Received: 20 April 2026; **Revised:** 19 August 2026; **Accepted:** 05 September 2026; **Published:** 08 September 2026
Archs Sci. (2026) Volume 76(2), Pages 7-13, Paper ID 2026202.

Technical Reappraisal of an Archived FCN–CycleGAN Workflow for Landscape-Plan Segmentation and Rendering

Yue Han
School of Art and Design, Fuzhou University of International Studies and Trade, Fuzhou 350202, Fujian Province, China
Corresponding author: Yue Han (e-mail: hanyue510242932@163.com).

Abstract This technical reappraisal reconstructs a two-stage deep-learning workflow for semantic analysis and visual rendering of landscape architectural plans. The intended pipeline uses a fully convolutional network (FCN) to assign labelled pixels to 11 landscape categories and then applies a cycle-consistent generative adversarial network (CycleGAN) to semantic-map patches. The surviving inventory records 328 plan images and 57,271 connected annotated elements. An unaudited internal-validation summary records pixel accuracy above 92% and mean Intersection over Union of 0.87; these figures are historical descriptors rather than independently verifiable performance results because the source-level partitions, per-class scores, repeated runs, code, and external test set are unavailable. Displayed CycleGAN checkpoints from Epochs 50–300 show greater texture density but persistent structural distortions. The surviving examples use prepared semantic masks, and no retained output verifies the complete FCN-prediction-to-CycleGAN pathway. The reappraisal therefore documents the intended integration, corrects an algebraically redundant mask-summation formulation, identifies a clean-mask-to-predicted-mask domain shift, and delineates the evidence required for reproducible evaluation. The record supports feasibility of the stages as an archived demonstration, but not reproducibility, end-to-end validation, or superiority over alternative methods.

Index Terms landscape architecture, semantic segmentation, fully convolutional network, CycleGAN, semantic image synthesis, technical reappraisal

I. Introduction

Landscape architectural plans combine vegetation, water, circulation, buildings, and site furniture in graphically complex representations. Automating their interpretation may reduce repetitive drafting work and support rapid comparison of preliminary visual treatments. Earlier landscape research has examined deep-learning-based plan recognition and rendering [1], the applicability of GAN-generated masterplans and professional acceptance [2], style transfer between traditional-garden and contemporary visual elements [3], data-driven landscape visual-quality assessment [4], and CycleGAN-based site identification and rendering [5]. A broader review identifies layout generation, image post-processing, simulation, and evaluation as major machine-learning applications in landscape architecture [6]. More recent landscape-visualization research has begun to employ latent diffusion models and domain-specific semantic controls [7], which places the FCN–CycleGAN pipeline studied here within an earlier but still instructive generation paradigm.

GANs learn a data distribution through competition between a generator and a discriminator [8]. For dense prediction, the FCN formulation converts classification networks into pixel-wise predictors [9]; encoder–decoder alternatives include U-Net [10], SegNet [11], and DeepLabv3+ [12]. In image-to-image translation, pix2pix learns a conditional mapping from paired examples [13], whereas CycleGAN learns bidirectional mappings from unpaired domains by adding cycle consistency [14]. Contrastive unpaired translation provides a later alternative that seeks stronger patch-level content correspondence [15]. Methods designed specifically for semantic-layout-to-image synthesis include spatially-adaptive normalization (SPADE) [16], semantic region-adaptive normalization (SEAN) [17], and adversarial semantic synthesis (OASIS) [18]. These methods are more direct comparators when a categorical map is the generator input.

The present article is a technical reappraisal rather than a report of a newly executed experiment. It reconstructs an archived application workflow comprising: (1) FCN-based segmentation into 11 semantic classes; (2) unpaired translation from semantic-map patches to target-style plan patches; and (3) overlap-weighted reconstruction of a full-resolution rendering. Its contributions are a consolidated description of the surviving dataset taxonomy, an internally consistent specification of the intended pipeline, correction of the mask-reassembly formulation, and an evidence-bounded analysis of the archived results and their failure modes.

The study addresses three questions. First, what information about segmentation performance survives in the internal-validation record? Second, what visual changes are observable across the available CycleGAN checkpoints? Third, which

methodological and graphical conditions limit interpretation? Because the record contains no external test collection, complete source-level partition record, baseline comparison, ablation experiment, or controlled expert study, no new estimate of predictive performance or generalization is made.

II. Related Work

A. Machine Learning in Landscape Architecture

Zhou and Liu [1] reported a closely related deep-learning workflow for recognizing and rendering landscape plans. Zhou and Xiang [2] subsequently evaluated landscape masterplan generation with image-analysis measures and user surveys, demonstrating the distinction between visual plausibility and professional acceptance. Huang et al. [3] studied style transfer as a means of integrating traditional-garden and modern design elements, while Fan et al. [4] addressed landscape visual-quality assessment through model fine-tuning. Huang [5] also described CycleGAN-enhanced site identification and rendering. Collectively, these studies establish plan recognition, rendering, style transfer, and visual-quality assessment as related but distinct tasks.

Table 1 compares the evidentiary scope of the archived workflow with closely related landscape-oriented research. The reappraisal focuses on what can be established from surviving records rather than treating architectural novelty or comparative superiority as an evaluable outcome.

B. Semantic Segmentation of Plan Graphics

FCN-8s fuses coarse semantic features with shallower spatial features through skip connections [9]. U-Net uses a symmetric encoder–decoder with feature concatenation [10]; SegNet reuses pooling indices during decoding [11]; and DeepLabv3+ combines atrous multi-scale context with a boundary-refining decoder [12]. These architectures are relevant baselines because landscape-plan objects vary greatly in size and contain thin paths, small facilities, text, transparent vegetation symbols, and designer-specific textures. Class imbalance further complicates learning. Reweighting, resampling, or focal-style objectives [19] can reduce domination by frequent or easy pixels, but the retained FCN run used ordinary multiclass cross-entropy and contains no class-balancing ablation.

C. Unpaired Translation, Semantic Synthesis, and Evaluation

Pix2pix requires aligned input–target pairs and uses a conditional GAN with a PatchGAN discriminator [13]. CycleGAN instead uses adversarial and cycle-consistency objectives to learn from unpaired domains [14]. Contrastive unpaired translation [15] provides an important comparison because patchwise feature correspondence may better preserve local content. SPADE [16] modulates generator activations with the semantic layout, SEAN [17] adds region-adaptive style control, and OASIS [18] redesigns adversarial supervision around semantic classes. These approaches are particularly relevant because the archived generator receives a semantic map rather than an unconstrained natural image. For landscape plans, preservation of circulation geometry, boundaries, and semantic identities is at least as important as texture realism.

No single image metric establishes design quality. Structural Similarity (SSIM) compares luminance, contrast, and structure [20]; LPIPS compares deep perceptual features [21]; Fréchet Inception Distance (FID) compares distributions in a learned feature space [22]; and Kernel Inception Distance (KID) offers a kernel-based alternative with an unbiased estimator [23]. These measures have different assumptions and should be accompanied by class-preservation checks and blinded professional assessment. The present archive contains only visual comparisons and grayscale histograms, so these stronger evaluations are specified as necessary future work rather than reported as completed analyses.

III. Methodology

A. Archived Material, Annotation, and Evidentiary Scope

The surviving inventory describes 328 landscape-plan images representing urban squares, municipal parks, residential courtyards, and waterfront recreational spaces. Images were normalized to 2000×2000 pixels. According to the archived documentation, three professional landscape architects produced spatially aligned PNG masks by vectorizing landscape elements in GIS software and rasterizing the layers. The masks encode 11 labelled categories: four softscape classes and seven hardscape classes. Table 2 reports the recorded RGB codes, connected-component counts, and category-level image counts. The components sum to 57,271.

The surviving files do not establish a complete chain of custody for the plans, annotations, checkpoints, and generated outputs, or identify which materials were created by the present author and which were obtained from an earlier project. This reappraisal therefore makes no claim of dataset ownership, and public release of the underlying material remains conditional on documentation of its origin, authorization, and permitted use.

The component distribution is strongly imbalanced. Grass and shrub layers contain 20,687 connected components, whereas sports fields contain 85. Component counts do not equal pixel frequencies and therefore cannot substitute for per-class pixel totals. The exact source-level training/validation counts, random seed, annotation guidelines, and annotator-agreement or adjudication records are unavailable. The aggregate scores discussed below are therefore unaudited historical records whose independence and sampling uncertainty cannot be reconstructed.

Figure 1 summarizes the archived data and annotation record without reproducing third-party source plans. The archive identifies example designs associated with external landscape practices, but source URLs, acquisition dates, licences, and written reuse terms were not preserved. Consequently, the underlying plans and derivatives require a documented rights review before any public data release.

The archived augmentation protocol lists random rotations up to $\pm 30^{\circ}$, horizontal and vertical flips, translations up to

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

Table 1: Comparison with Closely Related Landscape-AI Research

Study	Landscape task	Main method	Evaluation emphasis	Scope relative to this reappraisal
Zhou and Liu [1]	Plan recognition and rendering	Deep-learning recognition and image generation	Technical plan outputs	Direct task-level precedent for a recognition-and-rendering workflow
Zhou and Xiang [2]	Masterplan layout generation and rendering	Pix2pix–BicycleGAN workflow	Image analysis and practitioner surveys	Includes quantitative and practitioner-centred evaluation
Huang et al. [3]	Integration of garden visual styles	Neural style-transfer workflow	Visual and application-oriented comparison	Addresses style transfer in a related garden-design domain
Fan et al. [4]	Landscape visual-quality assessment	Fine-tuned assessment model	Prediction of visual quality	Treats visual-quality assessment as a task distinct from generation
Huang [5]	Site identification and landscape rendering	CycleGAN-enhanced pipeline	Identification and rendering results	Examines another landscape application of CycleGAN
Present reappraisal	Eleven-class plan segmentation and rendering	FCN-8s and CycleGAN	Surviving aggregate records, visual examples, and methodological audit	Evaluates evidentiary completeness and internal consistency

Table 2: Annotated Landscape Categories and Archived Distribution Summary

Element category	RGB value	Components	Images
Softscape elements			
Individual trees	(85, 165, 85)	17,892	318
Grass and shrub layers	(30, 45, 30)	20,687	310
Woodland masses	(200, 60, 25)	511	59
Water bodies	(30, 145, 240)	695	162
Hardscape elements			
Paved nodes	(245, 180, 180)	2,198	270
Main pathways	(240, 10, 250)	1,317	322
Step zones	(95, 10, 245)	276	74
Parking areas	(170, 15, 5)	96	33
Sports fields	(170, 220, 60)	85	31
Public installations	(250, 5, 5)	12,966	328
Building footprints	(250, 110, 190)	548	155

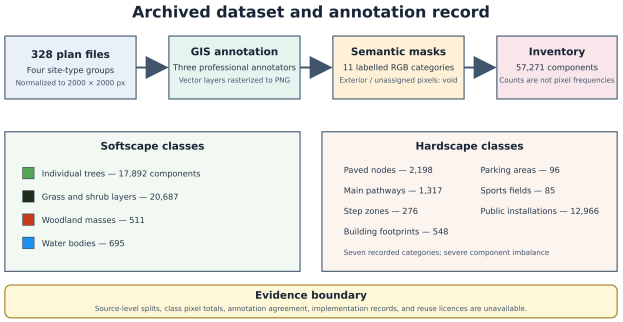

Figure 1: Schematic summary of the archived dataset and annotation record. The diagram reports the documented corpus size, annotation procedure, semantic groups, and evidentiary gaps; it does not reproduce third-party source plans.

± 50 pixels, and brightness and contrast adjustments up to $\pm 10\%$. Geometric transformations should be applied identically to an image and its mask. The surviving record does not establish whether the source-level split preceded augmentation or whether all derivatives of a source plan remained in one subset; potential leakage therefore cannot be ruled out retrospectively.

White exterior regions and pixels not assigned one of the 11 RGB codes are treated here as void rather than as a twelfth semantic class. In the mathematical reconstruction below, void pixels are excluded from the loss and confusion matrix. This convention makes the 11-class notation internally consistent, but the surviving material does not establish whether the original implementation used the same ignore rule.

B. Reconstructed Workflow

Figure 2 presents the only internally consistent sequential interpretation of the archived description. A landscape plan is mapped by FCN-8s to an 11-class semantic prediction. The predicted map is divided into overlapping patches, translated by $G : X \to Y$, and blended at the corresponding source coordinates. During CycleGAN training, domain X comprises prepared semantic-map patches and domain Y comprises unpaired target-style rendering patches.

The semantic map is the generator input, but the retained generator is not shown to use class-specific modulation or a separate conditioning channel. Moreover, the available visual examples use prepared semantic masks; no retained example demonstrates a prediction produced by FCN-8s and then rendered by CycleGAN. The diagram therefore describes the intended pipeline rather than a verified end-to-end execution.

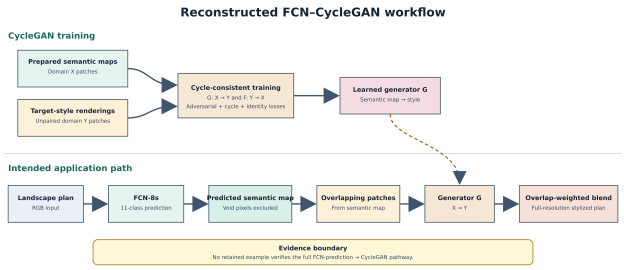

Figure 2: Reconstructed sequential workflow. Solid arrows define the intended application path from a landscape plan to an FCN prediction, overlapping semantic-map patches, CycleGAN translation, and overlap-weighted reconstruction. The training inputs to CycleGAN are prepared semantic maps and unpaired target-style renderings.

C. FCN Semantic Segmentation

The segmentation model is described as a VGG-based FCN-8s network [9], [24], summarized in Figure 3. For an input image $I \in \mathbb{R}^{H \times W \times 3}$, the network produces probabilities $\widehat{P} \in \mathbb{R}^{H \times W \times 11}$. The encoder contains five convolutional blocks with 3×3 kernels and channel widths of 64, 128,

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

256, 512, and 512. Each block is followed by 2×2 max pooling. The original fully connected layers are represented as convolutional layers, and 1×1 score layers map the deep, pool4, and pool3 features to 11-class logits. The deep score is upsampled by a factor of two and fused with the pool4 score; that result is upsampled by two, fused with the pool3 score, and upsampled by eight to the input resolution.

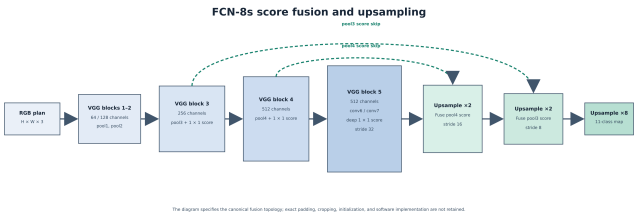

Figure 3: FCN-8s architecture reconstructed from the documented VGG-based design. The pool4 and pool3 score maps are fused successively with the upsampled deep score, followed by eightfold upsampling to the input resolution.

The retained protocol specifies normalization to 2000×2000 pixels but does not record the exact padding or cropping convention at successive downsampling stages. Let Ω_L denote the set of labelled, non-void pixels and $N = |\Omega_L|$. The reconstructed 11-class loss is

$$L_{\text{seg}} = -\frac{1}{N} \sum_{i \in \Omega_L} \sum_{c=1}^{C} y_{i,c} \log(\widehat{p}_{i,c}), \tag{1}$$

where $C = 11$, $y_{i,c}$ is the one-hot ground-truth label, and $\widehat{p}_{i,c}$ is the predicted probability for pixel i and class c. All confusion-matrix counts below are likewise restricted to Ω_L.

Pixel accuracy is

$$\text{PA} = \frac{\sum_{c=1}^{C} n_{cc}}{\sum_{c=1}^{C} \sum_{k=1}^{C} n_{ck}}, \tag{2}$$

where n_{ck} is the number of pixels from ground-truth class c predicted as class k. The Intersection over Union for class c and its unweighted class average are

$$\text{IoU}_c = \frac{n_{cc}}{\sum_k n_{ck} + \sum_k n_{kc} - n_{cc}}, \tag{3}$$

$$\text{mIoU} = \frac{1}{C} \sum_{c=1}^{C} \text{IoU}_c. \tag{4}$$

The archived training summary reports approximately 28,000 iterations, described as 40 epochs. It does not preserve the FCN optimizer, learning-rate schedule, batch size, weight decay, class weighting, initialization, framework version, hardware, checkpoint rule, or runtime. These missing implementation fields prevent exact reproduction and are not reconstructed from convention.

D. CycleGAN Style Transfer

CycleGAN learns $G : X \to Y$ and $F : Y \to X$, with discriminators D_Y and D_X [14]. Here, X denotes semantic-map patches and Y denotes unpaired target-style rendering patches. The archived manuscript states the following logistic adversarial objective:

$$L_{\text{GAN}}(G, D_Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x)))], \tag{5}$$

with an analogous reverse-direction term. Cycle consistency is

$$L_{\text{cyc}} = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1], \tag{6}$$

and identity preservation is

$$L_{\text{id}} = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|F(x) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G(y) - y\|_1]. \tag{7}$$

The recorded manuscript-level objective is

$$L_{\text{style}} = L_{\text{GAN}}(G, D_Y) + L_{\text{GAN}}(F, D_X) + 10 L_{\text{cyc}} + 0.5 L_{\text{id}}. \tag{8}$$

Figure 4 distinguishes one-way translations from cycle reconstructions. Each documented generator contains an initial 7×7 convolution, two stride-2 downsampling layers, nine residual blocks [25], two upsampling layers, and a final 7×7 convolution with a hyperbolic-tangent output. Each discriminator follows the 70×70 PatchGAN principle [13]: D_Y receives real y and generated $G(x)$ samples, while D_X receives real x and generated $F(y)$ samples.

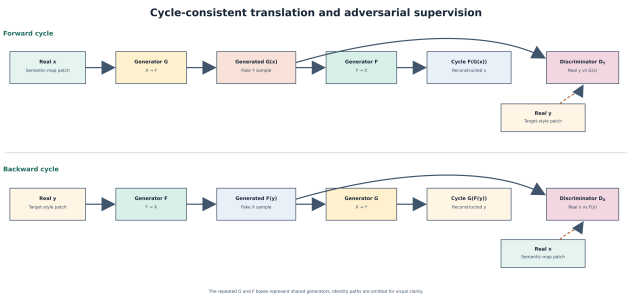

Figure 4: CycleGAN training structure for semantic-map domain X and target-style rendering domain Y. The solid horizontal paths show one-way translations and discriminator inputs; the return paths show the cycle reconstructions $F(G(x))$ and $G(F(y))$.

The archived configuration records random 256×256 crops from images resized to 1024×1024, a batch size of one, and Adam optimization [26] with learning rate 2×10^{-4}, $\beta_1 = 0.5$,

and $\beta_2 = 0.999$. Checkpoints at Epochs 50, 100, 200, and 300 are present. Framework and software versions, hardware, initialization, target-domain count, checkpoint-selection rule, and total training time are absent. Source code or logs are also unavailable to determine whether the implementation used the stated logistic loss or the least-squares variant common in CycleGAN software, or whether the stated coefficient 0.5 for L_{id} was multiplied by a separate cycle-loss weight. Equations (5)–(8) therefore reproduce the archived mathematical account rather than a code-verified objective.

E. Patch Reconstruction and the Role of Semantic Masks

Let I_p be a 256×256 semantic-map patch at location p, and let

$$S_p = G(I_p) \tag{9}$$

be its translated style patch. Multiplying the same S_p by every class mask and then summing the layers would reduce to S_p wherever the class masks are exhaustive because $\sum_c M_{c,p} = 1$. In void regions the class-mask sum is zero. Neither case produces a class-specific style, so mask summation is not part of the reconstructed method.

A generic normalized overlap-weighted reconstruction consistent with the archived description is

$$\widetilde{I}[i,j] = \frac{\displaystyle\sum_{p:(i,j)\in p} \omega_p(i,j) S_p[i_p, j_p]}{\max\left(\varepsilon, \displaystyle\sum_{p:(i,j)\in p} \omega_p(i,j)\right)}, \tag{10}$$

where $\omega_p(i,j)$ is a non-negative spatial weight, (i_p, j_p) is the local patch coordinate, and $\varepsilon > 0$ prevents division by zero. Pixels with no contributing patch require an implementation-defined fill value. Because the patch stride, boundary handling, and exact window function are unavailable, Equation (10) is a reconstruction of the blending principle rather than the exact historical implementation. Semantic masks remain useful for class-wise error analysis and semantic-preservation testing, but no class-specific generator or additional mask-conditioning mechanism is evidenced.

IV. Evaluation Protocol

The segmentation archive contains aggregate pixel accuracy and mIoU calculated with Equations (2)–(4). Because of the pronounced imbalance in Table 2, pixel accuracy must be interpreted with per-class IoU, precision, recall, and confusion information. Those class-level results, repeated-run variability, and exact validation sample counts are unavailable; the aggregate scores are therefore descriptive rather than comparative.

The rendering record contains side-by-side outputs from Epochs 50–300. The qualitative review considers contour preservation, texture continuity, transition smoothness, semantic-region stability, and visible artifacts. Figure 7 additionally presents normalized luminance distributions for one displayed sample. Pixels with all RGB channels above 245 are treated as background. The three histograms use identical bins and a common vertical scale, but the threshold and rasterized source panels still make the comparison illustrative only.

A complete evaluation would require immutable source-level splits; repeated runs; per-class segmentation results; FCN-8s, U-Net, SegNet, and DeepLabv3+ baselines; direct comparison of prepared-mask and FCN-predicted-mask inputs; CycleGAN, CUT, SPADE, SEAN, and OASIS rendering baselines; SSIM and LPIPS for aligned content; FID or KID for distributional comparison; semantic-preservation scores; and blinded ratings from landscape professionals. None of these experiments is recoverable from the available record, and no replacement values are introduced.

V. Results and Discussion

A. Archived Segmentation Record

The internal-validation summary records pixel accuracy above 92% and mIoU of 0.87. The underlying prediction files, validation partition, confusion matrix, and per-class values are unavailable. The figures therefore cannot be recalculated, and the record does not establish performance for rare categories such as sports fields, parking areas, and step zones. No FCN prediction panel survives that can be matched unambiguously to a ground-truth mask.

The accompanying archival notes identify four recurrent error types: confusion of dense hardscape textures with sports or water regions; merging of building footprints with paved spaces or voids; noisy responses around embedded labels; and fragmentation of semi-transparent or overlapping vegetation. Without paired prediction files, these observations remain qualitative error hypotheses rather than measured frequencies. Any disagreement between a recovered model output and an archived mask should be treated as an annotation-review case unless independent re-annotation establishes that the original label was incorrect.

B. Archived Style-Transfer Progression

Figure 5 displays the available checkpoints for archived sample No. 303. The sequence moves from relatively flat fills at Epoch 50 toward denser textures at later checkpoints, while local paths, planting boundaries, and built forms change visibly. The panel labelled “reference rendering” is retained from the archive; the documentation does not establish whether it was a paired target withheld from training or a qualitative design reference. It must not be interpreted as a ground-truth target without that provenance.

Figure 6 shows three additional archived cases. Later outputs generally exhibit richer vegetation and ground textures, but structures and transitions remain blurred in some regions. These examples demonstrate checkpoint-dependent visual change; they do not establish average improvement over a defined validation collection, semantic fidelity, or professional acceptability.

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

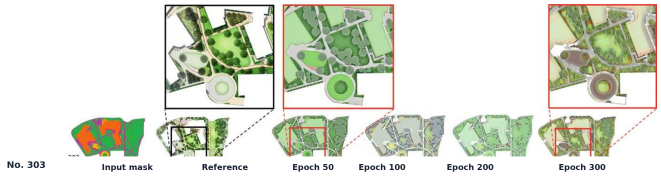

Figure 5: Available CycleGAN checkpoints for sample No. 303. The semantic input, archived reference panel, and outputs at Epochs 50, 100, 200, and 300 are shown. The relationship of the reference panel to the unpaired training set is undocumented.

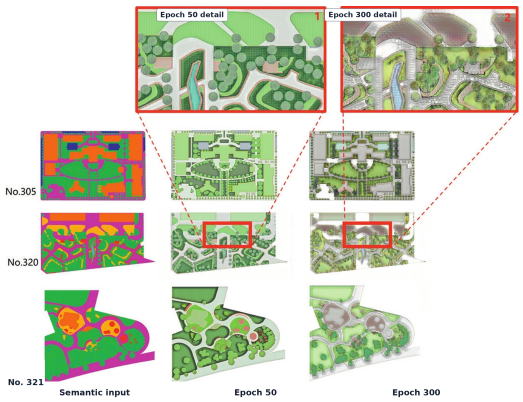

Figure 6: Three archived semantic inputs and their CycleGAN outputs at Epochs 50 and 300. Enlargements compare a local region of sample No. 320; callouts 1 and 2 correspond to Epochs 50 and 300, respectively.

C. Illustrative Luminance Diagnostic

Figure 7 compares the non-background luminance distributions of the archived reference, Epoch 50, and Epoch 300 panels for sample No. 308. The panels occupy overlapping intensity ranges, but their peaks differ. This establishes only that their rasterized luminance distributions are not identical. It does not measure colour fidelity, structural preservation, or perceptual equivalence because luminance histograms discard hue, saturation, and spatial arrangement.

D. Pipeline Coupling, Interpretation, and Practical Use

The surviving record documents an FCN segmentation stage and a CycleGAN translation stage, but it does not include an end-to-end example in which an FCN prediction is passed to the generator. The displayed CycleGAN inputs are prepared semantic masks. Training on clean annotations and applying the model to imperfect FCN predictions would introduce a domain shift in boundary shape, missing classes, isolated pixels, and colour coding. Its effect cannot be inferred from the available examples.

Accordingly, the evidence does not show that FCN-8s is preferable to newer segmentation models, that CycleGAN is preferable to alternative translators, or that the combined system improves design quality or professional efficiency. At most, it supports separate archived demonstrations of segmentation and semantic-map translation. Any practical use would require input-noise robustness tests and direct comparison of ground-truth-mask and FCN-predicted-mask rendering.

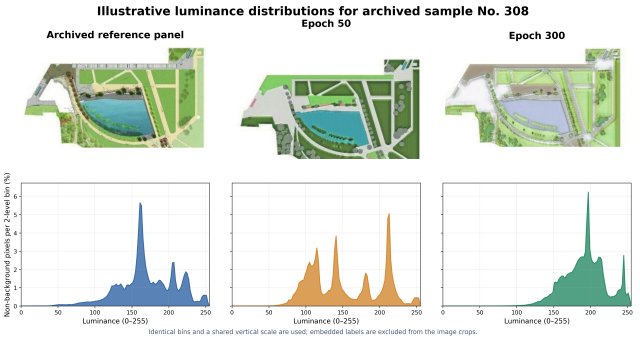

Figure 7: Illustrative luminance distributions for archived sample No. 308, calculated after removing embedded labels from the image crops. Pixels with $R, G, B > 245$ were excluded as background. Identical bins, normalization by the remaining pixel count, and a common vertical scale are used.

Earlier landscape research combines recognition and rendering [1], while later work provides quantitative and practitioner-centred evaluation [2]. This reappraisal therefore makes no claim of conceptual priority or state-of-the-art performance. The archived outputs are suitable only for discussing methodological risks in preliminary visualization. Generated plans would require inspection for circulation discontinuities, boundary leakage, structural distortion, and inappropriate texture before use in design communication.

E. Limitations

Seven limitations govern interpretation. First, exact source-level partitions, random seeds, and CycleGAN domain counts are unavailable. Second, source licences and annotation-adjudication records are incomplete. Third, the handling of void pixels cannot be confirmed from implementation records. Fourth, severe class imbalance is documented, but per-class metrics and class-balancing experiments are unavailable. Fifth, the scores describe one unaudited internal-validation summary without an external test set, baselines, ablations, or uncertainty estimates. Sixth, the rendering assessment relies on a small, potentially selected set of examples and an illustrative luminance diagnostic rather than structural, perceptual, distributional, and professional evaluation. Seventh, neither the end-to-end FCN–CycleGAN path nor the reconstructed overlap blending can be verified because prediction files, patch stride, boundary window, source code, framework, hardware, and runtime are unavailable. These limitations prevent claims of reproducibility, generalization, comparative superiority, or deployment readiness.

VI. Conclusion

This technical reappraisal reconstructs an archived FCN–CycleGAN workflow for landscape-plan segmentation and

rendering. The surviving inventory describes 328 plans and 57,271 connected annotations across 11 labelled semantic categories. An unaudited internal-validation summary records pixel accuracy above 92% and mIoU of 0.87, while the displayed CycleGAN checkpoints show increasing texture density accompanied by structural changes between Epochs 50 and 300. Algebraic analysis shows that summing a single translated patch through exhaustive class masks cannot create class-specific styles.

The surviving materials document the two stages separately but do not verify the combined prediction-to-rendering pipeline. Reproducible source-level partitions, verified permissions, annotation-agreement analysis, explicit void handling, per-class segmentation results, external testing, modern baselines, predicted-mask ablations, perceptual and semantic-preservation measures, and blinded professional assessment are required before accuracy, efficiency, or practical value can be established.

Data Availability

The data used is included within this paper.

Funding Statement

The work is not supported by any funding.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this study.

Declaration of Generative AI Use

During revision of this manuscript, the author used OpenAI's ChatGPT to assist with language editing, structural organization, LaTeX preparation, literature-search support, and the redrawing of schematic figures. The tool was not used to generate experimental data, execute the reported model training, or fabricate results. The author reviewed the AI-assisted material, verified the cited sources and technical statements to the extent permitted by the surviving record, and accepts full responsibility for the final content.

References

[1] Zhou, H., & Liu, H. (2021). Artificial intelligence aided design: Landscape plan recognition and rendering based on deep learning. Chinese Landscape Architecture, 37(1), 56–61.

[2] Zhou, H., & Xiang, S. (2024). Applicability evaluation and reflection on artificial intelligence-based "image to image" generation of landscape architecture masterplans. Landscape Architecture Frontiers, 12(2), 58–67.

[3] Huang, B., Mo, L., Tang, X., & Luo, L. (2024). Application of style transfer algorithm in the integration of traditional garden and modern design elements. PLOS ONE, 19(12), Article e0313909.

[4] Fan, R., Chen, Y., & Yocom, K. P. (2024). A new approach to landscape visual quality assessment from a fine-tuning perspective. Land, 13(5), Article 673.

[5] Huang, J. (2025). Cyclegan-enhanced site identification and rendering for landscape gardening. International Journal of High Speed Electronics and Systems. Advance online publication (Article 2540579).

[6] Shao, Y., Ma, N., Chen, M., Zhang, C., & Cui, Y. (2025). Machine learning in landscape architecture: A comprehensive review of advancements, applications, and future directions. Buildings, 15(21), Article 3827.

[7] Li, Z., Zhao, Q., Dong, L., & Sun, M. (2026). Research on AI-assisted generation of Jiangnan classical garden architectural scenes based on LLM and LDM. Frontiers in Built Environment, 12, Article 1825552.

[8] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672–2680.

[9] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431–3440). IEEE.

[10] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Medical image computing and computer-assisted intervention–MICCAI 2015 (Lecture Notes in Computer Science, Vol. 9351, pp. 234–241). Springer.

[11] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder–decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481–2495.

[12] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder–decoder with atrous separable convolution for semantic image segmentation. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), Computer vision–ECCV 2018 (Lecture Notes in Computer Science, Vol. 11211, pp. 833–851). Springer.

[13] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1125–1134). IEEE.

[14] Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2223–2232). IEEE.

[15] Park, T., Efros, A. A., Zhang, R., & Zhu, J.-Y. (2020). Contrastive learning for unpaired image-to-image translation. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), Computer vision–ECCV 2020 (Lecture Notes in Computer Science, Vol. 12354, pp. 319–345). Springer.

[16] Park, T., Liu, M.-Y., Wang, T.-C., & Zhu, J.-Y. (2019). Semantic image synthesis with spatially-adaptive normalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2337–2346). IEEE.

[17] Zhu, P., Abdal, R., Qin, Y., & Wonka, P. (2020). SEAN: Image synthesis with semantic region-adaptive normalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5104–5113). IEEE.

[18] Schönfeld, E., Sushko, V., Zhang, D., Gall, J., Schiele, B., & Khoreva, A. (2021). You only need adversarial supervision for semantic image synthesis. In Proceedings of the 9th International Conference on Learning Representations.

[19] Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2980–2988). IEEE.

[20] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600–612.

[21] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 586–595). IEEE.

[22] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Advances in Neural Information Processing Systems, 30, 6626–6637.

[23] Bińkowski, M., Sutherland, D. J., Arbel, M., & Gretton, A. (2018). Demystifying MMD GANs. In Proceedings of the 6th International Conference on Learning Representations.

[24] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations.

[25] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770–778). IEEE.

[26] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations.