UDC: 004.032.26:004.93
DOI: https://doi.org/10.30546/09090.2026.001.20.2012

AUTOENCODERS FOR REPRESENTATION LEARNING: FROM DIMENSIONALITY REDUCTION TO VOLUMETRIC GENERATION

Hamid GADIROV*
University of Groningen, Netherlands
** h.gadirov@rug.nl*

ARTICLE INFO
Article history:
Received:2025-10-23
Received in revised form:2025-11-18
Accepted:2025-12-16
Available online:2026-06-23
Keywords:
Representation Learning
Neural Embeddings
Autoencoders
Clustering
Volumetric Generation
2010 Mathematics Subject Classifications: 68T05, 68T10, 62H25, 62H30, 68U05

ABSTRACT
Autoencoders have re-emerged as vital tools in the machine learning landscape, driven by the growing demand for effective representation learning in high-dimensional, complex data. Their ability to capture non-linear structures makes them particularly well-suited for dimensionality reduction and unsupervised clustering, surpassing the limitations of classical methods. Beyond traditional tasks, AEs are increasingly applied in volumetric data domains—such as scientific ensemble analysis and volume rendering—where they help reveal latent structure and enable more expressive visualizations. This work surveys recent advancements in autoencoder-based representation learning, focusing on practical techniques that bridge the gap between dimensionality reduction, clustering, and volumetric generation.

1. INTRODUCTION

Dimensionality reduction (DR) and clustering are fundamental techniques for exploring and analyzing high-dimensional data. DR methods aim to project complex data into a lower-dimensional space (often 2D or 3D) while preserving meaningful structure, making patterns and clusters more visually apparent. Clustering techniques seek to group similar data points together without supervision. Traditional DR algorithms such as Principal Component Analysis (PCA) and non-linear visualization methods like t-SNE have been widely used for these purposes [9], but they can struggle with very high-dimensional data, complex non-linear manifolds, or scalability to large datasets. More recent manifold learning approaches, such as UMAP, improve scalability and local structure preservation, yet still rely on fixed similarity assumptions and hand-crafted distance metrics [10].

In recent years, autoencoders (AEs)—a class of deep neural networks for unsupervised representation learning—have emerged as powerful tools for dimensionality reduction and clustering [1,2]. Autoencoders learn a compressed latent representation of data by training the network to reconstruct its own input, enabling them to capture non-linear structures that are difficult to model with linear methods such as PCA. Variants such as denoising and variational autoencoders further improve robustness and expressiveness of learned representations [3, 4].

As a result, autoencoder-based embeddings often provide more informative low-dimensional representations for visualization and analysis, especially in high-dimensional settings.

Beyond serving as a pre-processing step, autoencoders can be integrated directly into clustering frameworks. By jointly optimizing reconstruction objectives together with clustering or metric-learning objectives, deep clustering methods simultaneously learn feature representations and cluster assignments [8]. This joint learning paradigm often produces latent spaces in which natural clusters are more compact and better separated, addressing limitations of classical clustering methods applied directly to raw data. Prior studies have shown that autoencoder-based representations can substantially improve clustering quality and visual interpretability, particularly for complex scientific and simulation datasets [13, 14].

In this work, we survey the use of autoencoders for dimensionality reduction and clustering, with a particular focus on representation learning techniques that bridge these two tasks. We discuss key ideas from foundational and recent studies and relate them to state-of-the-art (SOTA) developments in the field. We first review the role of autoencoders in dimensionality reduction, then examine how they enhance clustering performance, and finally discuss advanced methods that unify dimensionality reduction, clustering, and generative modeling.

2. RELATED WORK

The field of representation learning has evolved from linear projection methods to deep generative architectures. Dimensionality Reduction (DR) foundational techniques such as PCA and t-SNE [9] have long served the visualization community, but often struggle with the non-linear manifolds inherent in complex scientific data. Recent manifold learning approaches like UMAP [10] offer better scalability but remain sensitive to hyperparameter selection and distance metrics.

Autoencoders (AEs) and their variants—including Denoising AEs [3] and Variational Autoencoders (VAEs) [4]—have emerged as the primary alternative for non-linear DR. By optimizing for a reconstruction objective, these models learn compact latent spaces that preserve salient features [1, 2]. In the context of *Clustering*, deep clustering paradigms have shifted from sequential pipelines (representation learning followed by k-means) to joint optimization frameworks that integrate clustering-oriented losses directly into the bottleneck [8].

Most recently, these representations have been extended to Volumetric and Generative Modeling. With the rise of 3D Gaussian Splatting (3DGS) [23], AEs are now being utilized to tokenize unstructured primitives into structured latent grids, enabling downstream diffusion-based generation [24, 26, 27]. Our work builds upon these foundations by evaluating how tailored AE architectures can be optimized for specific scientific ensemble and volumetric tasks.

3. METHODOLOGY AND DISCUSSION

In this section, we detail the architectural frameworks and optimization objectives employed to learn expressive latent representations for high-dimensional data. We further analyze how these methodological choices—ranging from bottleneck regularization to joint clustering losses—directly impact the quality, stability, and interpretability of the resulting embeddings across scientific and generative tasks.

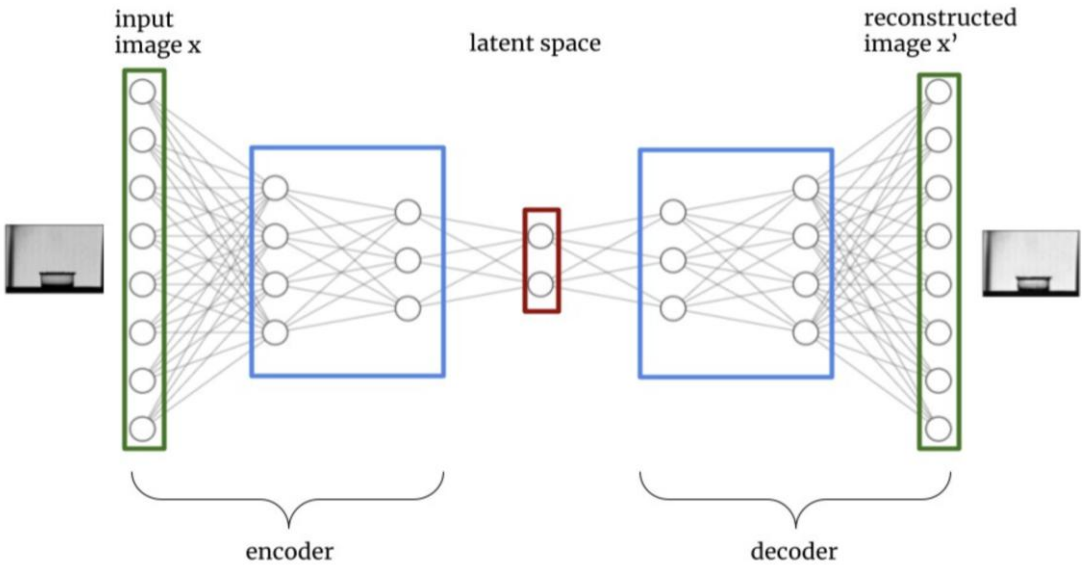

***Figure 1**: General structure of an autoencoder for dimensionality reduction (original illustration by the author, previously appearing in [14]). The encoder (left) compresses high-dimensional data into a latent code, and the decoder (right) reconstructs the data from this code. By training on reconstruction, the autoencoder learns a low-dimensional representation that preserves important features of the input. Deep autoencoders often use convolutional layers for image data and have symmetric encoder/decoder architectures.*

Autoencoders for Dimensionality Reduction. Autoencoders provide a powerful non-linear alternative to traditional dimensionality reduction techniques. A basic autoencoder consists of an encoder network that compresses the input data into a lower-dimensional latent code, and a decoder network that reconstructs the original data from this code [2]. By training to minimize reconstruction error, the autoencoder is forced to learn a compact representation that captures the most salient features of the data, effectively performing data-driven compression. Figure 1 illustrates a typical autoencoder architecture for dimensionality reduction. Formally, the encoder f parameterized by ϕ maps the high-dimensional input x to a latent code $z = f_\phi(x)$, while the decoder g parameterized by θ attempts to generate a reconstruction $g_\theta(z)$ that minimizes a distance metric, typically the Mean Squared Error (MSE):

$$\mathcal{L}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^{n} \|x_i - g_\theta(f_\phi(x_i))\|^2$$

This encoder–decoder structure can be realized using fully connected or convolutional layers, depending on the data modality, and is often designed symmetrically. The dimensionality of the latent space constitutes a key hyperparameter: a smaller latent code enforces stronger compression. Owing to their non-linear activation functions, autoencoders are able to model complex data manifolds that cannot be captured by linear techniques such as PCA [1, 2].

One prominent application of autoencoders in dimensionality reduction is their use as a preprocessing step prior to visualization methods such as t-SNE or UMAP [9, 10]. Gadirov et al. [13] demonstrated that extracting features using autoencoders before applying non-linear projection techniques can substantially improve visualization quality for large scientific ensemble datasets. In their study, directly applying UMAP to raw high-dimensional ensemble data resulted in dispersed projections with limited structural coherence. In contrast, projecting autoencoder-derived latent embeddings produced clearer spatial organization and more distinct cluster separation. This effect was observed across multiple datasets, including 3D soil

simulation ensembles and droplet impact dynamics, where autoencoder-based preprocessing led to two-dimensional embeddings that better reflected meaningful structures and classes present in the data.

A key finding of Gadirov et al. [13] was that the choice of autoencoder architecture and hyperparameters has a significant impact on dimensionality reduction quality. The authors evaluated a range of autoencoder variants, including standard deep autoencoders, sparse autoencoders, variational autoencoders (VAEs) [4], and β-VAEs, as well as different latent dimensionalities. The resulting two-dimensional projections were assessed using cluster quality measures such as neighborhood hit and silhouette score. The experiments revealed considerable variation in performance across configurations, with no single autoencoder variant consistently outperforming the others. For example, on one ensemble dataset, a β-VAE with moderate regularization (β in the range of 2–8) and a 32-dimensional latent space yielded the highest silhouette scores, whereas on another dataset a convolutional autoencoder tailored to spatio-temporal structure performed better. These results highlight that both insufficient and overly strong regularization can be detrimental: weak bottlenecks may fail to disentangle clusters, while excessive regularization (e.g., very large β values) can suppress meaningful variation and degrade visualization quality.

To mitigate the high cost of exhaustive architecture search, Gadirov et al. [13] proposed a guided selection strategy based on partially labeled data. They introduced projection quality metrics—neighborhood hit and silhouette score—that require ground-truth labels for evaluation, and demonstrated that labeling only a small fraction of the dataset (approximately 1%) is sufficient to reliably compare different autoencoder configurations. By visualizing the performance of each model in a two-dimensional metric space, a Pareto frontier of top-performing autoencoders can be identified, representing optimal trade-offs between competing quality criteria. This approach enables informed, data-driven selection of suitable autoencoder models without relying on arbitrary defaults. Overall, these findings demonstrate that autoencoders can significantly enhance dimensionality reduction for complex, high-dimensional data, provided that architecture design and hyperparameter choices are carefully tuned. When properly configured, autoencoder-based dimensionality reduction can reveal latent cluster structure that may remain obscured when applying projection methods directly to raw data.

Autoencoders for Clustering. Beyond serving as feature extractors for downstream clustering algorithms, autoencoders can be more tightly integrated with clustering objectives in so-called deep clustering approaches. Broadly, two main paradigms can be distinguished when using autoencoders for clustering tasks.

Sequential approaches first train an autoencoder in an unsupervised manner to learn compact latent representations, and subsequently apply a conventional clustering algorithm—such as k-means or spectral clustering—on the learned embeddings. This strategy decouples representation learning from clustering and often yields better results than clustering raw high-dimensional data. The improvement stems from the autoencoder's ability to filter noise and capture non-linear structure, producing embeddings that are more amenable to distance-based clustering. This approach has been shown to improve cluster compactness and separation in a variety of domains, particularly for image and scientific data where raw feature spaces are highly complex [2, 8, 13].

Joint (simultaneous) approaches integrate clustering directly into the autoencoder training process by augmenting the reconstruction objective with an additional clustering-oriented loss. In this setting, the model jointly optimizes representation learning and cluster formation,

encouraging the latent space to explicitly reflect cluster structure. The clustering loss may promote confident assignments, penalize overlap between clusters, or enforce similarity constraints among related samples [8]. Compared to sequential methods, joint learning often yields latent spaces that are more explicitly cluster-friendly, resulting in improved clustering performance and more interpretable embeddings [14].

While joint clustering approaches are powerful, they introduce additional training complexity and hyperparameters, such as weighting factors between reconstruction and clustering losses, cluster initialization strategies, and regularization strength. A well-balanced reconstruction objective remains crucial: without it, models may converge to trivial solutions in which all samples collapse into a single cluster. By preserving sufficient information to faithfully reconstruct the input, the reconstruction term prevents such degeneration while allowing the latent representation to form compact, well-separated clusters. Many modern deep clustering approaches rely on carefully balancing these competing objectives to achieve stable and meaningful results [2, 4].

Autoencoder-based clustering has also proven effective in related unsupervised learning settings such as anomaly detection and one-class classification. Methods based on reconstruction error or latent-space compactness leverage autoencoders to distinguish normal data from outliers, effectively treating anomalies as separate or weakly populated clusters [33, 34, 35]. These approaches further highlight the close connection between representation learning, clustering, and density estimation in autoencoder-based models.

Recent work has emphasized robustness and stability as critical challenges in deep clustering. Because clustering lacks explicit supervision, results can be sensitive to architectural choices, initialization, and training dynamics. Ensemble-based strategies have therefore been proposed to mitigate these issues by aggregating multiple autoencoder embeddings or clustering solutions. In the context of scientific ensemble visualization, Gadirov et al. demonstrated that combining multiple autoencoder configurations and evaluating them using cluster quality metrics leads to more reliable and expressive latent representations [13]. Similarly, reconstruction-based ensemble methods have been shown to improve robustness in anomaly detection and time-dependent simulation data by reducing sensitivity to individual model biases [18]. Overall, these developments confirm that autoencoders form a central building block in modern clustering methodologies. Whether used sequentially as feature extractors or jointly optimized with clustering objectives, autoencoder-based models consistently outperform classical clustering techniques on complex, high-dimensional data. When combined with ensemble strategies or task-specific regularization, they offer a powerful and flexible framework for unsupervised data analysis.

Advanced Methods and Recent Results. Recent research has continued to extend autoencoder-based clustering by incorporating additional objectives, weak supervision, and task-specific constraints, particularly in the context of scientific visualization, generative modeling, and uncertainty-aware analysis. The soft silhouette loss [14] is a continuous relaxation of the classical silhouette coefficient, allowing it to be optimized directly during training without requiring hard cluster assignments. This loss encourages compact intra-cluster structure while penalizing overlap between clusters, effectively guiding the autoencoder toward cluster-friendly embeddings. The contrastive loss further reinforces this behavior by pulling similar samples closer together and pushing dissimilar samples apart, drawing on principles from metric learning [8]. Because the target scientific ensemble datasets contain only limited labeled data, pseudo-labels are generated using a pretrained EfficientNet model to provide weak supervisory signals for unlabeled samples. The autoencoder is then trained jointly with reconstruction,

silhouette-based clustering, and contrastive objectives, and the resulting latent representations are evaluated using non-linear projection methods such as UMAP [10]. Experimental results demonstrate consistent improvements in clustering quality and visual separability over baseline autoencoders, particularly in cases where clusters exhibit partial overlap.

These findings suggest that while standard autoencoders often capture much of the intrinsic structure in high-dimensional scientific data, additional clustering-aware objectives can refine latent spaces and improve interpretability. However, the magnitude of improvement is typically incremental rather than dramatic, indicating that the effectiveness of joint clustering objectives depends strongly on data complexity and the informativeness of the added supervisory signals. In practice, sequential pipelines—autoencoder-based dimensionality reduction followed by clustering—may already yield strong performance, while joint approaches provide further gains when cluster boundaries are ambiguous or subtle [13, 14].

Beyond clustering, recent advances increasingly blur the line between representation learning, anomaly detection, and generative modeling. Reconstruction-based anomaly detection methods leverage autoencoders to identify samples that deviate from learned data distributions, effectively treating anomalies as low-density or weakly populated clusters [33, 34]. In scientific ensemble analysis, such ideas have been extended to time-dependent simulations, where reconstruction error and latent-space deviations are used to detect anomalies and regime changes across ensemble members [18]. These approaches highlight how autoencoder-based clustering concepts naturally extend to uncertainty estimation and outlier analysis.

Generative extensions of autoencoders have further expanded their applicability, particularly in the realm of 3D vision. Variational autoencoders (VAEs) and GAN-based models [4, 5] now form the backbone of volumetric data generation. In particular, the shift toward 3D Gaussian Splatting (3DGS) [23] has created a new frontier for VAEs; because raw Gaussian primitives are unstructured, recent architectures like DiffGS [24] and Atlas Gaussians [26] utilize VAEs to project these primitives into organized latent spaces. This allows diffusion-based models to operate on a continuous functional representation rather than discrete points, enabling the generation of high-fidelity 3D scenes [25, 27, 28]. These methods rely on structured latent representations that implicitly cluster spatial and semantic information, reinforcing the importance of expressive, well-organized embeddings.

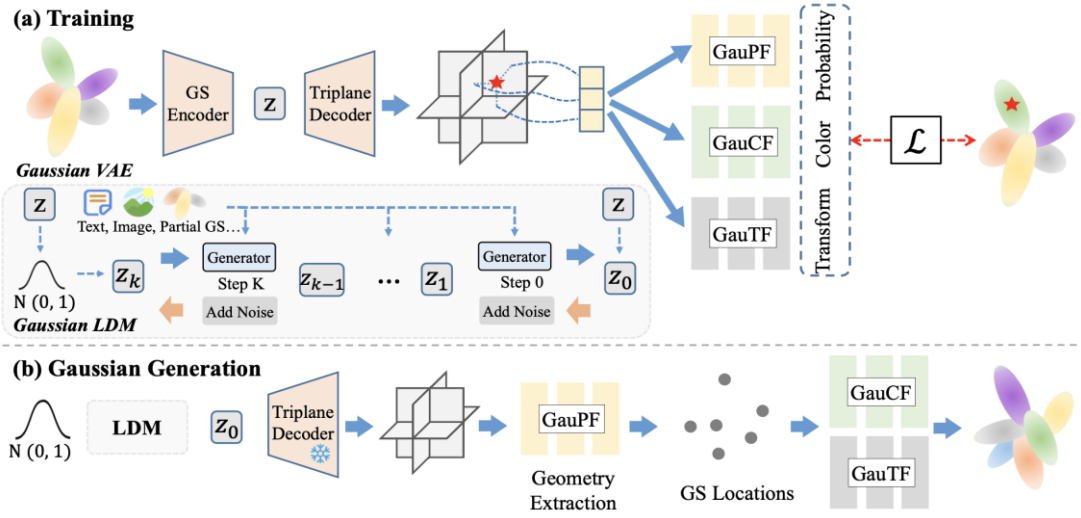

Figure 2. *Diffusion-based generative modeling of 3D Gaussian splatting representations (original illustration by the author, previously appearing in [24]).*

From a practical and methodological perspective, many of these advances are underpinned by well-established techniques, tools, and datasets that make large-scale autoencoder-based analysis feasible and reproducible. Regularization strategies such as dropout are commonly employed during training to mitigate overfitting and improve generalization in deep autoencoder architectures, especially when dealing with limited or noisy scientific data [6]. Modern implementations of autoencoder models and clustering pipelines are typically realized using automatic differentiation frameworks such as PyTorch, which enable flexible experimentation with custom loss functions, multi-objective optimization, and large neural architectures [7]. Foundational principles from scientific visualization further inform the design and evaluation of learned embeddings, particularly with respect to visual clarity, structure preservation, and interpretability [11, 12, 17]. In the context of ensemble visualization, prior work has demonstrated how autoencoder-based feature extraction can support more expressive projections and region-based analysis of high-dimensional simulation data [15, 16]. Standard benchmark datasets such as MNIST continue to serve as useful testbeds for validating representation learning and clustering behavior before deploying methods on more complex scientific datasets [30]. Finally, containerization technologies such as Docker play an important role in ensuring reproducibility, portability, and consistency of experimental environments across platforms, an aspect that has become increasingly critical for large-scale machine learning and visualization pipelines [31, 32].

Within scientific visualization, learning-based flow estimation and temporal interpolation methods such as FLINT (figure 3) and HyperFLINT further demonstrate how autoencoder-inspired architectures support structured latent representations for complex spatio-temporal data [19, 20]. Similarly, learning-based performance prediction for volume rendering [21] and radiance caching for volume path tracing [22] rely on compact latent encodings that organize high-dimensional simulation or rendering states in a cluster-aware manner. While these methods are not clustering algorithms per se, they exemplify how autoencoder-style representation learning enables downstream tasks that benefit from latent structure and separability.

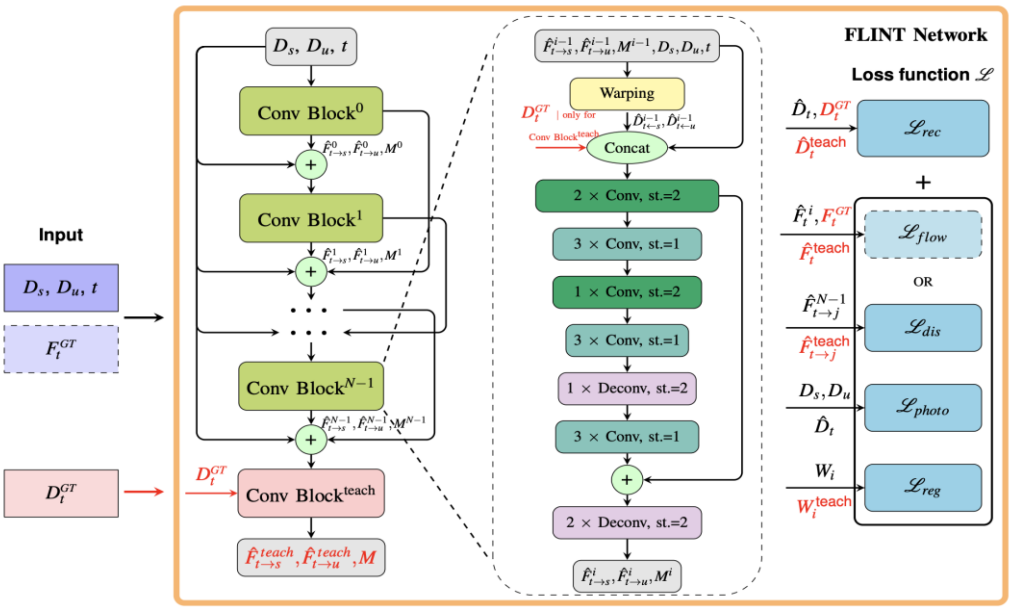

Figure 3. *Autoencoder-based structure within the FLINT network architecture (original illustration by the author, previously appearing in [19]). The encoder–decoder design enables learning compact latent representations that support flow estimation and temporal super-resolution in scientific ensemble data.*

Recent advancements indicate that autoencoder-based models remain central to state-of-the-art approaches for clustering, visualization, anomaly detection, and generative modeling in high-dimensional settings. By integrating clustering-aware losses, ensemble strategies, weak supervision, and generative objectives, modern methods achieve greater robustness, interpretability, and task alignment than classical techniques. At the same time, these gains introduce additional complexity in terms of model design, hyperparameter tuning, and training stability. Despite these challenges, the breadth of recent applications—from scientific ensemble analysis to volumetric generation and real-time rendering—underscores the continued relevance and versatility of autoencoders as a foundational tool for representation learning and clustering.

4. CONCLUSION

Autoencoders have proven to be highly effective for both dimensionality reduction and clustering, often enabling more insightful analysis of complex, high-dimensional datasets than traditional methods. By learning compact and informative representations, autoencoders overcome the limitations of linear dimensionality reduction techniques and provide embeddings that better reflect the intrinsic structure of the data [1, 2]. These learned representations can be used directly for visualization or as input to downstream clustering algorithms, yielding improved cluster separation and interpretability across a wide range of domains, including image analysis, scientific simulations, and ensemble data visualization [9, 10, 13].

Beyond their use as feature extractors, autoencoders can be extended with clustering-aware objectives that explicitly encourage the formation of compact and well-separated latent groups. Recent work has shown that incorporating auxiliary losses—such as silhouette-based clustering objectives or contrastive terms—can refine latent representations and improve clustering quality, particularly in challenging scientific datasets with overlapping structures [14]. Ensemble-based strategies and reconstruction-driven regularization further enhance robustness, reducing sensitivity to initialization and noise while preserving meaningful data variance [13, 18]. Several key insights emerge from the surveyed literature. First, the effectiveness of autoencoder-based dimensionality reduction and clustering is strongly data-dependent: optimal architectures, latent dimensionalities, and regularization strengths vary across datasets, motivating guided model selection strategies based on partial labels or cluster-quality metrics [13]. Second, while clustering-oriented losses can improve latent-space separability, the resulting gains are often incremental when the underlying unsupervised embedding is already strong, highlighting the importance of careful loss balancing to avoid over-regularization or variance collapse [4]. Third, robustness and stability remain critical concerns in unsupervised learning; ensemble approaches and reconstruction-based constraints have proven effective in mitigating these issues across clustering and anomaly detection tasks [18, 33, 34]. Finally, there are hybrid paradigms that combine unsupervised representation learning with weak or partial supervision—such as pseudo-labels—to steer embeddings without requiring full supervision [14].

In summary, autoencoders have become a cornerstone of modern high-dimensional data analysis. For dimensionality reduction, they uncover low-dimensional structures that preserve meaningful patterns and relationships in complex data. For clustering, they provide a flexible framework for learning representations that are inherently cluster-friendly, leading to improved accuracy, robustness, and interpretability. Ongoing advances—including multi-objective training, ensemble learning, and integration with generative and spatio-temporal models—continue to expand their scope and applicability [18–22, 29]. As datasets grow in size and complexity, autoencoder-based approaches are expected to remain at the forefront of representation learning, enabling scalable visualization, reliable clustering, and deeper insight into the latent structure of scientific and industrial data.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[2] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks", Science, 2006.

[3] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders", ICML, 2008.

[4] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes", ICLR, 2014.

[5] I. Goodfellow et al., "Generative adversarial nets", NeurIPS, 2014.

[6] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting", JMLR, 2014.

[7] A. Paszke et al., "Automatic differentiation in PyTorch", NeurIPS Workshops, 2017.

[8] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping", CVPR, 2006.

[9] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE", JMLR, 2008.

[10] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction", arXiv:1802.03426, 2018.

[11] A. Telea, "Data visualization: Principles and practice", CRC Press, 2014.

[12] A. Telea and J. J. van Wijk, "Simplified representation of vector fields", IEEE TVCG, vol. 5, no. 2, pp. 143–154, 1999.

[13] H. Gadirov, G. Tkachev, T. Ertl, and S. Frey, "Evaluation and selection of autoencoders for expressive dimensionality reduction of spatial ensembles", Advances in Visual Computing, pp. 222–234, 2021.

[14] H. Gadirov, L. Manuel, and S. Frey, "SENSE: Self-supervised neural embeddings for spatial ensembles", Journal of Baku Engineering University, vol. 9, no. 2, pp. 113–136, 2025.

[15] H. Gadirov, "Autoencoder-based feature extraction for ensemble visualization", Master's thesis, University of Stuttgart, 2020.

[16] H. Gadirov, "Machine learning for scientific visualization: Ensemble data analysis", PhD thesis, University of Groningen, 2025.

[17] O. Fernandes, S. Frey, G. Reina, and T. Ertl, "Visual representation of region transitions in multi-dimensional parameter spaces", Eurographics Italian Chapter Conference, 2019.

[18] H. Gadirov, M. Westra, and S. Frey, "TRACE: Reconstruction-based anomaly detection in ensemble and time-dependent simulations", arXiv:2601.08659, 2026.

[19] H. Gadirov, J. B. T. M. Roerdink, and S. Frey, "FLINT: Learning-based flow estimation and temporal interpolation for scientific ensemble visualization", IEEE TVCG, vol. 31, no. 10, pp. 7970–7985, 2025.

[20] H. Gadirov et al., "HyperFLINT: Hypernetwork-based flow estimation and temporal interpolation for scientific ensemble visualization", Computer Graphics Forum, vol. 44, no. 3, 2025.

[21] Z. Yin, H. Gadirov, J. Kosinka, and S. Frey, "ENTIRE: Learning-based volume rendering time prediction", arXiv:2501.12119, 2025.

[22] D. Bauer, Q. Wu, H. Gadirov, and K.-L. Ma, "GSCache: Real-time radiance caching for volume path tracing using 3D Gaussian splatting", IEEE TVCG, 2025.

[23] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3D Gaussian splatting for real-time radiance field rendering", ACM Transactions on Graphics (SIGGRAPH), vol. 42, no. 4, 2023.

[24] J. Zhou, W. Zhang, and Y. S. Liu, "DiffGS: Functional gaussian splatting diffusion", Advances in Neural Information Processing Systems (NeurIPS), vol. 37, pp. 37535–37560, 2024.

[25] Q. Gao, Y. Wang, and Z. Liu, "Can3Tok: Canonical 3D tokenization and latent modeling of scene-level 3D Gaussians", Proc. IEEE/CVF International Conference on Computer Vision (ICCV), 2025.

[26] H. Yang, S. Liu, and K. Chen, "Atlas Gaussians diffusion for 3D generation", Proc. International Conference on Learning Representations (ICLR), 2025.

[27] Y. Yang, Z. Ye, and J. Sun, "Prometheus: 3D-aware latent diffusion models for feed-forward text-to-3D scene generation", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025.

[28] Z. Dai, T. Liu, and Y. Zhang, "Efficient decoupled feature 3D Gaussian splatting via hierarchical compression", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025.

[29] R. Li and Y. M. Cheung, "Variational multi-scale representation for estimating uncertainty in 3D Gaussian splatting", Advances in Neural Information Processing Systems (NeurIPS), vol. 37, 2024.

[30] Y. LeCun and C. Cortes, "The MNIST database of handwritten digits", 2005.

[31] D. Merkel, "Docker: Lightweight Linux containers for consistent development and deployment," Linux Journal, vol. 2014, no. 239, 2014.

[32] H. Gadirov, "Report on container technology for the ATLAS TDAQ system," (No. CERN-STUDENTS-Note-2016-100).

[33] L. Ruff et al., "Deep one-class classification", Proc. International Conference on Machine Learning (ICML), 2018.

[34] T. Schlegl et al., "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery", Proc. Information Processing in Medical Imaging (IPMI), 2017.

[35] G. Hinton, Y. Bengio, and A. Courville, "Representation learning: A review and new perspectives", IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 35, no. 8, pp. 1798–1828, 2013.