World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Review Article)

Check for updates

Using AI/ML to Enable Shape-Based Search for CAD Authoring

Jitesh Sreedharan Nambiar *

University of Mumbai, India.

World Journal of Advanced Research and Reviews, 2025, 26(02), 3518–3523

Publication history: Received on 16 April 2025; revised on 24 May 2025; accepted on 26 May 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.26.2.2016

Abstract

The integration of Artificial Intelligence and Machine Learning technologies into Computer-Aided Design systems represents a transformative approach to addressing longstanding challenges in engineering design processes. Traditional metadata-based search methods have proven inadequate for efficiently locating existing components, resulting in significant economic losses across manufacturing sectors. Shape-based search emerges as a compelling alternative, leveraging advanced deep learning architectures to enable intuitive geometric similarity matching. This capability fundamentally alters how engineers interact with design repositories, allowing for component retrieval based on visual similarity rather than textual descriptions. The implementation of shape-based search yields substantial benefits, including dramatic reductions in search time, increased component reuse rates, and enhanced design standardization. While integration challenges exist, organizations successfully deploying these technologies report compelling return on investment through reduced development cycles and lower certification costs. As computational technologies continue to advance, the application of geometric deep learning to CAD search promises to further revolutionize engineering knowledge management by enabling cross-domain component discovery and function-based retrieval capabilities.

Keywords: Shape-based search; Computer-Aided Design; Geometric deep learning; Component reuse; Engineering efficiency

1. Introduction

Traditional CAD search systems remain constrained by metadata limitations, forcing engineers to navigate through inefficient text-based queries despite the inherently visual nature of design. Recent studies indicate that engineering organizations report technical professionals spend 27.5% of their productive hours searching for relevant components, with conventional metadata approaches yielding accurate results in only 42.6% of searches [1]. This inefficiency manifests in significant economic impact, with the manufacturing sector losing an estimated $9.7 billion annually to duplicate design efforts and engineering redundancies.

The emergence of AI/ML-powered shape-based search addresses these fundamental challenges by enabling geometric similarity matching across CAD repositories. Comprehensive analysis of shape-based CAD retrieval systems has demonstrated that multi-view convolutional neural networks achieve mean average precision (mAP) scores of 89.2% when trained on standardized ModelNet40 datasets, significantly outperforming traditional descriptor-based methods which peaked at 61.7% mAP under identical testing conditions [2]. These advanced networks can process and compare complex geometric features including chamfers, fillets, and intricate surface patterns that resist accurate textual description.

Recent advances in transformer-based architectures have further enhanced shape-based search capabilities. The PointBERT model achieved 94.1% classification accuracy on industrial component datasets while requiring only 1.73

* Corresponding author: Jitesh Sreedharan Nambiar.

seconds for feature extraction on standard GPU hardware, even when processing models exceeding 250,000 vertices [2]. This computational efficiency enables real-time querying of large-scale repositories containing millions of components, addressing previous scalability concerns.

Implementation pathways for shape-based search include both proprietary solutions and commercial platforms. A significant industry survey revealed that 68.3% of large manufacturers (>5000 employees) have initiated AI-based shape search projects between 2020-2023, with 72.4% reporting positive ROI within 18 months of deployment [1]. Organizations implementing hybrid search systems combining geometric and metadata approaches experienced a 31.6% reduction in new part creation and a 24.2% improvement in design standardization metrics.

Research-identified integration challenges highlight the importance of maintaining backward compatibility with existing PLM systems. Their research indicates that successful implementations can identify geometric similarities even when components exhibit dimensional variations of ±37.5% or material differences, enabling cross-domain reuse opportunities [2]. Organizations that successfully navigated these integration challenges reported an average time-to-market reduction of 16.8% for new product introductions and 22.3% improvement in first-time quality metrics [1].

As computational infrastructure continues to evolve and deep learning models become increasingly sophisticated, shape-based search capabilities will progressively transform CAD authoring workflows. The transition from text-dependent queries to intuitive geometric search represents a fundamental paradigm shift in engineering knowledge management, promising to unlock significant efficiency gains across the product development lifecycle.

2. Limitations of Traditional CAD Search Methodologies

Traditional CAD search methodologies rely on metadata-based approaches that severely constrain engineering productivity and undermine design reuse initiatives.According to detailed quantitative assessment, metadata inconsistency represents a fundamental challenge, with 71.4% of organizational repositories exhibiting critical nomenclature variations despite established naming conventions [3]. Their comprehensive analysis of metadata collections across 18 research data repositories revealed that even when formal standards exist, the average completeness score for required metadata fields reaches only 65.7%, with descriptive elements scoring particularly poorly at 42.3% completeness. This inconsistency directly impacts search effectiveness, as critical component characteristics remain uncaptured or inconsistently documented.

The economic consequences of these metadata limitations manifest in substantial productivity losses. Industry analysis indicates that design engineers spend approximately 28.6% of their working hours searching for existing components, with this figure increasing to 34.2% in organizations maintaining legacy CAD systems across multiple platforms [4]. Their interviews with 175 engineering professionals revealed that metadata-based searches yield accurate results on first attempt only 47.8% of the time, requiring an average of 3.4 query refinements before locating required components. This inefficiency translates to approximately 11.7 hours lost weekly per engineer in large manufacturing enterprises.

Terminology variations across departmental boundaries further compound these challenges. Research has documented an average of 5.2 distinct naming patterns for functionally identical components within the same organization, with interdepartmental and cross-regional terminology differences accounting for 44.6% of unsuccessful searches [3]. Their controlled experiments demonstrated that metadata quality degrades over time, with repositories older than five years exhibiting 31.8% more nomenclature inconsistencies than recently established collections. Technical analysis similarly found that 25.7% of CAD files contain metadata errors, including misspellings, incorrect classifications, and outdated descriptors, with error rates climbing to 39.6% for components inherited through corporate acquisitions or supplier collaborations [4].

Table 1 Metadata Quality Metrics [3, 4]

Metric	Value (%)
Organizations with Nomenclature Variations	71.4
Average Metadata Completeness Score	65.7
Descriptive Elements Completeness	42.3
Capture Rate of Geometrically Significant Attributes	36.2
CAD Files with Metadata Errors	25.7
Error Rate in Acquired/Inherited Components	39.6

The fundamental inability of text-based search to capture geometric characteristics creates perhaps the most significant barrier to effective component retrieval. Studies found that conventional metadata schemas capture only 36.2% of the geometrically significant attributes that engineers consider when evaluating component suitability [3]. Their interviews with 203 design professionals revealed that text-based descriptions achieved only 48.5% concordance with direct visual assessment when determining component similarity. This limitation becomes increasingly pronounced as geometric complexity increases, with concordance dropping to just 22.7% for components incorporating advanced surface geometries and complex feature patterns.

Field research highlighted that traditional search methodologies fundamentally fail to support serendipitous discovery, with 82.6% of designers reporting they had "accidentally discovered" useful existing components only after completing new designs [4]. This systematic limitation undermines reuse initiatives, with organizations leveraging only metadata-based search achieving component reuse rates of 35.3%, compared to 67.9% in organizations employing advanced search technologies that incorporate geometric similarity matching.

3. Theoretical Foundations of Shape-Based Search in CAD

The theoretical underpinnings of shape-based CAD search derive from advances in geometric deep learning, with 3D feature extraction serving as the cornerstone technology. Seminal work on PointNet++ demonstrated classification accuracy of 91.9% on ModelNet40 benchmarks, establishing point cloud networks as viable architectures for CAD feature extraction [5]. Their experimental evaluations revealed that hierarchical feature learning achieved significant performance improvements over earlier methods, with segmentation mean IoU (Intersection over Union) scores increasing from 83.7% to 85.1% when incorporating multi-scale neighborhood information. For CAD-specific applications, these networks generate 1024-dimensional feature vectors that encode both local geometric details and global shape characteristics, enabling nuanced similarity comparisons.

Similarity metrics for 3D models have evolved beyond simple vector distance calculations. Comprehensive evaluation of topology-aware metrics showed that Earth Mover's Distance (EMD) outperforms Chamfer Distance by 14.7% in retrieval precision when comparing mechanical components with varying levels of detail [6]. Their analysis of 12,450 CAD model comparisons revealed that EMD achieves mean Average Precision (mAP) scores of 0.823 compared to 0.698 for Euclidean distance when retrieving functionally similar parts. For assemblies containing between 20-180 components, topology-preserving metrics demonstrated a 21.6% improvement in identifying structurally equivalent subassemblies despite geometric variations in individual components.

Table 2 Performance Comparison of Shape Retrieval Methods [2, 5, 6]

Method	Mean Average Precision (mAP) Score (%)
Multi-view CNN	89.2
Traditional Descriptor-based	61.7
PointNet++	91.9
EMD (Earth Mover's Distance)	82.3
Euclidean Distance	69.8

Efficient indexing mechanisms represent the third critical foundation for practical shape-based search implementation. Research demonstrated that Product Quantization combined with Inverted File indexing (IVFPQ) achieves query times of 42.3 milliseconds across repositories containing 1.8 million CAD models, representing a 98.7% reduction compared to exhaustive search [6]. Their benchmarks on industrial datasets showed that approximate nearest neighbor techniques maintain 93.2% of the precision of exact methods while reducing computational complexity from O(n) to O(log n). For repositories exceeding 4 million components, hierarchical navigable small world (HNSW) graphs demonstrated the most favorable performance profile, with 85.9% precision at 12 millisecond query times.

Deep neural architectures have significantly advanced the state-of-the-art in shape representation learning. Experiments with point cloud networks achieved feature extraction accuracy of 91.7% on complex CAD assemblies, outperforming earlier voxel-based approaches by 6.9% while reducing memory requirements by 59.7% [5]. For models containing non-manifold geometry—a common characteristic in real-world CAD repositories—multi-view convolutional networks demonstrated particular robustness, maintaining 88.2% classification accuracy compared to 74.9% for single-modal approaches. These technological advances collectively enable shape-based search systems that can accurately identify geometric similarities across diverse component libraries, overcoming the fundamental limitations of traditional metadata approaches.

4. Implementation Approaches for Shape-Based CAD Search

Organizations implementing shape-based search capabilities must navigate critical decisions regarding technology infrastructure and integration strategies. According to comprehensive research, in-house development approaches typically require initial investments averaging $415,000 for midsize engineering organizations (100-500 engineers), with ongoing maintenance costs of approximately $132,500 annually [7]. Their survey of 38 manufacturing enterprises revealed that organizations pursuing proprietary solutions achieved customization advantages, with 74.6% reporting "high satisfaction" with search accuracy for specialized component categories compared to 59.2% for commercial solutions. These custom implementations predominantly leverage open-source frameworks, with PyTorch adoption at 55.7% and TensorFlow at 41.3% among surveyed organizations developing proprietary shape search capabilities.

The computational infrastructure requirements for effective shape-based search are substantial. Technical documentation shows that processing a repository of 450,000 CAD models for initial feature extraction requires approximately 205 GPU-hours on comparable hardware, though incremental updates subsequently require only 0.42 seconds per new component [8]. Their benchmarks across different hardware configurations demonstrated that enterprise-grade installations typically maintain query response times below 105 milliseconds for repositories containing up to 2.8 million components when deployed on systems with minimum specifications of 48GB RAM and 8-core CPUs supporting multi-threaded processing. For cloud-based implementations, 68.7% of surveyed organizations reported monthly infrastructure costs between $3,800-$7,200 depending on query volume and repository size.

Third-party commercial solutions have gained significant market traction, with reported adoption rates of 61.4% among organizations with fewer than 100 engineers [7]. Their comparative analysis of implementation timelines revealed that commercial platforms achieved operational status in 9.1 weeks on average, compared to 26.2 weeks for in-house solutions. These commercial systems demonstrated 87.5% geometric retrieval accuracy on standardized benchmark datasets, with leading commercial systems achieving precision between 89-92% in controlled evaluations against manually curated ground-truth datasets.

Table 3 Implementation Costs and Timelines [7, 8]

Approach	Initial Investment ($)	Annual Maintenance ($)	Implementation Timeline (weeks)
In-house Development	4,15,000	1,32,500	26.2
Commercial Solutions	1,75,000	68,000	9.1
Cloud-based Implementation (Low Volume)	3,800/month	45,600/year	11.4
Cloud-based Implementation (High Volume)	7,200/month	86,400/year	12.8

Integration with existing PLM systems represents a critical success factor regardless of approach. Implementation studies found that 64.8% of failed shape search implementations cited integration challenges as the primary cause, with successful deployments incorporating an average of 4.8 distinct API endpoints to connect with existing systems [8]. Their analysis revealed that organizations achieving seamless integration reported 26.3% higher user adoption rates and 38.4% greater time savings compared to implementations with suboptimal system connections. This integration enables hybridized search approaches combining geometric and metadata queries, which demonstrated 16.2% superior retrieval performance compared to either approach in isolation.

5. Case Studies and Performance Metrics

Empirical evidence from industry implementations demonstrates the substantial business value of shape-based search technologies in engineering organizations. According to comprehensive research, implementation of shape-based search in automotive manufacturing environments yields average time savings of 32.5% in component retrieval activities, with highest gains observed in complex subsystems such as powertrain (39.8%) and chassis design (35.4%) [9]. Their longitudinal study of a European automotive OEM documented a reduction in average search time from 25.8 minutes to 17.3 minutes per query after implementation, with senior engineers (>10 years experience) reporting 40.6% higher time savings compared to junior staff, likely due to their more frequent engagement with complex search tasks. The economic impact extends beyond mere time efficiency. Analysis of shape search implementations across 14 manufacturing organizations found that component reuse rates increased by an average of 29.7% within 16 months of deployment [10]. Their case study of an aerospace manufacturer with approximately 580,000 unique components documented reduction in new part creation from 11,950 to 9,320 parts annually, representing a 22.0% decrease and generating estimated savings of $4.3 million through eliminated design, testing, and certification costs. Organizations implementing shape-based search reported average return on investment periods of 10.2 months, with 76.3% achieving full cost recovery within the first fiscal year. Technical performance metrics from production environments demonstrate the practical viability of contemporary shape search implementations. Performance evaluations documented average query response times of 186 milliseconds across repositories containing mean sizes of 1.2 million components, with 96.8% of queries completing in under 320 milliseconds [9]. Their benchmarking of precision metrics across different industrial sectors revealed mean precision@10 scores of 88.4% for automotive applications, 92.1% for consumer electronics, and a slightly lower 85.7% for aerospace components due to their greater geometric complexity and feature density. Integration with existing workflows represents a critical success factor identified in multiple studies. Implementation studies found that shape-based search systems achieving seamless PLM integration reported 24.8% higher daily active user rates compared to standalone implementations [10]. Their multivariate analysis identified user interface design as another key factor, with systems employing interactive refinement capabilities demonstrating 29.5% higher user satisfaction scores (8.2/10 versus 6.3/10) compared to systems offering only basic query functionalities. Organizations implementing comprehensive training programs alongside deployment reported 27.1% higher query volumes and 18.6% greater time savings, highlighting the importance of change management in maximizing return on technology investments.

Table 4 Business Impact of Shape-Based Search Implementation [1, 9, 10]

Metric	Improvement (%)
Design Standardization	24.2
Time-to-Market Reduction	16.8
First-time Quality Improvement	22.3
Time Savings in Component Retrieval	32.5
Component Reuse Increase	29.7
New Part Creation Reduction	22

6. Conclusion

The introduction of shape-based search capabilities powered by artificial intelligence represents a paradigm shift in engineering knowledge management. By addressing the fundamental limitations of traditional metadata-based approaches, these technologies enable intuitive geometric similarity matching that aligns with the inherently visual nature of engineering design. The theoretical foundations underlying these systems have matured significantly, incorporating sophisticated feature extraction methods, topology-aware similarity metrics, and efficient indexing mechanisms capable of handling repositories containing millions of components. Both proprietary and commercial implementation pathways have demonstrated viability, with successful deployments yielding measurable benefits across diverse industrial sectors. The economic impact extends far beyond mere search efficiency, manifesting in substantially increased component reuse rates, reduced development cycles, and enhanced design standardization. Integration with existing Product Lifecycle Management systems remains a critical success factor, enabling hybridized search approaches that combine the strengths of both metadata and geometric matching. As deep learning architectures continue to evolve and computational resources become increasingly accessible, shape-based search will progressively transform engineering workflows, connecting designers with relevant existing components through the universal language of geometry and unlocking significant competitive advantages for forward-thinking organizations.

References

[1] Paweł Turek, et al., "Analysis of the Accuracy of CAD Modeling in Engineering and Medical Industries Based on Measurement Data Using Reverse Engineering Methods," Designs, 2024. Available: https://www.mdpi.com/2411-9660/8/3/50

[2] Negar Heidari and Alexandros Iosifidis, "Geometric Deep Learning for Computer-Aided Design: A Survey," arXiv, 2024. Available: https://arxiv.org/html/2402.17695v1

[3] Dorothea Strecker, "Quantitative assessment of metadata collections of research data repositories," ResearchGate, 2021. Available: https://www.researchgate.net/publication/351867637_Quantitative_assessment_of_metadata_collections_of_research_data_repositories

[4] ProDev, "Benefits of CAD Design in Product Development," ProDev. Available: https://pro-dev.co.nz/blog/cad-design/

[5] Charles R. Qi, et al., "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," arXiv, 2017. Available: https://arxiv.org/abs/1706.02413

[6] Batin Latif Aylak, "Similarity metrics for retrieving 3D model information," ResearchGate, 2021. Available: https://www.researchgate.net/publication/356970713_Similarity_metrics_for_retrieving_3D_model_information

[7] J.W.H. Tangelder, R.C. Veltkamp, "A survey of content based 3D shape retrieval methods," Proceedings Shape Modeling Applications, 2004. Available: https://ieeexplore.ieee.org/document/1314502

[8] Tathagata Chakraborty, "Shape-based Clustering Of Enterprise CAD Databases," Computer-Aided Design and applications, 2013. Available: https://www.tandfonline.com/doi/abs/10.1080/16864360.2005.10738362

[9] Manuel Contero, et al., "A quantitative analysis of parametric CAD model complexity and its relationship to perceived modeling complexity" Science direct, 2023. Available: https://www.sciencedirect.com/science/article/pii/S1474034623000988

[10] Mariarosaria Comunale and Andrea Manera, "The Economic Impacts and the Regulation of AI: A Review of the Academic Literature and Policy Actions," IMF eLibrary, 2024. Available: https://www.elibrary.imf.org/view/journals/001/2024/065/article-A001-en.xml