A State-of-Art Survey on Generative AI Techniques for Floor Planning

ANKUR JAIN, VIT Bhopal University, India
BINDUPAUTRA JYOTIBRAT, VIT Bhopal University, India
ARUNIM GOGOI, VIT Bhopal University, India
RANA TALUKDAR, VIT Bhopal University, India
AVINASH KUSHWAHA, VIT Bhopal University, India
NAVJOT SINGH GILL, VIT Bhopal University, India

In the recent years the integration of Generative Artificial Intelligence (GenAI) has reshaped the concept of the architectural design, particularly in floor planning domain. This study provides an in-depth review of GenAI techniques, highlighting their applications, evolution, and challenges. The paper highlights advanced generative models that can auto-design and enable one to have real-time customization to satisfy user-specific requirements, including GANs, Vision Transformers, and diffusion models. The paper identifies critical research gaps, such as the lack of multi-objective optimization, dataset limitations, and the need for greater adaptability to cultural and contextual constraints. By analyzing contributions from 2020 to 2024, this review offers valuable insights in order to help practitioners, researchers, and industries apply GenAI for advanced floor planning and eventually open up sustainable, inclusive, and effective architectural solutions. This paper draws together an overall table capturing the variety of technologies and strategies employed in the considered studies. Such technologies encompass cutting-edge tools and frameworks like diffusion models, BERT, dual attention mechanism, among others, as well as the architectures used in deep learning. This paper further critiques the datasets upon which all these studies have been conducted to train their respective models. These datasets, among the most popular ones like ECOBEE dataset, the Place Pulse 2.0 dataset, the Cityscapes dataset, and the HuMob Task 1 and 2 datasets, have been of utmost importance in supporting behavioral and environmental research.

Additional Key Words and Phrases: Generative AI, Floor Planning, Architectural Design, Generative Adversarial Networks (GANs), Vision Transformers (ViTs), Sustainability, Real-Time Customization, Urban Planning

ACM Reference Format:
Ankur Jain, Bindupautra Jyotibrat, Arunim Gogoi, Rana Talukdar, Avinash Kushwaha, and Navjot Singh Gill. 2025. A State-of-Art Survey on Generative AI Techniques for Floor Planning. In *Proceedings of* . ACM, New York, NY, USA, 23 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In architectural design floor planning is very significant in as it aids in the arrangement of spaces to achieve structural, functional, and aesthetic objectives. Floor planning ensures that an interior design of a structure or the room arrangements are rationally planned to satisfy the functional requirements of the client. Structural elements such as

Authors' addresses: Ankur Jain, VIT Bhopal University, Bhopal, India, ankur.jain@vitbhopal.ac.in; Bindupautra Jyotibrat, VIT Bhopal University, Bhopal, India, bindupautra.23bai10963@vitbhopal.ac.in; Arunim Gogoi, VIT Bhopal University, Bhopal, India, arunim.23bai11381@vitbhopal.ac.in; Rana Talukdar, VIT Bhopal University, Bhopal, India, rana.23bai10186@vitbhopal.ac.in; Avinash Kushwaha, VIT Bhopal University, Bhopal, India, avinash.23bai10237@vitbhopal.ac.in; Navjot Singh Gill, VIT Bhopal University, Bhopal, India, navjot.23bai11355@vitbhopal.ac.in.

load-bearing walls and columns are also considered to assure stability in the building. An intelligently designed floor plan makes the most of available space, especially in densely populated cities where land is scarce.

Although traditional methods are so effective, they are not only labor-intensive and time consuming but also less flexible and more inclined to human error. Flexibility and real-time responses are the characteristics lacking with these methods. The following are some of the common defects found in floor plans as found in surveys conducted by Builder magazine: floor plans poorly integrating the placement of electric outlets or door swing direction. As many as 56% of the respondents are frustrated due to insufficiently located outlets or poorly sited ones, and 32% were of the opinion that door swing angles were wrong. The Lean Construction Principles show that the conventional methods often end up producing inefficient resource usage since these do not factor in the spatial optimization or dynamic requirements.

The ever-increasing complexity and size of architectural projects, along with the lack of innovation in the conventional floor plan, is compelling architects to call for more intelligent and effective solutions. AI bridges the gap between automated design processes and customer requirements. AI opens the door for architectural innovation by enabling complex algorithms that can comprehend functional and spatial restrictions, thereby enhancing design accessibility and accuracy. The purpose of this paper is to give practitioners, developers, companies, and researchers useful information on the state of AI integrated floor planning now as well as the likely technologies and algorithms that will open the door for further advancements in this rapidly evolving field. This also reveals new algorithms and technologies that change the way layouts are thought of and created, such as Variational Autoencoders (VAEs), Vision Transformers (ViTs), and Generative Adversarial Networks (GANs).

It has been proven that GenAI can transform floor planning technology, which would better and speed up the processes of architectural design. AI models can take into consideration user requirements, functional needs, and spatial constraints in order to produce layouts that are optimized for practicality and efficiency. In addition, GenAI allows for real-time customization, significantly shortening the design cycle as architects can make incremental changes in response to client feedback. GenAI enhances sustainability in floor layouts by assessing energy efficiency, material optimization, and environmental consequences at the design stage, thereby making it easier for architects to more successfully produce green-certified buildings.

GenAI transforms the way of floor planning that was never heard of before or seen with such efficiency. With layout production automation, GenAI may significantly reduce design time while leaving architects to focus on strategic and innovative projects. Its real-time customization integration allows for highly customized solutions that quickly adapt to the user's preferences and constraints. It can maximize the usage of space and create solutions that are visually beautiful and functional by evaluating room measurements and enhancing connectivity and accessibility. GenAI can also assess material and energy usage at the design stage to encourage environmentally friendly activities and adherence to green construction standards. It can also have access to powerful design tools, allowing smaller businesses and individual designers to use high-caliber floor planning resources.

Although GenAI has the promise of revolutionizing floor layout, thereare significant obstacles to overcome. For instance, high-quality datasets are necessary; otherwise, the system may be prone to bias and errors, especially if the data is skewed or incomplete. The designs that do not exactly satisfy user wants or cultural preferences may be the result of a lack of contextual awareness. The lack of transparency of AI decision-making systems can cause erosion of trust between clients and architects. It struggles in complicated architecture projects, when human skill is still important. High implementation of GenAI systems may hinder the widespread adoption, especially for smaller businesses and individual designs. The process of adoption is further complicated by algorithm biases, system faults, and a lack of security controls.

The use of AI in architectural design has been widely discussed in review papers. Among them, some have approached fully integrated solutions such as "House-GAN" or concentrated on the 3D production of interior spaces and complex geometrical facades. The contributions made as reviews or survey studies are summarized and discussed in the subsequent paragraphs, followed by an outline of the gap that is to be covered with research.

Simon Fraser University's research article "House-GAN: Relational Generative Adversarial Networks for Graph-constrained House Layout Generation" [33] is a noteworthy contribution to AI-driven floor planning. This work offers a novel method for using GANs to build house plans that accurately depict and describe intricate spatial interactions between rooms. This is a major step toward automation of floor planning with personal touch by using graph constraints to ensure that the layouts are useful, aesthetically pleasing, and customized to user inputs.

Japan Advanced Institute of Science and Technology carried out a literature review on Generative AI for Architectural Design [23], which discusses the progress that has been made in generative AI and how it is implemented to generate 3D models. The paper will discuss the most recent advancements in state-of-the-art techniques that have transformed architect's thinking and space creation techniques, such as GANs and VAEs. They showed how GANs and VAEs can be used to generate all kinds of 3-D models of facades, such as complex geometries of cubes, designs of churches, and city interiors. Another part of the study tracks the progress of the most prominent large-scale models in computer science with their development and influence throughout various disciplines.

The University of Nebraska-Lincoln, USA, came up with a new technical approach in their article "Generative AI in the Construction Industry: Opportunities & Challenges" [14]. The study provides a hypothetical GenAI model for the construction industry that aims to predict critical variables such as construction prices, delivery methods, project complexity, material quantities, and completion dates. The study also explores the possibilities of advanced large language models, such as GPT, PaLM, and Llama, as well as the potential benefits and challenges of their use in the construction industry.

In the study "Intelligent Layout Generation Based on Deep Generative Models: A Comprehensive Survey" [#shi2023intelligent], the existing state-of-the-art techniques in layout generation were examined in detail, focusing in particular on deep generative models. The paper attempts to explore the recent advances and challenges that are present in the field of layout generation. The study also stresses the importance of qualitative and quantitative approaches toward dataset selection and evaluation along with their relations to particular layout generation tasks. As a final note on furthering the inclusion of AI into layout design, the study concludes by pointing out existing difficulties in intelligent layout generation and providing possible routes for further research.

New developments such as machine learning models and generative design algorithms enable systems to produce excellent floor plans that meet functional needs, improve space utilization, and match user preferences. AI-driven solutions are essential because of their potential in both residential and commercial applications. GenAI optimizes and automates the design process with the latest models. Models like "House-GAN," which shows it can produce graph-constrained layouts that respect spatial hierarchy and pragmatic constraints like zoning laws and architectural specifications, demonstrate this capacity.

We are aware that although much research has been conducted on the use of generative AI in 3D modeling and urban planning, little of it has been specifically done on the use of GenAI in floor planning. Current research often focuses on the functional and aesthetic aspects of design but fails to fully address multi-objective optimization in floor planning, such as balancing sustainability, usability, and space utilization. The adaptability of generative models to various cultural, legal, and contextual contexts in floor planning is rarely explored in the literature currently in publication. A significant barrier to AI uptake is the lack of appropriately diverse, high-quality, and specially designed datasets regarding floor planning. In fact, most research studies test neither the limitation of the dataset nor offer solutions on dealing with it. How gen AI techniques can be scaled up to larger projects such as multi-building layouts or whole urban neighborhoods is something hardly discussed, despite floor plan discussion being very common in such circles.

With this survey paper, we are hoping to investigate floor planning as a separate application domain by combining the technologies, approaches, and algorithms that have been especially designed for this use. We also point out that generative models, such as GeoFormer, can optimize several goals simultaneously, providing important insights into how to achieve complete design objectives. The customization capabilities of generative AI tools and an emphasis on real-time interaction help to clarify how these technologies can be made more dynamic and user-centric. Moreover, we check how GenAI systems can be modified to address a wide range of factors, to guarantee their applicability and relevance in different geographical areas. This paper focuses on guiding future work related to developing datasets by taking account of dataset needs, availability, and ways to overcome problems such as bias data or scarcity. By filling the gap between single-building applications and extensive urban planning, we expect to make generative AI models an essential tool for future developments in urban and architectural design.

In conclusion, the above discussion has included a few chosen survey studies and systematic literature reviews (SLRs) that highlight significant contributions in a variety of interesting themes. The available material was thoroughly arranged and categorized in this article using the many viewpoints and perspectives covered in the Materials and Methods section.

The rest of the paper is organized in the following manner; the "Materials and Methods" section of Section 2 describes the methodology for reviewing the body of existing literature. The explanation of seven study topics is examined in Section 3, "Findings and Discussion." A thorough summary of the paper's findings is given in Section 4, "Meta-Analysis," which also discusses possible directions for further research in Section 5, "Open Research Questions." Lastly, the paper's sources are provided at the end, and Section 6 concludes the work.

2 MATERIALS AND METHODS

This study proposes the objectives of this paper and systematic literature review on generative floor planning systems and deep-learning approaches.

Objectives

- To review and analyze state-of-the-art generative AI models applied to smart city planning and architectural design.
- To evaluate how generative AI techniques enhance efficiency, resource management, and sustainability in smart cities.
- To identify challenges faced by generative AI models in practical urban planning scenarios, such as integrating real-time data, addressing scalability, and maintaining human oversight.
- To propose new areas for integrating generative AI with IoT systems, public data, and environmental planning.
- To highlight ethical considerations and regulations necessary for the responsible deployment of AI in urban planning and design.

Table 1. Criteria for Inclusion and Exclusion of Papers.

Inclusion criteria	Exclusion criteria
The search term, which is written exclusively in English, was found in the study's title, abstract, or keywords	Keynotes, blogs, and inadequate reference materials like dictionaries, thesaurus, and Wikipedia
Research published in books, conferences, and journals between 2020 and 2024	Duplicate studies, i.e., studies that are published by more than one publisher

Search Strategy

This section discusses the search strategy for searching and mapping the relevant literature. A systematic approach was adopted to conduct this survey on Generative AI (GenAI) applications in urban floor planning. The project began with identifying the research focus, which involved exploring advancements in GenAI for these domains. Google Scholar, Science Direct, IEEE Xplore and ArXiv are among the various websites that we scanned through to find the relevant literature on the topic. To guide through the literature search, keywords such as "GenAI", "Urban Planning," "Floor Planning," and others were employed. Next, we examined the role of various transformer models, including Vision Transformers, Generative Pre-trained Transformers (GPT), Stable Diffusion, and other related technologies.

We searched for relevant research papers, surveys, reviews, articles, and journals published between 2020 and 2024, focusing on the intersection of these transformers and our identified keywords. The keywords used in these papers were catalogued and organized into a structured format using spreadsheets. Each paper was thoroughly analysed to identify the technologies and datasets employed, and these details were further compiled into separate datasets for ease of reference.

This systematic data collection and analysis provided the foundation for this survey paper, which aims to comprehensively review the role of Generative AI in urban and floor planning while highlighting the trends, advancements, and potential future directions in the field.

The criteria for inclusion/exclusion of publication are defined in Table 1. The literature has been tabulated, analysed, and mapped based on criteria defined in Table 1.

Time Frame and Digital Repositories

As indicated in Table 4, the time frame for looking through the pertinent literature is 2020–2024 (both years included). AI has made amazing strides and is now being used in many other fields, including architecture. It has been very helpful in automating processes like facade design, floor planning, weather impact analysis, and many more. Since then, a sizable body of literature has emerged, which this study maps. To search the literature, we have chosen a number of research areas, including Google Scholar, IEEE Xplore, ScienceDirect, Arxiv, and many more.These repositories were chosen because they offer pertinent publications, analytics, and outcomes.

Theoretical Framework and Initial Results

A list of the strings we used to search and map the literature is provided in Table 3. Several web search engines (described above) were used to search the search strings. The chosen digital repositories were searched using the search terms listed in Table 3. The outcomes are listed in Table 3. The publications fall into two categories: conferences and journal papers. Only esteemed conferences—those sponsored by ACM, IEEE, or Springer—are taken into account. The Figure **??** displays the ratio. Likewise, Figure **??** displays the chosen paper's frequency by year. Papers from 2020 to 2024 were

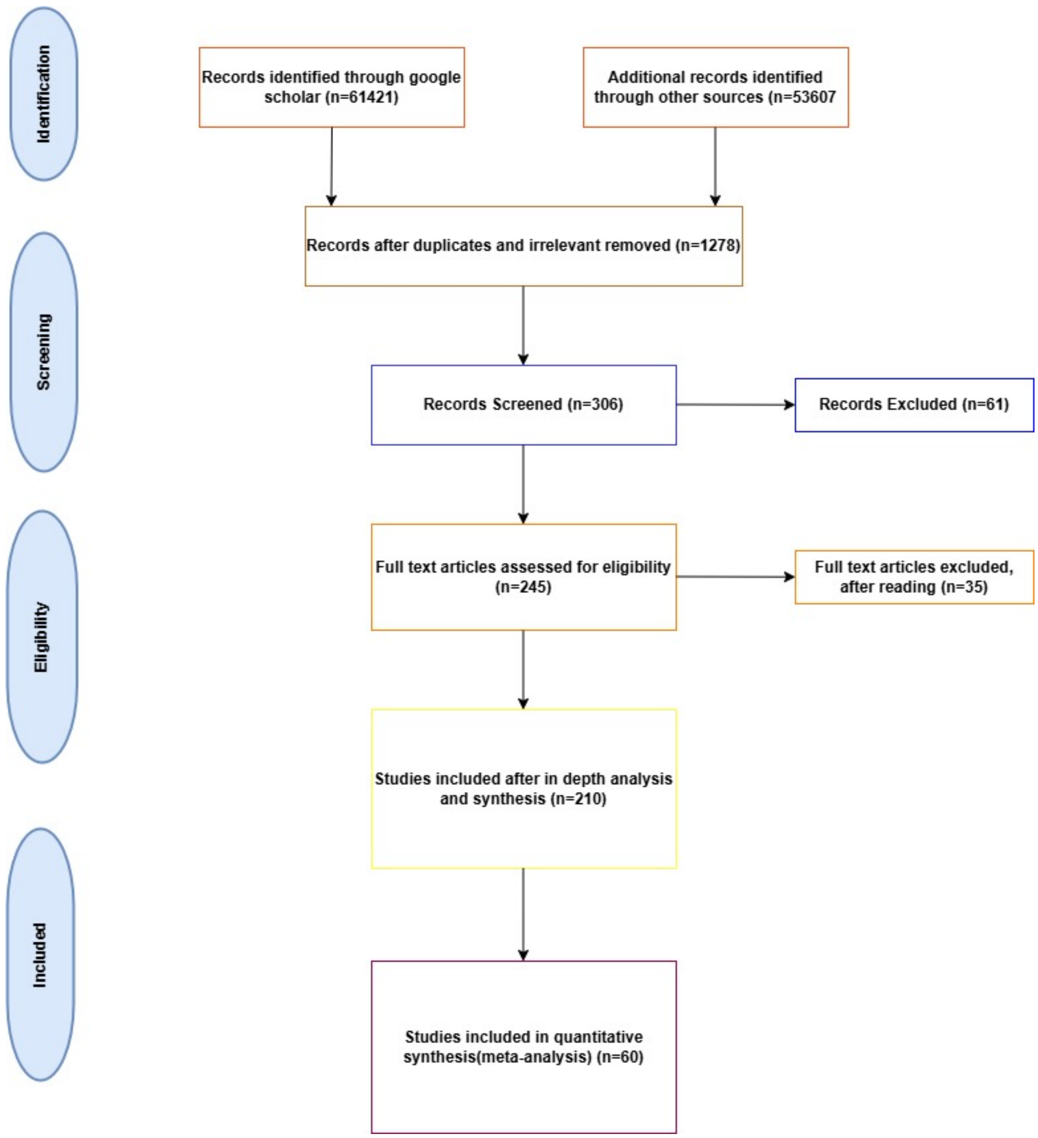

Fig. 1. The identification process of primary studies

chosen. Publications on AI-driven architectural systems have grown steadily. Table 4 presents the summary (most relevant papers) of the publications along with years, publishers and the type of the reviewed paper. We have tried to select only well-reputed journals and conferences.

3 FINDING AND DISCUSSIONS

This section addresses the key research questions (RQs) surrounding the application of transformer models in architectural floor planning systems. The motive of raising the questions is discussed in Table 2. It also provides an exhaustive review of selected publications from a pool of studies, focusing on the methodologies employed, datasets utilized, and evaluation metrics applied. Dedicated subsections delve into the findings, research gaps, and potential future directions for advancing floor planning through transformer-based approaches.

RQ1: What Is the Current Status of using GenAI in Floor Planning?

Transformer-based models have emerged as a potential tool in the architectural arena due to their ability to handle intricate spatial and multimodal data, thereby fully changing the scenario of floor planning. The adaptability and durability of these models in everything ranging from conceptual design to intricate layout optimization has proven to be significant. With such sophisticated frameworks being applied to numerous creative, analytical, and simulation tasks, it has seen the incorporation of Vision Transformers (ViTs) and Generative Pre-trained Transformers (GPTs).

Transformer-based systems gradually but significantly replaced traditional architectural techniques. The high usage of the models in tasks such as preliminary layout generation, where speed and flexibility are key, is an example of studies such as Generative AI for Architectural Design. Moreover, tools such as ComfortGPT and GeoFormer demonstrate how flexible these models are in handling matters of energy efficiency and city issues, which therefore means that they are crucial to use in modern architectural practice.

According to a report, more than 70% architects were looking at AI tools as architects to make their design process more efficient. Generative AI has significantly impacted architectural workflows. The Generative Design tool set provided by Autodesk is widely used for optimum layouts in any design. Real-life cases, such as The Edge in Amsterdam, showcase AI-based systems that can help create energy efficiency and maximize the use of space. Artificial intelligence systems used in designing facades make them responsive to climatic conditions. This is achieved by increasing energy efficiency through automated shading and regulating thermal effects. However, more challenges include high implementation costs, which limit training resources, thus having dataset biases that prevent this full realization of the real potential of GenAI in the field of floor planning.

Table 2. Research questions

RQ#	RESERCH QUESTION	MOTIVATION
RQ1	What Is the Current Status of Transformer-Based Systems in Floor Planning?	To study and map the current state of transformer-based systems in floor planning
RQ2	How Are Machine Learning and Transformer Models Used for Generating and Interpreting Floor Plans?	To understand how machine Learning and Transformer Models can be used for Generating and Interpreting Floor Plans
RQ3	What Types of Datasets Are Used for Floor Planning Systems?	To specify different datasets used for Floor Planning Systems
RQ4	What Are the Popular Approaches for Recognizing and Optimizing Floor Plans?	To understand the various approaches that can be taken for recognizing and optimizing floor plans
RQ5	Which Aspects of Floor Planning Are Targeted?	To understand which aspects of floor planning are the highest priority
RQ6	What Evaluation Metrics Are Used in Floor Planning Systems?	To understand what evaluation metrics Are used in Floor Planning Systems
RQ7	Which Models Have Demonstrated Better Performance for Specific Floor Planning Scenarios?	To specify what models have performed better for certain floor planning scenarios

RQ2: How Are Machine Learning and Transformer Models Used for Generating and Interpreting Floor Plans?

Machine learning (ML) and transformer models have now become indispensable tools in producing and interpreting floor plans, giving innovations in realism, efficiency, and adherence to architectural principles. These models perform two critical roles: generation and interpretation.

In generation, transformer-based models like Generative Pre-trained Transformers (GPTs) and diffusion models are used to generate optimized layouts. Studies like From Text to Transformation: Comprehensive Review of Large Language Models' Versatility and Generative AI for Visualizations: State of the Art and Future Directions demonstrate how these systems can transform textual descriptions or constraints into fully realized floor plans. It is through such models that layout generation can be automated, but in ways that take account of user-specific needs, spatial constraints, and functional requirements, all of which make them precious for dynamic design processes.

To interpret, models like ViTs are used to study and extract structural details from given layouts. For example, Improving Facade Parsing with Vision Transformers discusses the use of them in finding load-bearing walls, window placement with maximum space and utility pathway, and the Detection and Classification of Surface Defects on Hot-Rolled Steel Using Vision Transformers describes its use for detecting anomalies or defects. Transformers, in their forms of physical models, will therefore permit spatiotemporal simulation, as, for instance, in Integration of an Improved Transformer with Physical Models for Urban Flooding Depths. These capabilities are valuable for architects so far as spatial arrangements and structural integrity will have concerned them.

RQ3: What Types of Datasets Are Used for Floor Planning Systems?

Floor planning software comprises several datasets to accomplish tasks like architectural designs and layout schemes in cities. Crucial indoor building and architectural datasets include the Annotated Boston Floorplans, CubiCasa5K, RPLAN, Lifull Dataset, and Volumetric Design Database, which can be sampled for detailed floor plan annotations or full-fledged layouts.Then, large-scale building layouts and geographical data are concentrated from datasets of aerial and satellite pictures: the Massachusetts Building Dataset, Satellite II Dataset, and WHU Aerial Building Dataset. Datasets such as Semantic Segmentation Maps for Facade, Point Cloud Representations, and 2D and 3D Voxel Models are required for volumetric and structural representation in terms of 3D modeling. Sketch and conceptual design data, such as Sketches and Conceptual Diagrams and Architectural Floor Plans and Layouts, support pre-design stages of floor planning. There are Urban Planning Data, Building and Infrastructure Data, and Transportation and Mobility Data available online that enable integrated architecture with higher infrastructural needs. Furthermore, synthetic datasets, such as Neural Radiance Fields (NeRF) and Signed/Unsigned Distance Functions (SDF and UDF), resolve realistic 3D experience and shape modeling. These datasets together improve the accuracy and efficiency of floor planning systems.

RQ4: What Are the Popular Approaches for Recognizing and Optimizing Floor Plans?

Orientation and optimization of floor layouts are their prime concern in any floor planning system. It projects two types of research, namely, "An Illumination-Guided Dual Attention Vision Transformer for Low-Light Image Enhancement" and "Advanced Post Earthquake Building Damage Assessment", which underline and optimize the most commonly used. It would also refer the current application of SAR Coherence Time Matrix with Vision Transformer, which aims to structural details and spatial patterns detection in floor plans. Studies like "Enhancement of Road Traffic Flow in Sustainable Cities through Transformer Models" can usefully utilize this result in striking a balance among multiple objectives-based consideration such as space utilization, aesthetic appeal, and functionality or "Integration of an Improved Transformer with Physical Models for the Spatiotemporal Simulation of Urban Flooding Depths" to regulatory standards, material costs, or energy efficiency. These ensure compliance and sustainability. All these approaches will offer complete solutions to architectural problems.

RQ5: Which Aspects of Floor Planning Are Targeted?

Table 3. Studies found in the selected repositories.

String	Digital Repository	Studies Found	Studies Selected
DALL-E AND (Architecture OR "Urban Design" OR Residential OR "Design Architecture")	IEEE Xplorer	4	1
	Springer Nature	1838	3
	Total	1842	4
BERT AND (Infrastructure OR Urban OR "Urban Regeneration" OR "Transportation Planning")	ScienceDirect	1418	1
	Taylor & Francis	2135	1
	PLOS One	2092	1
	Google Scholar	226	1
	IEEE Xplorer	80	1
	Total	5951	5
Vision Transformer AND ("Generative AI" OR "Urban Planning" OR ("Generative AI" AND "Urban Planning"))	ScienceDirect	2112	5
	Total	2112	5
CLIP (Contrastive Language–Image Pretraining) AND ("Urban Planning" OR Transformer)	Google Scholar	25	1
	ScienceDirect	653	3
	arXiv	13700	1
	Total	14378	5
Text-to-Text Transfer Transformer AND ("Generative AI" OR "Urban Planning" OR ("Generative AI" AND "Urban Planning"))	IEEE Xplorer	40	1
	ScienceDirect	6943	1
	arXiv	153	1
	Springer Nature	22065	1
	MDPI	52	1
	Total	29253	5
Bootstrapping Language-Image Pre-Training AND ("Generative AI" OR "Urban Planning" OR ("Generative AI" AND "Urban Planning"))	arXiv	4	1
	ScienceDirect	163	2
	Springer Nature	205	1
	Google Scholar	1170	1
	Total	1542	5
Generative Pre-trained Transformers AND ("Generative AI" OR "Urban Planning" OR "Floor Planning" OR ("Generative AI" AND "Urban Planning"))	ScienceDirect	410	2
	Google Scholar	15400	3
	Total	15810	5
Generative AI AND ("Architectural Design" OR "Building Layout" OR "Floorplan Generation" OR "Deep Generative Models" OR "Urban Digital Twins" OR "Construction Industry" OR "Creative Industries" OR "AI in Architecture")	Frontiers Media	142	1
	arXiv	27	2
	MDPI	38	2
	ScienceDirect	1918	3
	Sciendo	3117	1
	Google Scholar	17300	1
	Total	22542	10
Urban Planning Software AND ("Free and Open Source" OR "3D Visualization" OR "Computer-Aided Design" OR "Environmental Planning" OR "Integrated Planning and Design")	ScienceDirect	2368	3
	Total	2368	3

When it comes to making an ideal floor plan there are various aspects that need to be taken consideration, however, space utilization, structural integrity and aesthetics are the targeted aspects to make a floor planning system that can generate ideal floor plans according to the user's need.

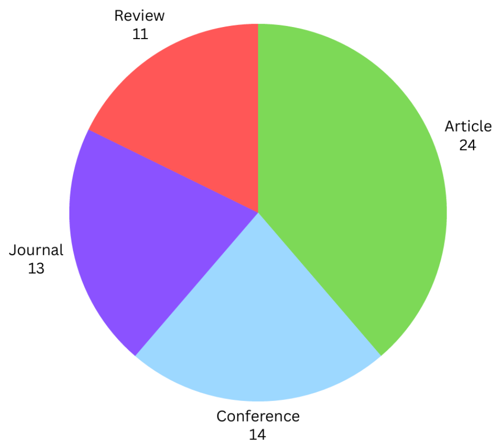

Fig. 2. Studies published in article, conferences, journals and review

- Space Utilization: In the designing of a floor plan proper utilization of available space has to be done so that space can be allocated for different purposes and wastage of space should not be done. For achieving this in the floor planning system studies like DALLE-URBAN: Capturing the urban design expertise of large text to image transformers.
- Structural integrity: It has to be ensured that the layout obtained can be constructed in the confines of the material and the engineering principle. It has to take into consideration factors such as load bearing walls, energy efficiency, column placement and so on. Achieving this in the floor planning system would require studies such as Generative AI model trained by molecular dynamics for rapid mechanical design of architected graphene and so on.
- Aesthetic appeal: Aesthetic appeal is the aspect that needs to be taken care of very carefully, as this is a factor where user always prefer with functionality, here the system has to design spaces that are pleasing to the eyes and also harmonious. To do so in the floor planning system studies like Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding can help us do so.

Identifying these aspects in floor planning ensures that the entire process of creating a space is holistic and the result is functional and aesthetically pleasing, as well as sustainable, accessible, and safe.

RQ6: What Evaluation Metrics Are Used in Floor Planning Systems?

Evaluation measures are essential in evaluating the performance of models to be used for floor layout. Quantitative metrics are used to evaluate measure of efficiency concerning wonderfully accurate design generation, structural coherence for uniformity of space arrangements in compliance with architectural principles, and energy efficiency for measuring energy consumption savings as gained through the optimal layouts. It could be found that qualitative measures can be utilized to measure satisfaction regarding paradigms that models can best adapt to regarding dealing with style or constraint as well as through the analysis of survey and feedback data pertaining to the spatial design functionality or aesthetics. It would be worthy of considering how these qualitative measures integrate with quantitative measures into a comprehensive assessment method of the floor planning system.

RQ7: Which Models Have Demonstrated Better Performance for Specific Floor Planning Scenarios?

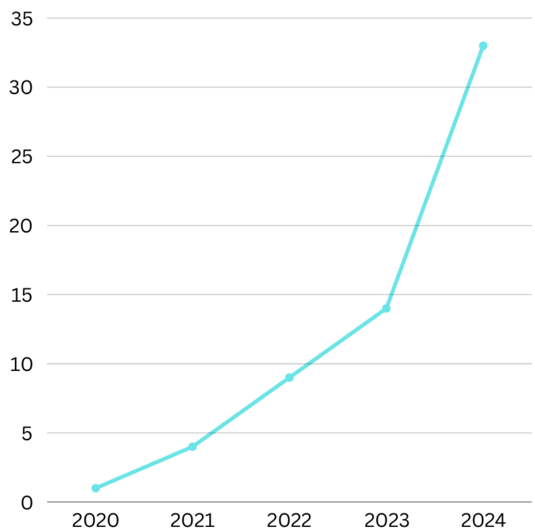

Fig. 3. Number of studies published from 2020 to 2024

It has been revealed by comprehensive research that Vision Transformers and Generative Pre-trained Transformers are models that provide the best results. Vision Transformers are especially helpful in analyzing high-resolution architectural blueprints and floor plans from a global point of view. Improving Facade Parsing with Vision Transformers; An Illumination-Guided Dual Attention Vision Transformer for Low-Light Image Enhancement; and Line Integration and Multi-Scale Knowledge Transfer Vision Transformer for 3D Vessel Shape Segmentation demonstrate nice examples of this. However, we have perfectly demonstrated the usefulness of Generative Pre-trained Transformers models in configuring inventive, user-specific layouts to a very large extent for most preliminary design stages studies" like Generative AI for Architectural Design: A Literature Review , Generative AI for Visualizations: State of the Art and Future Directions or DALL E-URBAN: Capturing the Urban Design Expertise of Large Text-to-Image Transformers ". The performance was further enhanced with a hybrid approach combining various transformers with domain-specific models, hence securing perfect alignment of generated floor plans with structural and environmental as well as regulatory considerations so that they are ideal for sustainability-driven projects.

Table 5. Technologies used in GenAI models

Year	Study	Technology
2020	[16]	Generative Design of Floor Plans
2021	[8]	Diffusion Models for Image Generation
2021	[63]	Reinforcement Learning , Multimodal Learning
2021	[25]	BERT, Dynamic Clustering, Flow Information, Machine Learning
2021	[19]	Generative Adversarial Networks (GANs), machine learning, human–artificial intelligence collaboration

Year	Study	Technology
2022	[22]	BERT, Transformers, Text Mining, Knowledge Extraction
2022	[9]	Artificial Intelligence, Urban Sustainability, Environmental Modeling
2022	[58]	Open source tools, Geographic Information Systems (GIS), Python, API, 3D city models, urban morphology, spatial analysis
2022	[30]	Generative design, building layout, path generation algorithms
2022	[43]	Text-to-Image Systems
2022	[1]	Few-Shot Learning, Multimodal Learning
2022	[44]	DALL-E, Urban Design, Text-to-Image Models
2022	[38]	Dynamic Temporal Graph Network, Machine Learning, Graph Networks
2022	[54]	Automated design methods, floorplan generation, machine learning, AI-based methods
2023	[31]	BERT, GPT, NLP
2023	[15]	Vision-Language Models
2023	[53]	NLP, Computer Vision, Deep Learning Frameworks
2023	[39]	DALL-E, StableDiffusion, Midjourney, Text-to-Image Models, Ideation, Sketching, Image Combination, Inpainting, Outpainting
2023	[7]	Sentiment Knowledge Enhanced Pre-Training, Latent Dirichlet Allocation, Natural Language Processing (NLP), Word Cloud
2023	[46]	Deep Generative Models, Layout Generation, Machine Learning, AI-Based Design Optimization
2023	[47]	GPT, Data Representation
2023	[62]	Text-to-Image Diffusion Models
2023	[10]	Big Data, AI, Machine Learning, Smart City Models
2023	[17]	Computer Vision , Transformer-Based Framework , Pre-trained Vision Model
2023	[32]	Transformer-Based Neural Networks
2023	[52]	Generative Model, Optimization Algorithms
2023	[42]	Transformer-Based Models, Textual Data Application
2023	[36]	Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Design Automation, Building Design Optimization
2024	[6]	Transformer-Based Architecture, Predictive Modeling
2024	[29]	Optimization Algorithms, Data-Driven Modeling
2024	[59]	GAN
2024	[37]	NLP, Computer Vision, Vision-Language Models
2024	[49]	Vision Transformers, Computer Vision
2024	[18]	Vision Transformer (ViT), Swin Transformer, Swin-Unet, Knowledge Distillation, TRSF-Net, VT-UNet, U-Net Transformer Variants.
2024	[55]	Dual Attention Mechanism, Vision Transformer (ViT), Illumination-Guided Network, Frequency-Domain Optimization.

Year	Study	Technology
2024	[51]	Vision Transformer (ViT), Revision-based Transformer Facade Parsing (RTFP), Line Acquisition, Filtering, and Revision (LAFR).
2024	[27]	Contrastive Language-Image Pre-Training (CLIP), Bidirectional Cross-Attention, Multimodal Keyword Spotting
2024	[26]	CLIP-based few-shot learning, Multi-modal prototypes, Vision Transformer (ViT), Transformer models
2024	[35]	Contrastive Language–Image Pretraining (CLIP), Multimodal models, Image quality prediction
2024	[50]	CLIP, Object manipulation, Contrastive learning, Pretraining models
2024	[11]	Image-to-Image Translation, Rectified Flow Transformers
2024	[28]	CLIP, Object detection, Aerial images, Vision Transformer (ViT)
2024	[12]	Text-to-Image Models, Floor Plan Generation
2024	[2]	Deep Learning, Generative Models, AI Design Tools
2024	[5]	Deep Learning Framework
2024	[20]	Transformer Models
2024	[48]	Optimization Algorithms, Spatial Data Analysis
2024	[61]	NLP, Predictive Modeling for Space Utilization
2024	[13]	Transformer Models
2024	[40]	Few-Shot Learnin, Convolutional Neural Networks
2024	[4]	Optimization Algorithms
2024	[57]	Vision Transformer (ViT), Synthetic Aperture Radar (SAR) Coherence, Time Matrix Analysis.
2024	[14]	Generative AI, AI-based optimization, construction design, smart city planning
2024	[3]	GPT
2024	[21]	LLM-based system
2024	[24]	NLP, Space Optimization Techniques
2024	[23]	Generative AI, architecture design, deep learning
2024	[64]	AI in architecture, smart building design, generative AI, urban development
2024	[34]	AI in architecture, smart building design, generative AI, urban development
2024	[56]	Generative AI, 3D modeling, urban data generation, smart city models, digital twins

Table 4. Summary of the included literature.

Year	Publisher	Type	Study
2020	arXiv	Conference	[16]
2021	arXiv	Conference	[8]
	ResearchGate	Conference	[63]
	IEEE	Conference	[25]
	Frontiers Media	Review	[19]
2022	arXiv	Conference	[43], [1]
	Elsevier	Article	[58], [22]
		Review	[54]
	IEEE	Conference	[44]
	MDPI	Article	[9]
	Sciendo	Article	[60]
	Taylor and Francis	Journal	[38]
2023	arXiv	Article	[15]
		Conference	[62]
	ACM	Conference	[47]
	Elsevier	Article	[45]
		Journal	[17]
		Review	[36]
	MDPI	Review	[41]
	PLOS One	Article	[7]
	Springer Nature	Article	[39], [30], [53]
		Conference	[10]
		Journal	[32]
	Wiley	Journal	[52]
2024	arXiv	Conference	[11], [28]
		Review	[21], [24], [23], [56]
	CellPress	Article	[49], [50]
	Elsevier	Article	[6], [29], [18], [55], [51], [27], [26]
		Journal	[20], [48], [13], [40], [57]
	Graphic Horizons	Conference	[12]
	IEEE	Journal	[5]
		Review	[3]
	KeAi	Journal	[61]
	MDPI	Article	[35]
		Journal	[14]
		Review	[34]
	ResearchGate	Article	[37]
		Journal	[4]
	Springer Nature	Conference	[2]
	Visual Informatics	Article	[59]
	We Make Spaces	Review	[64]

Table 6. Datasets used by the various papers

Year	Paper	Dataset	Remarks
2023	[17]	Place Pulse 2.0 Dataset	It includes 110,988 street view images captured between 2007 and 2012, covering 56 cities in 28 countries on six continents.

Year	Paper	Dataset	Remarks
		Cityscapes Dataset	The Cityscapes dataset is a large-scale dataset that contains a diverse set of stereo video sequences recorded in real-world street scenes from 50 different cities. It provides high-quality pixel-level annotations and encompasses a wide range of urban elements, including diverse roads, sidewalks, buildings, vehicles, and various other elements.
2023	[47]	HuMob Task 1 Dataset ("business-as-usual" human mobility)	This dataset focuses on normal human mobility patterns, making it useful for modeling everyday movement behaviors. It reflects regular, predictable mobility trends and serves as a baseline for human trajectory prediction.
		HuMob Task 2 Dataset ("emergency period" human mobility)	The emergency period dataset introduces challenges due to abrupt changes in mobility patterns caused by external events. It tests the model's ability to adapt to irregular, non-routine movement behaviors.
		Validation and Test Dataset Splits	The splits are crucial for robust evaluation, ensuring models generalize to unseen data and maintain reliability across both normal and emergency scenarios.
2024	[6]	ECOBEE dataset	It contains data from over 100k air-conditioned building thermostat users in North America since 2015 with hundreds of parameters collected at 5-min intervals.
2022	[1]	M3W (MultiModal MassiveWeb) Dataset	This dataset is a large collection of interleaved images and text scraped from approximately 43 million webpages. The dataset maps the positions of images relative to the text and formats the extracted text to include placeholders for images.

Year	Paper	Dataset	Remarks
		ALIGN Dataset	This dataset consists of 1.8 billion images paired with alt-text. It is used as a source for image-text pairs to enhance model training.
		LTIP (Long Text & Image Pairs) Dataset	This custom-collected dataset comprises 312 million image-text pairs, focusing on longer and higher-quality descriptions compared to ALIGN.
		VTP (Video & Text Pairs) Dataset	This dataset includes 27 million short videos (22 seconds on average) paired with textual descriptions, providing temporal information for training the Flamingo model.
2024	[28]	DIOR-C and DIOR-Cloudy	Datasets were specifically developed for Source-Free Object Detection (SFOD) tasks, where the target domain differs significantly from the source domain.
2021	[19]	NBA Player Movement, Anime Colorized, Icon Dataset, Chinese Typeface Dataset, 3D Terrain Sketch-to-Terrain Dataset, Text-Based Mario Brothers Level, Fashion Design Dataset	Each dataset is chosen to enhance specific creative applications, such as urban planning, graphic design, game development, and fashion, among others. These datasets are integral to improving the generative capabilities of GANs for collaborative and creative problem-solving.

Table 7. Accuracy

Year	Paper	Models	Results/Performance
2024	[35]	CLIP, Multimodal models and Image quality prediction	73.16%
2024	[28]	CLIP, Object detection, Aerial images, Vision Transformer (ViT)	44.10%
2023	[17]	CV,Transformer-Based Framework and Pre-trained Vision Model	79.89%
2023	[47]	GPT and Data Representation	43.43%
2023	[39]	DALL-E, StableDiffusion and Midjourney	99.90%
2022	[1]	Few-Shot Learning and Multimodal Learning	86.60%
2022	[58]	Open source tools, GIS	52.80%

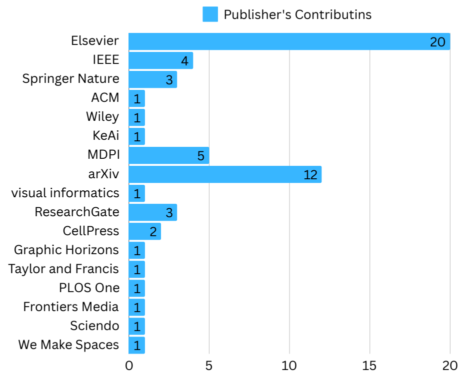

Fig. 4. Publisher's Contribution

4 META ANALYSIS

This section discusses a thorough analysis of the assembled literature, wherein citation analysis, contributions of publishers, and contributions by country is only a few aspects studied. Transformation based fast evolving structures have fundamentally changed several disciplines-here are architeture, environmental simulations, and spatial modelling. Some key contributions of current research on the important aspects for meta-analysis in terms of architectural floor planning systems have been highlighted. In several of their distinct complexities of floor layout and spatial design, all these works manifest the revolutionary potential of deep learning, particularly transformer architectures.

4.1 Applications of Transformers in Architectural and Spatial Design

They have been very good at identifying physical problems using versatile transformers, and UPDExplainer is an example study that uses street-view imagery to understand spatial dynamics. Fine-grain data collection methods that extract information about the urban village from high-resolution satellite images are meant for minuscule spatial segmentation. These will change the initial perspective for improved focus on floor-planning systems integrated with fine-grained spatial information in design. Besides, a design indication shows the importance of floor design-size variables when paired with physical models as well as simulation modeling of environmental impacts, for instance, degrees of flooding. This is how such approaches ensure resiliency and adaptability in architectural layouts design, particularly for regions vulnerable to environmental hazards.

4.2 Advancing Precision in Spatial Modeling

Research on generative pre-trained transformers (GPT) in automated data mining highlights the possible improvement of spatial precision in constructions. Examples of this research include ComfortGPT and its related projects. Such models can either forecast desired environmental setpoints or optimise spatial arrangements according to user preferences and energy efficiency objectives through integrating massive datasets. From here, well-advanced and intelligent floor plan generating systems will meet very different needs-from comfort and usage to occupant energy efficiency. With dynamic human movement data sets being added to floor pattern designs, it also marks a milestone with GeoFormer in predicting human mobility patterns based on GPT frameworks. Combinations between graphical or space layouts and people's behavior patterns can have overwhelming possibilities for improving the effectiveness of space considerably.

4.3 Generative Models for Architectural Innovation

High-resolution images, architectural layouts, can be efficiently synthesized by means of diffusion-based approaches and generative models like DALL-E. This study reveals the potential of text-to-image transformers for architectural systems: the research that confirms the possibility of such transformers in generating conceptual ideas and generating floor plans. In addition, the automated design process allows architects the flexibility to explore and render a vast range of design alternatives at unprecedented speed.

4.4 Multimodal and Vision-Language Models in Architecture

Various multimodal models like Flamingo have been promising in bringing verbal and visual data together in low-resource situations, like under-documented architectural styles or preservation histories. With these models, contextualised floor layouts are created that bear in mind historical and cultural nuances. Similarly, CLIP-guided methods for localizing short-term temporary construction items provide relevant answers for real-time architectural project management.

5 OPEN RESEARCH QUESTIONS

This portion explores the potential open research questions and challenges that currently exist. Here we use existing literature to outline the key research gaps in floor planning systems.

5.1 Improving Accuracy and Robustness in Floor Planning Models

Accuracy of floor planning systems becomes the crucial element in implementing these systems in real-world architecture. Due to technological constraints, variations in environmental conditions, and inherent complexity in input data, it is not easy to produce precise floor plans that will please the customers. Many studies have addressed problems like the ambiguity in spatial arrangements in layouts and neglecting user-specific requirements, which can be solved by optimizing algorithms to balance efficiency and detail with loss of performance.

5.2 Multimodal and Multilingual Floor Plan Representation

This is the most common problem that the system supports for various representations of floor plans: text description, 2D blueprint, and 3D model. The other one is multilingual support, which can cause a considerable problem in the availability of the system. Existing systems are just not flexible enough for this purpose, emphasizing the need for standardization and adaptable design frameworks.

5.3 Resource-Efficient Floor Plan Generation

Wherever there is floor planning, the major constraint is the requirement of computation, particularly on low-power devices. An effective scaling of the floor planning without degrading its performance is quite difficult; and yet, the researchers must find a balance concerning processing speed, power consumption, and storage needs, especially for portable devices.

5.4 Data Privacy and Security in Collaborative Architectural Systems

It is quite needed to upkeep the privacy and secrecy of design data so much more with the fast growing dependency of floor planning systems on such cloud-based tools and collaboration platforms. Some prominent security issues would include unauthorized access, design plagiarism, and data leakage. Therefore, it becomes very important to prioritize encryption, user authentication, and secure data transfer protocols for increasing confidence in such systems.

The priority for privacy and secrecy of design data is becoming increasingly pressing with the adoption of floor planning systems that are based upon cloud tools and collaboration platforms. This is relevant in alleviating unauthorized access as well as design plagiarism and data leakage. Prioritizing encryption, user authentication, and the use of secure data transfer protocols go a long way in increasing confidence for such systems.

5.5 Adaptation to Environmental and Structural Variability

In reality, floor planning systems will have to learn adaptation to many things in their environments-from varying levels of illumination and noise to space limits, and many more. One of the most challenging problems is ensuring the reliability of these systems in diverse situations. This implies that planned actions will have to obey different local laws and structural norms, in addition to improving the reliability of vision-based algorithms.

6 CONCLUSIONS

Even after extending transformer-based methods to the architectural floor layout, they still face several challenges. To demonstrate the importance of handling missing data in spatial representation, some studies include the utilization of dynamic temporal networks in forecasting traffic flow and pollution as examples. This is an important limitation for architectural systems that synthesize a variety of data into designing adaptable and resilient floor plans. Future research must also focus on enhancing the interpretability of transformer models, as indicated by studies such as the empirical evaluation of neural network topologies. It is not possible to achieve the right balance between explainability and complexity of the model while emphasizing the need for user confidence and acceptance in architecture domains.

The study has most comprehensively covered the impact of Generative AI (GenAI) on the phenomenal revolution in architecture, especially in floor planning, which includes the innovations in generative models such as Generative Adversarial Networks (GANs), Vision Transformers, and diffusion models.Beyond display, the research revealed how the technologies can increase the automated and refined processes of design, eventually solving some of the inadequacies and inefficiencies of the approaches that were once adopted. It then synthesized reflections on methodology, datasets, and evaluation criteria evidenced in the literature from 2020 to 2024.

As this study shows, GenAI could completely revolutionize architectural workflows, offering the kinds of speed, accuracy, and efficiency never seen before. They generate layouts optimized against user preferences, aesthetic standards, and functional requirements with real-time flexibility. The degree and relevance of GenAI applications in the area of floor planning have been greatly advanced by including multimodal inputs such as textual descriptions, structural factors, and images. Noteworthy contributions, such as House-GAN for small layouts and generative algorithms, show scalability across multiple contexts-from individual buildings to whole neighborhoods.

Yet, this study acknowledges the limitations and unmet research needs, while looking forward to a time when GenAI will be able to be broadly applicable in architectural design. Important limitations include the quality of data, algorithm bias, interpretability, and costs of implementation. More importantly, areas that have existed and will continue to develop all include the principles of sustainability and adaptability across diverse cultures and legal contexts. These are to ensure that the effective designs created by AI are fair and contextual.

This research brings to light the broader implications of GenAI in democratising access to sophisticated design tools, whereby individual designers and smaller firms could harness this potential to come up with innovative and high-quality architectural solutions. These technologies have the potential to transform the architectural industry by reducing costs, improving engagement among stakeholders, and automating labor-intensive processes. Furthermore, the application of GenAI to urban-scale projects showcases how it can be activated to address urgent global issues surrounding sustainable development, land shortage, and urban density.

In this area, future research can focus on many interesting aspects. Collaborative AI systems, enabling design and architecture fluid interaction via generative tools in real time, could redefine design workflows. Explainability, ethical frameworks, and trust-building for GenAI will foster accountability, data privacy, and intellectual property issues in society. Another use for generative models could be to work on complicated or unusual architectural projects, which can enhance their usability and employability. In conclusion, this review promises to be a rich resource for practitioners, academicians, and industries interested in investigating AI within architectural design. It maps out a path for future exploration and also sparks a call to action towards a more holistic approach of placing technology advances side by side with human creativity and moral responsibility, summarizing where GenAI currently stands with respect to today in floor planning. To be sure, the explosion of architectural techniques in GenAI is much more than every one of these. It could as well ensure global growth that is fair and sustainable by looking into more of what GenAI can do.

ACKNOWLEGMENT

The authors would like to acknowledge the financial support is given by VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore Madhya Pradesh - 466114

REFERENCES

[1] Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katherine Millican, Malcolm Reynolds, Roman Ring, Eliza Rutherford, Serkan Cabi, Tengda Han, Zhitao Gong, Sina Samangooei, Marianne Monteiro, Jacob L Menick, Sebastian Borgeaud, Andy Brock, Aida Nematzadeh, Sahand Sharifzadeh, Mikoł aj Bińkowski, Ricardo Barreira, Oriol Vinyals, Andrew Zisserman, and Karén Simonyan. 2022. Flamingo: a Visual Language Model for Few-Shot Learning. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 23716–23736. https://proceedings.neurips.cc/paper_files/paper/2022/file/960a172bc7fbf0177ccccbb411a7d800-Paper-Conference.pdf

[2] Yuxin Bao and Changying Xiang. 2024. Exploration of Conceptual Design Generation Based on the Deep Learning Model – Discussing the Application of AI Generator to the Preliminary Architectural Design Process. In *Creativity in the Age of Digital Reproduction*, Giancarlo Di Marco, Davide Lombardi, and Mia Tedjosaputro (Eds.). Springer Nature Singapore, Singapore, 171–178.

[3] Staphord Bengesi, Hoda El-Sayed, MD Kamruzzaman Sarker, Yao Houkpati, John Irungu, and Timothy Oladunni. 2024. Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers. *IEEE Access* 12 (2024), 69812–69837. https://doi.org/10.1109/ACCESS.2024.3397775

[4] P. M. Diaz C. Gnana Kousalya, V. Senthil kumar and M. Julie Emerald Jiju. 2024. Feature selection on real-time application using hybrid of flamingo search algorithm and improved non-dominated sorting genetic algorithm III. *Journal of Experimental & Theoretical Artificial Intelligence* 0, 0 (2024), 1–19. https://doi.org/10.1080/0952813X.2024.2383649 arXiv:https://doi.org/10.1080/0952813X.2024.2383649

[5] Zhuoqun Chai, Mengxi Liu, Qian Shi, Yuanyuan Zhang, Minglin Zuo, and Da He. 2024. Fine-Grained Urban Village Extraction By Mask Transformer From High-Resolution Satellite Images in Pearl River Delta. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2024).

[6] Kai Chen and Ali Ghahramani. 2024. ComfortGPT: a transformer-based architecture for predicting preferred temperature setpoints leveraging big data. *Building and Environment* 248 (2024), 111085. https://doi.org/10.1016/j.buildenv.2023.111085

[7] Kehao Chen and Guiyu Wei. 2023. Public sentiment analysis on urban regeneration: A massive data study based on sentiment knowledge enhanced pre-training and latent Dirichlet allocation. *PLOS ONE* 18, 4 (04 2023), 1–26. https://doi.org/10.1371/journal.pone.0285175

[8] Prafulla Dhariwal and Alexander Nichol. 2021. Diffusion Models Beat GANs on Image Synthesis. In *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (Eds.), Vol. 34. Curran Associates, Inc., 8780–8794. https:

//proceedings.neurips.cc/paper_files/paper/2021/file/49ad23d1ec9fa4bd8d77d02681df5cfa-Paper.pdf

[9] Qi Duan, Lihui Qi, Renyu Cao, and Peng Si. 2022. Research on Sustainable Reuse of Urban Ruins Based on Artificial Intelligence Technology: A Study of Guangzhou. *Sustainability* 14, 22 (2022). https://doi.org/10.3390/su142214812

[10] Muhammet Sıddık Emek and Ismail Burak Parlak. 2023. A New Method of Smart City Modeling Using Big Data Techniques. In *Intelligent and Fuzzy Systems*, Cengiz Kahraman, Irem Ucal Sari, Basar Oztaysi, Selcuk Cebi, Sezi Cevik Onar, and A. Çağrı Tolga (Eds.). Springer Nature Switzerland, Cham, 772–779.

[11] Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, and Robin Rombach. 2024. Scaling Rectified Flow Transformers for High-Resolution Image Synthesis. In *Forty-first International Conference on Machine Learning*. https://openreview.net/forum?id=FPnUhsQJ5B

[12] Angélica Fernández-Morales. 2024. The Possibilities of Text-to-Image Tools for the Generation of Floor Plans. In *Graphic Horizons*, Luis Hermida González, João Pedro Xavier, Antonio Amado Lorenzo, and Ángel J. Fernández-Álvarez (Eds.). Springer Nature Switzerland, Cham, 297–307.

[13] Hadi Ghaemi, Zakieh Alizadehsani, Amin Shahraki, and Juan M. Corchado. 2024. Transformers in source code generation: A comprehensive survey. *Journal of Systems Architecture* 153 (2024), 103193. https://doi.org/10.1016/j.sysarc.2024.103193

[14] Prashnna Ghimire, Kyungki Kim, and Manoj Acharya. 2023. Generative ai in the construction industry: Opportunities & challenges. *arXiv preprint arXiv:2310.04427* (2023).

[15] Jeremy Gwinnup and Kevin Duh. 2023. A Survey of Vision-Language Pre-training from the Lens of Multimodal Machine Translation. arXiv:2306.07198 [cs.CL] https://arxiv.org/abs/2306.07198

[16] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 6840–6851. https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf

[17] Chuanbo Hu, Shan Jia, Fan Zhang, Changjiang Xiao, Mindi Ruan, Jacob Thrasher, and Xin Li. 2023. UPDExplainer: An interpretable transformer-based framework for urban physical disorder detection using street view imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 204 (2023), 209–222. https://doi.org/10.1016/j.isprsjprs.2023.08.017

[18] Michael J. Hua, Junjie Wu, and Zichun Zhong. 2024. Multi-scale Knowledge Transfer Vision Transformer for 3D vessel shape segmentation. *Computers Graphics* 122 (2024), 103976. https://doi.org/10.1016/j.cag.2024.103976

[19] Robert T. Hughes, Lin Zhu, and Tomasz Bednarz. 2021. Generative Adversarial Networks–Enabled Human–Artificial Intelligence Collaborative Applications for Creative and Design Industries: A Systematic Review of Current Approaches and Trends. *Frontiers in Artificial Intelligence* 4 (2021), 604234. https://doi.org/10.3389/frai.2021.604234

[20] Hengxu Jin, Haipeng Lu, Yu Zhao, Zhizhou Zhu, Wujie Yan, Qiqi Yang, and Shuliang Zhang. 2024. Integration of an improved transformer with physical models for the spatiotemporal simulation of urban flooding depths. *Journal of Hydrology: Regional Studies* 51 (2024), 101627.

[21] Pravneet Kaur, Gautam Siddharth Kashyap, Ankit Kumar, Md Tabrez Nafis, Sandeep Kumar, and Vikrant Shokeen. 2024. From Text to Transformation: A Comprehensive Review of Large Language Models' Versatility. arXiv:2402.16142 [cs.CL] https://arxiv.org/abs/2402.16142

[22] Yohan Kim, Seongdeok Bang, Jiu Sohn, and Hyoungkwan Kim. 2022. Question answering method for infrastructure damage information retrieval from textual data using bidirectional encoder representations from transformers. *Automation in Construction* 134 (2022), 104061. https://doi.org/10.1016/j.autcon.2021.104061

[23] Chengyuan Li, Tianyu Zhang, Xusheng Du, Ye Zhang, and Haoran Xie. 2024. Generative AI for Architectural Design: A Literature Review. *arXiv preprint arXiv:2404.01335* (2024).

[24] Chengyuan Li, Tianyu Zhang, Xusheng Du, Ye Zhang, and Haoran Xie. 2024. Generative AI Models for Different Steps in Architectural Design: A Literature Review. arXiv:2404.01335 [cs.LG] https://arxiv.org/abs/2404.01335

[25] Man Li. 2021. BERT-based Dynamic Clustering of Subway Stations Using Flow Information. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*. 2762–2765. https://doi.org/10.1109/ICDE51399.2021.00321

[26] Yuanchang Liang, Prahlad Vadakkepat, David Kim Huat Chua, Shuyi Wang, Zhigang Li, and Shuxiang Zhang. 2024. Recognizing temporary construction site objects using CLIP-based few-shot learning and multi-modal prototypes. *Automation in Construction* 165 (2024), 105542. https://doi.org/10.1016/j.autcon.2024.105542

[27] Dong Liu, Qirong Mao, Lijian Gao, and Gang Wang. 2024. Leveraging Contrastive Language–Image Pre-Training and Bidirectional Cross-attention for Multimodal Keyword Spotting. *Engineering Applications of Artificial Intelligence* 138 (2024), 109403. https://doi.org/10.1016/j.engappai.2024.109403

[28] Nanqing Liu, Xun Xu, Yongyi Su, Chengxin Liu, Peiliang Gong, and Heng-Chao Li. 2024. CLIP-Guided Source-Free Object Detection in Aerial Images. arXiv:2401.05168 [cs.CV] https://arxiv.org/abs/2401.05168

[29] Milad Masrouri, Kamalendu Paul, and Zhao Qin. 2024. Generative AI model trained by molecular dynamics for rapid mechanical design of architected graphene. *Extreme Mechanics Letters* 72 (2024), 102230. https://doi.org/10.1016/j.eml.2024.102230

[30] Khalid Nassiri and Moulay Akhloufi. 2022. Transformer models used for text-based question answering systems. *Applied Intelligence* 53 (08 2022). https://doi.org/10.1007/s10489-022-04052-8

[31] Khalid Nassiri and Moulay Akhloufi. 2023. Transformer models used for text-based question answering systems. *Applied Intelligence* 53, 9 (01 May 2023), 10602–10635.

[32] Pritthijit Nath, Asif Iqbal Middya, and Sarbani Roy. 2023. Empirical assessment of transformer-based neural network architecture in forecasting pollution trends. *International Journal of Data Science and Analytics* (2023), 1–17.

[33] Nelson Nauata, Kai-Hung Chang, Chin-Yi Cheng, Greg Mori, and Yasutaka Furukawa. 2020. House-gan: Relational generative adversarial networks for graph-constrained house layout generation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16*. Springer, 162–177.

[34] Damilola Onatayo, Adetayo Onososen, Abiola Oluwasogo Oyediran, Hafiz Oyediran, Victor Arowoiya, and Eniola Onatayo. 2024. Generative AI Applications in Architecture, Engineering, and Construction: Trends, Implications for Practice, Education Imperatives for Upskilling—A Review. *Architecture* 4, 4 (2024), 877–902. https://doi.org/10.3390/architecture4040046

[35] Chibuike Onuoha, Jean Flaherty, and Truong Cong Thang. 2024. Perceptual Image Quality Prediction: Are Contrastive Language–Image Pretraining (CLIP) Visual Features Effective? *Electronics* 13, 4 (2024). https://doi.org/10.3390/electronics13040803

[36] Jean Parente, Eugénio Rodrigues, Bárbara Rangel, and João Poças Martins. 2023. Integration of convolutional and adversarial networks into building design: A review. *Journal of Building Engineering* 76 (2023), 107155. https://doi.org/10.1016/j.jobe.2023.107155

[37] Katerina Pastra and Yorick Wilks. 2004. Vision-Language Integration in AI: a reality check. In *ECAI*, Vol. 16. 937.

[38] Tao Hu Peixiao Wang, Yan Zhang and Tong Zhang. 2023. Urban traffic flow prediction: a dynamic temporal graph network considering missing values. *International Journal of Geographical Information Science* 37, 4 (2023), 885–912. https://doi.org/10.1080/13658816.2022.2146120 arXiv:https://doi.org/10.1080/13658816.2022.2146120

[39] Joern Ploennigs and Markus Berger. 2023. AI art in architecture. *AI in Civil Engineering* 2, 1 (17 Aug 2023), 8. https://doi.org/10.1007/s43503-023-00018-y

[40] Chunping Qiu, Xiaoyu Zhang, Xiaochong Tong, Naiyang Guan, Xiaodong Yi, Ke Yang, Junjie Zhu, and Anzhu Yu. 2024. Few-shot remote sensing image scene classification: Recent advances, new baselines, and future trends. *ISPRS Journal of Photogrammetry and Remote Sensing* 209 (2024), 368–382. https://doi.org/10.1016/j.isprsjprs.2024.02.005

[41] Abir Rahali and Moulay A Akhloufi. 2023. End-to-end transformer-based models in textual-based NLP. *AI* 4, 1 (2023), 54–110.

[42] Abir Rahali and Moulay A. Akhloufi. 2023. End-to-End Transformer-Based Models in Textual-Based NLP. *AI* 4, 1 (2023), 54–110. https://doi.org/10.3390/ai4010004

[43] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, Jonathan Ho, David J Fleet, and Mohammad Norouzi. 2022. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 36479–36494. https://proceedings.neurips.cc/paper_files/paper/2022/file/ec795aeadae0b7d230fa35cbaf04c041-Paper-Conference.pdf

[44] Sachith Seneviratne, Damith Senanayake, Sanka Rasnayaka, Rajith Vidanaarachchi, and Jason Thompson. 2022. DALLE-URBAN: Capturing the urban design expertise of large text to image transformers. In *2022 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 1–9.

[45] Yong Shi, Mengyu Shang, and Zhiquan Qi. 2023. Intelligent layout generation based on deep generative models: A comprehensive survey. *Information Fusion* (2023), 101940.

[46] Yong Shi, Mengyu Shang, and Zhiquan Qi. 2023. Intelligent layout generation based on deep generative models: A comprehensive survey. *Information Fusion* 100 (2023), 101940. https://doi.org/10.1016/j.inffus.2023.101940

[47] Aivin V Solatorio. 2023. GeoFormer: Predicting Human Mobility using Generative Pre-trained Transformer (GPT). In *Proceedings of the 1st International Workshop on the Human Mobility Prediction Challenge*. 11–15.

[48] Shahriar Soudeep, Most Lailun Nahar Aurthy, Jamin Rahman Jim, MF Mridha, and Md Mohsin Kabir. 2024. Enhancing road traffic flow in sustainable cities through transformer models: Advancements and challenges. *Sustainable Cities and Society* (2024), 105882.

[49] Vinod Vasan, Naveen Venkatesh Sridharan, Sugumaran Vaithiyanathan, and Mohammadreza Aghaei. 2024. Detection and classification of surface defects on hot-rolled steel using vision transformers. *Heliyon* 10, 19 (2024), e38498. https://doi.org/10.1016/j.heliyon.2024.e38498

[50] Jon Walbrin, Nikita Sossounov, Morteza Mahdiani, Igor Vaz, and Jorge Almeida. 2024. Fine-grained knowledge about manipulable objects is well-predicted by contrastive language image pre-training. *iScience* 27, 7 (2024), 110297. https://doi.org/10.1016/j.isci.2024.110297

[51] Bowen Wang, Jiaxin Zhang, Ran Zhang, Yunqin Li, Liangzhi Li, and Yuta Nakashima. 2024. Improving facade parsing with vision transformers and line integration. *Advanced Engineering Informatics* 60 (2024), 102463. https://doi.org/10.1016/j.aei.2024.102463

[52] Han Wang, Min Liu, and Weiming Shen. 2023. Industrial-generative pre-trained transformer for intelligent manufacturing systems. *IET Collaborative Intelligent Manufacturing* 5, 2 (2023), e12078.

[53] Xiao Wang, Guangyao Chen, Guangwu Qian, Pengcheng Gao, Xiao-Yong Wei, Yaowei Wang, Yonghong Tian, and Wen Gao. 2024. Large-scale Multi-Modal Pre-trained Models: A Comprehensive Survey. arXiv:2302.10035 [cs.CV] https://arxiv.org/abs/2302.10035

[54] Ramon Elias Weber, Caitlin Mueller, and Christoph Reinhart. 2022. Automated floorplan generation in architectural design: A review of methods and applications. *Automation in Construction* 140 (2022), 104385. https://doi.org/10.1016/j.autcon.2022.104385

[55] Yanjie Wen, Ping Xu, Zhihong Li, and Wangtu Xu(ATO). 2025. An illumination-guided dual attention vision transformer for low-light image enhancement. *Pattern Recognition* 158 (2025), 111033. https://doi.org/10.1016/j.patcog.2024.111033

[56] Haowen Xu, Femi Omitaomu, Soheil Sabri, Sisi Zlatanova, Xiao Li, and Yongze Song. 2024. Leveraging generative AI for urban digital twins: a scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. *Urban Informatics* 3, 1 (2024), 29. https://doi.org/10.1007/s44212-024-00060-w

[57] Yanchen Yang, Chou Xie, Bangsen Tian, Yihong Guo, Yu Zhu, Shuaichen Bian, Ying Yang, Ming Zhang, and Yimin Ruan. 2024. Advanced Post-earthquake Building Damage Assessment: SAR Coherence Time Matrix with Vision Transformer. *International Journal of Applied Earth Observation and Geoinformation* 133 (2024), 104133. https://doi.org/10.1016/j.jag.2024.104133

[58] Winston Yap, Patrick Janssen, and Filip Biljecki. 2022. Free and open source urbanism: Software for urban planning practice. *Computers, Environment and Urban Systems* 96 (2022), 101825. https://doi.org/10.1016/j.compenvurbsys.2022.101825

[59] Yilin Ye, Jianing Hao, Yihan Hou, Zhan Wang, Shishi Xiao, Yuyu Luo, and Wei Zeng. 2024. Generative AI for visualization: State of the art and future directions. *Visual Informatics* 8, 2 (2024), 43–66. https://doi.org/10.1016/j.visinf.2024.04.003

[60] Liu Yong and Hao Chibiao. 2022. A generative design method of building layout generated by path. *Applied Mathematics and Nonlinear Sciences* 7, 2 (2022), 825–848.

[61] Chaobo Zhang, Jie Lu, and Yang Zhao. 2024. Generative pre-trained transformers (GPT)-based automated data mining for building energy management: Advantages, limitations and the future. *Energy and Built Environment* 5, 1 (2024), 143–169.

[62] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 3836–3847.

[63] Wanrong Zhu, Xin Eric Wang, Tsu-Jui Fu, An Yan, Pradyumna Narayana, Kazoo Sone, Sugato Basu, and William Yang Wang. 2021. Multimodal Text Style Transfer for Outdoor Vision-and-Language Navigation. arXiv:2007.00229 [cs.CL] https://arxiv.org/abs/2007.00229

[64] Michal Šourek. 2024. Artificial Intelligence in Architecture and Built Environment Development 2024: A Critical Review and Outlook. *Energy and Built Environment* (2024). https://www.remspace.cz