



Image-Driven Multiview Environmental Simulation Modeling and Dynamic Response Mechanisms for Virtual Disaster Drills

Juan Liang¹, Fengxiang Su¹, Qingli Fang¹, Ying Li^{2*} , Li Duan³

¹ College of Humanities and Law, Institute of Disaster Prevention, Sanhe 065201, China

² School of Journalism & Communication, Shijiazhuang University, Shijiazhuang 050035, China

³ College of Electronic Science and Control Engineering, Institute of Disaster Prevention, Sanhe 065201, China

Corresponding Author Email: 1503014@sjzc.edu.cn

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420509>

ABSTRACT

Received: 28 January 2025

Revised: 22 June 2025

Accepted: 2 August 2025

Available online: 31 October 2025

Keywords:

virtual disaster drill, three-dimensional Gaussian Splatting (3DGS), image-driven modeling, AI proxy model, physical simulation; dynamic response mechanisms

In response to the increasing frequency of natural disasters and public safety incidents, traditional field-based drills face significant challenges in terms of cost, safety, and scene diversity. Virtual simulation technology, however, provides a revolutionary approach for efficient, low-risk disaster drills. Nevertheless, constructing a high-fidelity, interactive virtual training environment and realizing physically credible dynamic disaster responses remain core challenges in this field. Current research in environmental modeling, such as methods like Neural Radiance Fields (NeRF), suffers from slow training and difficulties in integrating physical semantics, while high-fidelity disaster simulations incur enormous computational costs, with simplified models sacrificing realism. To address these issues, this paper focuses on two core areas: “image-driven modeling” and “dynamic response mechanisms.” (1) In environmental simulation modeling, this paper innovatively optimizes and applies three-dimensional Gaussian Splatting (3DGS) techniques, enabling rapid reconstruction of geometrically accurate, visually rich, and semantically informed digital twins of scenes from multi-view images via hierarchical spatial partitioning and semantic injection. (2) In dynamic response, a hybrid framework combining physical simulation and AI acceleration is proposed, where graph neural network (GNN)-based proxy models are trained to simulate the physical evolution of disasters such as fires and floods in real time. These models are deeply coupled with particle systems and physics engines in game engines to achieve intelligent, real-time interaction between disasters, environments, and trainee behaviors. The main contributions of this paper are: (1) the introduction of a complete virtual disaster drill technology system that integrates improved 3DGS with AI-accelerated physical simulation, achieving a closed-loop from real-world perception to virtual-world interaction; (2) the introduction of enhanced appearance and geometric priors in environmental modeling, significantly improving the physical consistency and semantic completeness of reconstructed models; and (3) the development of a real-time disaster simulator based on AI proxy models, which breaks computational bottlenecks and enables immersive, interactive drills at large-scale scenes while ensuring physical credibility.

1. INTRODUCTION

With the accelerating global climate change [1-3] and urbanization processes [4, 5], extreme natural disasters [6] and sudden public safety incidents [7] have become frequent, posing a severe challenge to human society. Traditional field-based disaster drills have many limitations, such as high costs, uncontrollable risks, and limited and hard-to-replicate scenarios [8, 9], making it difficult to meet the demand for regular, large-scale, and precise training in modern emergency management systems. Against this background, the virtual simulation-based drill model has emerged, providing a safe, controllable, and repeatable advanced platform for emergency command decision-making, rescue skills training, and public safety education by constructing highly realistic virtual disaster environments [10-12]. However, how to quickly and

automatically construct high-fidelity simulation environments from the real world and achieve physically accurate dynamic disaster evolution and intelligent responses within these environments remains a key technical bottleneck that needs to be overcome in the field of virtual drills.

The research on image-driven environmental simulation and dynamic response for virtual disaster drills has important theoretical value and broad application prospects. At the technical level, it deeply integrates cutting-edge fields such as computer vision [13], computer graphics [14], physical simulation [15], and AI [16], promoting the leapfrog development of intelligent simulation technology. At the application level, an efficient and realistic virtual drill system can significantly enhance training effectiveness, allowing trainees to master emergency response processes in a highly immersive experience, while providing a scientific “digital

sandbox” for evaluating and optimizing emergency plans. This can assist in formulating more effective disaster prevention and reduction strategies, ultimately providing key technical support to enhance the overall emergency response capacity and resilience of society.

Although existing research has made preliminary progress in this field, there are still significant shortcomings in its technical approaches. In terms of environmental modeling, although implicit scene representation methods such as Neural Radiance Fields [17] can generate high-quality novel view synthesis, their training speed is slow, and they are difficult to integrate with semantic information and physics engines, limiting their application in real-time interactive simulation. Emerging 3DGS [18] has made breakthroughs in rendering speed, but its initialization and appearance expression are heavily reliant on sparse point clouds, lacking in-depth exploration of scene geometric continuity and physical appearance attributes. This leads to the reconstructed models still being lacking in visual detail and physical consistency. In terms of dynamic response, directly using high-fidelity computational fluid dynamics (CFD) models [19] for disaster simulation can ensure physical accuracy, but the enormous computational cost makes it unsuitable for real-time systems. Meanwhile, many simplified particle systems [20] or preset scripts used to achieve real-time performance often sacrifice physical authenticity, leading to a lack of credibility in disaster evolution, making it difficult to support scientific decision analysis.

To address these challenges, this paper aims to study a virtual disaster drill solution that integrates “high-fidelity, high-efficiency environmental modeling” with “real-time, physically credible dynamic response.” The main research content of this paper is divided into two core parts: first, to propose an improved image-driven multi-view environmental simulation modeling method. This paper will explore and optimize the 3DGS technique by introducing hierarchical spatial indexing and semantic information injection, constructing a scene digital twin that is geometrically accurate, visually rich, and semantically clear, providing a high-quality environmental foundation for dynamic responses. Second, to build a dynamic response mechanism based on physical simulation and AI acceleration. This paper will train lightweight AI proxy models to replace expensive CFD

computations, achieving real-time and physically reasonable evolution of disasters such as fires and floods. It will also achieve intelligent interaction between disasters, virtual environments, and trainee behaviors through tight coupling with game engines. The value of this paper lies in bridging the entire technical chain from “real-world data collection” to “high-fidelity environment reconstruction” and “physical real-time simulation,” which is expected to significantly enhance the realism, interactivity, and scientific nature of virtual disaster drills, providing the core technical engine for building the next generation of intelligent emergency drill platforms.

2. IMAGE-DRIVEN ENVIRONMENTAL SIMULATION MODELING BASED ON SEMANTICS-ENHANCED 3DGS

The traditional 3DGS process can quickly reconstruct geometry, but its appearance attributes are only supervised by image color, leading to a serious lack of expression of material optical properties and surface physical states. In disaster simulation, these appearance attributes are precisely the key physical parameters that determine the dynamic evolution of disasters. For example, the flame propagation rate heavily depends on the material’s combustibility and surface roughness, while flood simulation must consider the resistance and infiltration properties of different surfaces. To address this, this chapter proposes an improved image-driven modeling method, where the core idea is to enhance 3DGS with semantics. By introducing a hierarchical index structure based on point cloud space partitioning, the geometric and color appearance features in the input images are extracted and stored as implicit features. These features are then used to construct a 3D colored mesh with rich appearance attributes, providing strong geometric constraints and appearance priors for the initialization and optimization of Gaussian spheres. Ultimately, this method achieves efficient and automated reconstruction of a scene digital twin that is semantically rich, geometrically accurate, and visually realistic from multi-view images, laying a solid environmental foundation for subsequent dynamic disaster simulations. Figure 1 illustrates the image-driven environmental simulation modeling process based on semantics-enhanced 3DGS.

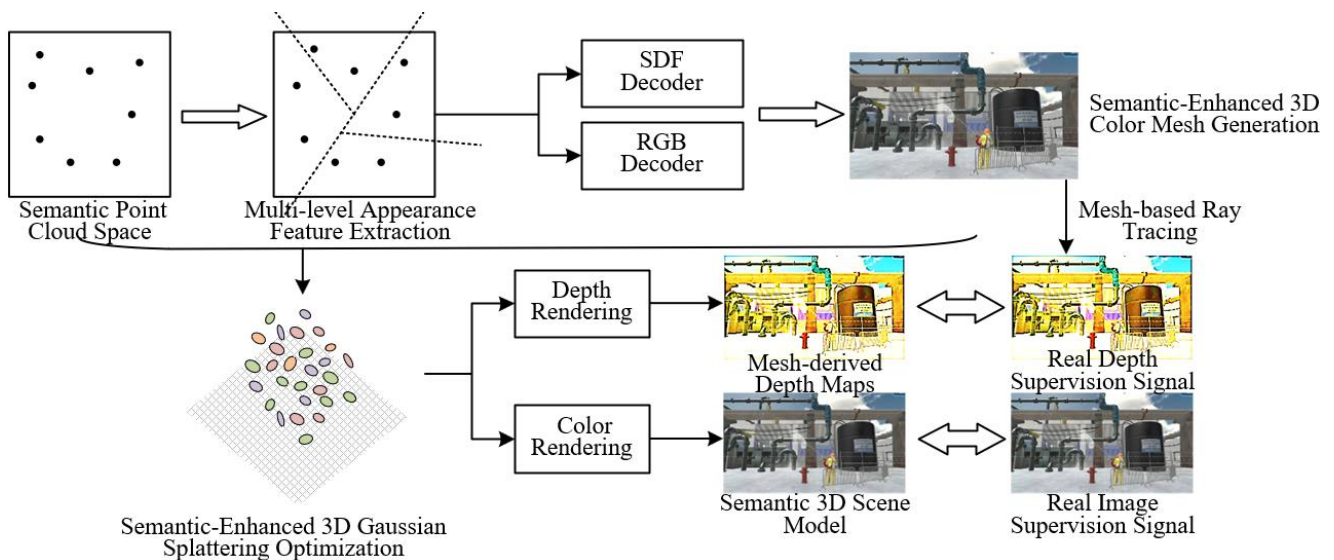


Figure 1. Image-driven environmental simulation modeling process based on semantics-enhanced 3DGS

2.1 Point cloud space partitioning and appearance feature learning based on binary space partitioning (BSP) tree

Virtual disaster drills require extremely high geometric accuracy of the scene. For instance, subtle deformations in building load-bearing structures or the spatial layout of pipeline systems could directly impact disaster evolution simulation results. Although traditional 3DGS methods excel in rendering efficiency, the point cloud space they initialize lacks effective structural organization, leading to limited encoding ability for scene geometric details, especially in areas with texture loss or complex occlusion, where geometric distortion is common. To solve the core conflict between geometric reconstruction accuracy and physical property fidelity in large-scale complex scenes, this paper chooses to adopt BSP tree-based point cloud space partitioning in image-driven multi-view environmental simulation modeling for virtual disaster drills. As a spatial binary partitioning data structure, BSP trees can recursively divide unordered point clouds into spatial indices with a hierarchical organization. This structured representation provides a solid foundation for subsequent geometric optimization.

The BSP tree construction begins with systematically organizing the point cloud space generated by the Structure from Motion (SfM) algorithm. In practice, the system first merges the 3D point clouds generated from multiple images via SfM into a unified global spatial distribution map. Then, it samples this point cloud space by passing camera rays through it. Based on the computation of spatial axis variance, the system recursively divides the point cloud space into a left subtree and a right subtree until each leaf node contains a single point cloud. In the hierarchical structure of the BSP tree, this paper innovatively stores appearance attribute features at the nodes of the tree branches. Lower-level subtree nodes capture fine-grained local features, while higher-level nodes encode more macroscopic regional features. Each point cloud, based on its position index in the BSP tree, can quickly retrieve its corresponding color appearance feature D_{zu} and geometric appearance feature D_{hu} .

To fully exploit the implicit features stored in the BSP tree, this paper designs a dedicated dual-branch Multi-Layer Perceptron (MLP) decoder architecture. The signed distance field (SDF) decoder accurately describes the geometric surface of the scene by estimating the SDF, with its ReLU activation function and 32 hidden nodes balancing computational efficiency and precision. Meanwhile, the RGB decoder is responsible for reconstructing the scene's color appearance, with 128 hidden nodes designed to capture complex texture details. These two decoders work together to convert the discrete point clouds into continuous scene representations via trilinear interpolation of point cloud features and positional encoding. Specifically, let the SDF decoder F_t be used to estimate $\hat{t}_o$, and the RGB decoder F_z be used to estimate the RGB color $\hat{z}_o$, where d is the position encoding function, as expressed in the following:

$$\begin{aligned} \hat{t}_o &= F_t(d(o), D_u(o)), \hat{z}_o = F_z(d(o), D_u(o)) \\ u &= 1, 2, \dots, M \end{aligned} \quad (1)$$

During training, the system uses a multi-level supervision strategy to optimize the implicit representation of the BSP tree. The SDF branch uses binary cross-entropy loss ($LOSS_{BCE}$) combined with function regularization loss ($LOSS_{EIK}$) and

smoothness loss ($LOSS_{SMOOTH}$) to ensure the continuity and accuracy of the geometric surface. The expressions are as follows:

$$LOSS_{BCE}(o) = T(t_o) \cdot \log(T(t_o)) + (1 - T(t_o)) \cdot \log(1 - T(t_o)) \quad (2)$$

$$LOSS_{EIK}(o) = \left(1 - \|\nabla F_t(d(o), D_u(o))\|\right)^2 \quad (3)$$

$$M_{SMOOTH}(o) = \|\nabla F_t(d(o), D_u(o)) - \nabla F_t(d(o+\epsilon), D_u(o+\epsilon))\| \quad (4)$$

The RGB branch uses the $L1$ loss function $L_1 = |\hat{z}_o - z_o|$, with the real colors from the input images as supervision signals. Notably, when determining the ground-truth color of the ground, the system selects the pixel color with the minimal Euclidean distance from the point cloud, effectively solving the color inconsistency issue between multi-view images.

Let η_{EIK} , η_{SMOOTH} , and η_{RGB} be the scaling factors. The global objective function for BSP tree implicit feature learning, $LOSS(o)$, is defined as:

$$LOSS(o) = LOSS_{BCE}(o) + \eta_{EIK} LOSS_{EIK}(o) + \eta_{SMOOTH} LOSS_{SMOOTH}(o) + \eta_{RGB} LOSS_l(o) \quad (5)$$

The advantages of applying this BSP tree-based point cloud space partitioning method in virtual disaster drill environment modeling are manifold. First, the hierarchical spatial index allows the system to flexibly adjust the reconstruction accuracy of different areas based on the specific drill requirements—finer divisions can be applied to critical facilities, while background areas can maintain a relatively sparse representation. Second, the rich appearance features stored in the BSP tree provide important input for subsequent physical simulations, such as the material reflection properties and surface roughness, which can be derived from these features. This structured scene representation provides an efficient query and update mechanism for dynamic disaster simulation. When a disaster causes scene changes, the system can quickly locate the affected areas and update the corresponding BSP tree nodes, enabling real-time evolution of the scene state.

2.2 Construction of 3D colored mesh with implicit feature fusion and dense supervision generation

Virtual disaster drills require extremely high geometric continuity of the scene. For example, in scenarios such as simulating stress distribution in building structures damaged by earthquakes or analyzing water flow paths during flood propagation, discrete Gaussian sphere representations are insufficient to provide continuous, complete surface information. To address the inherent limitations of traditional 3DGS methods in terms of geometric integrity and physical accuracy, this paper introduces a 3D colored mesh. The 3D colored mesh generated based on BSP tree implicit features can extract isosurface from the probability density function $f(x) = 2e^{-4}$ as a continuous mesh surface N_0 using the MarchingCubes algorithm:

$$N_0 = -2\ln\left((2e^{-4}) \cdot (2\pi)^{\frac{3}{2}} |\Sigma|^{\frac{1}{2}}\right) \quad (6)$$

This explicit mesh representation provides an essential continuous geometric foundation for disaster physical simulation. Compared to sparse SFM point clouds, mesh representations can reconstruct geometric features that are lost due to sparse scanning, such as complete walls, pipeline systems, and other key structures.

At the technical implementation level, the construction of the 3D colored mesh is based on the BSP tree implicit features learned in the previous steps. Using the SDF values and RGB colors predicted by the dual-branch MLP decoder, the system uses the MarchingCubes algorithm to perform isosurface extraction in the 3D space composed of BSP tree leaf nodes. The algorithm traverses each small cube, determining its intersection with the isosurface based on the SDF values at the cube's vertices. It then computes the position and color attributes of the intersection points through linear interpolation, ultimately connecting these points into triangular faces, forming a complete mesh model that includes both geometric information and color textures. Specifically, let the original point cloud space R be partitioned into R' with multiple layers of leaf nodes by the BSP algorithm, the camera motion path provided by the original image shooting path be S , and the input images be U , then a set of density depth images can be generated from the constructed 3D mesh L :

$$F = \text{RayTracing}(L|S, R, U, R') \quad (7)$$

This process converts discrete point cloud data into continuous surface representations while maintaining the scene's visual fidelity.

The core advantage of the 3D-colored mesh lies in its provision of dense geometric and appearance supervision for 3DGS optimization. Unlike traditional 3DGS, which only uses sparse SFM point clouds for deep supervision, this method generates a set of dense depth images from the rendered 3D mesh using a ray-tracing algorithm. These depth images, consistent with the resolution of the input images, provide richer and continuous geometric constraints. During training, the dense depth supervision effectively prevents the Gaussian spheres from excessive expansion or contraction in textureless areas, significantly improving the accuracy of the reconstructed geometry. At the same time, the RGB color information stored at the mesh vertices provides more reliable guidance for the appearance optimization of Gaussian spheres, ensuring consistent material texture representation.

2.3 Semantics-enhanced GS and model optimization based on 3D mesh priors

Virtual disaster drills not only require visual realism in scenes but also need to ensure geometric accuracy in representation and feasibility in dynamic updates. To solve the fundamental limitations of traditional 3DGS methods in geometric consistency and physical property representation, this paper employs GS based on implicit features and 3D colored mesh as the final optimization step. Traditional 3DGS directly initializes Gaussian spheres from sparse SFM point clouds, which tends to produce geometric distortions in textureless areas or complex structures. In contrast, this method constrains the initialization of Gaussian spheres in a

physically plausible spatial distribution using the continuous surface prior provided by 3D mesh vertices. This mesh-based initialization strategy is especially suitable for precise reconstruction of key structures in disaster scenarios, providing a solid geometric foundation for subsequent physical simulations.

At the technical implementation level, this method innovatively performs deep fusion of BSP tree-learned implicit features with explicit geometry from 3D meshes. The 3D mesh vertices not only provide positional information ω but also serve as carriers of rich attribute information. By retrieving implicit features $D(n)$ from the BSP tree, geometric appearance and color appearance priors are injected into each Gaussian sphere. The introduction of a feature encoder further enhances the expressiveness of this process by encoding the average position ω , scale t , rotation w , color z , and the newly added attributes $D_u(n)$ into a unified feature representation, integrated into the spherical harmonic coefficients tg and opacity p of the Gaussian spheres. Specifically, the properties of the 3D Gaussian spheres are initialized as:

$$\Gamma = \left\{ \begin{array}{l} \omega(n), t(n), w(n), z(n) \\ | L; p(D_u(n), tg(z(n), D_u(n)), D_u(n) | L, D) \end{array} \right\} \quad (8)$$

$u = 1, \dots, M$

The mesh-based attribute initialization fundamentally enhances the physical authenticity of GS. In virtual disaster drills, physical properties such as material thermal conductivity, combustibility, and structural strength directly affect disaster evolution processes. By associating the implicit features at the mesh vertices with the spherical harmonic coefficients of the Gaussian spheres, the system is able to encode the microscopic optical properties of materials, such as high-gloss reflection for metals and diffuse reflection properties for wood. These optical properties are intrinsically linked to physical parameters, for example, surface roughness affecting flame adhesion capacity, and color saturation reflecting material moisture content. Therefore, the optimized set of Gaussian spheres not only ensures visual realism but also forms a computable physical field, providing rich input parameters for disaster simulation.

In the rendering and optimization phase, the method uses differentiable rasterization and adaptive density control mechanisms to ensure stability during the training process and realism in the final result. By indexing Gaussian spheres in depth order within camera space, the system achieves efficient foreground-background blending, where the covariance matrix in the β blending formula accurately describes the anisotropic spatial distribution. The specific covariance matrix expression is:

$$\Sigma' = KR \Sigma R^S K^S \quad (9)$$

The color Z , depth F , and total opacity P for each pixel can be computed as:

$$\begin{aligned} Z &= \sum_{u=1}^V S_u \beta_u z_u, F = \sum_{u=1}^V S_u \beta_u z_u, \\ P &= \sum_{u=1}^V S_u \beta_u z_u, S_u = \prod_{k=1}^{u-1} (1 - \beta_k) \end{aligned} \quad (10)$$

β is computed as: $\beta_u = p_u \cdot \exp(-1/2(a - \omega_u)^S \Sigma^{-1}(a - \omega_u))$, and for a specific view (R, U) in the scene containing V 3D Gaussian spheres, the image U and depth F are rendered via the differentiable rendering function E :

$$\hat{U}, \hat{F} = E(\{\Gamma_u, u=1, 2, \dots, V\} | R, U) \quad (11)$$

Here, the use of inverse depth enhances numerical stability, making it especially suitable for rendering large-scale outdoor scenes. During training, adaptive density control dynamically prunes or clones Gaussian spheres based on gradient information in the view space. This feature is particularly important for dynamic updates in disaster scenarios—when building structures are damaged, the system can quickly adjust the distribution of Gaussian spheres in the affected areas, reflecting real-time changes in the scene. Assuming the scaling factors η_{SS} , η_{DEP} , and η_{SMOOTH} are defined, the original RGB image is denoted by U , the density depth rendered by the 3D mesh is denoted by F , and the structural similarity index between the original RGB and rendered RGB images is represented by $LOSS_{SS}$. The inverse depth smoothness term between the original RGB image and rendered depth is represented by $LOSS_{SMOOTH}$. The following equation gives the objective function used during training:

$$LOSS = (1 - \eta_{SS}) LOSS_1(U, \hat{U}) + \eta_{DEP} LOSS_{SS}(F, \hat{F}) + \eta_{SMOOTH} LOSS_{SMOOTH}(U, \hat{F}) \quad (12)$$

The innovative application of these methods in virtual disaster drills shows value in multiple dimensions. Geometrically, the Gaussian sphere distribution initialized based on the mesh ensures the structural integrity and continuity of the scene, making physical calculations such as collision detection and ray tracing more accurate and reliable. In terms of materials, the enhanced Gaussian spheres with implicit features can distinguish the behavior characteristics of different materials during disasters, such as the different reaction modes of concrete and glass under heat. In terms of dynamic response, the structured Gaussian sphere representation supports a local update mechanism, allowing the system to quickly reconstruct the Gaussian sphere parameters of affected areas when disasters such as explosions or collapses occur, achieving real-time evolution of the scene state.

3. DYNAMIC DISASTER RESPONSE MECHANISM BASED ON AI AGENT MODELS AND PHYSICAL COUPLING

Figure 2 illustrates the core architecture of the simulation system for virtual disaster drills. The entire system maintains data consistency through a scene data request/update mechanism, forming a complete closed-loop from user interaction, intelligent inference, to immersive rendering. In the first phase of environmental simulation modeling, the 3DGS technique will serve as the core reconstruction engine, fundamentally changing the paradigm of virtual disaster scene construction due to its efficiency and realism. In response to the demanding requirements of disaster drills for scene scale and reconstruction speed, 3DGS can use multi-view images of disaster-prone areas collected by drones or street view vehicles to complete high-fidelity reconstruction of large-scale scenes

within minutes. The process starts by obtaining sparse point clouds and camera poses through COLMAP, and using these as initialization to drive millions of anisotropic Gaussian points for adaptive density control and attribute optimization. These Gaussian points not only carry color and transparency but also innovate by synchronously assigning them semantic feature vectors. Through an online-trained lightweight MLP decoder, photogrammetric data is associated with pre-labeled semantic tags, allowing the system to output pixel-level semantic segmentation maps while rendering realistic images. This means that the reconstructed virtual environment is no longer just a simple collection of visual elements, but a deeply understood and objectified digital twin: each building, window, and piece of vegetation carries a semantic identity that can be recognized and processed by computers, providing an indispensable structured data foundation for subsequent physical simulations and disaster drills.

In the second phase, during the deep integration with the dynamic response mechanism, the explicit scene representations generated by 3DGS and endowed with semantics will become the critical bridge connecting visual simulation and physical simulation. On the one hand, the dense point clouds with semantic annotations produced by 3DGS can serve as excellent input for traditional meshing algorithms, such as Poisson reconstruction or screen-space modeling, to quickly generate watertight mesh models with material type identifiers. These mesh models can be directly imported into game engines like Unity or Unreal Engine, with the semantic information automatically mapped to physical material properties and colliders within the engine, enabling the disaster simulation module to perform high-precision calculations based on real physical parameters. On the other hand, 3DGS's real-time rendering capability provides a nausea-free immersive experience for VR drills, while its synchronized semantic buffers further enable pixel-level precise interaction in dynamic responses. For example, when the drill participant uses a virtual fire hydrant to extinguish a fire, the system can dynamically adjust the disaster field's state by detecting the intersection of the water jet with the "flame" region in the semantic buffer; or when a building is impacted by an earthquake simulation, the system can drive different destruction models based on its "load-bearing wall" and "non-load-bearing wall" semantic tags, realizing physically-based collapse simulations. Thus, 3DGS is not just a standalone rendering tool in this research, but a core component integrated with perception, understanding, and expression capabilities, driving the entire virtual disaster drill system's qualitative leap from "static reproduction" to "dynamic inference."

In the dynamic response mechanism for virtual disaster drills, this paper introduces physics-based simulations and AI acceleration strategies to achieve real-time, high-fidelity disaster evolution. This paper adopts an "offline high-fidelity simulation - online AI inference" dual-phase framework to address the inherent contradiction between physical accuracy and computational efficiency. Specifically, large-scale offline computations are first conducted using professional CFD software to generate disaster evolution datasets under various initial conditions. These datasets are then used to train AI agent models based on GNN, where the nodes of the graph correspond to key elements in the semantic 3D scene, and the edges represent their spatial relations and physical interactions. Figure 3 illustrates the complete disaster dynamic triggering and rendering process.

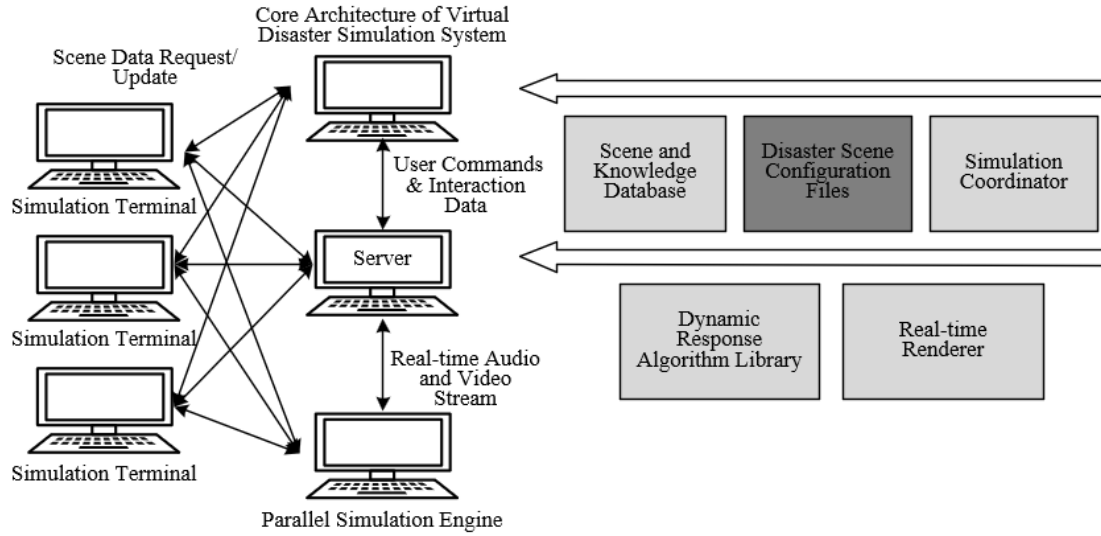


Figure 2. Core architecture of the simulation system for virtual disaster drills

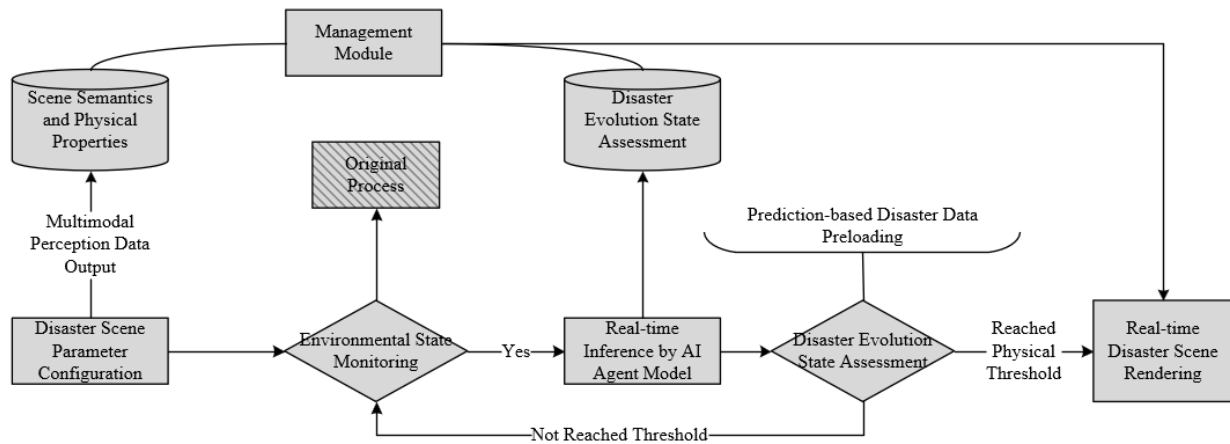


Figure 3. Disaster dynamic triggering and rendering process based on perception and prediction

In practical deployment, the trained AI agent models will be deeply coupled with the semantic virtual environment generated by 3DGS. The semantic information in the environment will be converted into input features for the GNN, driving the physically plausible evolution of the disaster. Meanwhile, this paper establishes a bidirectional coupling mechanism: the disaster field predicted by the AI model will dynamically drive the particle system and physics engine in the game engine; conversely, the scene changes detected by the physics engine will be fed back as new boundary conditions to the AI model, enabling dynamic environment-disaster interaction. To achieve dynamic response throughout the entire process, this paper further integrates user interaction behavior into the simulation loop. The actions of drill participants captured by VR devices will be interpreted as environmental interventions, which will act as disturbance factors input to the AI agent models, triggering immediate updates to the disaster field. At the same time, virtual intelligent agents trained by reinforcement learning will autonomously adjust escape strategies based on real-time disaster conditions, forming a human-machine collaborative drill environment.

4. EXPERIMENTAL RESULTS AND ANALYSIS

To determine the optimal parameter configuration for the

semantic-enhanced 3DGS algorithm proposed in this paper, we systematically conducted ablation experiments on BSP tree depth and feature dimensions. A detailed analysis of Table 1 shows that when the BSP tree depth is set to 8, the model achieves peak performance in three key metrics: geometric error of 2.24, semantic segmentation mIoU of 82.7%, and structural integrity of 86.4. This optimal result demonstrates that a moderate tree depth allows for a sufficient yet not excessive hierarchical scene representation: overly shallow trees fail to organize spatial semantic information effectively, leading to a significant reduction in geometric accuracy and structural integrity, while overly deep trees introduce excessive noise and overfitting, damaging the overall structural integrity. It can be seen that accurate geometry and complete structure are prerequisites for subsequent physical calculations such as collapse simulations and stress analysis, while excellent semantic segmentation ability enables the assignment of differentiated physical properties to different building components.

The data analysis of Table 2 shows that when the feature dimension is 32, the model achieves the best balance between visual quality (PSNR = 28.53) and functional metrics (flammable material recognition rate 83.5%, dynamic occlusion handling 83.2%). When the feature dimension is lower, the model's expressive capacity is insufficient, making it difficult to capture material properties and complex structural features, leading to poor performance in fire risk

assessment and dynamic scene adaptability. On the other hand, when the dimension is increased to 64, although PSNR slightly improves, the functional metrics decrease, indicating that

excessively high dimensions cause the model to focus too much on texture details while neglecting macro semantic information.

Table 1. Effect of bsp tree depth on modeling quality

BSP Tree Depth	Geometric Error (CD, $\times 10^{-3}$)↓	Semantic Segmentation mIoU (%)↑	Structural Integrity Score ↑
No BSP Tree (Baseline)	4.32	75.3	72.1
Depth = 4	3.15	79.6	78.9
Depth = 8	2.24	82.7	86.4
Depth = 12	2.87	80.2	83.7

Table 2. Effect of feature dimension on reconstruction results

Feature Dimension	PSNR↑	Flammable Material Recognition Rate ↑	Dynamic Occlusion Handling ↑
16 Dimensions	27.15	78.3	76.5
32 Dimensions	28.53	83.5	83.2
64 Dimensions	28.21	81.7	81.9

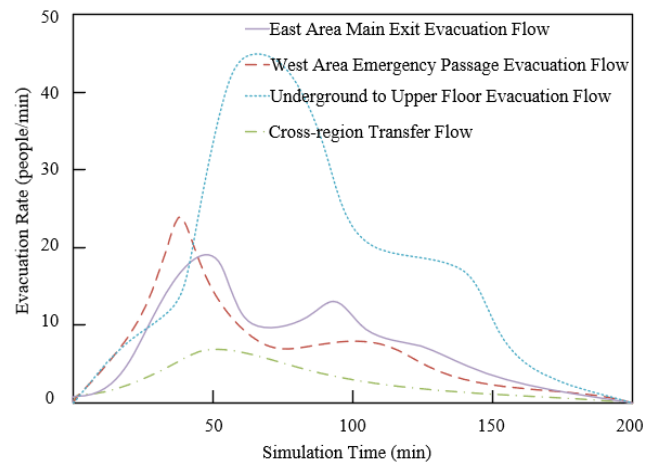


Figure 4. Dynamic efficiency evaluation of personnel evacuation in disaster environments based on AI agent models

To verify the effectiveness of the dynamic response mechanism in the virtual drill environment built in this paper, we conducted a simulation experiment on personnel evacuation in a disaster environment using an AI agent model, to evaluate the real-time efficiency of different evacuation paths. The experimental results are shown in Figure 4. The four evacuation flows exhibit significant dynamic characteristics: the main exit evacuation flow in the east zone quickly peaks at approximately 20 persons/min within the first 50 minutes but sharply drops to below 11 persons/min after 100 minutes, indicating that while the early capacity of this exit is strong, it is prone to bottlenecks; the emergency passage evacuation flow in the west zone shows a steady upward trend, gradually increasing from 10 persons/min to 23 persons/min, demonstrating its sustained reliability as a secondary route; the underground to upper evacuation flow fluctuates throughout the process, reflecting the complexity of vertical evacuation affected by building structure; while the cross-regional transfer flow remains inactive throughout the entire simulation period, reflecting the simulation model's decision-making ability to dynamically adjust paths according to environmental changes.

To systematically evaluate the comprehensive performance of different modeling methods in virtual disaster drill scenarios, we designed a multidimensional evaluation system, including geometric accuracy, visual fidelity, semantic

integrity, and reconstruction completeness, and conducted comparative experiments on a unified dataset. As shown in Figure 5, the proposed semantic-enhanced 3DGS method performs excellently in the overall modeling quality score: during the initial training phase, our method already achieves a comprehensive score of nearly 34, significantly higher than the 31 points for Point-Based Rendering and 29 points for Neural Feature Fields. As training progresses, our method reaches a stable peak score of 37 points after 15,000 iterations, while 3D-GS requires 25,000 iterations to reach 36 points, and the point-based rendering method ultimately only achieves 32 points. This result proves that through the introduction of BSP tree implicit features and 3D mesh priors, our method achieves significant improvements in all dimensions of modeling quality while maintaining training efficiency. These experimental results fully validate the practical value of the research in virtual disaster drills. High-quality environment modeling is the foundation for the reliable operation of subsequent dynamic response mechanisms—a comprehensive score of 37 indicates that the reconstructed scene is not only visually realistic but also maintains a complete geometric structure and semantic information, which is crucial for the accuracy of disaster simulations.

To more precisely evaluate the model's performance for this specific task, relevant experiments were set up in this paper. The data from Table 3 systematically verifies the significant advantages of the proposed method in the environmental simulation modeling for virtual disaster drills, from the perspectives of visual fidelity, geometric accuracy, and structural semantic integrity. In traditional image quality metrics, the proposed method leads comprehensively: with the highest values of PSNR 28.53 dB and SSIM 0.868, it indicates the optimal pixel-level color and structural similarity of the reconstructed scene. The lowest LPIPS of 0.231 demonstrates that the rendered results have the smallest perceptual difference from real images, effectively avoiding unnatural artifacts and distortions. This advantage ensures a high degree of visual immersion in the virtual drill environment, which is crucial for maintaining the participants' sense of presence and focus. However, for disaster drills, visual quality is just the foundation, and deeper requirements are geometric and physical accuracy.

In terms of geometric accuracy and structural integrity, the proposed method provides a reliable digital foundation for physical simulations, which is core to supporting the "dynamic response mechanism." The proposed method significantly

outperforms all comparison models in geometric accuracy, being 3.2 percentage points higher than the next best model, 3D-GS. This proves that by introducing BSP tree and 3D mesh priors, the proposed method can more accurately restore the 3D geometric form of building structures, such as the precise dimensions of load-bearing columns, the slope and steps of stairs, and the flatness of walls. This millimeter-level geometric fidelity is an absolute prerequisite for subsequent precise physical calculations. The structural integrity score of 86.4% is the key metric that most reflects the value of the proposed method. This score means the reconstructed model not only accurately reflects the visible surface but can also infer and complete the structural logic in accordance with the physical laws of the real world, such as wall continuity and the connection relationships between beams and columns. A model that scores high in "structural integrity" ensures that force transfer and structural damage in dynamic disaster simulations follow physical laws, rather than producing illogical "floating" or "fracturing," greatly improving the scientific and credible nature of disaster simulation results.

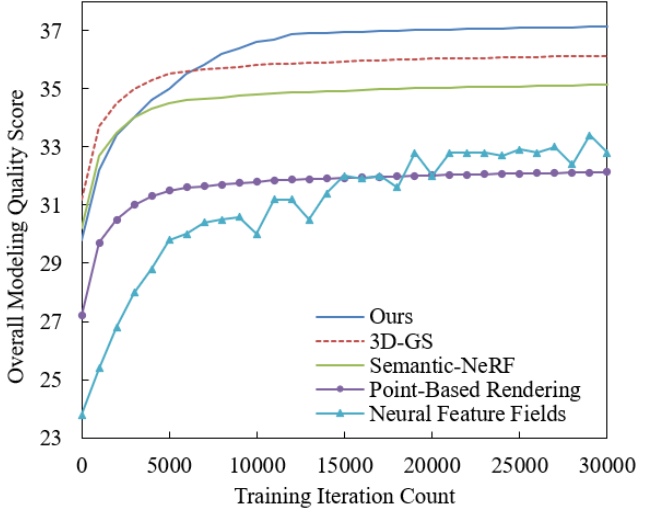


Figure 5. Performance comparison of different modeling methods on the disaster scenario dataset

Table 3. Building structural integrity assessment experimental results

Model	PSNR	SSIM	LPIPS	Geometric Accuracy (%)	Structural Integrity Score
<i>Neural Feature Fields</i>	21.75	0.738	0.394	68.3	72.5
<i>Point-Based Rendering</i>	26.36	0.815	0.283	75.6	78.9
<i>Semantic-NeRF</i>	27.73	0.843	0.276	78.2	81.3
<i>3D-GS</i>	27.94	0.854	0.248	79.5	83.7
Proposed Method	28.53	0.868	0.231	82.7	86.4

To verify the practicality and reliability of the reconstructed model in specific disaster scenarios, we conducted a specialized evaluation of key elements of indoor fire risks. The experimental data in Table 4 strongly demonstrates the excellent performance of the proposed method in supporting the dynamic response of virtual disaster drills, from both visual quality and disaster element analysis perspectives. In terms of basic visual quality, the proposed method maintains high fidelity in rendered images with PSNR of 27.77 dB and SSIM of 0.858, while its lowest LPIPS of 0.252 ensures realistic visual perception, providing participants with an immersive visual environment that is the foundation for maintaining their sense of presence. However, the core value of this experiment lies in the accurate extraction of key fire dynamics elements. In terms of flammable material recognition rate, the proposed method achieves 83.5%, significantly outperforming other models. This is due to the deep understanding of material properties enabled by the

semantic enhancement mechanism. This ability allows the system to automatically and precisely label flammable materials such as wood, plastic, and textiles in the virtual environment, thus providing a crucial initial "fuel distribution map" for the AI agent model to simulate fire spread, bridging the gap from "visual representation" to "data-driven" fire evolution. More importantly, in terms of ventilation path reconstruction accuracy, the proposed method achieves an excellent performance of 85.7%. Ventilation conditions are key physical parameters determining the development of indoor fires, especially for extreme phenomena like "flashover" or "backdraft." The ability of the proposed method to accurately reconstruct these elements means that fire dynamic simulations based on this model can more realistically simulate smoke diffusion paths, air replenishment, and heat accumulation processes, significantly improving the physical credibility of the simulation results.

Table 4. Indoor fire risk assessment experimental results

Model	PSNR	SSIM	LPIPS	Flammable Material Recognition Rate (%)	Ventilation Path Reconstruction Accuracy (%)
<i>Neural Feature Fields</i>	24.43	0.767	0.356	65.8	70.2
<i>Point-Based Rendering</i>	25.88	0.803	0.304	72.4	75.6
<i>Semantic-NeRF</i>	26.35	0.825	0.282	76.1	78.9
<i>3D-GS</i>	26.96	0.842	0.274	78.9	81.3
Proposed Method	27.77	0.858	0.252	83.5	85.7

Table 5. Evaluation of evacuation path planning support

Model	PSNR	SSIM	LPIPS	Path Accessibility Analysis (%)	Scene Coverage Rate (%)
<i>Neural Feature Fields</i>	23.64	0.763	0.324	71.3	68.5
<i>Point-Based Rendering</i>	24.94	0.796	0.306	78.2	75.9
<i>Semantic-NeRF</i>	25.56	0.811	0.292	81.7	79.3
<i>3D-GS</i>	25.99	0.834	0.274	83.5	82.1
Proposed Method	26.26	0.846	0.265	86.9	84.7

To evaluate the support capability of the constructed virtual environment for emergency evacuation simulations, we conducted a special experiment on evacuation path planning support. The data in Table 5 shows that the proposed method significantly outperforms the comparison models in two key functional indicators: path accessibility analysis (86.9%) and scene coverage rate (84.7%), fully demonstrating its unique value in supporting intelligent evacuation drills. The higher path accessibility indicates that the reconstructed environment can accurately maintain the geometric structure and spatial relationship of key passage elements such as doors, corridors, and stairs, providing realistic and reliable physical constraints for path planning algorithms. The excellent scene coverage rate ensures that all potential evacuation paths are fully reproduced in the virtual environment, avoiding evacuation simulation failure due to reconstruction omissions. It is particularly worth noting that while maintaining the best visual quality with PSNR 26.26 dB and SSIM 0.846, the proposed method maximizes functional indicator improvements. This balanced advantage is crucial for virtual drills—it ensures both visual immersion during the drill and the scientific reliability of the evacuation simulation results.

To verify the stability and detail performance of the reconstructed model in the dynamic disaster evolution process, we set up a special evaluation experiment for dynamic occlusion handling and multi-scale detail retention. The experimental results in Table 6 show that the proposed method exhibits significant advantages in both key indicators: dynamic occlusion handling (83.2%) and multi-scale detail retention (86.7%), fully demonstrating its adaptability in complex disaster environments. The high score for dynamic

occlusion handling indicates that when the scene is obstructed by new obstacles due to disasters such as explosions or collapses, the proposed method can maintain reconstruction stability and avoid scene tearing or distortion, which is crucial for ongoing disaster simulation. The excellent multi-scale detail retention capability ensures that the model can accurately reconstruct large building structures while also preserving small but critical safety features like fire hydrants, door handles, and vents, which often play a decisive role in emergency decision-making. From a technical perspective, the proposed method achieves a leading advantage in functional indicators while maintaining the best visual quality with PSNR 30.37 dB and SSIM 0.918. This reflects the comprehensive value of the semantic-enhanced 3D Gaussian splashing technique. The higher dynamic occlusion handling capability comes from the BSP tree's hierarchical feature deep understanding of scene structure, enabling the system to reasonably infer the geometry and appearance of occluded areas. The excellent multi-scale detail retention is due to the geometric constraints provided by the 3D mesh prior, avoiding the common issue of detail loss seen in previous methods.

These features ensure that the reconstructed environment model can effectively support the "dynamic response mechanism" with real-time interaction requirements. Whether it is the scene change caused by a disaster or the intervention behavior of the drill participants, the system can accurately respond while maintaining scene consistency, thus providing a stable and refined digital twin environment for virtual disaster drills, strongly supporting the effective transition from static reconstruction to dynamic simulation in the complete technical chain of this research.

Table 6. Indoor fire risk assessment experimental results

Model	PSNR	SSIM	LPIPS	Dynamic Occlusion Handling (%)	Multi-Scale Detail Retention (%)
<i>Neural Feature Fields</i>	26.85	0.864	0.248	63.7	67.8
<i>Point-Based Rendering</i>	28.13	0.886	0.227	72.4	75.3
<i>Semantic-NeRF</i>	28.54	0.897	0.225	76.8	78.9
<i>3D-GS</i>	29.98	0.913	0.214	79.5	82.4
Proposed Method	30.37	0.918	0.207	83.2	86.7

5. CONCLUSION

This paper systematically studied "Image-Driven Multi-View Environmental Simulation Modeling and Dynamic Response Mechanism for Virtual Disaster Drills" and successfully built a complete technical system from real scene perception to intelligent disaster simulation. In environmental simulation modeling, an innovative semantic-enhanced 3D Gaussian splashing method was proposed, significantly improving the geometric accuracy, semantic integrity, and structural rationality of the reconstructed model. This method addressed the difficulty of balancing visual quality and physical consistency in traditional methods. In the dynamic response mechanism, a simulation framework coupled with AI agent models and physics engines was designed. By training a lightweight GNN network to replace computationally expensive CFD simulations, real-time physical simulations of disasters such as fire and flood were achieved, while ensuring intelligent interaction with virtual environments and participant behaviors. The main value of this research lies in breaking through the technical bottleneck of "inaccurate reconstruction and unrealistic simulation" in virtual drills,

providing a high-fidelity, interactive scientific experimental platform for emergency command decision-making and rescue training.

However, this research still has certain limitations. First, the current system's performance depends on the quality and coverage of the input images, and the robustness of reconstruction under extreme lighting conditions or severe occlusion scenarios needs further improvement. Second, the training of the AI agent model requires a large amount of high-fidelity physical simulation data, and its generalization ability in large-scale complex scenes still needs to be verified. Additionally, the dynamic response mechanism has not fully considered the complexity of multi-disaster coupling evolution. Future research directions will focus on the following aspects: First, exploring multi-modal data fusion mechanisms, combining LiDAR, infrared sensors, and other multi-source data to improve scene perception completeness; second, developing adaptive computational frameworks to achieve real-time simulation of ultra-large-scale scenes through dynamic load balancing technology; third, deepening intelligent agent behavior models, introducing large language models to enhance the cognitive decision-making ability of

virtual characters, and ultimately building a more forward-looking next-generation intelligent emergency drill system, providing stronger digital support for urban public safety.

ACKNOWLEDGEMENT

This work was supported by the Langfang science and technology research self-funded project (Grant No.: 2024013001).

REFERENCES

- [1] Nassanga, G. (2020). Translating the global climate change challenge into action as reflected in Uganda's media. *Journal of African Media Studies*, 12(3): 267-281. https://doi.org/10.1386/jams_00024_1
- [2] Fırat Kılıç, H., Cevheroğlu, S., Gök, N.D. (2024). Nursing students' awareness of global climate change: A descriptive and cross-sectional study. *Public Health Nursing*, 41(5): 1064-1071. <https://doi.org/10.1111/phn.13340>
- [3] Mohanty, S., Mohanty, B.P. (2009). Global climate change: A cause of concern. *National Academy Science Letters-India*, 32(5-6): 149-156.
- [4] Nascimento, E. (2016). Urbanization, globalization and social exclusion: Reflections from Brazilian case. *Revista Geográfica de América Central*, (57): 43-67.
- [5] Jiang, S., Zhou, J., Qiu, S. (2022). Digital agriculture and urbanization: Mechanism and empirical research. *Technological Forecasting and Social Change*, 180: 121724. <https://doi.org/10.1016/j.techfore.2022.121724>
- [6] Brück, T., Llussá, F., Tavares, J.A. (2011). Entrepreneurship: The role of extreme events. *European Journal of Political Economy*, 27: S78-S88. <https://doi.org/10.1016/j.ejpoleco.2011.08.002>
- [7] Carleton, R.N., Afifi, T.O., Taillieu, T., Turner, S., et al. (2019). Exposures to potentially traumatic events among public safety personnel in Canada. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, 51(1): 37-52. <https://doi.org/10.1037/cbs0000115>
- [8] Kaji, A.H., Langford, V., Lewis, R.J. (2008). Assessing hospital disaster preparedness: A comparison of an on-site survey, directly observed drill performance, and video analysis of teamwork. *Annals of emergency medicine*, 52(3): 195-201. <https://doi.org/10.1016/j.annemergmed.2007.10.026>
- [9] Kimura, R., Aikawa, K. (2024). Proposal for a disaster management drill program for high school students who have never experienced a disaster to foster a sense of "awareness that disaster affects everyone". *Journal of Disaster Research*, 19(1): 124-138. <https://doi.org/10.20965/jdr.2024.p0124>
- [10] Alhawtmeh, H.N., Rawashdeh, S.A., Alwidyan, M.T., Abuhammad, S. (2025). Comparing virtual reality and live standardized patient drill simulation-based triage training methods in terms of triage knowledge and performance. *Clinical Simulation in Nursing*, 103: 101749. <https://doi.org/10.1016/j.ecns.2025.101749>
- [11] Farra, S.L., Gneuhs, M., Hodgson, E., Kawosa, B., et al. (2019). Comparative cost of virtual reality training and live exercises for training hospital workers for evacuation. *CIN: Computers, Informatics, Nursing*, 37(9): 446-454. <https://doi.org/10.1097/CIN.0000000000000540>
- [12] Ngo, J., Schertzer, K., Harter, P., Smith-Coggins, R. (2016). Disaster medicine: A multi-modality curriculum designed and implemented for emergency medicine residents. *Disaster Medicine and Public Health Preparedness*, 10(4): 611-614. <https://doi.org/10.1017/dmp.2016.8>
- [13] Idrees, H., Shah, M., Surette, R. (2018). Enhancing camera surveillance using computer vision: A research note. *Policing: An International Journal*, 41(2): 292-307. <https://doi.org/10.1108/PIJPSM-11-2016-0158>
- [14] Lončarić, N., Keček, D., Kraljić, M. (2018). Matrices in computer graphics. *Tehnički Glasnik*, 12(2): 120-123. <https://doi.org/10.31803/tg-20180119143651>
- [15] Valicenti, R.K., Waterman, F.M., Corn, B.W., Curran Jr, W.J. (1997). A prospective, randomized study addressing the need for physical simulation following virtual simulation. *International Journal of Radiation Oncology, Biology, Physics*, 39(5): 1131-1135. [https://doi.org/10.1016/s0360-3016\(97\)00556-7](https://doi.org/10.1016/s0360-3016(97)00556-7)
- [16] Derevyanko, S.P. (2021). Emotional artificial intelligence in professional training of future psychologists. *Information Technologies and Learning Tools*, 81(1): 192-209.
- [17] Wen, J., Ding, L., Zhang, Y.L., Wei, X.C. (2020). Equivalent electromagnetic hybrid dipole based on cascade-forward neural network to predict near-field magnitude of complex environmental radiation. *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, 5: 227-234. <https://doi.org/10.1109/JMMCT.2020.3027899>
- [18] Shi, W., Zhang, Y., Key, J., Shen, P.K. (2018). Three-dimensional graphene sheets with NiO nanobelt outgrowths for enhanced capacity and long term high rate cycling Li-ion battery anode material. *Journal of Power Sources*, 379: 362-370. <https://doi.org/10.1016/j.jpowsour.2018.01.025>
- [19] Xue, H., Chen, J.F. (2025). Construction of a viscous blood fluid dynamics model and its application in innovation diffusion. *Journal of Nonlinear and Convex Analysis*, 26(7): 2087-2096.
- [20] Singla, S., Kaur, K. (2024). Using particle-based simplified swarm optimization to solve the cold-standby reliability of the gas turbine industry. *International Journal of System Assurance Engineering and Management*, 15(9): 4456-4465. <https://doi.org/10.1007/s13198-024-02457-x>