

Leveraging Machine Learning Techniques in Converting 2D Floorplans Images to 3D Models: Applications, Challenges, and Future Directions

I. Bazie, B. Nzazi, J. Biaba, T. Tashev, W. Kambale, K.
Kyamakya, N. Kasoro, S. Kasereka

*The 16th International Conference on Emerging Ubiquitous Systems
and Pervasive Networks (EUSPN 2025) October 28-30, 2025,
Istanbul, Türkiye*



Outline

Introduction

Research Methodology

Machine Learning Techniques

Applications

Challenges and Limitations

Comparative Analysis

Future Research Directions

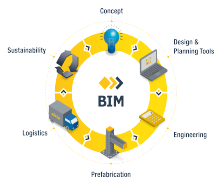
Conclusion

Context & Motivation

- ▶ Growing demand for automation in architecture and real estate
- ▶ Traditional CAD approaches lack **scalability and adaptability**
 - ▶ **Labor-intensive:** Manual reconstruction requires expert knowledge.
 - ▶ **Limited scalability:** Cannot handle diverse architectural styles.
 - ▶ **Static workflow:** Difficult to integrate with dynamic applications.

Critical Applications

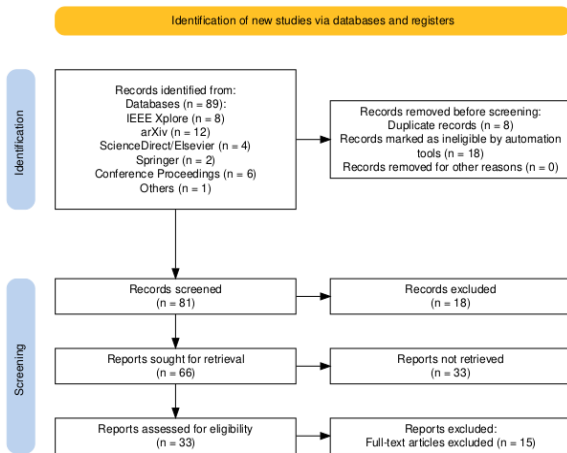
- ▶ Real estate visualization for interactive property tours
 - ▶ **BIM Integration:** Automated semantic 3D model generation
 - ▶ **VR/AR Experiences:** Immersive architectural visualization
 - ▶ **Digital Twins:** Smart city and building management
- ▶ ML techniques enable intelligent automation
- ▶ Computational challenges remain significant barriers



Objectives

- ▶ Synthesize recent **machine learning approaches** for 2D-to-3D floorplan conversion
- ▶ Identify core ML techniques used in **real-world applications**
- ▶ Evaluate commonly used **benchmark datasets and metrics**
- ▶ Discuss major challenges: input ambiguity, architectural diversity, computational limitations
- ▶ Propose future research directions: **diffusion models, foundation models, neuro-symbolic AI**

PRISMA 2020 Methodology



Search Strategy & Selection

Databases Queried

- ▶ **Primary sources:** Google Scholar, IEEE Xplore, ScienceDirect, Springer, MDPI, arXiv
- ▶ **Additional repositories:** ResearchGate, TheCVF Open Access
- ▶ **Keywords:** "2D floorplan to 3D model", "neural rendering", "architectural modeling"

Selection Results

Initial corpus: 89 records → **Final selection:** 33 studies

- ▶ 22 novel ML architectures for floorplan-based 3D modeling
- ▶ 8 comparative evaluations

ML Approaches Overview

CNN-Based

- ▶ Spatial feature extraction
- ▶ **3DPlanNet**: 95.3% wall accuracy
- ▶ **ResNet+FPN**: 0.81 IoU

GAN-Based

- ▶ Realistic texture synthesis
- ▶ **Plan2Scene**: Superior realism
- ▶ **Pix2Vox++**: 0.84 IoU, 7× faster

CNN and GAN Based ML Approached Visual Results

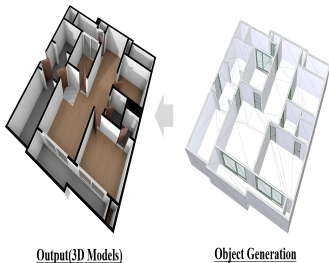


Figure: CNN-Based (3DPlanNet):
 Ensemble approach achieving 95.3% wall detection accuracy with semantic segmentation

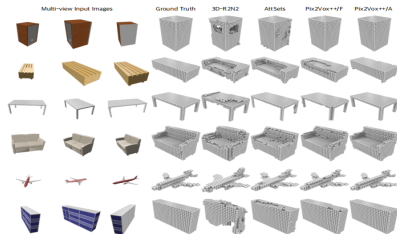


Fig. 7 Example of multi-view 3D object reconstruction on ShapeNet at 32^3 resolution.

Figure: GAN-Based (Pix2Vox++):
 Graph Neural Network integration with 0.84 IoU, $7\times$ faster

Transformer & Hybrid Approaches

Transformers

- ▶ Global context modeling
- ▶ **Rectified Flow**: 68.5-80% geometry win rates
- ▶ Challenge: High computational overhead

Hybrid Methods

- ▶ **3DPlanNet**: CNN ensemble
→ 110K+ models
- ▶ **I-Design**: LLM + CLIP →
0% out-of-bound vs 57.6%
baseline
- ▶ Flexibility but inherit
component limitations

Transformer & Hybrid Approaches Visual Results



Figure: I-design: Personalized LLM interior designer for 3D room layouts from natural language

Real Estate & BIM Applications

Real Estate Visualization

- ▶ Interactive property showcases
- ▶ **Plan2Scene**: Superior realism in user studies
- ▶ **3DPlanNet**: 95.3% wall accuracy at scale

Building Information Modeling

- ▶ IFC-compatible 3D BIM generation
- ▶ **ResNet + FPN** framework
- ▶ State-of-the-art IoU: 0.81 on CubiCasa5k

Impact

Enabling automated creation of semantic 3D models for architectural workflows and digital twin integration

VR/AR & Urban Applications

Virtual & Augmented Reality

- ▶ Immersive architectural experiences
- ▶ **U²-Net**: 15s reconstruction time
- ▶ Accuracy: 76.7% for complex layouts

Urban Modeling

- ▶ 3D town synthesis from single images
- ▶ **Rectified flow transformers**
- ▶ 92% geometry win rates

Limitation

Computational intensity remains barrier for real-time applications

VR/AR & Urban Applications Visual Results

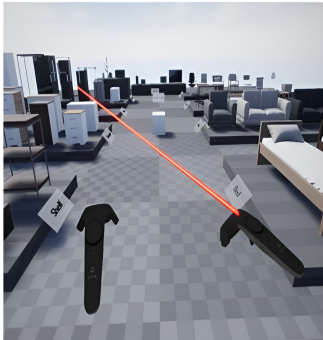


Figure: VR application: VR immersive architectural walkthrough from a single image



Figure: 3D town generation: 3D urban model from single image

Data-Related Challenges

Dataset Limitations

- ▶ **Scarcity:** Limited large-scale paired 2D-3D datasets
- ▶ **Inconsistency:** Proprietary datasets with varying annotation standards
- ▶ **Diversity:** Poor generalization to non-Western floorplans

Specific Examples

- ▶ **Japanese style:** Flexible tatami-mat divisions challenge wall-detection
- ▶ **African style:** Circular/polygonal rooms contradict rectangular assumptions

Technical & Deployment Challenges

Model Limitations

- ▶ **CNNs:** Struggle with long-range dependencies
- ▶ **GANs:** Training instabilities, mode collapse
- ▶ **Transformers:** High computational overhead

Deployment Barriers

- ▶ **Computational:** Multi-GPU setups required
- ▶ **Manual intervention:** 3DPlanNet needs corrections
- ▶ **Integration:** ML-rendering engine challenges

Impact

These challenges limit real-time applications and widespread commercial deployment

State-of-the-Art Performance

Method	Approach	Accuracy	Speed	Limitations
Plan2Scene	GAN + GNN	Superior realism	-	Semantic errors, limited surfaces
3DPlanNet	CNN ensemble	95.3% (walls)	Fast	Manual correction needed
U ² -Net VR	CNN segmentation	76.7% (complex)	15s	Limited furniture types
Rectified Flow	Transformer	68.5-80% win	-	Duplicated facades
I-Design	LLM + CLIP	0% out-of-bound	-	Dense scene failures
DiffDreamer	Diffusion	PSNR: 23.56	Slow	Limited content diversity
Pix2Vox++	Multi-scale CNN	0.84 IoU	7× faster	Memory intensive

Table: Performance comparison of ML approaches

Key Insight

Trade-offs exist between accuracy, speed, and computational requirements

Emerging Technologies

Diffusion Models for 3D Generation

- ▶ DiffDreamer: Conditional diffusion for scene extrapolation
- ▶ Challenge: Computational intensity for real-time applications
- ▶ Need: Lightweight architectures for practical deployment

Foundation Models & Vision-Language Integration

- ▶ Florence-2, GPT-4V: Spatial + semantic reasoning
- ▶ I-Design: Multi-agent LLM for conversational design
- ▶ Potential: Revolutionary architectural workflow interfaces

Advanced AI Approaches

Few-Shot Learning

- ▶ Reduce reliance on large annotated datasets
- ▶ Enable adaptation to new floorplan styles with minimal supervision
- ▶ Critical for regions with scarce architectural datasets

Neuro-Symbolic Approaches

- ▶ Integrate building code compliance with generative models
- ▶ Automatic redesign when regulatory violations detected
- ▶ Combine neural flexibility with symbolic reasoning

Key Findings

- ▶ **CNNs** excel in structural feature extraction and semantic segmentation
- ▶ **GANs** produce high-fidelity textures and volumetric reconstructions
- ▶ **Transformers** enhance global context modeling for complex layouts
- ▶ **Hybrid methods** offer flexibility but inherit component limitations

Persistent Challenges

Limited annotated datasets, computational complexity, semantic consistency across architectural styles

Future Outlook






Promising Directions

- ▶ **Diffusion models:** Consistent 3D scene generation
- ▶ **Foundation models:** Conversational design interfaces
- ▶ **Neuro-symbolic AI:** Building code compliance integration
- ▶ **Few-shot learning:** Data-efficient reconstruction
- ▶ **Explainable AI:** Trust and validation for architects

Impact Potential

Enable robust, scalable, user-centric 3D reconstruction pipelines for widespread adoption in real estate, VR, and smart urban planning

Some References

-  Vidanapathirana, M. et al., "Plan2Scene: Converting Floorplans to 3D Scenes," *IEEE/CVF CVPR*, 2021.
-  Park, S. and Kim, H., "3DPlanNet: Generating 3D Models from 2D Floor Plan Images," *Electronics*, vol. 10, no. 22, 2021.
-  Barreiro, A.C. et al., "Automatic Reconstruction of Semantic 3D Models from 2D Floor Plans," *MVA*, 2023.
-  Zheng, K. et al., "Constructing a 3D Town from a Single Image," *arXiv:2505.15765*, 2025.
-  Çelen, A. et al., "I-Design: Personalized LLM Interior Designer," *arXiv:2404.02838*, 2024.

End

Thank you for your attention!

Follow me on LinkedIn by scanning this QR code.



selain.kasereka@unikin.ac.cd & selain.kasereka@tu-sofia.bg