Sarah Mokhtar, Renaud Danhaive, Caitlin Mueller

Neural Implicit Fields for Performance-Informed Geometries in Building Design

Abstract: A high-performing built environment is essential for a livable future in cities that will face increased pressures from densification and climate change. Designing such buildings requires not only access to complex simulation workflows, but also advanced means of geometry representation, so that concepts can be easily synthesized, developed, and improved in early-stage design. Fundamental to this challenge is the ability to generate and semantically manipulate structured representations of buildings in iterative evaluation workflows through real-time performance feedback. The introduction of continuous implicit fields has emerged as an alternative approach to explicit 3D representations that is topology-, complexity- and resolution-agnostic, capable of representing high-level details in a compact, efficient, and structured manner as well as promotes implicit-explicit and 3D-2D cross-modality without compromising differentiability that is essential for design interactivity and optimization. The opportunities that such representations present for design and simulation in building applications are presented through a novel framework in this paper. The focus is on evaluating their usability for downstream design applications. The findings of this study highlight the potentials for universal representations that are interchangeable across simulation disciplines, spanning over a high-fidelity latent space that enables geometric interpolation and optimization across its performance landscape.

1 Introduction

1.1 Motivation

1.1.1 Open challenge for high-performance design in architecture: representing the built environment

A high-performing built environment is essential for a livable future in cities that will face increased pressures from densification and climate change. Designing such environments requires not only access to complex simulation workflows to capture their perceptual and performative qualities, but also advanced means of geometry representation, so that concepts can be easily synthesized, developed, and improved in early-stage design. The built environment is inherently complex as it is composed of a wide range of morphologically distinct structures, dominated by buildings that vary substantially in scale, detail, and complexity. One key challenge lies in the ability of a computational representation to capture high-fidelity features in an efficient and

https://doi.org/10.1515/9783111162683-030

structured manner for buildings of arbitrary complexity, while concurrently balancing between the diversity potential and realism of the learned representation space. To support performance-driven design interaction and targeted optimization, the representation should additionally enable generation and meaningful manipulation of building geometries in iterative evaluation workflows that are coupled with real-time performance feedback.

1.1.2 Learning 3D shapes for design and simulation

The integration of deep learning models for simulation workflows has attracted substantial attention in design and engineering fields due to their ability to learn highly dimensional, non-linear, and complex phenomena. These workflows, however, require effective methods of capturing the complexities and irregularities of 3D shape data. Explicit representations have been widely used in 3D learning including voxels and point clouds, which are restricted to low resolutions and lack explicit topology; octrees, which lack differentiability; meshes, which in learning approaches rely on template deformations restricting topology variations; and low-order representations, such as depth-maps and hand-crafted feature descriptors, which cannot accurately capture the 3D structure (Ahmed et al. 2018). The introduction of continuous implicit fields has emerged as an alternative approach to 3D representation that is topology-, complexity- and resolution-agnostic, capable of representing high-level details in a compact and efficient manner and promotes implicit-explicit and 3D-2D cross-modality without compromising differentiability that is essential for design iteration and optimization (Remelli et al. 2020). It additionally provides the means to work with data in the wild, and through its continuous latent space enables the expansion of geometrical design explorations beyond hand parametrizations.

1.1.3 Neural implicit fields for building design

The capacity to develop high-fidelity compact representations for the built environment, enabled by recent advances in the field, motivates further explorations into the opportunities that such representations present for the design and simulation of building geometries. In contrast to other design and engineering applications that typically deal with objects of pre-defined scale, architectural applications necessitate an approach that can scale up to entire buildings. The dearth of openly accessible structured building geometry-to-performance datasets and the computational expense of creating synthetic ones additionally motivates the need for universal building representations that can interchangeably be used across simulation disciplines and interpolate geometrically across performance landscapes.

1.2 Related Work

1.2.1 Learning continuous shape parametrizations using neural implicit fields

Neural implicit fields have been developed to learn continuous shape parametrizations that map 3D coordinates to a signed distance field or occupancy function. Global shape-conditioned formulations (Park et al. 2019), in addition to providing high-quality single-shape reconstructions, enable latent-space interpolation within entire classes of shapes. This is achieved by conditioning an auto-decoder fully-connected neural network through latent code concatenation. The network learns to predict signed distance values for associated input coordinates, through encoding individual shapes in the latent variables and the entire class of shapes in the network weights. The training is achieved through backpropagation of a loss function accounting for the deviation between predicted and actual signed distance values, a sparsity loss for latent codes and often associated with weight regularization terms. The implicit representation can be discretized to explicit representations at arbitrary resolutions (Xie et al. 2022). The potentials of neural implicit fields have been highlighted with a focus on reconstruction of objects across scales, but their suitability for and integration in building design applications has not yet been explored.

1.2.2 Exploring and optimizing latent spaces

The continuous latent space embedded in the learning of neural implicit fields expands their applicability beyond just efficient and compact encodings. Learning latent space representations have been employed more generally to compress higher-dimensional spaces for semantic synthesis and organization through capturing key features and structural similarities (Grossmann et al. 2022) as well as for generative modeling through sampling a learned underlying probability distribution or performing latent vector arithmetic operations (Asperti and Tonelli, 2022). Exploring learned latent spaces more broadly and systematically has attracted less attention, but a few approaches have been explored including random or guided latent space walks (Li et al. 2021), interpolation and extrapolation (Park et al. 2019), as well as sampling dimensionality-reduced latent spaces using linear and non-linear embeddings (Jahan et al. 2021). Shape performance optimization operating on object-level learned implicit parametrizations has also been proposed to promote topological diversity while maintaining semantically correct generation of representations (Remelli et al. 2020). The wide array of latent space exploration approaches, particularly for neural implicit fields, offers a novel, yet largely unexplored, territory for smoother performance-informed design integration.

1.2.3 Design space exploration and latent space learning in architecture

Performance-driven design has been commonly approached in the architectural practice through coupling parametric modeling with performance simulation as a strategy to balance between quantitative and qualitative design goals. This process is associated with the challenge of generating a design space that is morphologically diverse and semantically meaningful. Fit-for-purpose handcrafted parametric models, however, are often not extendable, not scalable, biased, and constrained by the mode and resolution they were originally defined in.

Recent experimental work challenged some of these limitations through latent space learning including floor plan generation and semantic characterization (Chaillou 2021), performance-constrained generative modeling for structural design (Danhaive and Mueller 2021), structural morphology ideation using sketching (Ong et al. 2021), among others. While these have shown substantial promise, they are still constrained by their pre-defined dataset generation modes of production. Neural implicit fields learning expands such latent approaches by enabling text-image-3D cross-modality (Poole et al. 2022; Sanghi et al. 2022) through differentiable rendering and explicit-to-implicit pipelines (Remelli et al. 2020; Guillard et al. 2021), enabling the learning to act directly on expandable datasets of geometries rather than handcrafted parametrizations, while providing the flexibility to sample geometries at varying resolutions that match application-specific or simulation tool constraints.

1.3 Contribution

The opportunities that neural implicit field representations present for design and simulation in building applications are presented here through a novel design parametrization-to-exploration end-to-end framework. This is achieved through: (1) demonstrating the diversity and morphological qualities of a building latent space directly conditioned on geometry with no explicit parametrization, (2) developing a structured approach to latent space exploration for design that concurrently captures semantic characteristics of the design and performative landscape and provides control over shape generation, and finally (3) integrating performance-driven feedback through local gradient descent of differentiable analytical equations for building shape refinement and optimization.

2 Methodology

2.1 Design framework

The developed design parametrization-to-exploration framework consists of four main stages, shown in Fig. 1: (1) acquisition and sampling of a large and morphologically diverse dataset of building geometries, (2) training of continuous shape parametrizations, (3) definition of differentiable performance evaluation functions, and (4) latent space exploration. The framework provides the flexibility of continuously augmenting the trained neural implicit model with more building geometry data as it becomes available, and which can be sourced from datasets that are wild (e. g., from 3D scans or individually modeled building datasets), synthetic (e. g., sampled from hand-coded parametric models), or a combination of both. Defining differentiable objective functions for common building performance goals is challenging and, in many cases, not possible due to the black-box nature of most simulation tools. These can, however, be replaced with proxy models through direct abstractions or deep learning pipelines, which are becoming increasingly available.

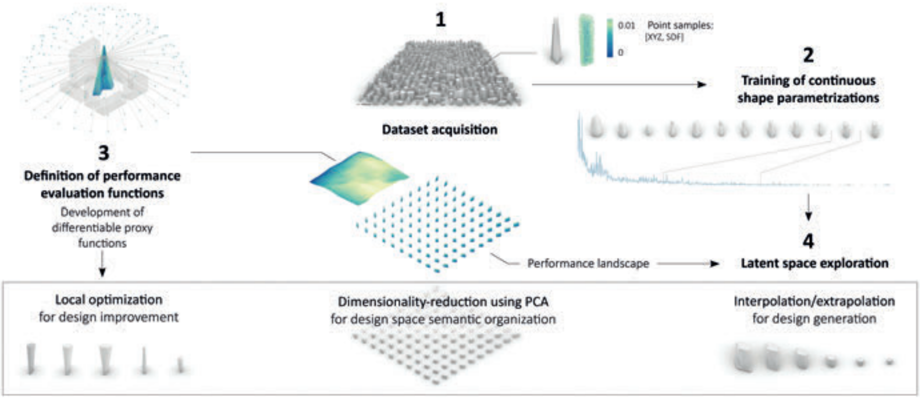

Fig. 1: Design framework for performance-driven design with learned implicit representations. A differentiable performance metric is evaluated over a human-legible latent space to support comparing, exploring, and generating highly diverse geometries for buildings massings without the need for hand-coded parametrizations.

2.2 Synthetic dataset generation

Large open-source morphologically diverse structured datasets of building geometries are scarce. While neural implicit fields learning does not rely on explicit parametrization, in the absence of suitable and accessible geometric datasets for training, a synthetic dataset was developed for this paper. The latter can interchangeably be replaced by and/or augmented with wild data. Thousands of individual building shapes were generated synthetically through sampling a parametric model with variables including size, height, shape, corner type, massing carving, lift-up conditions as well as vertical and horizontal variations. Figure 2 shows the 2,000 building geometries used for training and validation and highlights a sub-selection for detailed inspection of some of the generated forms. Considerations were made to ensure a balance between diversity and representation potentials of the dataset with respect to real urban contexts. Building geometries were centered in 400×400 m bounding boxes and the smallest detail across all ranged from 1 to 4 m. For each building mesh, point samples were selected based on a sub-set rejection sampling approach defined as a function of the distance d to the mesh (Davies et al. 2020): where the probabilistic acceptance criterion is defined as $e^{-\beta|d|}$, where β is 80. Starting with 100M Latin Hypercube samples with the bounding box, a smaller 1M points were re-sampled and augmented with 500k near-surface and 500k random additional samples for which corresponding signed distance values were computed.

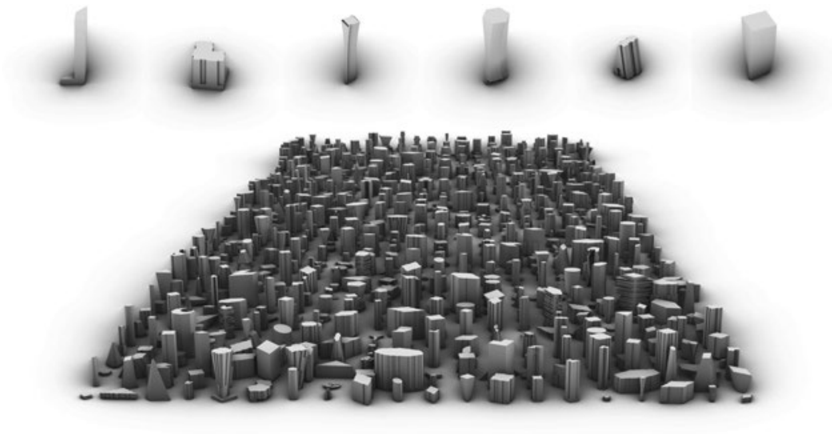

Fig. 2: Synthetic building dataset.

2.3 Training the continuous signed distance field

Learning shape-conditioned representations was achieved through training an auto-decoder network comprising of 5 fully-connected layers of 512 hidden neurons and ReLU activation functions. A latent code length of 256, initialized based on a normal distribution with standard deviation of 0.01, is used to condition the network. The loss was defined in two terms: a L1 loss accounting for the sum of absolute deviation between predicted and actual signed distance values, and a latent sparsity term defined as the mean-square of latent code values. The parameters selected for training were selected based on a trade-off between model performance as identified by informal experimentations and computational feasibility. To assess the learned latent space's ability to represent out-of-training samples, for each geometry in the validation set, the network weights are fixed, and their latent codes are optimized. Meshes are reconstructed using Marching Cubes (MC), and the reconstruction quality is primarily evaluated through visual inspection.

2.4 Differentiable performance evaluation – view factor

Physics-based simulations are essential to performance-informed building design, and while they often rely on tools and methods for which differentiability is not straightforward, the development of differentiable proxy models offer an alternative. A differentiable view factor approximation function was developed based on a naïve implementation of Moller Trumbore ray-triangle intersection algorithm coupled with differentiable relaxation of non-differentiable operations. This function is adapted to capture three distinct proxy performance evaluations: (1) sky view factor, a metric that describes the geometrical relationship between surfaces and the sky, and typically provides insights into exposure to daylight and urban microclimate performance, (2) cumulative incident solar radiation, and (3) access to views. The first calculates intersections between rays cast from 145 points sampled on the sky hemisphere based on the Tregenza subdivision scheme to a uniformly sampled offset of the building mesh surface. The second, additionally, associates the 145 points with weighing factors that correspond to cumulative radiation values extracted for the summer season in Boston using Radiance. The third replaces the sky points with points sampled uniformly on a horizontal surface representing a park view.

Figure 3 shows the performance distribution of all geometries in the dataset for the three evaluation metrics and identifies best, worst and average performing shapes for each. A faster and non-differentiable version of the performance objective function was used for evaluating building shapes outside of optimization loops for computational efficiency.

2.5 Latent space exploration

Latent space explorations were structured into three main strategies: interpolation for design generation, sampling dimensionality reduced embeddings using principal component analysis (PCA) for design space semantic organization and local optimization for design improvements. Pairwise interpolations between building shapes in the training sets are achieved through linear interpolations across latent vectors. To explore the latent space in its PCA dimensions, uniform samples are defined across

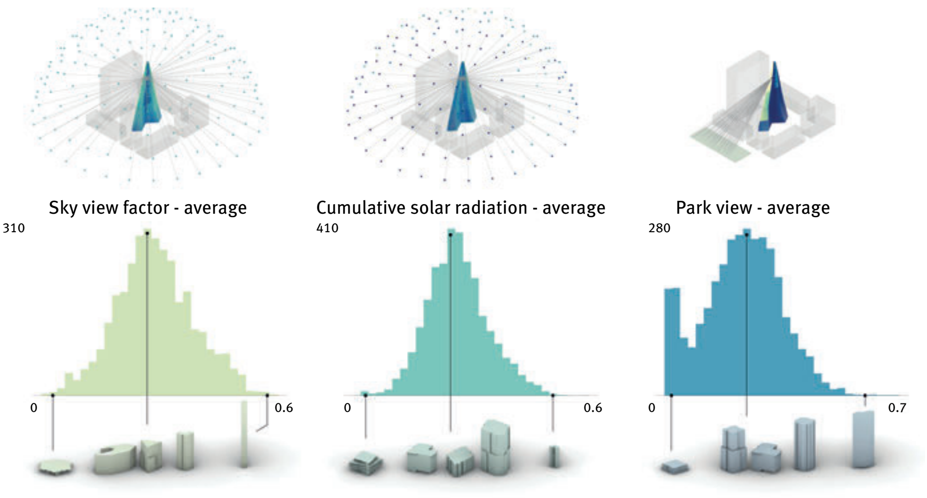

Fig. 3: Performance distribution of dataset across evaluation metrics.

two select principal components (PC) while keeping the mean value for others, followed by inverse mapping to recover their corresponding latent parameters. For local optimization, a starting latent code corresponding to a building of choice is initialized and for each gradient descent step: the mesh is reconstructed using MC through sampling the decoder, surface sample points are extracted and evaluated against the differentiable view factor evaluation, loss terms are computed, and the derivative of the explicit-implicit representation is derived based on the formulation by (Remelli et al. 2020) in terms of surface normals. Two soft constraints are integrated to promote feasible designs: a volume constraint term that accounts for the deviation between the volume of the original and optimized design, and a conservative regularization term that encourages designs that are close to the training dataset. The latter is defined as a l_2 norm of the difference between the latent vector and its closest k latent vectors from the training set. The combined loss term is defined by the following expression:

$$L_{\text{comb}} = \frac{\sum_{s=1}^{n} VF(s)}{n} + \alpha \cdot \sum_{z^i \in Z_k} \frac{\|z - z^i\|_2^2}{|Z_k|} + \beta \cdot |V_{\text{orig}} - V_{\text{cur}}|$$

It combines three terms: a performance term that captures the mean performance loss across n point samples, a deviation from k latent vectors z trained latent space term and a volume constraint term, for which their contribution to the total loss is tuned using α and β.

3 Results

3.1 Model training and reconstruction performance

The neural implicit field was trained on 1,565 geometries from the synthetic dataset based on a 90/10 split, and for each batch, 16,384 points are randomly selected out of the originally sampled 2M. The progression of learning is shown in Fig. 4, highlighting the decrease in loss as the model learns and the reconstruction quality improvement through a building sample. The ability of the network to reconstruct building forms from their associated latent codes is shown in Fig. 5 which shows pairs of building ground truth meshes from the training set and validation sets with their corresponding reconstructed shapes at MC resolution of 128.

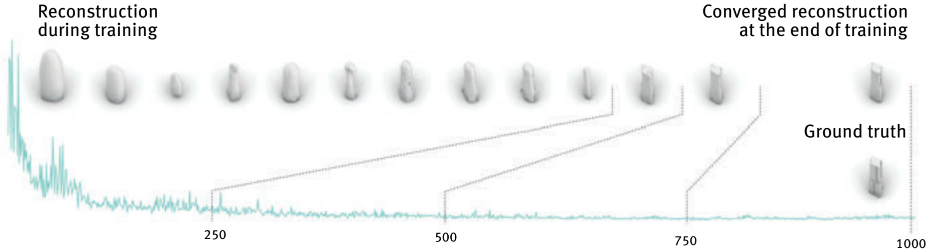

Fig. 4: Loss progression during training and corresponding reconstruction.

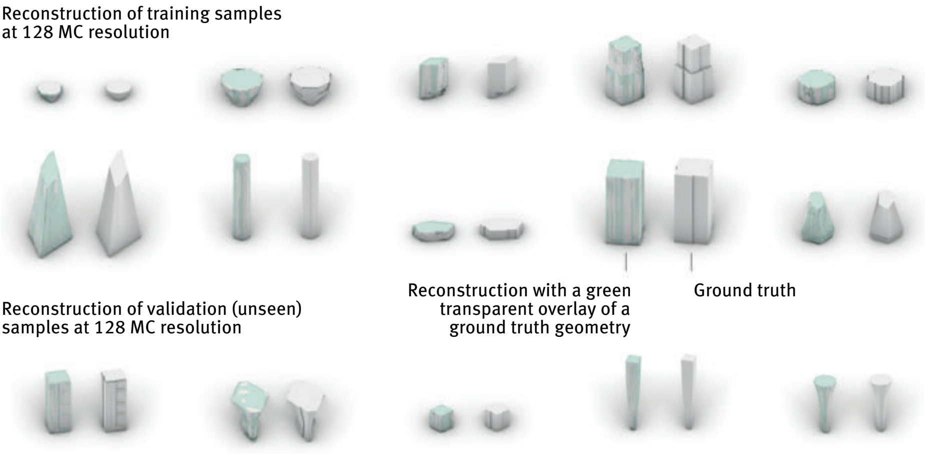

Fig. 5: Reconstruction quality of training and validation samples.

To highlight the deviation between pairs, a green transparent overlay of the ground truth geometry is added on top of reconstruction for clarity. By visual inspection of reconstructions, boundaries of building shapes are consistently identified, with erroneous artifacts associated with surface smoothness and smaller details, particularly boundary shape protrusion variations, that are challenging for the model to learn. These observations apply to both training and out-of-training samples, which is an indication that the learnt latent space is continuous and has a representation ability that extends beyond the data it was trained on. While reconstruction quality is an initial indicator of the model's learning success, for performance-informed design applications, the smoothness and representation of the learnt latent space is the primary goal.

3.2 Exploring and evaluating the latent space

3.2.1 Interpolation for design generation

Pairwise latent vector interpolation between diverse building shapes, shown in Fig. 6, was performed to inspect the representation fidelity and space continuity of the learned latent space. For each pair of buildings from the training set, shown at the two ends of A-R sub-figures, four intermediate buildings are generated from the model. Smooth interpolations are achieved across diverse building features that vary in scale, complexity, and detail. These include global features such as shape, vertical variation, but also smaller-scale and more challenging features such as the lift-up conditions in E, K and L as well as massing carvings in A and O. In addition to demonstrating the quality of the latent space, interpolation can be used as a design technique in itself; for example, a design team could input two candidate design geometries and use latent space interpolation to generate intermediate designs that capture features of both.

3.2.2 PCA for design space semantic organization

To capture the key semantic building features learned by the model, uniform sampling across 18 PC axes (dimensionally reduced from a latent space size of 256) was performed at a resolution of 10×10 and visualized in Fig. 7. Zoom-ins identifying the corner building shape conditions at the range extremities help identify the semantic meaning of each set of PCs.

It is unsurprising that the first set of components represent height, size, proportion, and rotation, which are key global building design features that vary considerably across the built environment. The shape and vertical variation are captured in subsequent components, and fine and more local details such as massing carvings, lift-up conditions and horizontal variations only identified in later components. The performance landscape can be captured by evaluating these samples with respect to the

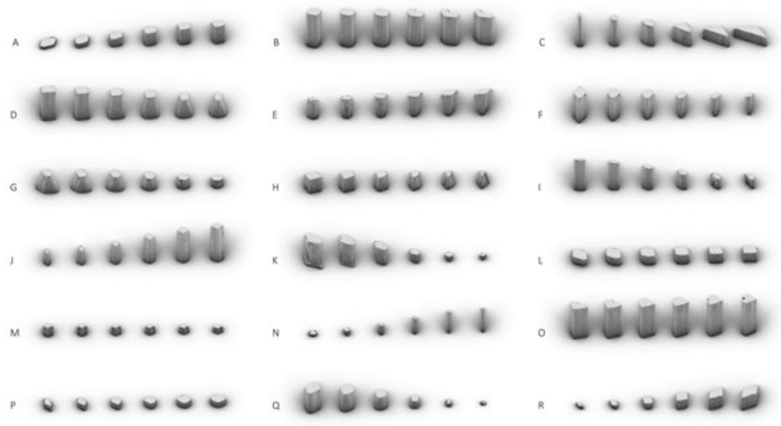

Fig. 6: Design generation using latent vector interpolations.

three metrics, and as shown in Fig. 8. It provides a smooth representation of performance variation across key semantic parameters of building design as identified by the trained model. This shows a compelling and direct application of neural implicit fields for performance-informed design, in which meaningful interpretation of design parameter-to-performance associations can help make informed decisions without any explicit design space geometrical parametrization.

3.2.3 Local optimization for design improvements

By integrating differentiability to the performance evaluation function, the continuous latent space and associated shape embeddings offer the ability to improve the shapes of buildings by operating on object-level low-dimensional and semantically meaningful implicit parametrizations. These learned parametrizations are essential for exploring large continuous shape possibility spaces to ensure feasible solutions to an optimization process coupled with task-specific volumetric and/or spatial constraints. Figure 9 shows the shape evolution of a select building at 50 iteration intervals during an optimization for two metrics and three soft constraints conditions. As the gradient flow is largely dependent on the performance evaluation of mesh reconstructions using MC at every iteration, any reconstruction failure due to the incomplete definition of surface boundary or volume collapse leads to interruptions in the optimization process. In the absence of volumetric constraints, while this leads to higher performing geometries at the onset, this shape collapse is observed early in the process. The two penalty terms: volume penalty and distance from training set were tested to overcome this. More consistent building shape propositions are observed, but with the now tougher challenge of concurrently improving performance values. The sensitivity of the shape

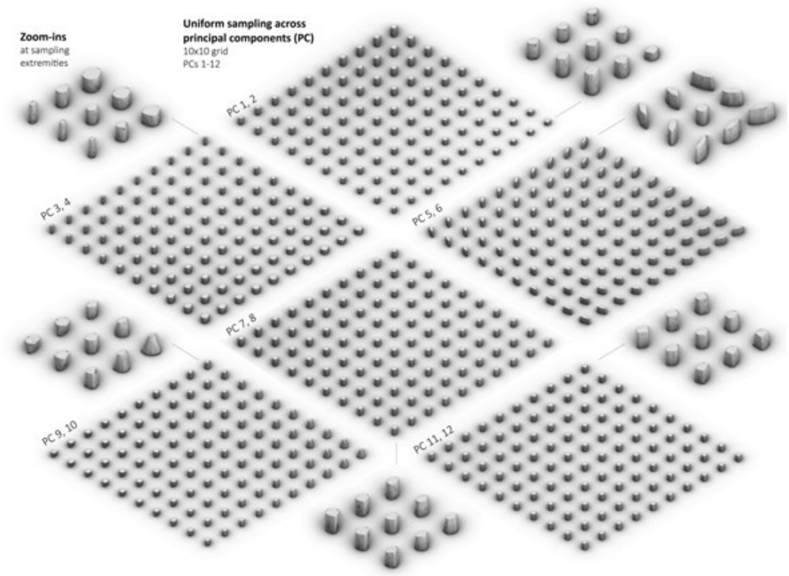

Fig. 7: Uniform sampling across PCA.

optimization to the performance evaluation function is shown in the formal variations observed for the park view and sky view factor objective functions. While the first promotes taller slender buildings oriented to maximize unobstructed park views, the second maximizes exposure to the sky by orienting the vertical slope of building façade up and rotating the shape to minimize surfaces obstructed by the high-rise building in the surrounding context.

4 Conclusion

Motivated by the representation challenges associated with performance integration in architectural design, this paper presented a novel design parametrization-to-exploration end-to-end framework using neural implicit fields. The findings indicate that morphologically diverse design spaces and semantic associations between parameters and performance metrics can be made without explicit parametrizations. Results show that smooth and meaningful latent design spaces are possible but that a balance between reconstruction quality, computational efficiency and latent space quality is challenging. For design and simulation applications, research into neural implicit fields should expand the focus from just reconstruction quality and scalability to latent space optimization and performance integration. This work expands the possibilities for performance-informed building design through: (1) offering a continuous design

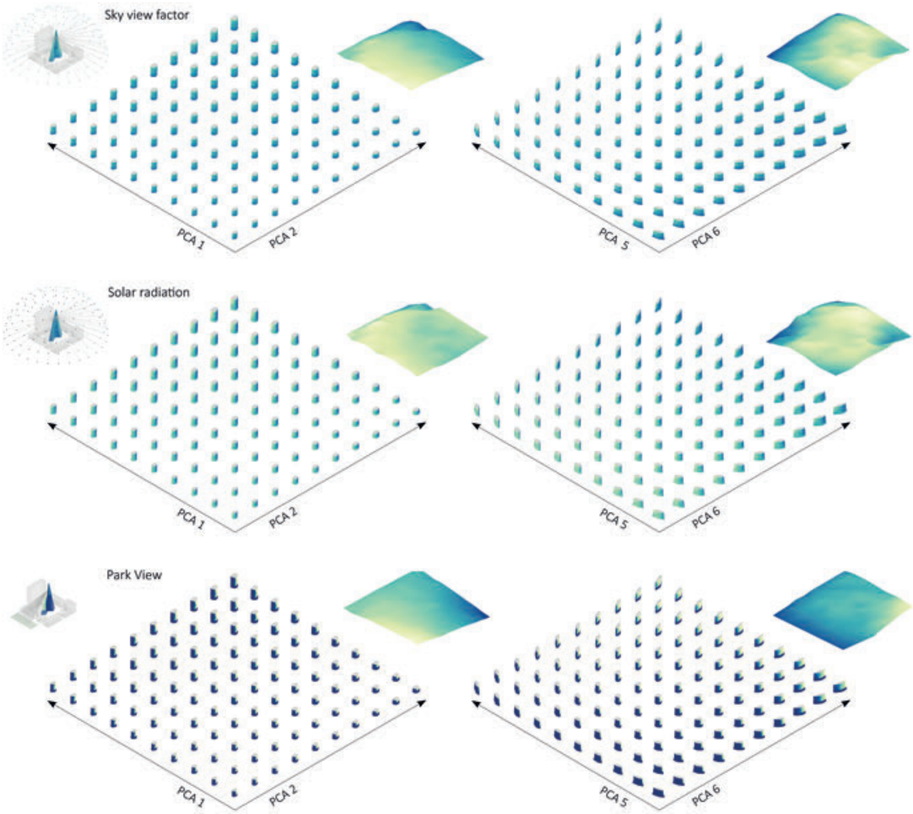

Fig. 8: Performance landscape across PCA 1, 2, 5 and 6 for three metrics.

space parametrization of building shapes without the need for handcrafting parameters making them expandable to wild and synthetic datasets alike, (2) expanding the integration of optimization workflows to complex geometries through coupling efficient fixed-length and semantically-meaningful representations with differentiable performance evaluation, and (3) enabling representation learning of buildings shapes across scales and resolutions, across performance assessment frameworks and potentially across modes (sketches, 3D geometries, etc.) through differentiable rendering. The development of universal topology-, complexity- and resolution-agnostic building representations that can interchangeably be used across creative design and simulation disciplines have the potential to expand design space explorations to the global performance landscape of design possibilities.

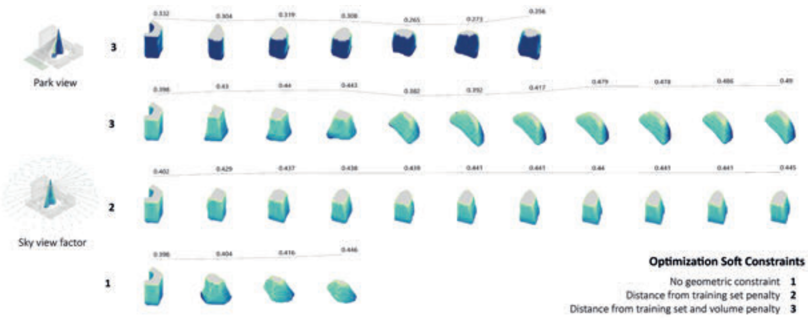

Fig. 9: Shape optimization progression. Reported values give the performance component of the loss terms.

References

Ahmed, E. et al. (2018). A survey on deep learning advances on different 3D data representations, https://arxiv.org/abs/1808.01462 [Preprint].

Asperti, A. and Tonelli, V. (2022). Comparing the latent space of generative models, *Neural Computing and Applications,* pp. 1–18.

Chaillou, S. (2021). AI and architecture: An experimental perspective, In *The Routledge Companion to Artificial Intelligence in Architecture,* pp. 420–41.

Danhaive, R. and Mueller, C. (2021). Design subspace learning: Structural design space exploration using performance-conditioned generative modeling, *Automation in Construction,* 127(103664).

Davies, T., Nowrouzezahrai, D. and Jacobson, A. (2020). On the effectiveness of weight-encoded neural implicit 3d shapes, https://arxiv.org/abs/2009.09808 [Preprint].

Grossmann, N., Gröller, E. and Waldner, M. (2022). Concept splatters: Exploration of latent spaces based on human interpretable concepts, *Computer & Graphics,* 105, p. 7384.

Guillard, B. et al. (2021). Sketch2Mesh: Reconstructing and Editing 3D Shapes from Sketches, In *Proceedings of the IEEE/CVF International Conference on Computer Vision,* pp. 13023–32.

Jahan, T., Guan, Y. and Van Kaick, O. (2021). Semantics-Guided Latent Space Exploration for Shape Generation, *Computer Graphics Forum,* 40(2), pp. 115–26.

Li, G. et al. (2021). Discovering density-preserving latent space walks in GANs for semantic image transformations, In *Proceedings of the 29th ACM International conference on multimedia,* pp. 1562–70.

Ong, B. W. X., Danhaive, R. and Mueller, C. (2021). Machine learning for human design: Sketch interface for structural morphology ideation using neural networks, In *Proceedings of the 7th International Conference on Spatial Structures,* pp. 1–12.

Park, J. J. et al. (2019). Deepsdf: Learning continuous signed distance functions for shape representation, In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* pp. 165–74.

Poole, B. et al. (2022). *Dreamfusion: Text-to-3d using 2d diffusion,* https://arxiv.org/abs/2209.14988

Remelli, E. et al. (2020). Meshsdf: Differentiable iso-surface extraction, *Advances in Neural Information Processing Systems,* 33, pp. 22468–78.

Sanghi, A. et al. (2022). Clip-forge: Towards zero-shot text-to-shape generation, In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* pp. 18603–13.

Xie, Y. et al. (2022). Neural fields in visual computing and beyond, *Computer Graphics Forum,* 41(2), pp. 641–76.