land use boundaries. Two nodes are connected with an edge if the underlying geographical elements touch each other. There are five categories of edges in total, including L-L, L-S, L-J, S-S, and S-J. In the stage of land use planning, the agent selects one L-J edge that decides the location to place a given land use type. In the figure as an example, we select the edge between L2 and J4, thus a block in

L2 and near J4 is sliced and assigned as a newly planned land use type. After each step, new urban geographical elements are added and contiguity relationships change, thus a different graph with a distinct topology will be constructed at the next step. b, Road planning. Graph in the road planning task is defined in the same way as the land use planning task. At each step, the agent selects one S node that is currently a land use boundary, and plans a road at its location. In the figure, for example, we select S10, which is the boundary of L2 and L4 from three feasible candidates (S9, S10, and S12), and replace it with a road segment. In the next step, the node attribute of S10 changes, that is, its type changes from boundary to road, and two feasible candidates are left (S9 and S12).



Extended Data Fig. 3 | Overview of the proposed framework. Our framework consists of two separate policy networks (b and d) that take actions for the land use planning task (e) and road planning task (f) respectively, and share a GNN state encoder (a) with a value network (c) that estimates the effect of the current planning result. a, A GNN-based state encoder takes the contiguity graph states as input and learns representations of nodes, edges, and the whole graph. b, A land use policy network utilizes the obtained edge embeddings and scores each edge

with a MLP for edge selection. d, Similarly, a road policy network scores each node from the obtained node embeddings. c, A value network containing one fully-connected layer predicts the performance of the current spatial plan with the graph embedding. e, The spatial planning environment receives actions on the contiguity graph, places land use or road accordingly in actual geographical space, and transforms it into states in a graph form.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection This study used OpenStreetMap to collect geographical data of communities. Python packages are adopted to collect the original data. The used packages for data collection are osmnx>=1.1.2 and geopandas>=0.11.1. The Python version used in this work is 3.8.0

Data analysis This study used Python packages to conduct data analysis. Required packages are osmnx>=1.1.2, geopandas>=0.11.1, and PyGAD>=2.17.0. The codes developed in the study can be found at <https://github.com/tsinghua-fib-lab/DRL-urban-planning>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

This work uses open-sourced data from OpenStreetMap (<https://www.openstreetmap.org>). The data supporting the results of this study is available on Zenodo (<https://doi.org/10.5281/zenodo.8175420>) and GitHub (<https://github.com/tsinghua-fib-lab/DRL-urban-planning>).

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="No human participants were involved in this study."/>
Population characteristics	<input type="text" value="No human participants were involved in this study."/>
Recruitment	<input type="text" value="No human participants were involved in this study."/>
Ethics oversight	<input type="text" value="No human participants were involved in this study."/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<input type="text" value="This study investigated spatial planning of land use and roads for urban communities. A computational approach based on deep reinforcement learning and graph neural networks was proposed. Experiments on real-world urban communities verified the capability of deep reinforcement learning to design geographical layouts and optimize spatial efficiency. The study provides us with powerful tools to automate the process of urban planning, pushing the boundaries of what is possible in urban science."/>
Research sample	<input type="text" value="This study used geographical data of two real-world urban communities in Beijing, China. The data was downloaded from the public OpenStreetMap platform (https://www.openstreetmap.org/)."/>
Sampling strategy	<input type="text" value="No sampling calculation was performed."/>
Data collection	<input type="text" value="Publicly available data from the public and widely acknowledged OpenStreetMap platform (https://www.openstreetmap.org/)."/>
Timing and spatial scale	<input type="text" value="The data includes locations for existing roads and land use in two urban communities in Beijing, Huilongguan and Dahongmen."/>
Data exclusions	<input type="text" value="No data were excluded from the analyses."/>
Reproducibility	<input type="text" value="For each urban community, we repeated the experiments for at least 5 times with different initial values of neural networks, and the experimental results were consistent."/>
Randomization	<input type="text" value="No randomization was required for this study."/>
Blinding	<input type="text" value="Blinding was not relevant to this study."/>

Did the study involve field work? Yes No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Included in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Included in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |