AI-Based Architectural, Mechanical, Electrical and Plumbing BIM Object Classification Model and Semantic Enrichment Framework

Research Article

Abstract

Building Information Modelling (BIM) is increasingly being adopted across the architecture, engineering, and construction industries for digitally simulating and managing infrastructure projects. Despite the growth in BIM utilisation, challenges persist in the classification and semantic enrichment of architectural, mechanical, electrical, and plumbing (MEP) objects, critical components in infrastructure modelling. While some studies have addressed object classification or semantic enrichment independently, there is limited research on integrating both, particularly for MEP components where semantic clarity and interoperability are essential for effective cross-disciplinary stakeholder collaboration. This paper introduces a novel AI-based framework for architectural MEP BIM object classification and semantic enrichment, incorporating multiple deep learning components. The proposed system leverages 3D Convolutional Neural Networks (CNN) for spatial feature extraction, Graph Neural Network Transformers for capturing relational features, and a CNN-based feature fusion model that synthesises geometric and relational attributes. These enriched features are subsequently used in a multi-layer neural network for robust classification, which is also used to enrich the semantic database. The framework was trained and evaluated using the IFCNet dataset, achieving a 97.95% training accuracy and 86.97% testing accuracy. These results demonstrate robust generalisation across architectural and MEP object types. Its joint use of geometric and relational features enabled better performance than methods that handle only classification or semantic enrichment. The framework provides a solid foundation for enhancing BIM interoperability, facilitating automated querying, reducing manual labelling efforts, and minimising errors due to poor semantic representation.

Keywords: Artificial Intelligence (AI); Building Information Modeling (BIM); AEC; MEP; Deep Learning; Semantic Enrichment

1 Introduction

The architecture, engineering and construction (AEC) industry is undergoing a significant transformation driven by the adoption of digital technologies such as building information modelling (BIM), which has emerged as a foundational tool worldwide, serving not only as a digital tool for construction but also as a key management approach. It is anticipated to be crucial in project management Wang and Chen (2023) and is recognised as a technology that can significantly advance the AEC industry Aladağ et al. (2023).

BIM involves the generation and management of digital representations encompassing the physical and functional characteristics of built assets Sacks et al. (2018), effectively digitising building information. Its potential to advance and transform the operation and maintenance of architectural, mechanical, electrical, and plumbing (MEP) systems is particularly noteworthy, providing a platform for facility managers to access, analyse, and process building information within a digital 3D environment Gao and Pishdad-Bozorgi (2019).

In today's competitive landscape, many AEC industries are embracing a *data-driven* approach Huang et al. (2021), recognising the value of data-triggered decisions supported by analytics over traditional instinct-based judgements. Advanced BIM solutions play a crucial role in this shift, transforming raw data into valuable, actionable information throughout the construction process Services (2025). The BIM project and asset information model serve as a centralised repository, accommodating graphical, alphanumeric, and documentation datasets linked to construction projects and assets across their entire lifecycle Sompolgrunk et al. (2024).

The integration of AI, particularly machine learning (ML) and deep learning (DL), with BIM is revolutionising the AEC industry. BIM provides a comprehensive digital representation of a building's physical and functional characteristics, serving as a foundation for design, construction, and facility management Abanda et al. (2020). Meanwhile, AI's ability to analyse complex datasets and uncover intricate patterns has demonstrated remarkable success in domains like image recognition and natural language processing (NLP) LeCun et al. (2015), and its application to BIM offers unprecedented opportunities for automation and enhanced decision-making Zhang et al. (2021).

A key advancement in this synergy is semantic enrichment, the process of augmenting BIM models with meaningful, context-aware data, such as non-graphical information (e.g., material properties, maintenance schedules, or regulatory compliance) Borkowski and Maroń (2024a); Borkowski (2023). This enrichment transforms BIM objects—such as walls, pipes, or ducts—into intelligent entities that facilitate seamless communication across project stakeholders and improve project outcomes Bloch (2022). Deep learning models, including Convolutional Neural Networks (CNNs) and Point Cloud-based Deep Neural Networks (PC-DNNs), excel at classifying BIM objects and capturing spatial contexts to enable accurate semantic enrichment Miky et al. (2024); Seydgar et al. (2024). These capabilities enhance construction scheduling, risk management, and operational efficiency by providing actionable insights from BIM data Akintoye et al. (2024); Pan and Zhang (2021).

Despite the recognised potential of BIM to enhance efficiency and collaboration in the AEC industry, a critical gap persists: the misalignment between graphical elements (e.g., 3D geometry) and non-graphical data (e.g., metadata), which limits interoperability and data usability Jiang et al. (2023). Current BIM workflows, particularly for MEP systems, rely heavily on manual processes for object classification and semantic annotation.

This lack of a robust semantic enrichment framework further exacerbates these challenges, as many BIM objects are devoid of comprehensive contextual metadata Grilo and Jardim-Goncalves (2010). This deficiency hampers advanced analytics, automation, and effective lifecycle management, resulting in fragmented data, miscommunication, and operational inefficiencies that contribute to project delays and increased costs Abdelalim et al. (2024). These issues are particularly pronounced in the MEP domain, where the complexity and interdependencies of the heating, ventilation, and air conditioning (HVAC), electrical, and plumbing systems are inadequately represented in existing BIM models, impeding efficient management across design, construction, and operation phases Shehadeh et al. (2024).

Although recent advancements in ML, including DL and NLP, offer promising solutions for handling complex BIM datasets, achieving accurate, automated classification and consistent semantic enrichment for diverse MEP systems remains a significant challenge. There is an urgent need for intelligent frameworks capable of reliably classifying BIM objects and enriching them with detailed semantic context to foster cross-disciplinary collaboration, optimise data utilisation, and enhance efficiency throughout the building lifecycle Borkowski and Maroń (2024b).

This work is motivated by the growing demands within the AEC industry to enhance efficiency, prioritise sustainability, and optimise cost-effectiveness, as current manual methods for classifying and semantically annotating MEP BIM objects are error-prone, time-consuming, labour-intensive, and lack standardisation. These limitations lead to significant interoperability issues across different BIM project stakeholders, softwares and applications Khosrowshahi and Arayici (2012), causing inefficiencies throughout the design, construction, and facility management phases. Furthermore, the absence of robust semantic enrichment restricts the full potential of BIM models, limiting their use in advanced applications such as predictive maintenance, energy optimisation, and intelligent decision support systems. To overcome these challenges, AI-driven solutions provide a transformative opportunity. By leveraging deep learning to automate MEP BIM object classification and semantic enrichment, AI can reduce human effort, improve data consistency and reliability, and enable standardised, context-aware BIM models.

To address the identified limitations in current BIM practices, this research proposes the development of an AI-driven framework for automated classification and semantic enrichment of BIM objects, with a particular emphasis on the complexities of architectural and MEP systems in AEC industries. The framework integrates an AI-based classification model and a semantic enrichment module through specialised neural network modules and combines them using an attention-based fusion mechanism to streamline the classification process and enhance the contextual metadata associated with BIM objects and enhance the results of semantic knowledge representation. By leveraging this advanced DL techniques, this solution aims to deliver accurate, efficient, and scalable automation, improving data consistency and interoperability across AEC workflows. This approach represents a pivotal advancement in the digital transformation of the AEC industry, promoting a more sustainable, collaborative, and intelligent built environment.

This research advances the integration of AI and data-driven methodologies in the AEC industry through the following key contributions:

1. The research develops a novel AI-based model that utilises deep learning techniques to automate the classification of complex architectural and MEP BIM objects. By processing both geometric and relational attributes using specialised neural networks and integrating them via an attention-based fusion mechanism, the model enhances the non-graphical metadata of BIM objects. This approach improves data accuracy, consistency, and semantic knowledge representation across BIM platforms, addressing critical limitations in manual processes.
2. Automates the classification of architectural and MEP BIM objects, helping to reduce manual effort, minimise costly errors, and mitigate misinterpretations. This leads to accelerated project delivery timelines and fewer change orders during construction, thereby improving overall project efficiency and cost-effectiveness.
3. The framework addresses data interoperability challenges, as well as the heterogeneity and interdependencies of MEP systems, by providing a context-aware approach to BIM object classification. This fosters better collaboration among stakeholders in the AEC industry, including architects, engineers, and contractors, by ensuring consistent, interoperable, and semantically enriched BIM data, ultimately streamlining cross-disciplinary workflows.

Collectively, these contributions tackle significant shortcomings in existing BIM practices. The proposed framework delivers immediate, practical solutions to industry challenges while establishing a foundation for future AI-driven advancements in BIM. By promoting efficiency, interoperability, and sustainability, this research paves the way for broader adoption and further exploration of AI applications in the AEC sector.

The remainder of this paper is organised as follows. Section 2 reviews related works on AI-driven classification and semantic enrichment of BIM objects. Section 3 provide details of the proposed methodology adopted in this research. Section 4 describes the experimental setup, datasets, data preprocessing, and hyperparameter configurations and presents the results including performance metrics and confusion matrix analysis. Section 5 interprets the results, stating the limitations and implications of the proposed framework. Finally, Section 6 concludes the study and outlines the directions for future work.

2 Related Work

The integration of AI into BIM has significantly advanced the classification and semantic enrichment of architectural and MEP objects, enhancing accuracy, interoperability, and decision-making in the AEC industry. This section reviews recent literature (2020–2025) on AI-based architectural and MEP BIM object classification and semantic enrichment frameworks, alongside related works on BIM object classification and semantic enrichment. The discussion is organised into four categories: (1) AI-based architectural and MEP BIM object classification and semantic enrichment frameworks, (2) BIM object classification, (3) semantic enrichment frameworks for BIM objects in AEC and MEP, and (4) broader reviews and supporting technologies that demonstrate the effectiveness of deep learning in this domain.

Recent studies have leveraged AI to address the unique challenges of architectural and MEP systems within BIM workflows. Utkucu et al. (2024) developed an ensemble learning approach combining semantic keyword search, rule-based inferencing, machine learning with geometric features, and deep learning with visual shape features. Their dataset of 3,410 instances across 40 classes achieved a 91% accuracy and an F1-score of 0.88, demonstrating the efficacy of ensemble models for architectural and MEP object classification in building performance evaluation, such as energy efficiency assessments. Similarly, Utkucu and Sacks (2023) a rule-based semantic enrichment method for architectural and MEP object classification using interdomain topological relationships and geometry conditions was proposed. By applying four rule sets across 32 knowledge graphs, they achieved over 90% accuracy, highlighting the potential of semantic web technologies to enhance architecural and MEP object classification and interoperability. Lilis et al. (2025) introduced the Geometric Relation Checking (GRC) tool, which automatically detects geometric relations between Industry Foundation Classes (IFC) objects to infer missing semantic relationships. Tested on MEP and architectural data, this tool enhances semantic graph connectivity, supporting applications like HVAC topology generation. Chen and Jiang (2025) proposed NE-Graph-BERT, a node-enhanced, self-supervised graph neural network (GNN) model for spatial recognition, achieving 97.08% precision and a 96.75% F1-score in classifying spatial relationships, significantly advancing semantic enrichment for MEP systems.

BIM object classification is essential for automating BIM workflows, and recent works have utilised AI to improve accuracy and scalability. Emunds et al. (2022) presented SpaRSE-BIM, a sparse CNN model for classifying IFC-based geometry, achieving state-of-the-art accuracy with improved computational efficiency. This approach mitigates inconsistencies in BIM object classification caused by unsupported export functionalities or manual errors, facilitating downstream applications. Rogage and Doukari (2024) developed a deep learning-based method for 3D object recognition to generate semantic BIM data, leveraging geometric representations to address inconsistent IFC standard applications. Tested on the National BIM Library and real-world BIM models, their approach supports model reuse and enrichment. Jang et al. (2025) proposed a GNN-based method to subcategorise low-LOD BIM objects into detailed subtypes based on functional requirements, such as insulation or load-bearing properties. Using a novel threshold-enhanced triangle intersection (TETI) algorithm, they achieved a weighted F1-score of 0.8766 across 42 subtypes, enhancing early-stage engineering analyses like cost estimation.

Semantic enrichment frameworks enhance BIM models with domain-specific information, critical for AEC and MEP applications. Mirarchi et al. (2024) employed machine learning to bridge the gap between graphical and semantic data in BIM models, achieving over 80% accuracy in transforming graphical information into computable data. This method improves decision-making and collaboration in construction processes. Xue et al. (2021) conducted a ten-year review of semantic enrichment in BIM and City Information Models (CIMs), highlighting their integration and the role of semantic enrichment in multidisciplinary applications like construction management and urban planning. Sacks et al. (2017) developed the SeeBIM system, which uses computer-readable rules to infer object types and relationships, handling complex geometries and ensuring thorough classification. Their approach was validated on a real-world highway bridge model with 333 components. Jiang et al. (2023) reviewed semantic enrichment methods, analysing their applicability across the facility lifecycle and identifying gaps in interoperability and LOD requirements.

Broader reviews and supporting technologies provide context for the effectiveness of deep learning in BIM. Sacks et al. (2020) explored the intersection of BIM, AI, and construction technology, identifying semantic enrichment as a central research challenge and emphasising the reliance of construction innovations on digital building information. Hu et al. (2022) reviewed knowledge extraction and discovery in BIM, covering knowledge description, discovery, storage, inference, and application, and suggesting future directions for integrating knowledge science with BIM. Yue et al. (2025) investigated deep learning-based point cloud completion for MEP components, achieving superior Chamfer Distance (CD) and F-score with the PoinTr model, improving classification accuracy for occluded point clouds. Wang et al. (2022) proposed a deep learning-based object verification approach for scan-to-BIM processes, enhancing reconstruction accuracy for MEP systems by filtering false positives, validated on real scan data from a water treatment facility.

The body of literature reviewed highlights notable progress in the application of AI for architectural and MEP BIM object classification and semantic enrichment, achieving impressive performance metrics (e.g., 97.08% precision as reported by Chen and Jiang (2025)). However, a critical gap remains in effectively addressing interoperability among diverse BIM stakeholders. This survey, which synthesises some of the most influential studies in the field, provides a valuable understanding of the prevailing methodologies and key findings that have shaped current knowledge. Drawing from these insights, this research aims to bridge existing gaps by proposing an integrated framework. Notably, most existing studies tend to concentrate on either classification or semantic enrichment in isolation, lacking a unified approach. A detailed summary of these related works is presented in Table 1.

3 Proposed Framework

This section presents details of the proposed deep learning-based framework for classifying and semantically enriching BIM objects in the AEC industry. The framework addresses deficiencies in architectural and MEP BIM object classification by integrating geometric and relational feature processing with semantic knowledge representation. It achieves accurate classification across 20 distinct BIM object categories and produces interoperable and semantically rich outputs, as illustrated in Figure 1.

As depicted in Figure 1, the workflow begins with the data preprocessing phase, where raw BIM data is analysed for compatibility. A balanced dataset of 600 instances per class is curated to ensure robust training across the 20 object categories. This step standardises the input data, preparing it for subsequent feature extraction and analysis.

Feature extraction is performed through two parallel neural network models. The geometric feature extraction model, shown in Figure 2, employs 3D CNNs to process discrete representations, such as point clouds, meshes, or voxel grids. It generates a 256-dimensional geometric descriptor that captures spatial attributes, including shape, dimensions, and material properties, with dropout applied for regularisation.

Concurrently, the relational feature extraction model, depicted in Figure 3, uses GNN with message-passing and multi-head graph attention mechanisms. This model processes BIM component graphs

Table 1: Summary of Related Works on AI-Based MEP BIM Object Classification and Semantic Enrichment

Year	Ref.	Dataset	Focus Area	Key Finding
2022	Emunds et al. (2022)	IFCNet	BIM Classification	Efficient IFC geometry classification
2023	Utkucu and Sacks (2023)	Not stated	MEP Semantic Enrichment	Over 90% accuracy with rule sets
2024	Rogage and Doukari (2024)	ModelNet10, Extended ModelNet10	BIM Classification	Deep learning for semantic BIM data
2024	Mirarchi et al. (2024)	Not stated	Semantic Enrichment	Over 80% accuracy in data transformation
2024	Utkucu et al. (2024)	Self compiled	MEP Classification	91% accuracy with ensemble models
2025	Lilis et al. (2025)	Not stated	Semantic Enrichment	Tool for geometric relation detection
2025	Chen and Jiang (2025)	SpaceGraph (self-generated) Semantic Enrichment with GNNs	97.08% precision, 96.75% F1-score	
2025	Yue et al. (2025)	Not stated	Supporting Technology	Improved MEP classification accuracy

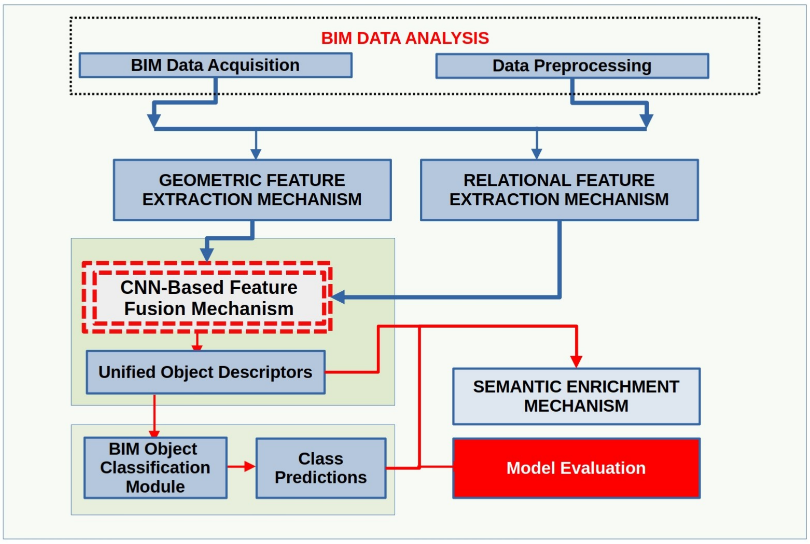

Figure 1: Architecture of the proposed framework for architectural and MEP BIM object classification and semantic enrichment.

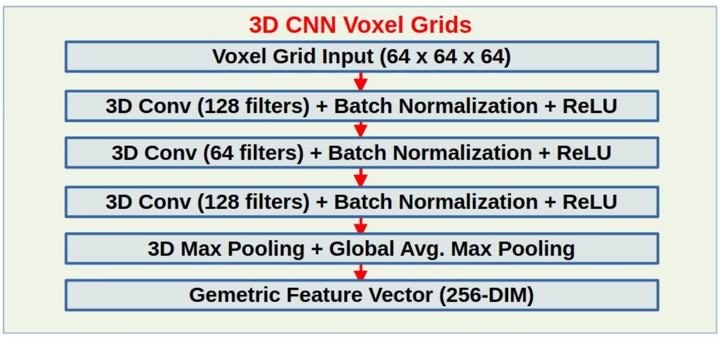

Figure 2: Geometric feature extraction branch architecture.

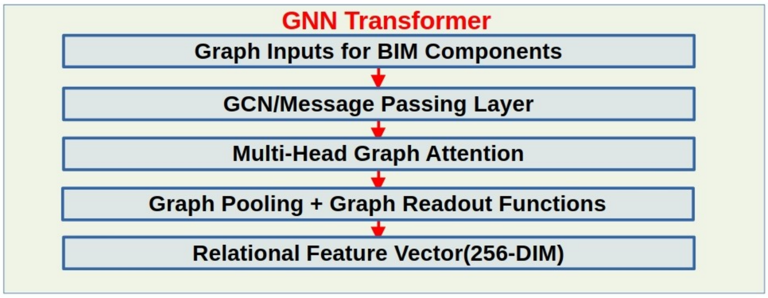

Figure 3: Relational feature extraction branch architecture.

to capture spatial relationships and connectivity, producing a 256-dimensional relational descriptor. The incorporation of graph pooling and readout functions, combined with dropout, ensures a robust representation of contextual dependencies.

The detailed functionalities and architecture of the feature fusion mechanism, the classification mechanism, and the semantic enrichment mechanism are discussed in Sections 3.1, 3.2, and 3.3, respectively.

3.1 The feature fusion mechanism

The feature fusion mechanism, a key component of the proposed framework (see Figure 1) processes the 256-dimensional vectors of the geometric and relational features through a structured pipeline, as detailed in Figure 4. The fusion mechanism comprises several stages: (a) feature projection layers that reduce the dimensionality of both geometric and relational features from 256 to 128 dimensions while applying ReLU activation and batch normalisation; (b) an attention mechanism that learns to weight the importance of geometric versus relational features for each sample, thus producing two attention weights through a softmax function and (c) a fusion layer that combines the weighted features into a final 256-dimensional representation, followed by the addition of dropout to prevent overfitting. This fusion approach allows the model to adaptively emphasise either geometric or relational features, depending on which is more relevant for classifying each specific BIM element.

3.2 The classification mechanism

In the proposed framework, classification is performed by a multi-layer neural network (see Figure 5) designed to effectively leverage the rich 256-dimensional fused feature representations to predict the architectural and MEP BIM object classes (e.g., Wall, HVAC Duct). The classifier processes the feature representations through fully connected layers to produce probability distributions over the 20 categories, ensuring accurate and reliable predictions.

3.3 The Semantic Enrichment Mechanism

The semantic enrichment mechanism serves as a sophisticated transformation mechanism that converts the raw BIM component data into semantically rich, structured, and queryable knowledge. This layer's effectiveness stems from its carefully orchestrated information flow, beginning with two critical input sources that provide complementary information about the BIM components.

As illustrated in Figure 6, the first input to the semantic enrichment mechanism originates from the feature fusion mechanism, which outputs unified object descriptors. These descriptors are detailed

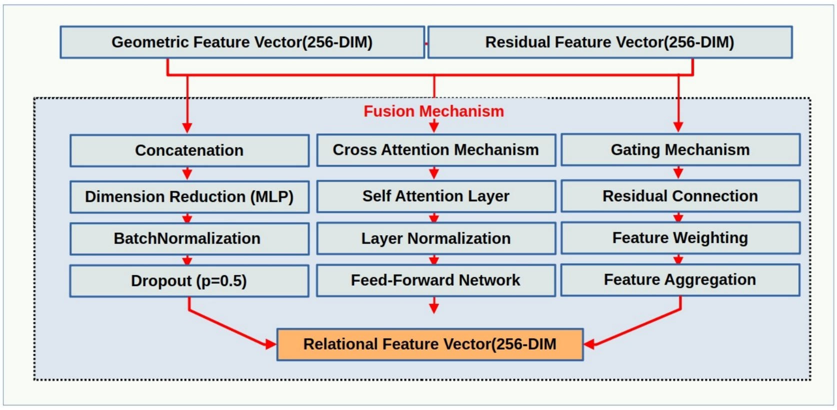

Figure 4: Architecture of the feature fusion mechanism of the proposed framework.

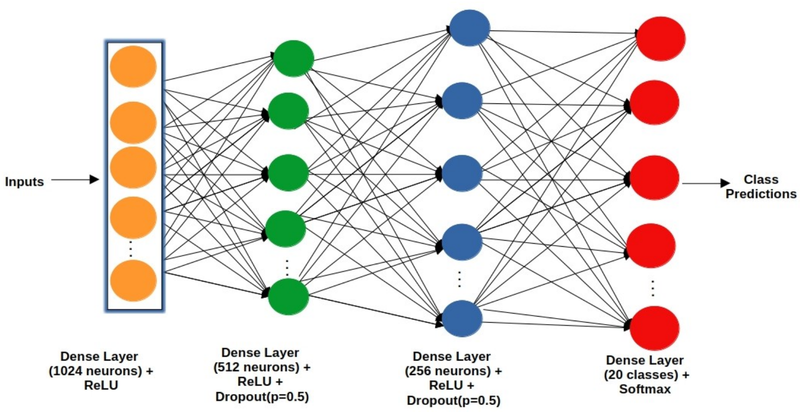

Figure 5: Architecture of the classification mechanism in the proposed framework.

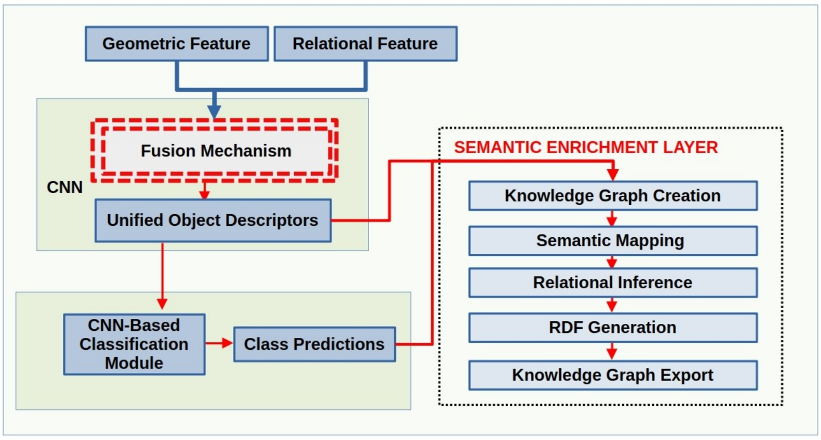

Figure 6: Architecture of the proposed semantic enrichment framework for BIM elements

numerical vectors that integrate both the geometric and relational characteristics of BIM elements. They are generated by merging data from two parallel extraction streams: one focused on geometric feature extraction (processing point clouds, meshes, and voxel grids through 3D CNNs and PointNet architectures) and the other on relational feature extraction (capturing spatial relationships, connectivity, and adjacency using GNN transformers). The unified descriptors are particularly robust because they encode multi-dimensional information, encompassing each BIM component's shape, dimensions, location, orientation, material attributes, and contextual interactions with neighboring elements. In effect, they act as a rich, machine-interpretable signature for each component, encapsulating subtleties that simple classification models might overlook. For instance, a descriptor might represent a wall as being 3 meters in height, made of concrete, load-bearing, and spatially linked to adjacent floors, ceilings, and mechanical infrastructures.

Additionally, the semantic enrichment module receives a second critical input stream from the classification module in the form of class prediction outputs. The unified object descriptors are then passed through fully connected neural network layers equipped with softmax activation functions to yield probability distributions over possible BIM component categories. These predictions provide semantic labels like *Wall*, *Door*, *HVAC Duct*, or *Structural Beam*, offering essential semantic context to support the enrichment process.

Finally, the framework's performance is evaluated on unseen BIM data to assess its generalization. By combining geometric and relational feature processing with semantic enrichment, the framework enhances the classification of architectural and MEP BIM object and also supports interoperable knowledge representation, facilitating advanced AEC applications.

3.3.1 Progressive transformation through the semantic enrichment pipeline

The initial processing stage combines the unified object descriptors and class predictions to construct a foundational knowledge graph structure, as shown in Figure 7. Each BIM element becomes a node in this graph with properties derived from its descriptor vectors and class prediction. The edges between nodes represent the explicit physical and spatial relationships detected during the relational feature extraction phase.

This knowledge graph creation process is computationally intensive, involving spatial indexing

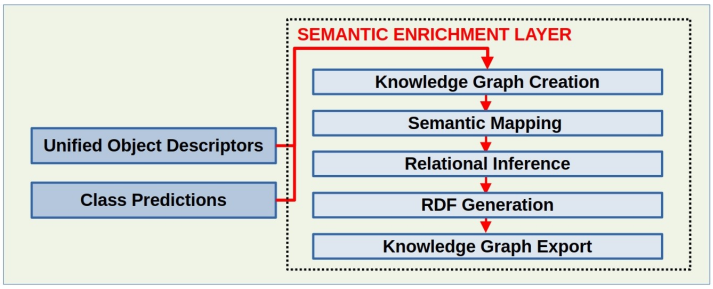

Figure 7: Progressive transformation of BIM data through the semantic enrichment pipeline

algorithms to efficiently identify relationship candidates, geometric reasoning to qualify the nature of these relationships (e.g., supports, connects-to, adjacent-to), and graph construction operations to build the initial structure. The process must handle potential inconsistencies between geometric evidence and classification results, often employing probabilistic approaches to resolve ambiguities. The resulting initial graph represents the physical building structure as a network of typed components with explicit connections but lacks the rich semantic context provided by industry standards and domain knowledge.

Once the foundational knowledge graph is constructed, as illustrated in Figure 7, it undergoes semantic mapping to align with established AEC industry ontologies. This process transforms generic nodes and relationships into semantically standardized entities by mapping them to concepts from ontologies like Industry Foundation Classes (IFC), Building Topology Ontology (BOT), or Omniclass.

As illustrated in Figure 6, the semantic enrichment layer progressively transforms these inputs through five interconnected processing components, each building upon the results of the previous stage to create increasingly sophisticated semantic representations. The mapping process of this layer employs a combination of techniques, including direct class-to-concept mapping, context-aware semantic assignment, and property-based inference. For example, a node initially classified as a *Wall* with certain geometric properties might be mapped to the more specific *ifcWall* concept with appropriate property sets. Similarly, a simple *connected-to* relationship between a pipe and a valve might be refined to *has-flow-control-element* based on domain ontologies. This mapping phase essentially transforms the graph from a project-specific representation to one that conforms to industry-standard semantics, enabling interoperability and standardized interpretation. With the graph now semantically aligned to standard ontologies, the relational inference process depicted in Figure 6 extends it by discovering implicit relationships not directly observable from the geometric or relational features. This phase applies logical reasoning, domain-specific rules, and pattern recognition to infer higher-order relationships and properties. The inference engine applies various reasoning strategies, including transitive inference (if A supports B and B supports C, then A indirectly supports C), functional inference (components of the same subsystem are functionally related), and typological inference (components of certain types typically have specific relationships). For instance, the system might infer that all electrical outlets connected to the same circuit breaker belong to a single electrical distribution system, even without explicit wiring connections in the model. Similarly, it might infer containment relationships between spaces and components based on spatial boundaries. This expansion phase significantly enriches the knowledge graph with semantically meaningful relationships that extend beyond what is explicitly modeled in the original BIM data.

After inference expands the semantic richness of the knowledge graph, the RDF generation component formalizes this knowledge into standardized RDF triples. Each triple follows the subjectpredicate-object structure, representing statements about BIM components and their relationships in a W3C-standardized format. This formalization process involves careful vocabulary selection, URI assignment for uniquely identifying each entity, datatype specification for literals, and logical structuring of statements. For instance, a wall supporting a beam might be represented as *ex:Wall_123 bot:supports ex:Beam_456*. The process must handle complex property chains, reification for statements about statements, and proper namespace management. The resulting RDF model provides a standardized way to express the semantic knowledge in a machine-readable format that aligns with semantic web technologies and can be processed by standard reasoning engines and query processors.

The formalized RDF representation is then packaged through the knowledge graph export component of Figure 6 to make it accessible for external applications and systems. This process involves serializing the RDF data into appropriate formats (Turtle), organizing the data into coherent datasets, and providing necessary metadata for proper interpretation. The export process implements versioning mechanisms to track changes over time, modularization strategies to separate different aspects of the building model (structural, mechanical, electrical), and compression techniques to manage large datasets efficiently. It may also include differential export capabilities to capture only changes since the previous version, reducing data transfer overhead for incremental updates.

In the final stage, the knowledge graph is exported and is available for sophisticated interrogation through semantic query interfaces. The semantic queries component implements SPARQL endpoints and other query mechanisms that allow users and applications to extract specific information from the semantically enriched BIM model.

This query capability transforms the enriched BIM data from a static representation into an interactive knowledge base that can answer complex questions about the building. Users can perform queries like *Find all HVAC components that serve meeting rooms on the third floor* or *Identify all fire egress routes that pass through areas with flammable materials*. The query engine implements optimization strategies for efficient execution, caching mechanisms for frequently used queries, and result formatting options for different consumption patterns. This querying capability is ultimately what makes the semantic enrichment valuable in practical applications, allowing stakeholders to extract precisely the information they need from the rich semantic model.

This seamless flow of information enables the translation between the geometric/physical world of building components and the semantic world of meaning, function, and relationships. The resulting semantically enriched BIM model supports advanced applications in design validation, facility management, regulatory compliance checking, and system analysis that would be impossible with traditional BIM approaches focused primarily on geometry and explicit relationships.

4 Experimental Evaluation

The proposed AI-driven framework aims to enhance the automated classification and semantic enrichment of architectural and MEP BIM objects. It leverages Scikit-Learn, Keras, TensorFlow (with GPU support), and the SNAP tool, seamlessly integrated within Google Colab—an extensive open-source Jupyter notebook platform designed for machine learning. Section 4.2 outlines the dataset and preprocessing techniques utilized to derive the findings. Subsequently, Sections 4.3 and 4.4 provide an in-depth discussion of the BIM objects evaluation metrics and result analysis.

4.1 Model training configuration

A total of 12,000 experimental samples—comprising 600 instances per class—were partitioned into 70% for training and 30% for testing. The geometric feature extractor, relational feature extractor, and feature fusion module were each trained over 50 epochs using the Adam optimizer with an initial learning rate of 0.001. A ReduceLROnPlateau scheduler was employed to dynamically reduce the learning rate based on validation performance. Training was conducted with a batch size of 32, and

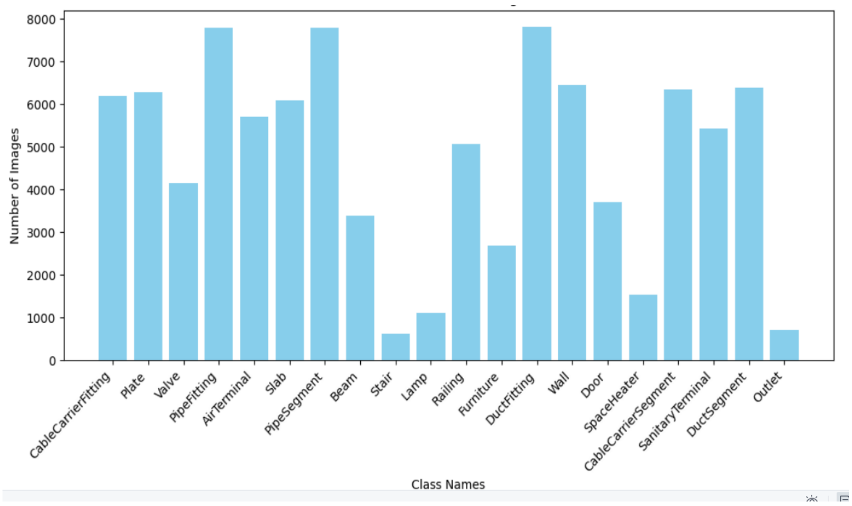

Figure 8: Training set distribution of BIM object classes showcasing significant imbalance across 20 classes.

model optimization was guided by minimizing the cross-entropy loss function throughout the training process.

Similarly, the classification model was trained for 50 epochs using 1,903 fused features for training and 476 for testing. The same optimization setup—Adam optimizer, an initial learning rate of 0.001, ReduceLROnPlateau scheduler, and a batch size of 32—was applied during this phase to ensure consistency across all training modules.

4.2 Dataset description and preprocessing

This study employs the IFCNet dataset Emunds et al. (2021) for the experimental evaluation of the proposed BIM object classification and semantic enrichment framework. IFCNet comprises 19,613 single-entity IFC files, spanning 65 classes that encapsulate both geometric and semantic information, making it a robust resource for BIM-related tasks. The dataset includes a diverse range of architecural and MEP BIM elements, from structural components (e.g., Beams, Walls) to Mechanical, Electrical, and Plumbing (MEP) elements (e.g., Duct Segments, Valves), providing a comprehensive benchmark for real-world architectural and construction applications.

The diversity and complexity of the IFCNet dataset make it an ideal benchmark for evaluating BIM object classification systems. Its inclusion of both geometric and semantic data ensures that the proposed framework is rigorously tested for real-world AEC applications, such as automated design validation and facility management.

The IFCNet dataset exhibits significant class imbalance, with sample counts per class ranging from 624 to 7,800, as detailed in Table 2. Dominant classes, primarily MEP components, include Duct Fitting (7,800 samples), Pipe Segment (7,788 samples), and Pipe Fitting (7,776 samples), while structural and specialized elements, such as Stair (624 samples), Outlet (708 samples), and Lamp (1,104 samples), are underrepresented. This imbalance, visualized for the training and testing sets in Figures 8 and 9, respectively, poses a challenge for supervised learning, as the model must effectively classify both prevalent and rare classes. The class distribution is consistent between the training and testing sets, ensuring reliable model evaluation.

To mitigate the effects of class imbalance, a balanced subset of 600 samples per class was

Table 2: Distribution of architecural and MEP BIM elements in the IFCNet dataset.

BIM Element Class	**Component Category**	**Samples**	**Percentage (%)**
Duct Fitting	MEP	7,800	8.20
Pipe Segment	MEP	7,788	8.18
Pipe Fitting	MEP	7,776	8.17
Duct Segment	MEP	6,372	6.70
Air Terminal	MEP	5,700	5.99
Sanitary Terminal	MEP	5,424	5.70
Valve	MEP	4,152	4.36
Space Heater	MEP	1,524	1.60
		46,536	**48.90**
Cable Carrier Segment	Electrical	6,348	6.67
Cable Carrier Fitting	Electrical	6,192	6.51
Lamp	Electrical	1,104	1.16
Outlet	Electrical	708	0.74
		14,352	**15.08**
Wall	Architectural	6,444	6.77
Railing	Architectural	5,064	5.32
Door	Architectural	3,708	3.90
Stair	Architectural/Circulation	624	0.66
		15,840	**16.65**
Plate	Structural/Architectural	6,276	6.60
Slab	Structural	6,084	6.39
Beam	Structural	3,384	3.56
		15,744	**16.55**
Furniture	Interior	2,688	2.82
Total		**95,160**	**100**

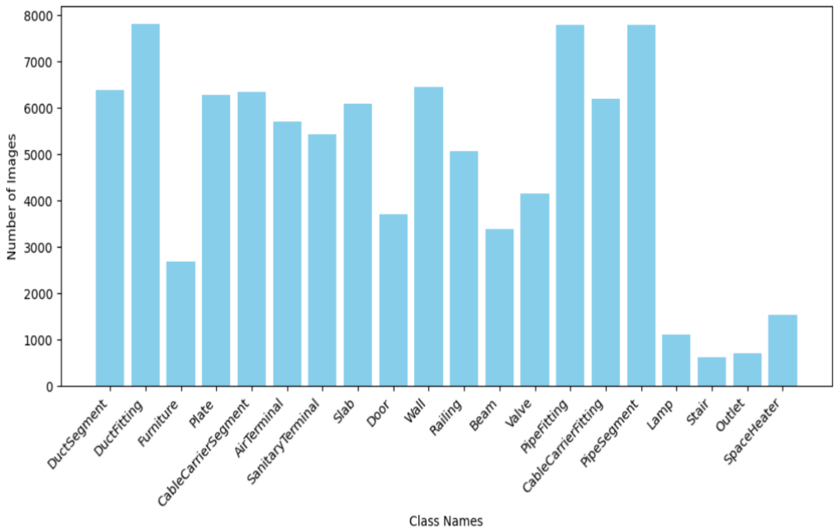

Figure 9: Testing set distribution of BIM object classes showcasing significant imbalance across 20 classes.

selected for model training, as shown in Figure 10. This subset, comprising 20 uniquely labeled BIM elements, is split into distinct training and testing directories to support model development and evaluation. By balancing the dataset, the framework can better learn to classify underrepresented classes, enhancing its robustness and generalization.

Figure 11 and Figure 12 also shows samples of BIM element dataset and augmentations. The figures shows representative samples from the BIM object classification dataset used in this research. Figure 11 depicts the original image samples from several BIM element classes demonstrating the variety of components encountered in BIM, including structural, mechanical, and electrical elements. Figure 12 depicts the augmented samples of the samples shown in Figure 11, showing the effectiveness of preprocessing applied to enhance model training robustness, including rotations, scaling variations, and normalization. Each image has been standardized to 64×64 pixels with RGB channels and normalized using ImageNet statistics. These augmentations help the model generalize across different orientations and scales while mitigating the effects of class imbalance in the original dataset. The augmented dataset provided 600 balanced samples per class for model training.

4.2.1 Data preparation for geometric and relational feature extraction

In the input data preparation phase, all 2D rendered images of BIM elements were resized to 64×64 pixels and converted into PyTorch tensors. Standard normalization was applied using ImageNet mean and standard deviation values to ensure consistent input scaling for RGB channels. Each processed image tensor was saved as a .pt file, accompanied by metadata that includes class sizes and dataset statistics, such as class distribution and total sample counts. For the two-branch architecture, distinct feature types were extracted as follows:

1. For geometric feature extraction, the normalized 3×64×64 image tensors served as the primary input, preserving the visual and shape characteristics of each BIM element. These tensors were organized into class-specific directories for both training and testing sets, facilitating input to 2D CNNs for geometric feature extraction.

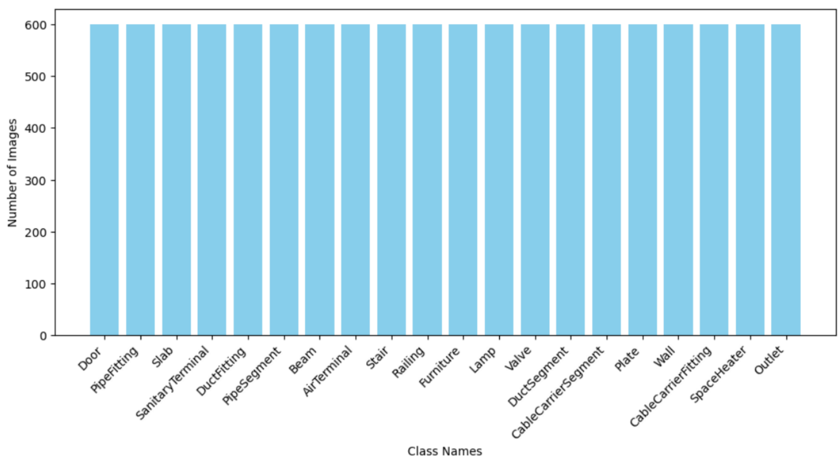

Figure 10: Balanced class distribution in the training data for BIM object classification across 20 classes.

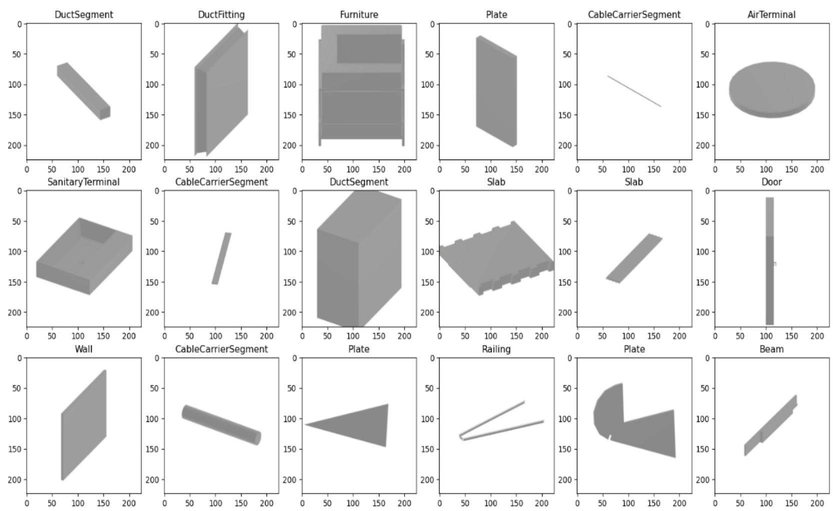

Figure 11: 3D visualization of 18 original training samples of BIM elements

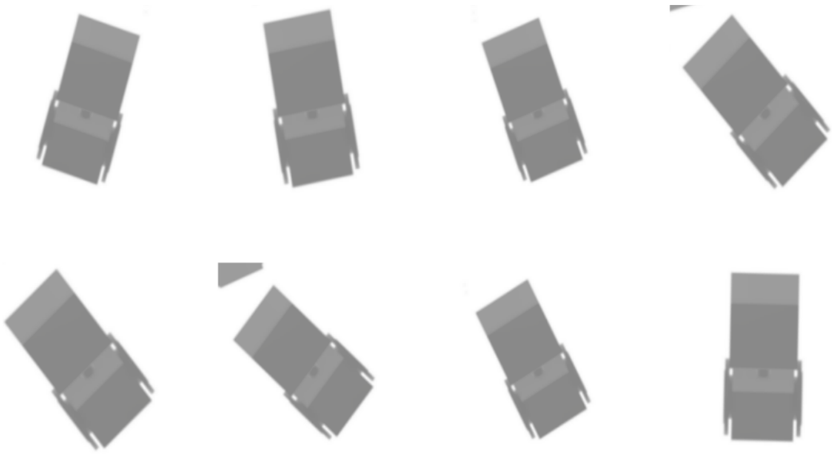

Figure 12: 3D visualization of augmented training samples of BIM elements

2. For relational feature extraction, binary edge maps were generated as a proxy for structural boundaries. This process involved converting RGB images to grayscale, computing gradient magnitudes, and applying a threshold of 0.1 to produce 1×64×64 binary edge map tensors. These edge maps highlight the structural outlines of each BIM element, which were subsequently used to construct graph representations for connectivity analysis in the relational branch using GNNs or Transformers.

Visualization of both feature types was performed, with representative samples from each of the 20 BIM classes presented in Figure 13 and Figure 14 The geometric features retain color and texture information, while the edge maps emphasize boundaries and structural elements, serving as inputs to the relational branch of the architecture.

4.3 Evaluation Metrics for BIM Object Classification Model

This section evaluates the performance of the proposed BIM object classification framework using metrics derived from a confusion matrix, a critical tool for assessing classifier effectiveness (Tharwat, 2021). The confusion matrix provides a comprehensive summary of prediction outcomes by tabulating correct and incorrect classifications across all classes, enabling the computation of key performance indicators such as accuracy, precision, recall, and the F1-score. These metrics collectively offer insights into the model's precision, reliability, and overall classification capability, guiding further refinements to enhance performance.

The confusion matrix organizes predictions into a two-dimensional table, capturing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). From these, accuracy is calculated as the ratio of correctly predicted instances (TP + TN) to the total number of samples, including both positive (P) and negative (N) cases, as shown in Equation (4.1). An accuracy of 1.0 indicates perfect classification, while 0.0 signifies complete misclassification (Vujovic, 2021). This metric provides a broad measure of the model's overall performance but may be less informative for imbalanced datasets.

$$\text{Accuracy} = \frac{TP + TN}{P + N} \tag{4.1}$$

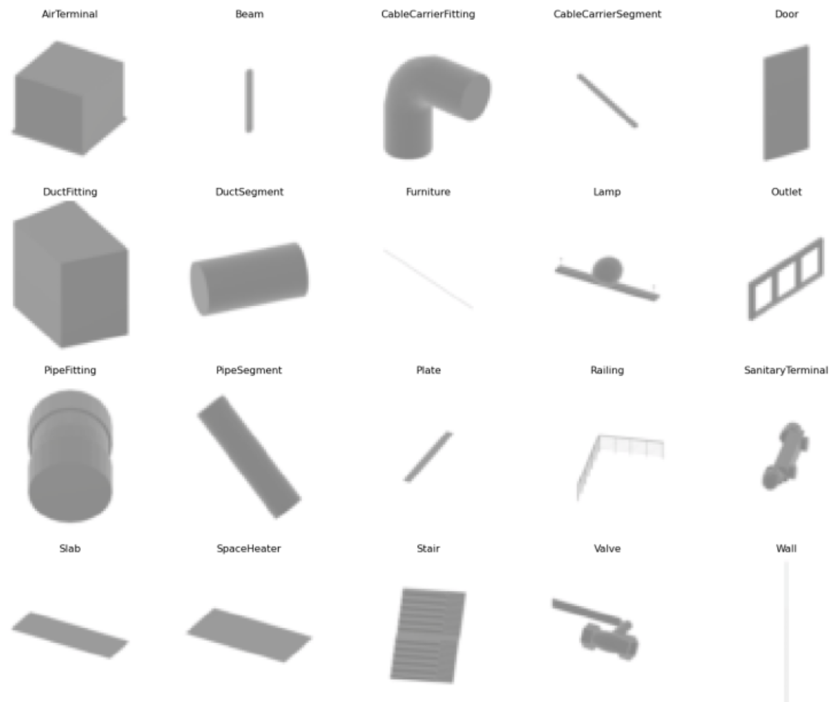

Figure 13: 3D geometric representations of training BIM elements

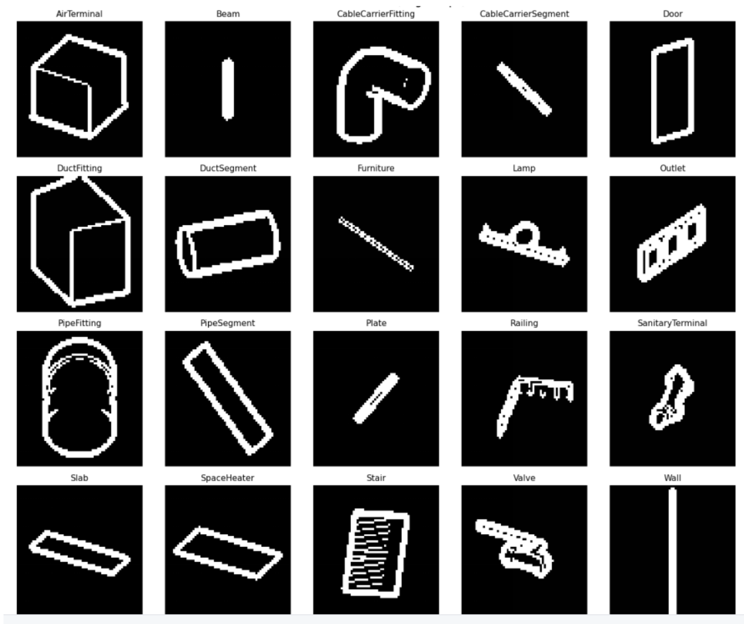

Figure 14: BIM element edge maps for relational features

Precision and recall offer more granular insights, particularly for imbalanced classes. Precision, defined in Equation (4.2), measures the proportion of correctly identified positive instances among all predicted positives (TP / (TP + FP)), reflecting the model's ability to avoid false positives. Recall quantifies the fraction of actual positive instances correctly detected (TP / (TP + FN)), as shown in Equation (4.3). Recall is crucial for ensuring the model identifies all relevant instances, especially in applications where missing positives is costly (Vujovic, 2021).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

The F1-score, calculated as the harmonic mean of precision and recall (Equation (4.4)), balances these two metrics to provide a single measure of performance, making it particularly valuable for datasets with uneven class distributions (Heydarian et al., 2022). By integrating precision and recall, the F1-score ensures a comprehensive evaluation of the model's ability to correctly classify BIM objects while minimizing both false positives and false negatives. Together, these metrics enable a thorough assessment of the framework's classification performance, highlighting its strengths and areas for improvement in real-world AEC applications.

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

4.4 Results and Analysis

The proposed framework for BIM object classification and semantic enrichment is implemented by selecting a balanced subset of 600 images per BIM object class. As shown in Figure 1, the model integrates geometric and relational feature extraction through dedicated neural network branches, followed by a CNN-based feature fusion module that combines these features into unified object descriptors. These descriptors enable accurate BIM object classification, producing class predictions that support subsequent semantic enrichment.

4.5 Geometric feature descriptor model evaluation

Figure 15 illustrates the trends of loss and accuracy throughout the training and validation phases of the geometric feature descriptor model over 50 epochs. The model exhibited consistent improvement, with validation accuracy increasing from 23.58% at the first epoch to 66.04% by epoch 45, where it peaked. By epoch 15, validation accuracy reached 56.03%, indicating substantial learning. Training accuracy rose steadily to 86.79% by epoch 50, with a final test accuracy of 65.66%, reflecting robust performance. The model exhibited some fluctuations in validation accuracy during training, suggesting potential overfitting at certain stages. However, the learning rate scheduler helped stabilize performance by reducing the learning rate when validation loss plateaued.

The confusion matrix given in Figure 16 reveals varying performance across different BIM element classes. Some classes like Outlet (precision 1.00), Railing (precision 0.82), and CableCarrierFitting (precision 0.78) showed strong precision, while others like Lamp (precision 0.27) and SpaceHeater (precision 0.27) were more challenging for the model to identify correctly.

The model achieved an overall validation accuracy of 65.69%, with average precision of 63.50%, average recall of 59.07%, and average F1-score of 59.92%, indicating moderate classification performance across the 20 BIM element classes. In the results for accuracy (see Table 3), the BIM element class, *Outlet* has the highest accuracy of 99.54 and the *DuctSegment* has the lowest accuracy of 93.78%, while the mean accuracy for all the element is 96.57 %.

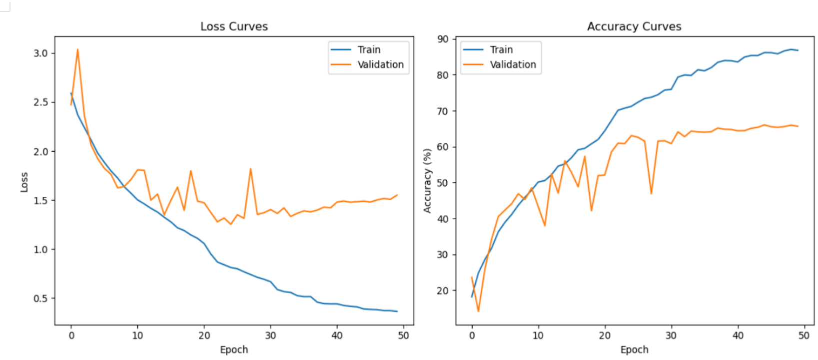

Figure 15: Trends in training and validation accuracy and loss across epochs for the geometric feature descriptor model.

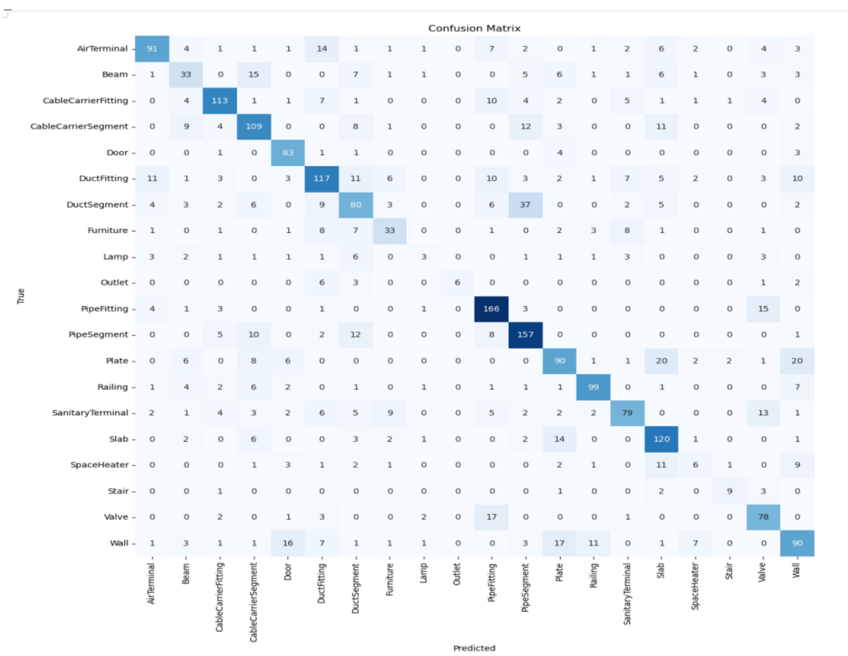

Figure 16: Confusion matrix for the geometric feature descriptor model performance evaluation

Table 3: Class-wise performance evaluation metrics for the geometric feature descriptor model used for geometric object representation

Class	Precision	Recall	F1-Score	Accuracy
AirTerminal	0.7647	0.6408	0.6973	0.9668
Beam	0.4521	0.3929	0.4204	0.9617
CableCarrierFitting	0.7847	0.7290	0.7559	0.9693
CableCarrierSegment	0.6488	0.6855	0.6667	0.9542
Door	0.6917	0.8925	0.7793	0.9802
DuctFitting	0.6393	0.6000	0.6190	0.9394
DuctSegment	0.5369	0.5031	0.5195	0.9378
Furniture	0.5690	0.4925	0.5280	0.9752
Lamp	0.2727	0.1111	0.1579	0.9865
Outlet	1.0000	0.3529	0.5217	0.9954
PipeFitting	0.7186	0.8557	0.7812	0.9609
PipeSegment	0.6767	0.8051	0.7354	0.9525
Plate	0.6122	0.5732	0.5921	0.9479
Railing	0.8182	0.7795	0.7984	0.9790
SanitaryTerminal	0.7248	0.5809	0.6449	0.9634
Slab	0.6316	0.7895	0.7018	0.9571
SpaceHeater	0.2727	0.1579	0.2000	0.9798
Stair	0.6923	0.5625	0.6207	0.9954
Valve	0.6904	0.7500	0.6724	0.9680
Wall	0.5844	0.5590	0.5714	0.9432

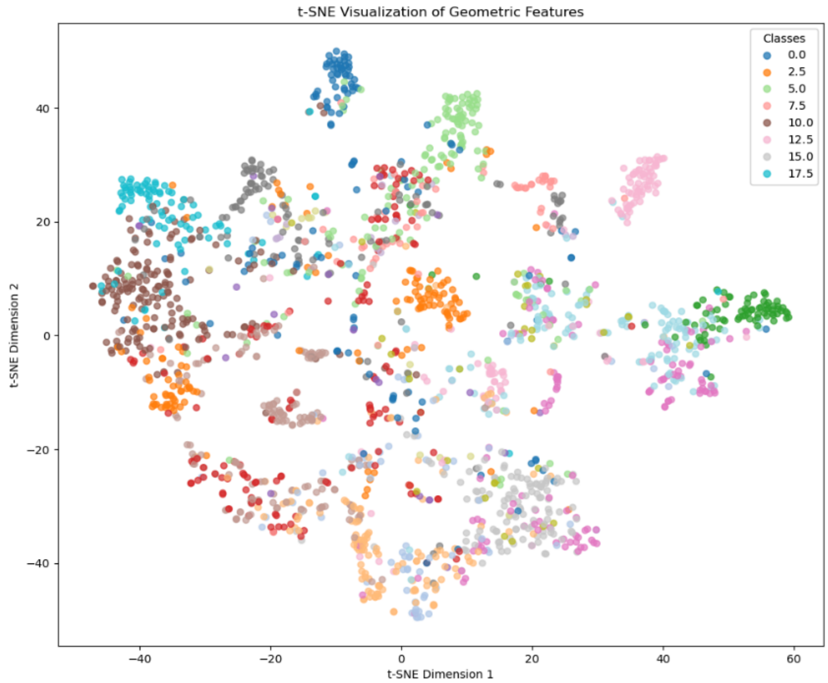

Figure 17: The t-SNE visualization of the 256-dimensional geometric features

The t-SNE visualization of the extracted geometric features depicted in Figure 17 also demonstrates how well the model separates different classes in the feature space. Classes with distinct geometric characteristics form clear clusters, while classes with similar appearances show some overlap. The final model evaluation reveals a weighted average F1-score of 0.65 across all classes. The model performs particularly well on Door (F1-score of 0.78), PipeFitting (F1-score of 0.78), and Railing (F1-score of 0.80) classes. However, performance is weaker on classes with high visual similarity, such as Lamp (F1-score of 0.16) and SpaceHeater (F1-score of 0.20). These results highlight areas where the integration of relational features might significantly improve classification performance.

4.6 Performance for the GNN Transformer model used for the relational feature descriptor

The GNN Transformer model used for the relational feature extraction demonstrated steady improvement throughout training (see Figure 18), starting with an initial validation accuracy of 24.09% after the first epoch. By epoch 10, the validation accuracy reached 51.07%, showing significant early learning. The best validation performance was achieved at epoch 27 with an accuracy of 60.32%. The training accuracy continued to improve throughout all 50 epochs, with the final training accuracy reaching 94.77% and a test accuracy of 58.85%.

The model exhibited more pronounced fluctuations in validation accuracy during training compared to the geometric branch, suggesting higher sensitivity to the relational features. The gap between training and validation accuracy widened in later epochs, indicating some overfitting despite the regularization measures.

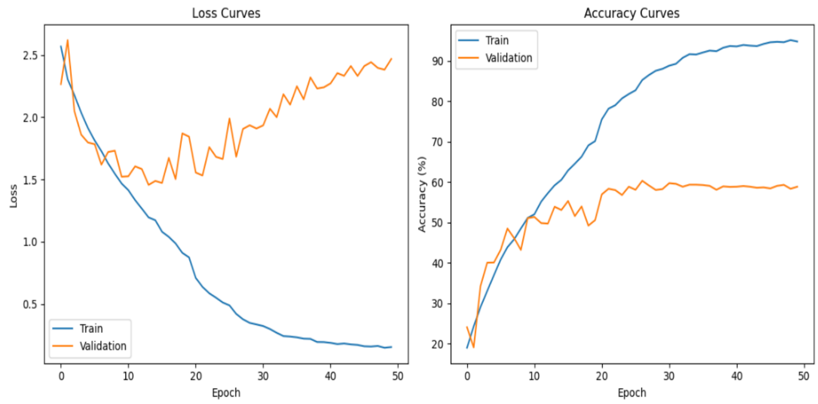

Figure 18: Trends in training and validation accuracy and loss across epochs for the relational feature descriptor model.

The confusion matrix of Figure 19 also reveals varying performance across BIM element classes. Some classes like Door (precision of 0.76), SanitaryTerminal (precision of 0.73), and Railing (precision of 0.78) showed strong precision, while others like SpaceHeater (precision of 0.10) and Lamp (precision of 0.21) were particularly challenging for the relational branch to identify correctly.

The model achieved an overall accuracy of 58.87%, with a macro average precision of 55.43%, macro average recall of 52.64%, and macro average F1-score of 53.36%, indicating moderate classification performance across the 20 BIM element classes. As shown in Figure 4, the BIM element class, "Stair" has the highest accuracy of 99.41 %, "DuctSegment" has the lowest accuracy of 93.17%, while the mean accuracy for all the element is 95.89 %.

The t-SNE visualization of the extracted relational features demonstrates how the model separates different classes in the feature space. The clustering of elements based on relational features provides complementary information to the geometric features, potentially capturing different aspects of the BIM elements.

The final model evaluation reveals a weighted average F1-score of 0.59 across all classes. The model performs particularly well on Door (F1-score of 0.76), PipeFitting (F1-score of 0.71), and Valve (F1-score of 0.68) classes. However, performance is weaker on classes with complex relationship patterns, such as SpaceHeater (F1-score of 0.10) and Beam (F1-score of 0.33). These results highlight areas where integration with geometric features might significantly improve classification performance

4.7 Performance of the CNN-based model used for feature fusion.

The CNN-based model use for the feature fusion demonstrated rapid learning (see Figure 21) achieving a validation accuracy of 63.66% after just the first epoch. The best validation performance was recorded at epoch 4 with an accuracy of 66.60%. Throughout training, the model showed a steady increase in training accuracy, reaching 94.48% by the final epoch, while validation accuracy fluctuated, suggesting some overfitting despite regularization measures.

As shown in Figure 21, the gap between training and validation accuracy widens significantly after epoch 10. While training accuracy continues to improve, validation accuracy plateaus. This pattern suggests the model was learning features specific to the training set, which did not generalize well to the validation set.

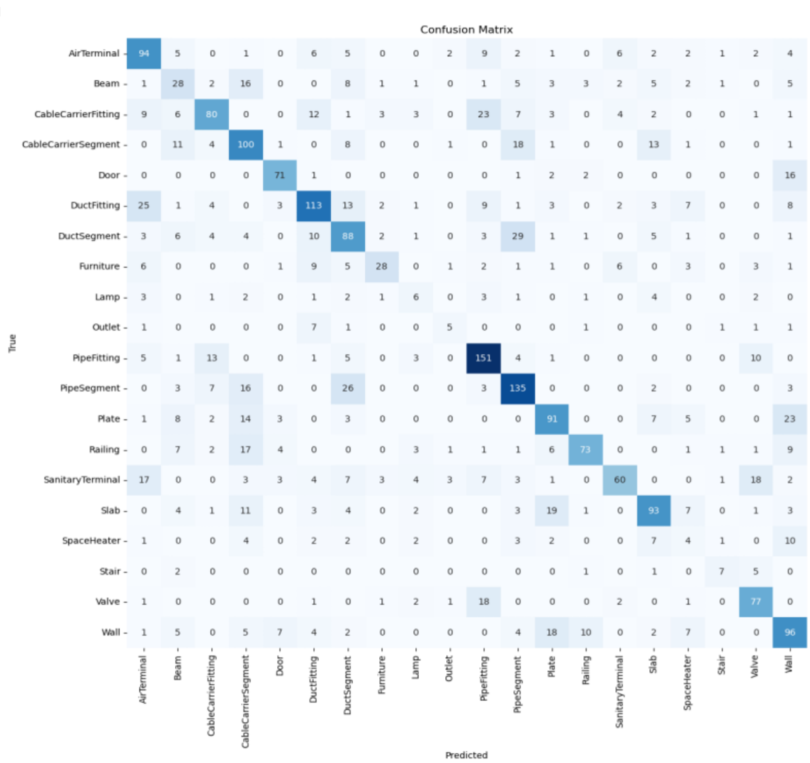

Figure 19: Confusion matrix for the relational feature descriptor model performance evaluation

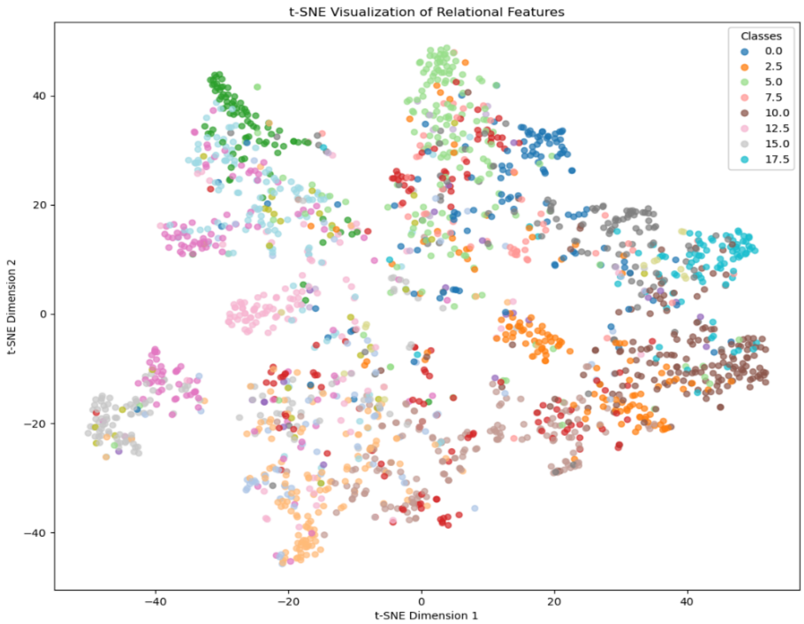

Figure 20: The t-SNE visualization of the 256-dimensional relational features

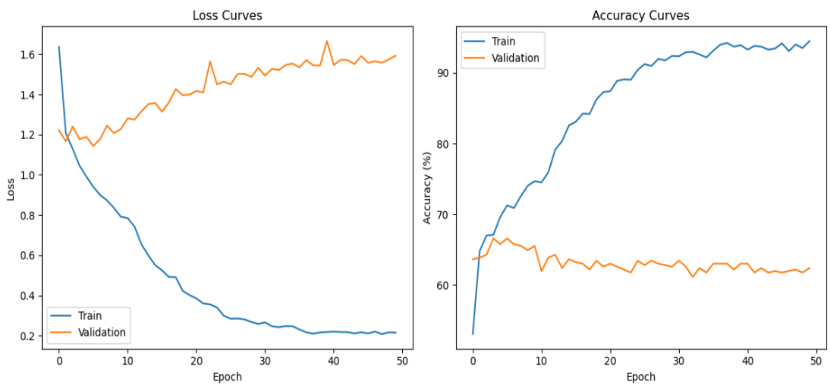

Figure 21: Trends in training and validation accuracy and loss across epochs for the feature fusion model.

Table 4: Class-wise performance metrics for the GNN Transformer model used for the relational feature descriptor

Class	Precision	Recall	F1-Score	Accuracy
AirTerminal	0.5595	0.6620	0.6065	0.9487
Beam	0.3218	0.3333	0.3275	0.9516
CableCarrierFitting	0.6667	0.5161	0.5818	0.9516
CableCarrierSegment	0.5181	0.6289	0.5682	0.9361
Door	0.7634	0.7634	0.7634	0.9815
DuctFitting	0.6494	0.5795	0.6125	0.9399
DuctSegment	0.4889	0.5535	0.5192	0.9315
Furniture	0.6829	0.4179	0.5185	0.9781
Lamp	0.2143	0.2222	0.2182	0.9819
Outlet	0.3571	0.2941	0.3226	0.9912
PipeFitting	0.6565	0.7784	0.7123	0.9487
PipeSegment	0.6193	0.6923	0.6538	0.9399
Plate	0.5948	0.5796	0.5871	0.9462
Railing	0.7849	0.5748	0.6636	0.9689
SanitaryTerminal	0.7317	0.4412	0.5505	0.9588
Slab	0.6370	0.6118	0.6242	0.9529
SpaceHeater	0.0976	0.1053	0.1013	0.9701
Stair	0.5833	0.4375	0.5000	0.9941
Valve	0.6364	0.7404	0.6844	0.9701
Wall	0.5217	0.5963	0.5565	0.9357

As illustrated in the accuracy curves of Figure 21, the fusion model attained a peak validation accuracy of 66.60%, slightly outperforming the geometric branch (66.04%) and the relational branch (60.32%) individually. This marginal gain implies that integrating both feature types yields complementary insights, thereby enhancing the classification of certain BIM elements.

The model recorded an overall classification accuracy of 62.39%, with a macro average precision of 56.06%, macro average recall of 59.63%, and a macro average F1-score of 56.58%. These results reflect moderate performance in classifying the 20 BIM element categories, with relatively balanced precision and recall across classes.

According to Table 5, class-wise accuracy varied across BIM elements. The highest accuracy was observed for the *Stair* class at 99.79%, while *DuctFitting* showed the lowest at 92.44%. The average accuracy across all classes stood at 96.24%.

The attention weights indicate that the model effectively learned to emphasize geometric features for elements with unique shapes (e.g., Doors and Railings), and relational features for components defined by their structural connections (e.g., Pipes and Ducts). This dynamic feature prioritization highlights the advantage of the fusion strategy compared to simple feature concatenation.

Table 5: Class-wise Performance Metrics for the Convolutional Neural Network (CNN) Model used for the Feature Fusion

Class	Precision	Recall	F1-Score	Accuracy
AirTerminal	0.6452	0.7407	0.6897	0.9622
Beam	0.5000	0.3333	0.4000	0.9685
CableCarrierFitting	0.7500	0.7241	0.7368	0.9685
CableCarrierSegment	0.6389	0.6970	0.6667	0.9517
Door	0.6000	0.7500	0.6667	0.9811
DuctFitting	0.6667	0.4348	0.5263	0.9244
DuctSegment	0.6207	0.5806	0.6000	0.9496
Furniture	0.2500	0.1538	0.1905	0.9643
Lamp	0.1667	0.2500	0.2000	0.9832
Outlet	0.2000	0.5000	0.2857	0.9895
PipeFitting	0.7949	0.7949	0.7949	0.9664
PipeSegment	0.5897	0.6765	0.6301	0.9433
Plate	0.6000	0.6154	0.6076	0.9349
Railing	0.8333	0.7692	0.8000	0.979
SanitaryTerminal	0.5250	0.6563	0.5833	0.937
Slab	0.5789	0.6471	0.6111	0.9412
SpaceHeater	0.4000	0.2500	0.3077	0.9811
Stair	0.5000	1.0000	0.6667	0.9979
Valve	0.7647	0.7647	0.7647	0.9832
Wall	0.5882	0.5882	0.5882	0.9412

4.8 Evaluation of the classification model's effectiveness in BIM object categorization.

The classification model demonstrated remarkably rapid convergence, as illustrated in Figure 23. It achieved a validation accuracy of 87.82% after just the first epoch, indicating that the fused features from the geometric and relational branches already provided highly discriminative information for BIM object classification. The peak validation accuracy of 88.66% was reached at epoch 25, while the final test accuracy settled at 86.97%.

The training trajectory revealed a consistent and stable learning pattern. The model quickly attained high accuracy and maintained steady performance throughout training, with the final training accuracy reaching 97.95%. The relatively narrow gap between training and validation accuracies suggests that the model generalized well to unseen data, indicating minimal overfitting despite the high-dimensional input features.

The confusion matrix, summarizing class-wise predictions for the CNN model, is presented in Figure 24. In addition, detailed class-wise performance metrics are provided in Table 6. The model delivered robust classification performance across most BIM element categories, achieving an overall validation accuracy of 86.97% and a weighted average F1-score of 0.87, reflecting balanced precision and recall.

A closer look at class-specific performance reveals:

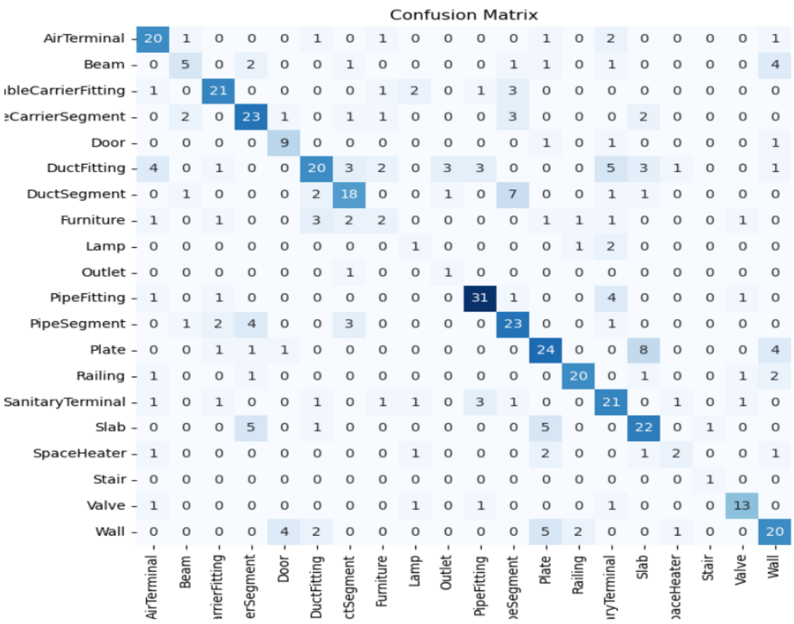

Figure 22: Confusion matrix for the feature fusion model performance evaluation

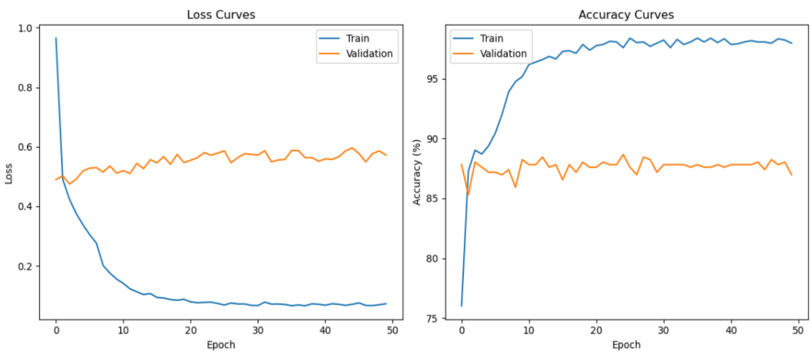

Figure 23: Trends in training and validation accuracy and loss across epochs for the CNN model used for the BIM object classification.

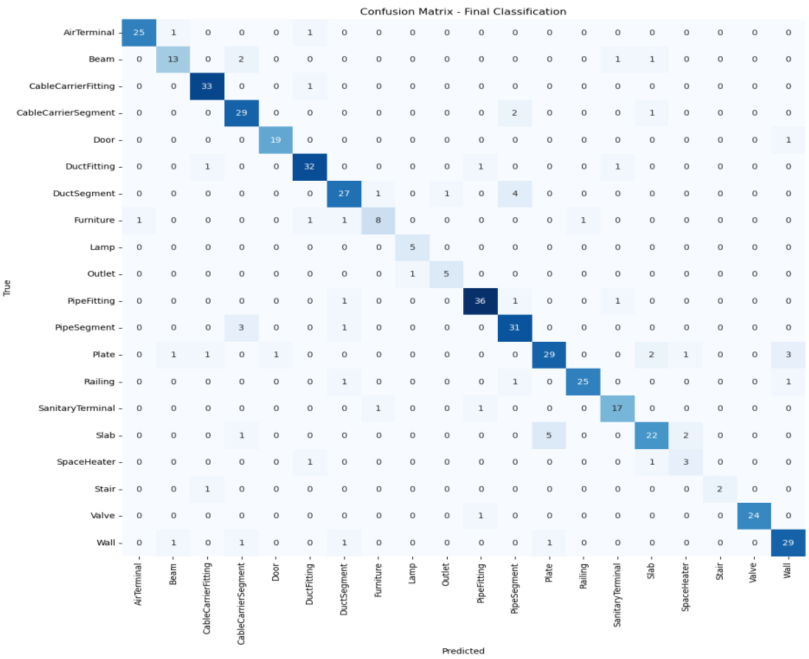

Figure 24: Confusion matrix showing classification performance across all BIM element classes for the CNN model used for the BIM object classification

1. High-performing classes (F1-score $>$ 0.90): AirTerminal (0.94), CableCarrierFitting (0.94), Door (0.95), Lamp (0.91), Valve (0.98)

2. Medium-performing classes (F1-score between 0.80 and 0.90): Beam (0.79), DuctFitting (0.90), DuctSegment (0.83), PipeFitting (0.92), PipeSegment (0.84), Railing (0.93), SanitaryTerminal (0.87), Wall (0.87)

3. Lower-performing classes (F1-score $<$ 0.80): Furniture (0.73), Plate (0.79), Slab (0.77), SpaceHeater (0.55), Stair (0.80)

These results highlight the effectiveness of the model in classifying complex BIM elements, while also identifying areas where performance could be further improved, particularly for underrepresented or visually ambiguous categories.

Significantly, the classification results offer key insights, demonstrating that the fused features yield superior classification performance (86.97%) compared to relying solely on the geometric branch (65.66%) or relational branch (58.85%). This validates the complementary nature of both feature types. Classes with distinctive shapes and relational patterns—such as Door, Valve, and AirTerminal—achieve the highest accuracy, benefiting from both feature types. Meanwhile, classes that posed challenges for the individual branches, including Lamp and Stair, show marked improvement, suggesting that their unique characteristics emerge only when both geometric and relational aspects are integrated. However, SpaceHeater remains the most challenging class (with F1-score of 0.55), likely due to its visual similarity to other elements and limited training examples.

Table 6: Class-wise performance metrics for the CNN-based model used for the BIM object classification.

Class	Precision	Recall	F1_Score	Accuracy
AirTerminal	0.9615	0.9259	0.9434	0.9259
Beam	0.8125	0.7647	0.7879	0.7647
CableCarrierFitting	0.9167	0.9706	0.9429	0.9706
CableCarrierSegment	0.8056	0.9063	0.8529	0.9063
Door	0.9500	0.9500	0.9500	0.9500
DuctFitting	0.8889	0.9143	0.9014	0.9143
DuctSegment	0.8438	0.8182	0.8308	0.8182
Furniture	0.8000	0.6667	0.7273	0.6667
Lamp	0.8333	1.0000	0.9091	1.0000
Outlet	0.8333	0.8333	0.8333	0.8333
PipeFitting	0.9231	0.9231	0.9231	0.9231
PipeSegment	0.7949	0.8857	0.8378	0.8857
Plate	0.8286	0.7632	0.7945	0.7632
Railing	0.9615	0.8929	0.9259	0.8929
SanitaryTerminal	0.8500	0.8947	0.8718	0.8947
Slab	0.8148	0.7333	0.7719	0.7333
SpaceHeater	0.5000	0.6000	0.5455	0.6000
Stair	1.0000	0.6667	0.8000	0.6667
Valve	1.0000	0.9600	0.9796	0.9600
Wall	0.8529	0.8788	0.8657	0.8788

4.9 Performance evaluation and outcomes of the semantic enrichment module in the proposed BIM framework.

Figure 25 presents a sample BIM semantic graph, illustrating the intricate interconnections between various BIM elements. This semantic graph is constructed by leveraging inputs derived from both the class classification module and the fused feature descriptor model, which integrate geometric and relational features into a unified representation. Additionally, the graph's data is systematically stored within a semantic database, enabling efficient querying to retrieve relational details and class information associated with the BIM elements, thereby supporting advanced applications in AEC workflows such as automated design validation and facility management.

4.10 Comparison with leading research for BIM objects classification and semantic enrichment

.

To assess the effectiveness of the proposed framework for MEP BIM object classification and semantic enrichment, this research conducted a comparative analysis against existing state-of-the-art approaches in the domain. Unlike many prior studies that primarily focus either on object classification or semantic enrichment in isolation, the approach adopted in this research integrates both geometric

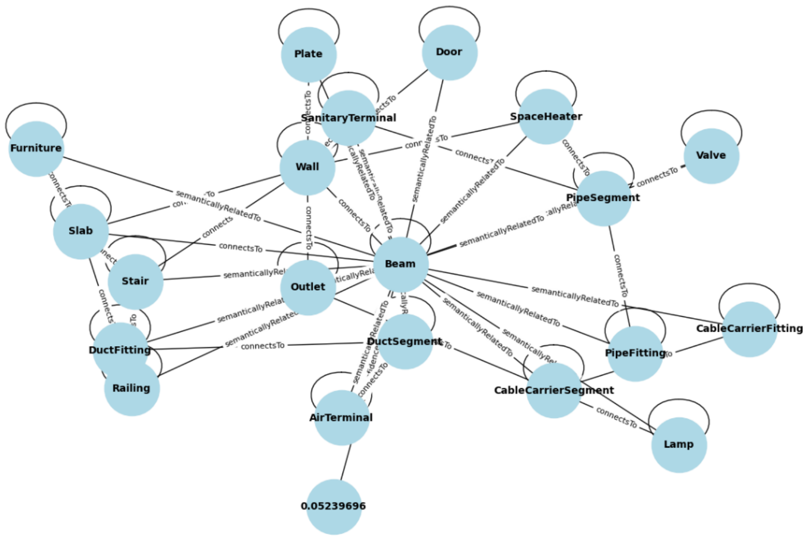

Figure 25: Knowledge Graph Representation of BIM Components and Their Semantic Relationships

and relational features, a strategy that remains underexplored—particularly in the context of MEP components.

Existing literature often emphasizes either BIM object classification without incorporating semantic enrichment, or focuses solely on semantic enrichment without considering object classification. This disjointed approach limits the ability to achieve full interoperability among AEC stakeholders, which is essential for seamless collaboration and data exchange throughout a building's lifecycle.

As a result, the methods proposed in earlier studies may fall short when applied to the dual task of accurately classifying MEP BIM objects and enriching the BIM database with semantically rich, interoperable information. In contrast, the proposed framework leverages the IFCNet dataset—specifically developed to support interoperability and domain-specific information exchange in BIM workflows—for thorough and realistic evaluation.

The model proposed in this researchsuccessfully classifies 20 distinct MEP BIM object categories, demonstrating superior accuracy and semantic enrichment capabilities when compared to existing solutions (see Table 7). These results underscore the practical relevance and technical advantage of the proposed framework in advancing BIM intelligence and interoperability in AEC applications.

5 Discussion

The Architecture, Engineering, and Construction (AEC) industry is increasingly embracing Building Information Modeling (BIM) as a foundational tool for simulating and managing real-world entities through semantically rich 3D models Mirarchi et al. (2024). As BIM adoption expands, its application has extended into diverse domains, including health monitoring, maintenance and repair, energy management, emergency response, and security planning. For instance, Panah and Kioumarsi (2021) underscore BIM's utility in supporting the health monitoring and maintenance lifecycle of buildings, while Gao and Pishdad-Bozorgi (2019) provide insight into its transformative potential in

Table 7: Comparison with leading research

Ref.	Dataset	Focus	Accuracy
Utkucu and Sacks (2023)	Not stated	Semantic Enrichment	Over 90%
Rogage and Doukari (2024)	ModelNet10, Extended ModelNet10	BIM Classification	*
Mirarchi et al. (2024)	Not stated	Semantic Enrichment	Over 80%
Utkucu et al. (2024)	Self compiled	MEP Classification	91%
Lilis et al. (2025)	Not stated	Semantic Enrichment	*
Chen and Jiang (2025)	SpaceGraph (self generated)	Semantic Enrichment	97.08%
Proposed framework	IFCNet	BIM classification and semantic enrichment	Training = 97.95%, testing = 86.97%

facilities operation and maintenance by enabling digital access to business-critical data within an interoperable 3D platform. Researchers such as Liu et al. (2019) and Maureira et al. (2021) further highlight BIM's capability to deliver comprehensive geometric and semantic information throughout a building's lifecycle.

The integration of Artificial Intelligence (AI), particularly deep learning, into BIM workflows has been recognized as a promising advancement in automating and enriching BIM data. Deep learning models have proven effective in the automatic classification of BIM objects, enhancing semantic enrichment processes and reducing reliance on manual labor Mirarchi et al. (2024).

Building upon these advancements, this study introduces an AI-based framework for the classification and semantic enrichment of Architectural, Mechanical, Electrical, and Plumbing (MEP) objects in BIM. The proposed model leverages a Convolutional Neural Network (CNN) to extract geometric features and a GNN-Transformer to capture relational dependencies between BIM objects. A dedicated feature fusion module combines these representations, enhancing the overall classification accuracy. The framework achieved a training accuracy of 97.95% and a testing accuracy of 86.97%, highlighting the effectiveness of combining both geometric and relational features for MEP object classification.

The high performance of the model validates the advantage of multimodal feature integration. Furthermore, the framework supports the creation of a semantically enriched BIM database, enabling automated querying of relational knowledge. This not only improves data consistency but also enhances interoperability across stakeholders. As demonstrated in Figure 25, the framework facilitates the development of a BIM knowledge graph capable of representing the complexity and interdependence of MEP systems. Such semantic enhancement is particularly valuable in supporting collaborative decision-making across the AEC value chain.

Despite these promising results, the model exhibits lower class-wise accuracy (below 80%) for certain object categories, including Beam, Furniture, Plate, Slab, SpaceHeater, and Stair (see Table 6). This is likely due to data imbalance or insufficient samples for these categories. Additionally, the computational cost of training the model remains high, presenting a barrier to practical scalability and real-time deployment.

Overall, the proposed framework represents a significant step toward automating BIM semantic enrichment and improving classification performance for MEP systems. By addressing interoperability challenges and reducing manual overhead, this approach contributes to the development of intelligent, data-driven workflows within the AEC industry.

6 Conclusion

This study presents an optimized and robust AI-driven framework for the classification and semantic enrichment of MEP BIM objects within the Architecture, Engineering, and Construction (AEC) industry. The framework integrates 3D CNNs for extracting spatial features, CNNs and a GNN Transformer for capturing relational dependencies among BIM components, a dedicated CNN-based fusion model to combine geometric and relational features, and a multi-layer neural network that utilises the fused features for accurate classification.

The proposed method demonstrated strong results with 97.95

Future work should consider efficiency-focused strategies like knowledge distillation to reduce training complexity. Testing on broader, more diverse datasets will also enhance the model's generalizability. Additionally, implementing federated learning could support distributed model training across different stakeholders while preserving data privacy and improving scalability.

Acknowledgment

A brief acknowledgement section may be given after the conclusion section just before the references. The acknowledgments of people who provided assistance in manuscript preparation, funding for research, etc. should be listed in this section.

References

Abanda, F. H., Musa, A., Clermont, P., Tah, J. H., and Oti, A. H. (2020). A bim-based framework for construction project scheduling risk management. *International Journal of Computer Aided Engineering and Technology*, 12(2):182–218.

Abdelalim, A. M., Shawky, K., Salem, M., Al-Adwani, M., and Sherif, A. (2024). Automation and optimization of bim execution plans for mega construction projects. *Preprints*.

Akintoye, A. et al. (2024). Integrating building information modelling (bim) and artificial intelligence (ai) for smart construction schedule, cost, quality, and safety management: challenges and opportunities. *Journal of Construction Engineering and Management*, 150(5):04024038.

Aladağ, H., Demirdöğen, G., Demirbağ, A. T., and Işık, Z. (2023). Understanding the perception differences on bim adoption factors across the professions of aec industry. *Ain Shams Engineering Journal*, 14(11):102545.

Bloch, T. (2022). Connecting research on semantic enrichment of bim-review of approaches, methods and possible applications. *J. Inf. Technol. Constr.*, 27:416–440.

Borkowski, A. S. (2023). Evolution of bim: Epistemology, genesis and division into periods. *Journal of Information Technology in Construction*, 28.

Borkowski, A. S. and Maroń, M. (2024a). Semantic enrichment of non-graphical data of a bim model of a public building from the perspective of the facility manager. *Big Data and Cognitive Computing*, 8(10):138.

Borkowski, A. S. and Maroń, M. (2024b). Semantic enrichment of non-graphical data of a bim model of a public building from the perspective of the facility manager. *Big Data Cogn. Comput.*, 8(10):138.

Chen, Y. and Jiang, H. (2025). Optimized graph neural networks for spatial recognition to support automatic building information model semantic enrichment - exploring node features enhancing strategies. *Engineering Applications of Artificial Intelligence*, 147:110365.

Emunds, C., Pauen, N., Richter, V., Frisch, J., and van Treeck, C. (2021). Ifcnet: A benchmark dataset for ifc entity classification. In *Proceedings of the 28th International Workshop on Intelligent Computing in Engineering (EG-ICE)*.

Emunds, C., Pauen, N., Richter, V., Frisch, J., and van Treeck, C. (2022). Sparse-bim: Classification of ifc-based geometry via sparse convolutional neural networks. *Advanced Engineering Informatics*, 53:101641.

Gao, X. and Pishdad-Bozorgi, P. (2019). Bim-enabled facilities operation and maintenance: A review. *Advanced Engineering Informatics*, 39:227–247.

Grilo, A. and Jardim-Goncalves, R. (2010). Value proposition on interoperability of bim and collaborative working environments. *Automation in Construction*, 19(5):522–530. Building Information Modeling and Collaborative Working Environments.

Heydarian, M., Doyle, T. E., and Samavi, R. (2022). Mlcm: Multi-label confusion matrix. *Ieee Access*, 10:19083–19095.

Hu, Z., Leng, S., Lin, J., et al. (2022). Knowledge extraction and discovery based on bim: A critical review and future directions. *Archives of Computational Methods in Engineering*, 29:335–356.

Huang, Y., Shi, Q., Zuo, J., Pena-Mora, F., and Chen, J. (2021). Research status and challenges of data-driven construction project management in the big data context. *Advances in Civil Engineering*, 2021(1):6674980.

Jang, S., Lee, G., Park, M., Lee, J., Suh, S., and Koo, B. (2025). Semantic elaboration of low-lod bims: Inferring functional requirements using graph neural networks. *Advanced Engineering Informatics*, 64:103100.

Jiang, S., Feng, X., Zhang, B., and Shi, J. (2023). Semantic enrichment for bim: Enabling technologies and applications. *Advanced Engineering Informatics*, 56:101961.

Khosrowshahi, F. and Arayici, Y. (2012). Roadmap for implementation of bim in the uk construction industry. *Engineering, construction and architectural management*, 19(6):610–635.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.

Lilis, G. N., Wang, M., Katsigarakis, K., Mavrokapnidis, D., Korolija, I., and Dimitrios, R. (2025). Bim-based semantic enrichment and knowledge graph generation via geometric relation checking. *Automation in Construction*, 173:106081.

Liu, Z., Lu, Y., and Peh, L. C. (2019). A review and scientometric analysis of global building information modeling (bim) research in the architecture, engineering and construction (aec) industry. *Buildings*, 9(10):210.

Maureira, C., Pinto, H., Yepes, V., and García, J. (2021). Towards an aec-ai industry optimization algorithmic knowledge mapping: An adaptive methodology for macroscopic conceptual analysis. *IEEE Access*, 9:110842–110879.

Miky, Y., Alshawabkeh, Y., and Baik, A. (2024). Using deep learning for enrichment of heritage bim: Al radwan house in historic jeddah as a case study. *Heritage Science*, 12(1):255.

Mirarchi, C., Gholamzadehmir, M., Daniotti, B., and Pavan, A. (2024). Semantic enrichment of bim: the role of machine learning-based image recognition. *Buildings*, 14(4):1122.

Pan, Y. and Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction*, 122:103517.

Panah, R. S. and Kioumarsi, M. (2021). Application of building information modelling (bim) in the health monitoring and maintenance process: A systematic review. *Sensors*, 21(3):837.

Rogage, K. and Doukari, O. (2024). 3d object recognition using deep learning for automatically generating semantic bim data. *Automation in Construction*, 162:105366.

Sacks, R., Eastman, C., Lee, G., and Teicholz, P. (2018). *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*. John Wiley & Sons.

Sacks, R., Girolami, M., and Brilakis, I. (2020). Building information modelling, artificial intelligence and construction tech. *Developments in the Built Environment*, 4:100011.

Sacks, R., Ma, L., Yosef, R., Borrmann, A., Daum, S., and Kattel, U. (2017). Semantic enrichment for building information modeling: Procedure for compiling inference rules and operators for complex geometry. *Journal of Computing in Civil Engineering*, 31(6):04017062.

Services, A. B. (2025). Building information modelling (bim): What is it and what part does it play in our industry? Accessed: 17-04-2025.

Seydgar, M., Poirier, E. A., and Motamedi, A. (2024). Comparative evaluation of deep neural network performance for point cloud-based ifc object classification. *IEEE Access*, 12:108303–108312.

Shehadeh, A., Alshboul, O., Taamneh, M. M., Jaradat, A. Q., and Alomari, A. H. (2024). Enhanced clash detection in building information modeling: Leveraging modified extreme gradient boosting for predictive analytics. *Results in Engineering*, 24:103439.

Sompolgrunk, A., Banihashemi, S., Golzad, H., and Le Nguyen, K. (2024). Strategic alignment of bim and big data through systematic analysis and model development. *Automation in Construction*, 168:105801.

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, 17:168–192.

Utkucu, D. and Sacks, R. (2023). Mep domain object classification through interdomain rule-based semantic enrichment on knowledge graphs. In *European Conference on Computing in Construction (EC3) and 40th International CIB W78 Conference*.

Utkucu, D., Ying, H., Wang, Z., and Sacks, R. (2024). Classification of architectural and mep bim objects for building performance evaluation. *Advanced Engineering Informatics*, 61:102503.

Vujovic, Z. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 12(6).

Wang, B., Wang, Q., Cheng, J. C., and Yin, C. (2022). Object verification based on deep learning point feature comparison for scan-to-bim. *Automation in Construction*, 142:104515.

Wang, T. and Chen, H.-M. (2023). Integration of building information modeling and project management in construction project life cycle. *Automation in Construction*, 150:104832.

Xue, F., Wu, L., and Lu, W. (2021). Semantic enrichment of building and city information models: A ten-year review. *Advanced Engineering Informatics*, 47:101245.

Yue, H., Wang, Q., Yan, Y., and Huang, G. (2025). Deep learning-based point cloud completion for mep components. *Automation in Construction*, 175:106218.

Zhang, L., Pan, Y., Wu, X., and Skibniewski, M. J. (2021). *Artificial intelligence in construction engineering and management*. Springer.