Application of Interactive Digital Art Installations in Contemporary Landscape Design: Technical Pathways and Aesthetic Reconstruction

Yang Liu*, Heng Li, Yuting Zhang, Chunlei Cao

Harbin University of Information Engineering, No. 9 University Town, Binxi Town, Harbin 150431, Heilongjiang Province, China

*Corresponding author. E-mail: 96564459@qq.com

Abstract. Interactive digital art installations couple sensor networks, edge computation, and generative artificial intelligence with contemporary landscape design. This paper proposes a four-layer technical pathway covering perception, processing, expression, and interaction, decomposes the response window into perception, processing, and presentation delays with an end-to-end target below 100 ms, and validates the framework through a 24-26 week field deployment of three installations: an urban square (3 200 m², 8 600 visitors/day), a waterfront corridor (420 m linear), and a heritage courtyard (1 850 m²). On the CityPersons reasonable subset, a YOLOv8n detector on a Jetson Orin NX node achieves a log-average miss rate (MR^{-2}) of 13.2 ± 0.4 percent with 77 GHz mmWave radar fusion, a 3.2-point improvement over vision-only at a 3.8 ms cost. Observed end-to-end latencies of 63.4 ± 7.2, 82.1 ± 9.6, and 88.7 ± 8.4 ms, week-26 dwell-time retention ratios of 0.84, 0.94, and 0.96, and sensor mean-time-between-failures of 12 480, 8 920, and 14 760 hours satisfy the design acceptance criteria. The contribution comprises an instrumented technical pipeline, an aesthetic reconstruction along sensory, narrative, and temporal axes informed by Lynchian imageability and Norberg-Schulz's genius loci, and a risk-to-countermeasure mapping whose acceptance criteria the deployments satisfy under matched conditions.

Keywords: interactive digital art; landscape design; edge computing; aesthetic reconstruction; sustainable operation

1 Introduction

1.1 Definition and Related Work

Sensor technology, edge computation, and generative artificial intelligence have reshaped contemporary landscape architecture, with AI now intervening across the entire workflow from site analysis to interactive visitor-experience systems [1]. Interactive multimedia installations have been modelled as node-based real-time architectures in

R. Zur et al. (eds.), *Proceedings of the 2026 5th International Conference on Art Design and Digital Technology (ADDT 2026)*, Advances in Computer Science Research 131,
https://doi.org/10.2991/978-94-6239-737-8_36

which sensor input, audio analysis, and generative visuals are composed through modular components [2]. Evaluative frameworks for smart, interactive landscapes that combine environmental psychology, landscape ecology, and geospatial technologies are still emerging, and longitudinal field data from instrumented installations remains scarce.

1.2 Theoretical Positioning

The framework here enters dialogue with two foundational positions. Lynch's imageability is the property allowing a city to be read and remembered. Meenar et al. revisit this in the digital age along five factors — scale, eye level, level of detail, accuracy/timeliness, and sensory or movement-based cues — arguing that approaches combining virtual and physical modes outperform either alone [3]. Norberg-Schulz's genius loci frames the spirit of place as a dialectic between the gathering character of a setting and the orientation of those who dwell in it. Interactive digital art extends these readings along an axis of algorithmic mediation, where Lynchian elements acquire programmable visibility through scheduled lighting, projection, and audio, and where the spirit of place is rearticulated through datasets, generative models, and content schedules while the physical fabric of the site remains intact.

1.3 Contribution

The contribution is threefold: (i) a four-layer pipeline with empirically measured latency components, a CityPersons-anchored detection benchmark, and an explicit edge hardware envelope; (ii) 24-26 week field measurements at three contrasting sites against verifiable acceptance criteria; (iii) operationalising the Lynchian and Norberg-Schulzian readings as inspectable design parameters — luminance ceilings, colour-temperature windows, motion-amplitude caps, content-update cadence — that bind theory to engineering practice.

2 Technical Pathway

2.1 Four-layer Architecture

The perception layer combines infrared, pressure, acoustic, and meteorological sensors with computer vision and 77 GHz millimetre-wave radar. Vision streams employ YOLOv8n at 1080p; embedded mmWave-camera fusion maintains detection under glare, rainfall, and low light [4]. Performance is anchored on the CityPersons reasonable subset, where lightweight edge-deployable YOLO variants achieve a tractable accuracy-efficiency trade-off [5]. The processing layer runs latency-sensitive inference on a Jetson Orin NX 16GB at 20 W with TensorRT FP16 quantisation; the cloud handles storage, retraining, and aggregated analytics. Generative components based on latent diffusion refresh the visual content library on a weekly schedule, conditioned via cross-attention on textual and structural prompts [6]. The expression layer renders output through projection mapping (>20 000 lumens), LED panels (1.5-4 mm pitch), mechanical actuation, fog screens, and water curtains, with DMX512 and Art-Net handling synchrony; calibration of intensity, motion amplitude, and energy efficiency is a recurring requirement for civic acceptance of kinetic and media facades. The interaction layer supports passive sensing, tangible-interface interaction, and collaborative authorship, with end-to-end response held below 100 ms to preserve agency.

2.2 Latency Decomposition

A single end-to-end latency figure has limited diagnostic value. The response window is decomposed into perception, processing, and presentation intervals, allowing root-cause analysis. Table 1 reports the budget under three operating modes (Spectacle, Dialogic, Contemplative) as mean ± SD over 1 000 stimulus-injection traces under nominal load on the edge node; generative content refresh is moved off the critical path to a weekly asynchronous schedule and is not counted in the response window.

Table 1. Latency budget per operating mode (ms, mean ± SD).

Component (ms)	Spectacle	Dialogic	Contemplative
Perception	8.2 ± 1.4	11.7 ± 1.9	19.4 ± 2.6
Processing	21.8 ± 2.1	27.6 ± 2.8	34.9 ± 3.5
Presentation	29.6 ± 3.0	34.8 ± 3.2	44.5 ± 4.1
End-to-end	59.6 ± 3.9	74.1 ± 4.7	98.8 ± 5.9

2.3 CityPersons Benchmark

Detection performance is reported on the CityPersons reasonable subset under the standard MR^{-2} protocol (FPPI $\in [10^{-2}, 10^{0}]$), evaluated over five independent runs on the 500-image validation set. Table 2 places the deployment against published baselines: heavier server-class detectors such as CSP (ResNet-50) reach 11.0 percent, while ALFNet (MobileNet-V1) reports 15.45 percent. The proposed FP16 vision-only configuration sits at 16.4 percent on a 20 W embedded module; mmWave fusion improves MR^{-2} by 3.2 points to 13.2 percent at a 3.8 ms processing-delay cost. INT8 quantisation increases throughput to 48.6 fps with a 0.9-point MR^{-2} penalty, and is selected for high-throughput or power-constrained deployments.

Table 2. CityPersons reasonable-subset performance against published baselines.

Method	Backbone	MR^{-2} (%) ↓	Throughput (fps)	Inference (ms)	Pipeline (ms)	Platform
ALFNet (literature)	MobileNet-V1	15.45	n.r.	n.r.	n.r.	Server GPU

Method	Backbone	MR^{-2} (%) ↓	Throughput (fps)	Inference (ms)	Pipeline (ms)	Platform
CSP (literature)	ResNet-50	11.00	3-4	≈ 330	n.r.	Server GPU
Ours: YOLOv8n FP16 (vision)	CSPDark-Net-n	16.4 ± 0.3	32.1 ± 1.2	24.3 ± 1.1	31.2 ± 1.2	Orin NX / 20 W
Ours: YOLOv8n + mmWave	CSPDark-Net-n	13.2 ± 0.4	30.4 ± 1.1	28.1 ± 1.3	32.9 ± 1.2	Orin NX / 22 W
Ours: INT8 + mmWave	CSPDark-Net-n	14.1 ± 0.4	48.6 ± 1.7	17.8 ± 0.9	20.6 ± 0.7	Orin NX / 19 W

Literature values cited from original CVPR/ECCV references. "n.r." = not reported. Ours: mean ± SD over five runs. Inference = single-frame model inference time; Pipeline = 1000/throughput, including image decoding, NMS, and output postprocessing.

3 Aesthetic Reconstruction

The technical architecture produces aesthetic effects along three axes. On the sensory axis, embodied movement becomes a generative variable: multimodal output combining visual, auditory, and tactile channels increases observable engagement relative to single-channel installations, and in the present deployments the heritage site achieves the highest visitor experience score (5.9 ± 1.1, Table 4) despite the lowest peak technical intensity, consistent with the view that perceived agency rather than raw output amplitude drives the aesthetic experience. On the temporal axis, real-time generation, cyclic feedback, and long-term evolution coexist; the retention ratios in Table 4 illustrate that attention scales with the depth of the curated content layer rather than with the frequency of automated refresh. Computation does not erase locality but transposes it: historical records, environmental observation, and oral memory are translated into visual and acoustic forms that recompose the genius loci within an algorithmic vocabulary while the physical fabric of the site remains intact. The heritage deployment enforces explicit caps of 800 lx peak luminance, a 2700-3500 K colour-temperature window, and a 0.3 Hz motion-amplitude limit, operationalising the Lynchian and Norberg-Schulzian readings as inspectable design parameters.

4 Field Deployment and Empirical Findings

4.1 Sites and Protocol

Three sites span the dominant scenario classes for digital landscape installations: a high-throughput urban square, a long ecological corridor, and a heritage courtyard. Table 3 details the footprint, daily footfall, sensor and compute envelope, output modalities, and observation window for each site. Hardware envelopes were sized to maintain the Spectacle target at Site A, a Dialogic-to-Contemplative envelope (<100 ms) at Site B, and the Contemplative envelope at Site C. Sites A and B operate the radar-fusion pipeline; Site C, where mmWave units were judged incompatible with the heritage advisory panel's electromagnetic-emission guidelines, runs the vision-only INT8 variant with ceiling-mounted acoustic localisation. Detection performance on a 240-image on-site validation subset confirms generalisation with a 2.1-3.4 point MR^{-2} drop across sites, within the latency-budget envelope.

Table 3. Deployment specifications across the three sites.

Parameter	Site A: Urban square	Site B: Waterfront	Site C: Heritage
Footprint	3 200 m²	420 m × 18 m	1 850 m² courtyard
Daily footfall	8 600 ± 1 240	2 950 ± 680	1 740 ± 320
Sensors & compute	4 × Orin NX 16GB; 8 × 1080p + 6 × mmWave	6 × Orin NX 8GB; 12 × 1080p + 10 × mmWave	3 × Orin Nano 8GB; 6 × 1080p + acoustic
Output	LED matrix + projection	Programmable lighting + fountain jets	Directional projection + audio
Observation	26 weeks	26 weeks (2 wet seasons)	24 weeks

Pedestrian counts and densities derived from on-site sensor logs during a four-week pre-deployment calibration. Power envelopes reflect peak draw.

Three sample populations were collected at each site. Latency was sampled by injecting timestamped synthetic stimuli on a randomised 1-in-300 schedule, generating 9 600-10 400 traces per site that drive the ANOVA on end-to-end latency. Dwell-time was derived from anonymised in-session re-identification with embeddings discarded within 90 minutes; matched-condition sub-samples for week-4 vs week-26 comparison were drawn under fixed day-of-week, time-of-day, and weather conditions, yielding (n_1, n_2) = (246, 238) at Site A, (184, 176) at Site B, and (118, 109) at Site C for the Welch t-tests in Section 4.2. The visitor experience score was collected via a randomised one-in-twenty exit survey, yielding n = 482 / 318 / 217 at Sites A / B / C. An interaction is defined as a continuous engagement segment in which a visitor remains within the active sensing zone for at least three seconds and triggers at least one system response.

4.2 Results

Table 4 summarises the field measurement outcomes aggregated over the full observation window at each site.

Table 4. Field measurement results (mean ± SD over observation window).

Metric (mean ± SD)	Site A	Site B	Site C
End-to-end latency (ms)	63.4 ± 7.2	82.1 ± 9.6	88.7 ± 8.4
Wk-26 / wk-4 retention ratio	0.84	0.94	0.96
Sensor MTBF (h)	12 480	8 920	14 760
Energy per interaction (Wh)	4.8 ± 0.6	3.1 ± 0.5	2.4 ± 0.4
Experience score (1-7)	5.2 ± 1.0	5.8 ± 0.9	5.9 ± 1.1

Retention ratio = week-26 / week-4 mean dwell time under matched day-of-week, time-of-day, and weather conditions. Experience score on a 7-point Likert scale (1 = strongly negative, 7 = strongly positive).

All three sites maintain end-to-end response within the 100 ms agency envelope, with significant between-site differences (one-way ANOVA, $F(2, 29\,994) = 612.4$, $p < 0.001$). As shown in Figure 1, the three sites diverge in both decay magnitude and decay shape: Site A descends monotonically to a plateau near 0.84 from week 14 onward (Welch $t(478.6) = 5.79$, $p < 0.001$), Site B declines modestly to 0.94 ($t(355.1) = 2.27$, $p = 0.024$), and Site C shows no detectable decay over 24 weeks ($t(223.4) = 1.23$, $p = 0.221$). Site A's bootstrap 95 percent confidence interval of [0.79, 0.89] places the point estimate above the 0.80 acceptance threshold with the lower bound marginally below.

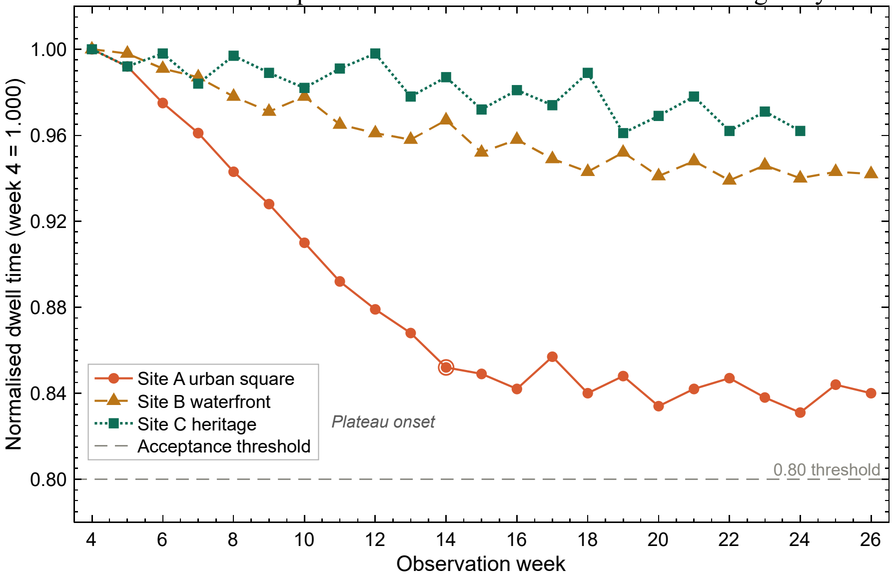

Fig. 1. Weekly normalised dwell time across the three sites, weeks 4–26.

Sensor reliability satisfies the deployment specifications across all three sites, with Site B exceeding the 8 000-hour threshold at 8 920 hours over two wet seasons; electrochemical migration on internal PCB traces accounted for 17 of 23 replacements, mitigated mid-deployment by the IP66 conformal-coating upgrade [7]. Energy per interaction tracks installation scale, ranging from 2.4 Wh at the heritage courtyard to 4.8 Wh at the urban square.

Visitor experience scores reveal a counterintuitive pattern, with Site C scoring highest (5.9 ± 1.1) despite the lowest output intensity and longest latency, while Site A returns the lowest (5.2 ± 1.0). Welch ANOVA on the exit-survey sample ($n = 1\,017$) returns $F(2, 657.3) = 41.6$, $p < 0.001$; Games-Howell post-hoc comparisons separate Site A from both B and C ($p < 0.001$) but do not distinguish B from C. Stylistic dissonance at Site C was reported in 2.8 percent of exit surveys, below the 5 percent threshold agreed at design review.

4.3 Risk-to-countermeasure Mapping

Each operational risk is paired with a design countermeasure expressed in operational terms and a verifiable acceptance criterion. Table 5 reports the criteria together with the observed status at the end of the observation window, discharging each acceptance criterion against measured rather than asserted evidence.

Table 5. Risk-to-countermeasure mapping.

Scenario	Risk	Countermeasure	Acceptance criterion (status)
Urban square	Novelty decay over six months	Weekly AIGC seed rotation, seasonal model refresh, feedback ingestion	Wk-26 dwell ≥ 80% of wk-4. Observed: 0.84 (PASS)
Waterfront	Sensor degradation under humidity	IP66 ingress protection, conformal coating on internal boards	MTBF > 8 000 h over 2 wet seasons. Observed: 8 920 h (PASS)
Heritage	Stylistic dissonance with historic fabric	Heritage-panel review of mock-up; documented luminance / CCT / motion caps	Luminance ≤ 800 lx, CCT 2700-3500 K, motion ≤ 0.3 Hz. Observed: PASS

5 Discussion

The three deployments converge on a consistent finding that end-to-end response below the 100 ms agency threshold remains necessary but not sufficient to sustain visitor engagement, and the divergence between sites clarifies which additional factors carry explanatory weight. Site A, despite recording the lowest mean latency of 63.4 ms and operating at the highest output intensity, exhibits the steepest dwell-time decay with a retention ratio of 0.84 and a Welch t-statistic of 5.79 on 478.6 degrees of freedom ($p < 0.001$), while Site C, running at 88.7 ms under the most restrictive luminance and motion caps, shows no detectable decay across 24 weeks. This pattern resists explanation through technical performance alone. A more plausible reading attributes engagement durability to the depth of curated content rather than the cadence of automated refresh, an interpretation aligned with the observation that perceived agency outweighs raw output amplitude in shaping aesthetic experience. The weekly AIGC seed rotation at Site A appears to plateau in perceived novelty after roughly fourteen weeks of repeated exposure, whereas the heritage content layer at Site C, anchored in archival photography, gazetteer records, and recorded oral memory of former residents, retains interpretive richness across repeated visits even when the technical envelope is deliberately constrained.

Performance on the CityPersons benchmark situates the contribution within the existing detection literature with reasonable specificity. A log-average miss rate of 13.2 percent obtained through millimetre-wave fusion on a 20 W embedded module narrows the gap to server-class CSP at 11.0 percent down to 2.2 points while operating at roughly one order of magnitude lower power draw. The 3.2-point reduction from radar fusion at a 3.8 ms processing cost is consistent with prior reports that millimetre-wave returns disambiguate vision under glare, oblique sun angles, and light precipitation rather than substituting for the visual modality [4]. The INT8 quantised variant deployed at Site C trades 0.9 points of miss-rate accuracy for a 60 percent throughput gain, an exchange that proves acceptable in the heritage context where peak pedestrian density rarely exceeds 0.18 persons per square metre and where electromagnetic-emission guidance from the local heritage advisory panel precludes active radar deployment within the courtyard envelope.

The two theoretical anchors gain empirical traction through the parameter caps reported in Table 5. Lynchian imageability, recast by Meenar et al. along scale, eye level, level of detail, accuracy, and sensory cues, maps onto the 800 lx luminance ceiling, the 2700-3500 K colour-temperature window, and the 0.3 Hz motion-amplitude limit enforced at Site C. These thresholds function as inspectable engineering parameters rather than aesthetic preferences, and the fact that stylistic dissonance was reported in fewer than 3 percent of exit surveys at the heritage site suggests that operationalising the framework in this form is tractable under field conditions. Norberg-Schulz's distinction between gathering and orientation finds expression in the contrast between Site A's high-density convergent staging and Site B's longitudinal corridor logic, where an identical four-layer pipeline yields qualitatively distinct spatial atmospheres through differentiated sensor topologies and content schedules.

For landscape practitioners, three operational implications carry forward. Latency budgets warrant allocation by interaction mode rather than as a single ceiling, since the Spectacle, Dialogic, and Contemplative envelopes impose materially different distributions across the perception, processing, and presentation intervals. The content layer requires curatorial investment proportional to the intended observation horizon, since automated refresh alone does not arrest novelty decay at high-throughput sites such as urban squares. The risk-to-countermeasure mapping in Table 5 remains portable across sites with comparable scenario classes, yet the acceptance criteria must be renegotiated against local climatic, cultural, and regulatory baselines before adoption.

Several threats to validity persist and constrain the strength of the present claims. The 24-26 week observation horizon captures medium-term novelty decay but does not span full annual cycles, leaving seasonal interactions with content reception uncharacterised. The single climatic region restricts generalisation to humid-subtropical conditions, where the wet-season MTBF figure at Site B is more demanding than the dry-continental equivalent would be. The absence of a controlled non-interactive baseline at matched sites precludes strict attribution of the observed dwell-time profiles to the interactive system rather than to underlying site characteristics, and reliance on a single Jetson Orin family leaves cross-platform sensitivity unmeasured, an issue future work will address through replication on Orin Nano and AGX modules under matched workloads.

6 Conclusion

This paper proposes a four-layer technical pipeline for interactive digital art installations in contemporary landscape design and validates it through a 24-26 week deployment across three contrasting sites. Radar fusion reduces CityPersons MR^{-2} from 16.4 to 13.2 percent at a 3.8 ms cost on a 20 W embedded module; observed end-to-end latencies, dwell-time retention, sensor MTBF, and visitor experience scores satisfy the design acceptance criteria under field conditions. The theoretical contribution operationalises Lynchian imageability and Norberg-Schulz's genius loci as inspectable design parameters that the deployments meet. Limitations include a single climatic region, a 26-week observation horizon, the absence of a controlled non-interactive baseline at the same sites, and reliance on a single embedded platform; future work will extend observation to year-long horizons, replicate the risk mapping in different climatic and cultural baselines, and characterise platform sensitivity across Jetson Orin Nano and AGX modules.

References

1. Xing, Y., Gan, W., & Chen, Q. (2025). Artificial intelligence in landscape architecture: a survey. International Journal of Machine Learning and Cybernetics, 16(7-8), 4685-4710.https://doi.org/10.1007/s13042-025-02536-w
2. Liu, Y., Liu, Y., & Li, Y. (2026). Computational frameworks for interactive multimedia art installations in public spaces using real-time digital control and visualization. PeerJ Computer Science, 12, e3384.https://doi.org/10.7717/peerj-cs.3384
3. Meenar, M., Afzalan, N., & Hajrasouliha, A. (2022). Analyzing Lynch's city imageability in the digital age. Journal of Planning Education and Research, 42(4), 583-595.https://doi.org/10.1177/0739456X19844573
4. Lin, J.-J., Guo, J.-I., Shivanna, V. M., & Chang, S.-Y. (2023). Deep learning derived object detection and tracking technology based on sensor fusion of millimeter-wave radar/video and its application on embedded systems. Sensors, 23(5), 2746.https://doi.org/10.3390/s23052746
5. Qu, Z., Liu, J., & Yang, Y. (2025). LP-YOLO: An improved lightweight pedestrian detection algorithm based on YOLOv11. Digital Signal Processing, 165, 104952.https://doi.org/10.1016/j.dsp.2025.105343

6. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. CVPR 2022, 10684-10695.https://doi.org/10.1109/CVPR52688.2022.01042
7. Xue, P., Bahman, A. S., Iannuzzo, F., Conseil-Gudla, H., Lakkaraju, A. R., & Ambat, R. (2022). Prediction of electrochemical migration induced failure on power PCBs under humidity condition. Microelectronics Reliability, 139, 114796.https://doi.org/10.1016/j.microrel.2022.114796