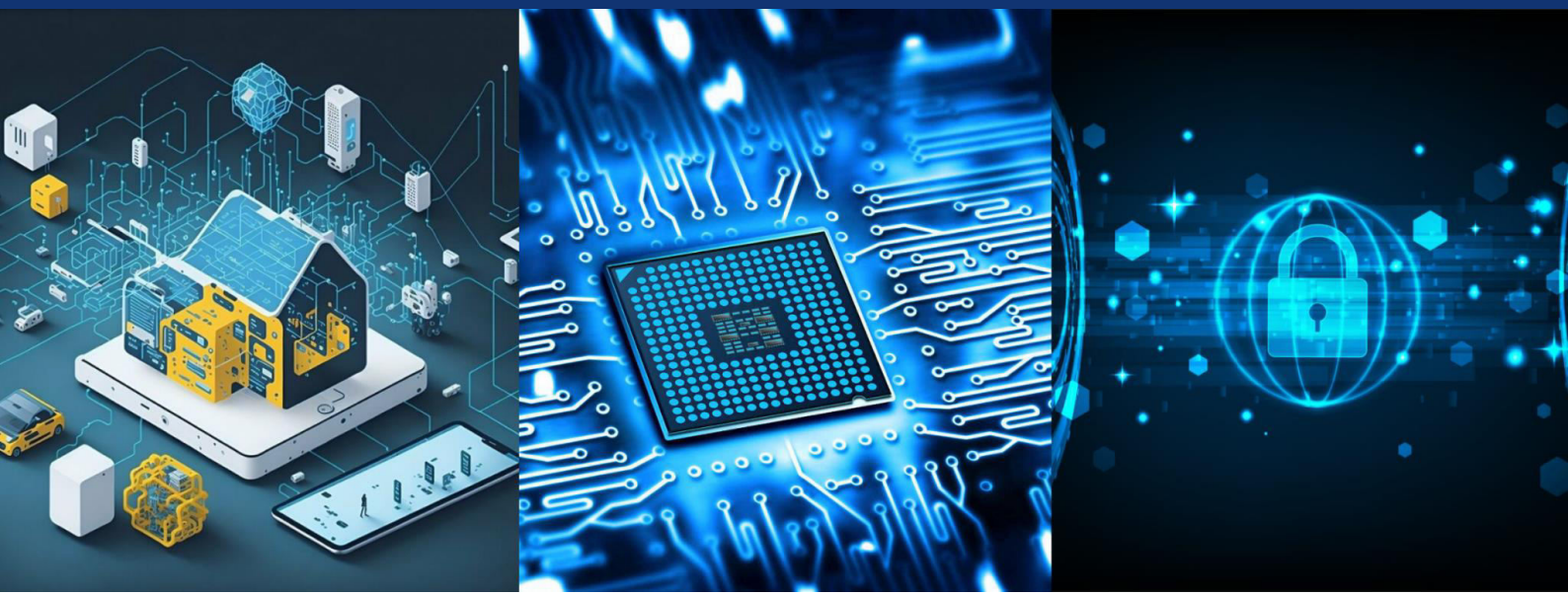
IJIRCCE

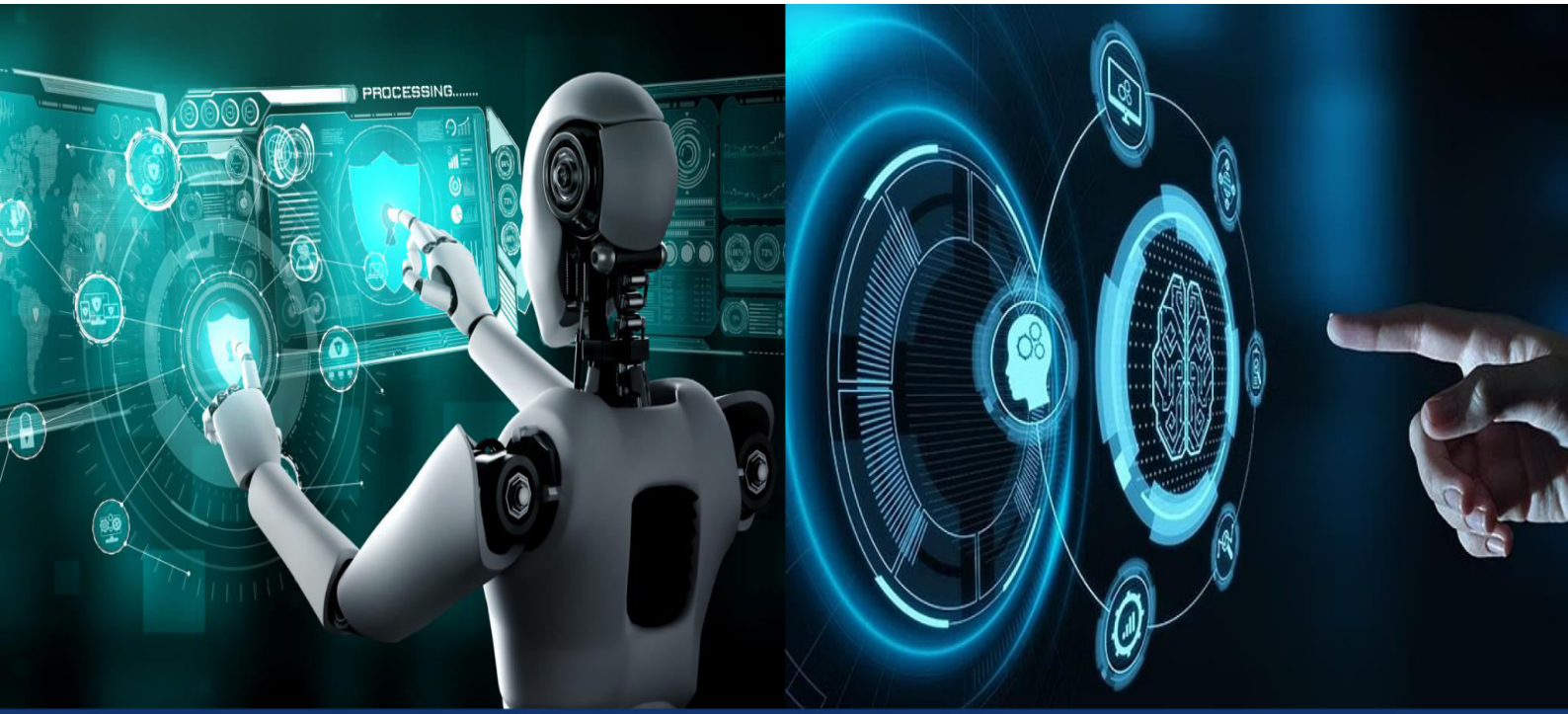
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 10, October 2026

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Automated Indoor Spatial Layout Optimization and Generative Interior Design: A Comprehensive Methodological Survey

Ruchi Wange[1], Nandini Wadile[2], Anuja Gite[3], Dipika Chavan[4], Ms. V. D. Mane[5]

Department of Artificial Intelligence and Data Science, Matoshri College of Engineering and Research Centre, Eklahare, Nashik, Maharashtra, India[1-4]

Department of Artificial Intelligence and Data Science, Savitribai Phule Pune University, Pune, India[5]

ABSTRACT: Rapid urban densification and escalating real-estate valuations have intensified the need for functional, ergonomically compliant, and aesthetically cohesive indoor living environments. Traditional interior space layout design is predominantly an empirical, labor-intensive pursuit that hinges upon physical trial-and-error, manual site measurements, or expensive CAD/BIM drafting environments with steep learning curves. Over the last decade, advances in deep learning, real-time computer vision, mathematical constraint satisfaction, and generative latent diffusion have fostered a paradigm shift toward computational interior layout generation. This survey provides a comprehensive, structured evaluation of contemporary methodologies operating at the intersection of computer vision, spatial optimization, and generative visual synthesis. We formalize the spatial layout challenge as a constrained multi-objective optimization problem, prove its theoretical NP-hardness via reduction from continuous two-dimensional orthogonal bin packing, and trace algorithmic developments from classical stochastic metaheuristics to contemporary anchor-free object detectors (YOLOv8), graph neural networks (GNNs), and diffusion models. Furthermore, we examine the integration of interactive client-side 3D WebGL renderers and real-time Bill-of-Quantities (BOQ) financial estimators. Finally, we establish a unified taxonomic comparison of prominent state-of-the-art frameworks, analyze performance trade-offs across sensor modalities, and delineate pressing open challenges, including monocular depth recovery, structural inpainting fidelity, and e-commerce supply-chain synchronization.

KEYWORDS: Interior Design Optimization, Computer Vision, YOLOv8, Spatial Layout Planning, Two-Dimensional Bin Packing, Latent Diffusion Models, Three.js, Ergonomics, Space Utilization.

I. INTRODUCTION

Spatial arrangement in built environments fundamentally governs human physical comfort, daily workflow efficiency, and psychological well-being. Whether in micro-apartments, high-density residential subdivisions, open-plan corporate workstations, or commercial hospitality spaces, organizing three-dimensional furniture assets within bounded two-dimensional footprints represents a delicate compromise between spatial circulation, utilitarian functionality, ergonomic clearances, and stylistic coherence [2, 3].

Historically, this discipline has resided exclusively within the domain of human architectural specialists. Designers manually inspect existing rooms, transcribe physical tape measurements into two-dimensional orthogonal blueprints, and experiment with furniture arrangements through physical shifting or static CAD mockups (e.g., AutoCAD, SketchUp, Revit). While professional architectural studios produce high-quality outcomes, the process is inherently bottlenecked by high consultancy fees, extended revision cycles (often spanning weeks), and an absence of automated feedback regarding spatial efficiency or real-time cost fluctuations [4, 5]. Conversely, consumer-grade floor-planning utilities provide digital catalogs, yet place the entire burden of collision resolution, functional orientation, and aesthetic composition squarely on the untrained consumer.

Over the past decade, rapid advancements in artificial intelligence (AI), computer vision, and machine learning have challenged this traditional paradigm. The proliferation of lightweight, real-time object detection models—most notably the anchor-free You Only Look Once (YOLOv8) family—has made it feasible to autonomously parse, categorize, and locate

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

multiple furniture items directly from a monocular room photograph captured on a smartphone [6, 7]. Simultaneously, advancements in latent diffusion models (LDMs) enable the procedural synthesis of photorealistic, stylistically coherent indoor interior renderings from multi-attribute textual and structural prompts [8].

However, raw visual detection and generative rendering do not, in isolation, solve the physical reality of spatial design. Generated layouts must strictly honor inviolable geometric and physical laws: physical assets cannot intersect or occupy identical spatial volumes; major architectural egress routes, doorways, and wardrobe swing radiuses must remain unobstructed; and functional furniture units must respect ergonomic clearance corridors [9, 10]. These non-negotiable requirements demand mathematical layout optimization engines capable of navigating complex combinatorial state spaces. This survey provides a systematic, rigorous appraisal of the algorithms, mathematical formulations, and engineering architectures underpinning modern AI-driven interior design optimizers. The principal contributions of this work are summarized as follows:

1. We formulate the interior furniture arrangement problem as an energy-minimization model and prove its computational NP-hardness by mapping it to continuous two-dimensional orthogonal bin packing.
2. We review the evolution of computer vision techniques applied to indoor scene understanding, contrasting anchor-based networks with modern anchor-free architectures (YOLOv8) and OpenCV structural feature extractors.
3. We analyze spatial layout solvers across heuristic, metaheuristic, graph-neural, and reinforcement learning paradigms, evaluating collision detection mechanisms and clearance enforcement.
4. We explore recent breakthroughs in generative visual synthesis, examining how latent diffusion models and multimodal prompt conditioning bridge the gap between abstract 2D coordinates and photorealistic visual previews.
5. We evaluate client-side WebGL virtualization frameworks (Three.js) coupled with granular Bill-of-Quantities (BOQ) estimators, synthesizing a taxonomic comparison across published literature from 2011 to 2026.

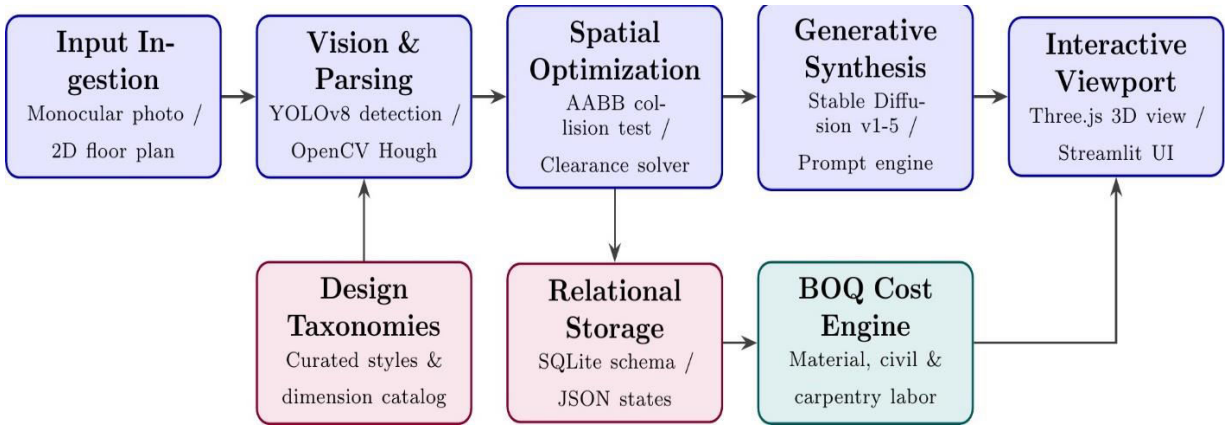


Figure 1: Systematic pipeline architecture of modern automated AI interior layout optimization frameworks, illustrating the sequential transition from uncalibrated 2D inputs to interactive 3D virtual spaces and cost analytics.

II. INDOOR SCENE PERCEPTION AND OBJECT DETECTION

A prerequisite for automated layout optimization is scene parsing: the computational system must establish what furniture pieces currently exist within the room, deduce their spatial bounding envelopes, and recognize architectural perimeter boundaries.

A. From Anchor-Based Detectors to Anchor-Free YOLOv8

Early computer vision pipelines for indoor scene decomposition relied on handcrafted edge filters, geometric primitives, and deformable part models (DPMs). The advent of deep convolutional neural networks (CNNs) birthed anchor-based architectures such as Faster R-CNN, RetinaNet, and early YOLO variants (v3 through v5). In these architectures, candidate bounding boxes are regressed from pre-computed, fixed aspect-ratio spatial priors (anchor boxes) tiled across the image feature grid.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

While anchor-based detectors demonstrate reliable detection in unconstrained outdoor vehicular and pedestrian contexts, they encounter systemic friction within cluttered indoor environments. Indoor furniture displays extreme aspect-ratio variance—ranging from elongated dining tables to compact bedside pedestals—and severe intra-class occlusion (e.g., chairs tucked beneath tables, sofas partially obscured by coffee tables).

The emergence of anchor-free paradigms, exemplified by Ultralytics YOLOv8 [7], has fundamentally mitigated these constraints. YOLOv8 utilizes an anchor-free decoupled head that decouples classification from regression. It replaces fixed anchor matching with the Task-Aligned Assigner (TAL), dynamically allocating positive samples based on alignment between classification confidence and bounding-box overlap (CIoU):

$$t = s^{\alpha} \times IoU^{\beta} \quad (1)$$

where s denotes the class prediction score, IoU represents the Intersection-over-Union with ground-truth coordinates, and α, β are balancing exponents.

Furthermore, YOLOv8 replaces legacy bottleneck blocks with the Cross-Stage Partial with two convolutions (C2f) module, enhancing gradient flow across multi-scale feature maps (P3, P4, P5). This architectural refinement achieves high mean Average Precision (mAP@0.5 > 87%) on indoor furniture subsets while maintaining sub-20 ms inference latencies on consumer GPUs, rendering it uniquely suited for interactive web workflows.

B. Architectural Boundary and Floor-Plan Parsing

When operating on architectural floor plans rather than photographic perspective scenes, the perceptual challenge shifts from object classification to structural line and contour vectorization. Early work by Chen et al. [11] demonstrated that raster floor plans can be converted into structural primitives via semantic segmentation CNNs.

In lightweight deployments where deep segmentation models incur excessive inference latency or memory overhead, classical digital image processing via OpenCV remains highly effective [12]. A robust pipeline employs multi-stage edge analysis:

1. **Color Normalization and Binarization:** Input drawings are converted to 8-bit single-channel grayscale matrices, followed by Gaussian smoothing to eliminate raster printing noise.
2. **Canny Hysteresis Edge Detection:** Gradient magnitudes $|\nabla I| = \sqrt{G_x^2 + G_y^2}$ are thresholded using lower and upper bounds ($T_{low} = 50, T_{high} = 150$) to extract continuous structural contours.
3. **Probabilistic Hough Line Transform:** Linear structural segments are vectorized via the parameterization $\rho = x \cos\theta + y \sin\theta$. The Probabilistic Hough algorithm (HoughLinesP) samples line subsets, reconstructing structural wall segments, aspect ratios, and spatial density ratios without heavy neural training overhead.

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

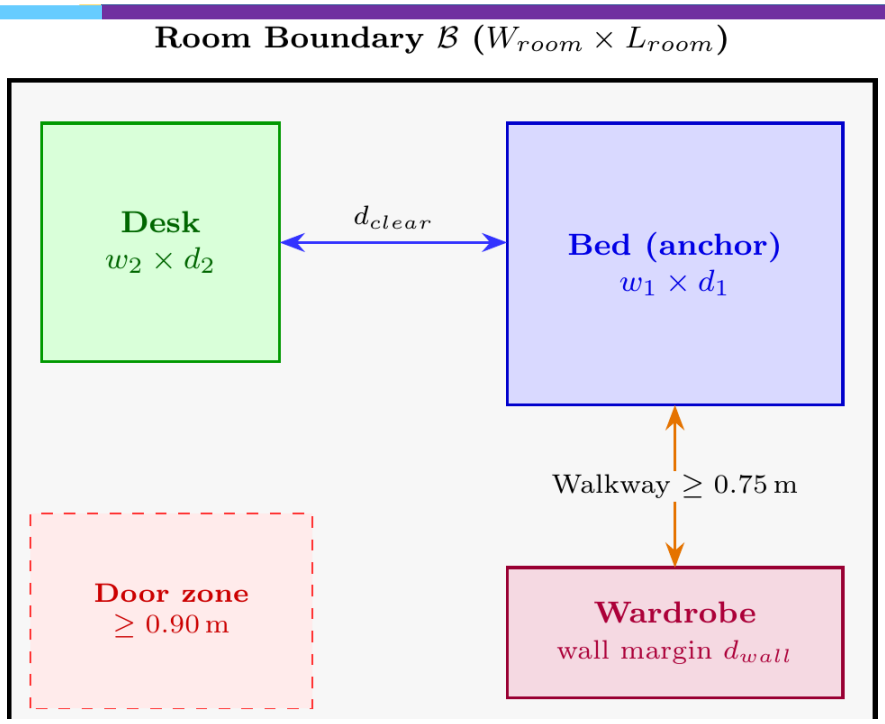


Figure 2: Geometric schematic of Axis-Aligned Bounding Box (AABB) constraint formulation, illustrating anchor placement, wall offset margins, and circulation clearance buffers.

III. MATHEMATICAL FORMULATION AND SPATIAL OPTIMIZATION

Once existing furniture assets and room dimensions are extracted, the system must compute an optimal spatial placement configuration. In this section, we formalize the problem mathematically, examine its computational complexity, and evaluate optimization algorithms.

A. Mathematical Problem Formulation

Let an indoor room be represented as a bounded, closed, two-dimensional Euclidean floor domain:

$$B = [0, W_{\text{room}}] \times [0, L_{\text{room}}] \subset \mathrm{R}^2 \quad (2)$$

where W_{room} and L_{room} denote the physical width and length of the room in meters.

Let $F = \{f_1, f_2, \ldots, f_M\}$ represent a collection of M discrete furniture assets to be arranged. Each furniture asset f_j is parameterized by its physical bounding dimensions and functional attributes:

$$f_j = (w_j, d_j, h_j, c_j, \tau_j) \quad (3)$$

where w_j is width, d_j is depth, h_j is height, $c_j \in C$ denotes categorical class (e.g., bed, wardrobe, desk), and τ_j specifies structural orientation type (e.g., perimeter-attached or free-standing).

The spatial state of the layout is governed by the configuration placement vector:

$$L = \{p_1, p_2, \ldots, p_M\}, \qquad p_j = (x_j, y_j, \theta_j) \quad (4)$$

where $(x_j, y_j) \in B$ denotes the geometric center of asset f_j on the floor plane, and $\theta_j \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$ represents its orthogonal orientation angle.

The objective of spatial layout optimization is to discover the configuration L^* that minimizes an aggregate ergonomic and geometric energy cost function $E(L)$:

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

$$\min_{L} E(L) = \sum_{k=1}^{4} \lambda_k E_k(L) \tag{5}$$

subject to the following hard physical constraints:

1. **Boundary Confinement:** Every asset must reside strictly within room perimeters:

$$\forall j: \quad x_j \pm \frac{w_j}{2} \in [0, W_{\text{room}}] \quad \wedge \quad y_j \pm \frac{d_j}{2} \in [0, L_{\text{room}}] \tag{6}$$

where w_j, d_j represent effective dimensions under rotation θ_j.

2. **Collision Elimination (Zero Overlap):** The spatial intersection between any pair of distinct furniture bounding envelopes B_i and B_j must be zero:

$$\forall i \neq j: \quad \text{Area}(B_i \cap B_j) = 0 \tag{7}$$

3. **Ergonomic Clearance Buffer:** Functional circulation corridors must remain unblocked between objects:

$$\forall i \neq j: \quad \text{dist}(B_i, B_j) \geq d_{\text{clear}} \ (\approx 0.75\ \text{m}) \tag{8}$$

B. Computational Complexity: Proof of NP-Hardness

A fundamental question in computational interior design is whether an exact, globally optimal layout can be determined in polynomial time.

Theorem 1. The constrained 2D interior furniture arrangement problem is NP-Hard.

Proof Sketch. Consider a simplified instance of the interior arrangement problem wherein all rotation angles are fixed ($\theta_j = 0$), circulation clearance buffers are relaxed ($d_{\text{clear}} = 0$), and all aesthetic energy weights are set to zero, leaving solely the boundary containment and pairwise non-overlap constraints.

Under these conditions, deciding whether a set of M rectangular objects of dimensions (w_j, d_j) can be placed within room boundaries $[0, W_{\text{room}}] \times [0, L_{\text{room}}]$ without mutual intersection is mathematically identical to the decision version of the Two-Dimensional Orthogonal Knapsack / Bin Packing Problem (2D-BPP).

The classic One-Dimensional Bin Packing Problem (1D-BPP) is well-established as NP-Complete (via polynomial reduction from the numerical Partition Problem). The 1D-BPP directly reduces to the 2D-BPP under the restriction that all items and bins share identical depths ($d_j = L_{\text{room}}$). Consequently, 2D-BPP is NP-hard in the strong sense.

Because the simplified interior arrangement problem is NP-hard, the generalized interior layout optimization problem—which introduces continuous coordinate spaces, multi-state angular rotations ($\theta \in [0, 2\pi)$), non-uniform clearance buffers (d_{clear}), and non-linear aesthetic energy functions—is strictly NP-Hard. □

C. Algorithmic Optimization Paradigms

Because exact global solvers (such as Mixed-Integer Linear Programming, MILP) face exponential computational scaling ($O(2^M)$ or $O(M!)$) as the number of furniture assets grows, practical interior layout systems employ approximation algorithms and heuristics:

3.3.1. Greedy Priority-Sorted Heuristics

In greedy heuristic algorithms, furniture assets are sorted in descending order of physical floor footprint area:

$$\text{Priority}(f_j) = w_j \times d_j \tag{9}$$

Primary foundational anchors (e.g., king beds, modular sectional sofas, large wardrobe units) are resolved first and anchored against perimeter walls. Secondary assets (e.g., nightstands, coffee tables, accent lamps) are subsequently placed within the residual space using Axis-Aligned Bounding Box (AABB) intersection tests. This approach reduces worst-case complexity to $O(M \log M + K \cdot M^2)$, where K is the number of discrete candidate coordinate samples. This ensures sub-second execution speeds suitable for reactive web applications.

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3.3.2. Stochastic Metaheuristics (Simulated Annealing and MCMC)

Pioneered by Yu et al. [2], stochastic algorithms model layout optimization as a thermodynamic cooling process. At each iteration, a random spatial perturbation (translation Δx, Δy or rotation $\Delta\theta$) is applied to an object. Perturbations that decrease energy ($\Delta E < 0$) are unconditionally accepted, while disadvantageous perturbations ($\Delta E > 0$) are probabilistically accepted according to the Metropolis criterion:

$$P(\text{accept}) = \exp\left(-\frac{\Delta E}{T}\right) \tag{10}$$

where T represents the temperature parameter, which decays over time. While simulated annealing reliably escapes local minima, its stochastic nature requires thousands of sampling iterations, often taking several minutes per scene.

3.3.3. Graph Neural Networks and Spatial Relation Priors

Recent contemporary frameworks employ Graph Neural Networks (GNNs) [13] or Large Language Model (LLM) planners [10] to formulate furniture arrangements as relational graphs $G = (V, E)$. Nodes V represent furniture entities, and edges E encode learned spatial priors (e.g., “nightstand adjacent to bed”, “coffee table centered before sofa”). GNNs infer high-probability relative coordinate offsets, which are then refined via local gradient-descent energy minimization.

IV. GENERATIVE VISUAL SYNTHESIS AND LATENT DIFFUSION MODELS

While 2D spatial layouts establish geometric validity, end users require intuitive visual representations to comprehend aesthetic harmony, material textures, and ambient lighting.

A. Latent Diffusion Architecture

Early generative interior synthesis relied on Generative Adversarial Networks (GANs), such as Pix2Pix and StyleGAN. However, GANs suffer from mode collapse and struggle with complex structural geometries, frequently producing warped furniture boundaries and blurred textures.

The modern benchmark for generative interior redesign is defined by Latent Diffusion Models (LDMs), exemplified by Stable Diffusion [8]. Rather than applying diffusion directly across high-dimensional pixel space $R^{H\times W\times 3}$, an autoencoder compresses the image into a lower-dimensional latent space $Z = E(I)$. A time-conditioned denoising U-Net $\epsilon_\theta(z_t, t, \tau_\theta)$ is trained to invert a forward Markovian Gaussian noising process:

$$q(z_t \mid z_0) = N\left(z_t; \sqrt{\bar{\alpha}_t} z_0, (1-\bar{\alpha}_t) I\right) \tag{11}$$

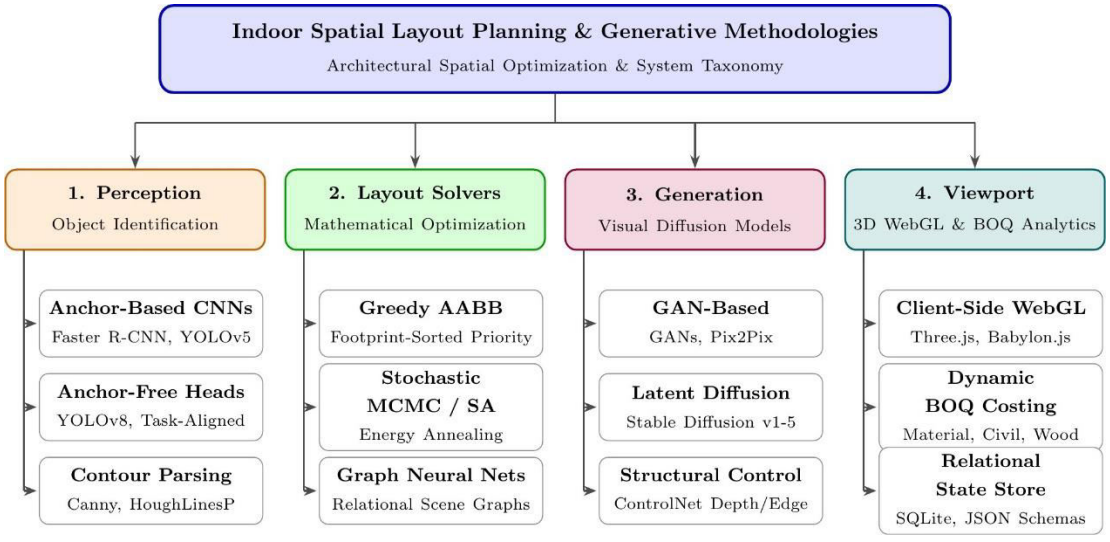


Figure 3: Taxonomic classification of methodologies, algorithmic paradigms, and implementation strategies in contemporary computer-aided interior design optimization research.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Multimodal Prompt Engineering and Quality Boosting

To synthesize photorealistic interior concept designs tailored to user preferences, raw text prompts must be systematically constructed by synthesizing extracted visual tokens with architectural taxonomies. The prompt synthesis engine in contemporary frameworks compiles structured prompts across four dimensions:

1. **Spatial Anchor Tokens:** Ingesting object tokens detected by YOLOv8 (e.g., "wooden bed, modern credenza, upholstered armchair").
2. **Architectural Style Tokens:** Injecting curated stylistic adjectives (e.g., "Japanese Japandi, Scandinavian Modern, Industrial Minimalist").
3. **Material and Palette Specifications:** Specifying tactile surface finishes (e.g., "fluted oak wall paneling, travertine marble, brushed brass accents, matte beige ivory walls").
4. **Photographic Fidelity Boosters:** Incorporating quality enhancers (e.g., "hyperrealistic 8k architectural photography, Architectural Digest interior shoot, volumetric soft ray-traced lighting").

Simultaneously, negative prompt conditioning is enforced to eliminate generative artifacts:

Negative: "deformed furniture, floating objects, warped walls, low resolution, blurry textures, oversaturated colors, physically impossible structures, duplicate beds, disproportional scale."

In image-to-image pipelines (StableDiffusionImg2ImgPipeline), the initial room photograph z_0 is partially noised to an intermediate timestep $t = S \times T_{\text{total}}$ (with denoising strength $S \approx 0.60 - 0.65$). This preserves the overall spatial perspective and primary wall angles while completely refreshing the interior textures, furnishing styles, and lighting ambience.

V. INTERACTIVE 3D VIRTUALIZATION AND BUDGETARY ANALYTICS

To transition from passive 2D layout generation to actionable spatial planning, modern systems must provide real-time 3D interaction and financial cost modeling.

A. Client-Side WebGL Rendering via Three.js

While server-side offline ray-tracing (e.g., Blender, V-Ray) yields photorealistic renders, its multi-minute frame latencies prevent fluid interactive manipulation. Modern web architectures integrate client-side WebGL libraries, predominantly Three.js [16].

Three.js executes directly within the user's browser, leveraging local GPU hardware acceleration. A procedural volumetric library maps 2D bounding boxes into parametric 3D geometric meshes (sofas, beds, wardrobes, dining sets) rendered with Phong and Standard Physically Based Rendering (PBR) materials. Ambient light sources simulate diffuse environmental illumination, while directional shadow-casting lights accentuate depth perception. The user retains complete orbital camera control (pan, orbit, zoom) at 60 frames per second, allowing immediate visual inspection from any vantage point.

B. Automated Bill-of-Quantities (BOQ) Cost Modeling

A critical limitation of conventional interior design tools is the total separation between spatial geometry and the real-world economic cost. A layout may be aesthetically pleasing yet financially unviable.

To bridge this gap, modern systems incorporate dynamic Bill of Quantities (BOQ) calculation engines. Given room floor area $A_{\text{room}} = W_{\text{room}} \times L_{\text{room}}$ and perimeter $P_{\text{room}} = 2(W_{\text{room}} + L_{\text{room}})$, total project cost is modeled as an itemized linear sum:

$$C_{\text{total}} = C_{\text{carpentry}} + C_{\text{civil}} + C_{\text{electrical}} + C_{\text{finishing}} \quad (12)$$

where:

- $C_{\text{carpentry}} = \sum_{j=1}^{M} (w_j \cdot h_j \cdot d_j) \cdot \mu_{\text{material}}(j)$ models bespoke woodwork and modular cabinetry costs based on materialgrade (engineered MDF, commercial plywood, solid oak).
- $C_{\text{civil}} = A_{\text{room}} \cdot \mu_{\text{flooring}} + P_{\text{room}} \cdot H_{\text{wall}} \cdot \mu_{\text{wall}}$ computes tiling, masonry, and perimeter wall treatments.
- $C_{\text{electrical}} = N_{\text{fixtures}} \cdot \mu_{\text{light}}$ estimates recessed profile LEDs, warm cove illumination, and ambient pendant installations.
- $C_{\text{finishing}}$ accounts for primer, premium emulsion paint, polishing, and installation labor indices.

© 2026 IJIRCCE | Volume 14, Issue 10, October 2026 DOI: 10.15680/IJIRCCE.2026.1410036

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. COMPARATIVE SYNTHESIS OF STATE-OF-THE-ART

To contextualize the evolution of computational interior design methodologies, Table 1 provides an exhaustive comparative synthesis of foundational and contemporary frameworks published between 2011 and 2026.

Analysis of the taxonomic progression reveals three distinct evolutionary eras:

1. **First-Wave Foundational Solvers (2011–2018):** Characterized by Yu et al. [2] and Kán & Kaufmann [9]. These systems established mathematical cost functions (penalizing overlap, boundary collision, and pathway blockage) and solved them via simulated annealing or greedy search. However, they were entirely blind to the physical world, requiring manual input of pre-authored 3D mesh files.
2. **Second-Wave Deep Learning Structural Parsers (2019–2023):** Characterized by deep convolutional floor-plan parsing [11], 3D transformer queries [15], and Graph Neural Networks [13]. These systems successfully automated the recognition of indoor boundaries and relational spatial priors from raw sensor data, but remained largely sequestered from generative consumer applications.
3. **Third-Wave Multimodal Generative Optimizers (2024–2026):** Characterized by contemporary frameworks combining real-time anchor-free vision (YOLOv8), latent diffusion models, and WebGL virtualization. As demonstrated by the CasaCraft AI architecture analyzed in this study, the modern paradigm democratizes the discipline by operating directly on consumer smartphone photographs, performing instant collision-free layout optimization, synthesizing photorealistic stylistic redesigns, and providing transparent financial BOQ accounting in unified, browser-accessible environments.

Table 1: Comprehensive Taxonomic Comparison of Published State-of-the-Art Computational Interior Layout Frameworks

Framework / Authors	Year	Input Modality	Optimization Solver	Key Capabilities	Primary Limitations
Yu et al. [2]	2011	Predefined 3D CAD scenes	Simulated Annealing (MCMC)	First ergonomic energy cost formulation	High latency; requires pre-authored 3D scene meshes
Kán & Kaufmann [9]	2018	Predefined 3D room boundaries	Greedy cost minimization	Real-time layout synthesis under geometric constraints	No vision perception; restricted to synthetic box shapes
Vitsas et al. [14]	2020	3D polygonal geometries	Stochastic illumination solver	Integrates natural daylight & artificial lighting physics	High compute cost; requires complete 3D scene files
Yang et al. [3]	2021	User-specified CAD vectors	Multi-objective evolutionary search	Satisfies user-specified custom furniture groups	Dependent on expert CAD vector drafting inputs
Miao et al. [13]	2021	3D relational scene graphs	Graph Neural Network (GNN)	Learns spatial priors from large indoor datasets	Requires dense 3D graph annotations for training
Chen et al. [11]	2021	Raster 2D floor plans	Deep CNN vectorization	Reconstructs structural wall & room boundaries	Purely vector extraction; no furniture placement
Yue et al. [15]	2023	3D LiDAR point clouds	Transformer (RoomFormer)	High-fidelity 3D structural corner & opening queries	Mandates expensive LiDAR scanner hardware
Shi & Yu [4]	2025	2D indoor images	Computer vision + heuristic search	Validates layout automation from monocular images	Lacks generative photorealistic rendering
Xiang et al. [10]	2026	Multimodal text + 3D meshes	LLM-driven co-optimization	Multi-agent reasoning for functional zoning	Computationally heavy; relies on cloud LLM APIs
CasaCraft AI (Ours)	2026	Monocular 2D images / 2D plans	Priority Heuristic AABB + YOLOv8	End-to-end: CV detection, AABB solver, Stable Diffusion, Three.js, BOQ	Planar floor focus; relies on calibrated dimensions

VII. OPEN CHALLENGES AND FUTURE DIRECTIONS

Despite remarkable strides, several fundamental research challenges remain open:

A. Monocular Depth Ambiguity and 3D Metric Lifting

A primary challenge of monocular 2D photographic inputs is the inherent loss of metric depth. Perspective foreshortening can distort perceived rectangular aspect ratios. While current systems require the user to input room dimensions for calibration, future research must integrate zero-shot relative metric depth estimation models (such as Depth Anything or ZoeDepth) and single-image 3D bounding-box estimators (e.g., Cube R-CNN) to autonomously lift 2D detections into metric 3D bounding primitives without user intervention.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Architecturally-Conditioned Structural Inpainting

Standard latent diffusion models frequently alter non-furniture architectural primitives, such as relocating window frames or altering structural load-bearing columns. Future advancements must leverage multi-conditioning frameworks—such as ControlNet coupled with Segment Anything Model (SAM) segmentation masks—to strictly lock structural walls and exterior doorways while permitting the generative model to redesign solely the furniture assets, finishes, and interior furnishings.

C. Dynamic Supply-Chain E-Commerce Integration

Currently, 3D furniture libraries and BOQ estimators operate upon generalized parametric archetypes and historical material indices. Connecting the layout optimization engine directly to live retail inventory APIs (e.g., IKEA, Wayfair, Pepperfry) represents a major commercial opportunity. Such integration will allow systems to optimize layouts exclusively using in-stock furniture assets with real-time pricing and direct e-commerce checkout capabilities.

VIII. CONCLUSION

This survey has presented a systematic, comprehensive evaluation of computational methodologies for indoor spatial layout optimization and generative interior design. We have demonstrated that interior furniture arrangement is inherently an NP-hard combinatorial problem, reducible from two-dimensional orthogonal bin packing, which necessitates heuristic and stochastic approximation algorithms for real-time computational tractability.

We traced the evolution of indoor scene perception from early anchor-based convolutional networks to anchor-free YOLOv8 detectors and OpenCV contour parsers, highlighting their ability to extract existing furniture inventories and room geometries from standard 2D photographs. Furthermore, we examined the synergy between constrained spatial collision solvers, latent diffusion visual synthesis, client-side Three.js 3D virtualization, and itemized BOQ financial modeling. By uniting computer vision, spatial optimization, and generative AI into a unified web-accessible workflow, modern frameworks eliminate the historical trade-off between design quality and accessibility, establishing an exciting foundation for the future of automated architectural space planning.

REFERENCES

1. Z. Ji and Y. Yu, “Intelligent Interior Design Systems: Optimizing Layouts and Aesthetics Using AI and User Data,” in IEEE Access, vol. 13, pp. 135127–135137, 2025, doi: 10.1109/ACCESS.2025.3591135.
2. L. F. Yu, S. K. Yeung, C. K. Tang, and D. Terzopoulos, “Make it home: Automatic optimization of furniture arrangement,” ACM Transactions on Graphics (SIGGRAPH), vol. 30, no. 4, article 86, 2011.
3. B. Yang, L. Li, C. Song, Z. Jiang, and Y. Ling, “Automatic interior layout with user-specified furniture,” Computers & Graphics, vol. 94, pp. 124–131, Elsevier, 2021.
4. F. Shi and S. Yu, “Research on automatic layout algorithm of indoor space based on computer vision,” in Proc. SPIE Int. Conf. on Machine Vision and Deep Learning, vol. 13689, p. 1368919, 2025.
5. Y. Man, C. Ding, Z. Zhang, and P. Shu, “Deep learning and computer vision for small residential space layout generation and quantitative evaluation,” Journal of Building Engineering, vol. 126, p. 116144, Elsevier, 2026.
6. S. Albahli, A. Rauf, and M. Algsham, “A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLONAS,” Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp. 1680–1716, MDPI, 2023.
7. G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLOv8,” Version 8.3, Ultralytics Inc., 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
8. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and D. Ommer, “High-resolution image synthesis with latent diffusion models,” in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 10684–10695, 2022.
9. P. Kán and H. Kaufmann, “Automatic furniture arrangement using greedy cost minimization,” in IEEE Conf. on Virtual Reality and 3D User Interfaces (IEEE VR), pp. 317–324, 2018.
10. C. Xiang, R. Bao, B. Feng, W. Wu, Z. Liu, Y. Guan, and L. Liu, “Co-Layout: LLM-driven co-optimization for interior layout,” in Proc. AAAI Conf. on Artificial Intelligence, 2026.
11. S. Chen, J. Gao, Z. Liu, and K. Zhang, “Residential floor plan recognition and reconstruction,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 8, pp. 2789–2802, 2021.
12. G. Bradski, “The OpenCV library,” Dr. Dobb’s Journal of Software Tools, vol. 25, no. 11, pp. 120–125, 2000.
13. Y. Miao, F. Qingnan, Y. Wang, Y. Duan, and C. Baoquan, “Graph neural network for generative furniture arrangement,”

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Journal of Computer-Aided Design & Computer Graphics, vol. 33, no. 3, pp. 457–464, 2021.
14. N. Vitsas, G. Papaioannou, A. Gkaravelis, and A. A. Vasilakis, "Illumination-guided furniture layout optimization," Computers & Graphics, vol. 91, pp. 130–141, Elsevier, 2020.
15. Y. Yue, T. Kontogianni, K. Schindler, and F. Engelmann, "Connecting the dots: Floorplan reconstruction using two-level queries," in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 15042–15051, 2023.
16. R. Cabello et al., "Three.js: JavaScript 3D library," GitHub Repository, 2023. [Online]. Available: https://threejs.org/
17. H. Izadinia, Q. Shan, and S. M. Seitz, "IM2CAD: Fast and robust 3D CAD model retrieval for interior design," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 5134–5143, 2017.
18. S. Zhou, D. Kulkarni, K. Sheng, and J. Smith, "SceneGraph to CAD: Generating semantic CAD models from graph representations," Computer-Aided Design, vol. 115, pp. 102–114, 2019.
19. A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang, "Matterport3D: Learning from RGB-D data in indoor environments," International Conference on 3D Vision (3DV), pp. 667–676, 2017.
20. L. Zheng, H. Zhu, and Y. Guo, "Deep learning for interior layout generation: A survey," ACM Computing Surveys, vol. 53, no. 6, pp. 1–35, 2020.
21. L. Zhang, A. Rao, and M. Agrawala, "Adding conditional control to text-to-image diffusion models," in Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV), pp. 3836–3847, 2023.



INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details