

A Review of Clustering Algorithms: Evolution, Comparison, and Challenges from Traditional Clustering to Deep Clustering

Zihao Fan*

School of Computer Science and Engineering, Central South University, Changsha 410083, China

**Corresponding author: Zihao Fan.*

Abstract

Against the background of big data, clustering algorithms, as a core task of unsupervised learning, have been widely applied, and the related theories and technologies have continued to iterate and develop. Existing reviews mostly focus on the classification and comparison of traditional clustering algorithms, while insufficient attention has been paid to the evolutionary relationship between traditional clustering and deep clustering. Therefore, this paper reviews clustering algorithms and their evolutionary logic, systematically sorts out the evolutionary route of clustering algorithms from traditional methods to deep models, describes the relatively classic and main clustering algorithms, and compares and analyzes their advantages and disadvantages. It also shows the key transition from classic clustering to deep clustering. In addition, this paper describes the core challenges that clustering algorithms may face in the future and analyzes further research directions according to current research progress.

Keywords

clustering algorithm, deep clustering, unsupervised learning

1. Introduction

With the arrival of the big data era in recent years, machine learning technology has been developing rapidly. Clustering algorithms, as a major category of machine learning algorithms, are also progressing quickly. Clustering algorithms have been applied to image segmentation, information retrieval, data mining, market analysis, and many other fields [1]. Clustering is a core task of unsupervised learning. Different from supervised learning, which provides a label pattern [2], the goal of clustering is to automatically group unlabeled data so that the similarity of objects within a class is maximized and the similarity of objects between classes is minimized as much as possible. At present, existing reviews mostly focus on the classification and comparison of traditional algorithms or discuss deep clustering separately. In recent years, deep clustering has developed rapidly, while related literature reviews are relatively limited. Therefore, this paper focuses on sorting out the evolutionary context of clustering algorithms, analyzes traditional clustering algorithms and deep clustering algorithms, explores the connection between them, and predicts future research directions and core challenges. This can help readers understand the related knowledge and evolutionary logic of clustering algorithms and provide a reference for future research and application.

2. Traditional Clustering Algorithms

2.1 The Traditional K-means Clustering Algorithm

There are many types of traditional clustering algorithms, and they are now often divided according to different schools, such as density-based, distribution-based, and graph-based methods. Among them, the most popular and widely used one is the classic k-means clustering algorithm based on centroid partitioning. The k-means algorithm was independently proposed by Steinhaus in 1955, Lloyd in 1957, Ball and Hall in 1965, and McQueen in 1967 in their own different research fields. More than 50 years have passed since k-means was proposed, and it is still one of the most widely used traditional clustering algorithms [3].

The essence of the k-means algorithm is to solve a mathematical optimization problem. Its goal is to divide unlabeled data into k clusters, so that the sum of squared distances from each data point to the centroid of the cluster to which it belongs is minimized, that is, the within-cluster sum of squares is minimized. Euclidean distance is generally used as the index for measuring the similarity between unlabeled data, and similarity is inversely proportional to the distance between data. The core principle of the k-means algorithm is iterative optimization. First, K initial samples are randomly selected from the data set as clustering centers $\{C_1, C_2, \dots, C_k\}$. Then, in the assignment stage of each iteration, the Euclidean distance between each data object and each clustering center is calculated, and the related data object is assigned to the cluster corresponding to the nearest clustering center according to the nearest-neighbor criterion. In the update stage, the arithmetic mean of all data objects in each cluster is recalculated and used as the new clustering center of that cluster. The above assignment and update processes are repeated until the clustering centers no longer change significantly, or the number of iterations reaches the preset upper limit [4]. The related formulas are expanded in reference [4].

The k-means clustering algorithm is simple, efficient, easy to understand, and highly efficient, so it is very suitable for processing large-scale data. These advantages make it one of the most popular algorithms. However, it also has certain limitations, including the need to preset the value of k , uncertainty caused by abnormal values in data, sensitivity to noise and outliers, and clustering validity [5]. In view of the shortcomings of the k-means algorithm, scholars in different fields have proposed related variants and improvements. For example, M. Emre Celebi and other researchers improved the shortcomings by using initialization methods with different time complexities and superlinear initialization methods [6]. At the same time, optimization can also be carried out based on different functions and algorithms [7]. These studies have provided great help for the further application of the k-means algorithm.

2.2 Other Classic Clustering Algorithms

Although the k-means algorithm is efficient and easy to understand, it mainly solves problems based on the centroid partitioning method, which is not applicable to some different situations and may instead affect the complexity of the algorithm. Meanwhile, other classic clustering algorithms have also made related improvements on the basis of k-means. Therefore, for different problems, different schools have different answers to the question of what a good clustering result is. As shown in Table 1, this paper organizes and compares other classic clustering algorithms from different schools and explores their differences from k-means.

Representative density-based clustering algorithms include DBSCAN and OPTICS, and their core idea is to detect non-spherical or arbitrarily shaped clusters. Among them, DBSCAN does not need to preset the number of clusters and has the concept of noise, while OPTICS does not need to provide a density threshold [8]. However, when the sample density is uneven, the clustering quality may be low. The most representative distribution-based clustering method is the Gaussian mixture model (GMM). Its data follow a mixed probability distribution, and it belongs to soft clustering [9]. Hierarchical clustering includes agglomerative clustering and divisive clustering. It also does not need to preset the value of k and can use various distance calculation methods, but for this reason it may have low efficiency when facing large amounts of data [10]. The core of graph-based or spectral clustering is to group nodes into meaningful clusters according to graph structures or features. For example, spectral clustering clusters graph data according to the eigenvalues and eigenvectors of a similarity matrix, which is helpful for revealing data information with complex structures. However, such methods have relatively high requirements for the selection of hyperparameters [11]. STING is a representative grid-based clustering algorithm. This kind of method uses a grid structure, divides the space

by rectangular blocks, and realizes clustering according to the distribution information of the blocks [12]. This hierarchical grid structure is convenient for efficiently processing data. However, such algorithms need to select appropriate grid-size parameters in advance; otherwise, the clustering granularity and accuracy may be affected [13].

From the comparative analysis of various traditional clustering algorithms, it can be seen that the quality of each classic clustering algorithm is not absolute. In the face of different problems and situations, each clustering algorithm has its own advantages and disadvantages. This diversity is exactly the significance of comparing and analyzing classic clustering algorithms.

Table 1: Comparison of Other Classic Clustering Algorithms

School	Representative algorithms	Core idea	Key difference from k-means	Applicable scenarios
Density-based	DBSCAN, OPTICS, Mean Shift	Clusters are formed by high-density regions, and low-density regions are regarded as boundaries.	K does not need to be preset; noise can be automatically identified; no density threshold is required in some methods.	Spatial data and irregularly shaped clusters
Distribution-based	GMM (Gaussian mixture model)	Data are assumed to follow a mixed probability distribution.	Soft clustering is used, and membership probabilities are output.	Density estimation and soft-partition scenarios
Hierarchical clustering	AGNES, DIANA	A tree-like nested structure, namely a dendrogram, is constructed.	Hierarchical cluster relationships can be output, and K does not need to be preset.	Biological classification and hierarchical text organization
Graph/spectral-based	Spectral clustering	Graph cut and eigendecomposition are used.	It is suitable for non-convex distributions and is based on a similarity graph rather than distance.	Image segmentation and social networks
Grid-based	STING	The space is divided into grid cells.	Large-scale data can be processed efficiently.	Spatial data mining

3. Deep Clustering Algorithms

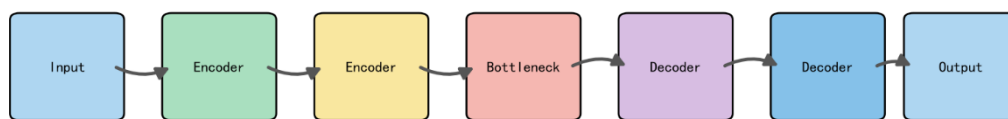
The current period is an information-explosion era of big data. Machine learning is developing rapidly, and traditional clustering algorithms have gradually become unable to meet the needs of current clustering research. Therefore, the development of deep clustering algorithms is especially important. Deep clustering is clustering based on deep learning. In essence, it is a relatively advanced clustering method that improves clustering results through the powerful representation ability of deep learning [14]. At present, data information is large in scale, strong in real-time performance, and complex in structure. When classic clustering algorithms face such complex and massive data, they often cannot achieve ideal results. The reasons include that classic clustering generally pays attention only to local data information, traditional Euclidean distance has difficulty evaluating sample similarity, and the data distribution in high-dimensional space is sparse [15]. In contrast, deep clustering, which has developed vigorously in recent years, has excellent nonlinear mapping ability. When combined with other models, it can obtain stronger ability to process complex data [16]. This chapter describes four technical schools of deep clustering and conducts related analysis, showing the change from traditional clustering to deep clustering and from designing distances to learning representations.

3.1 Autoencoder-based Deep Clustering

The autoencoder (AE) algorithm is one of the most important deep clustering algorithms. It is a classic unsupervised learning method, and its good feature learning ability is helpful for identifying the significant features of data. As shown in Figure 1, an autoencoder has two parts, an encoder and a decoder. During operation, it first compresses high-dimensional input, such as an image, into a low-dimensional core feature vector through the encoder, and then allows the decoder to reconstruct an output that is as consistent as possible

with the input based on this feature. In this way, the clustering objective is optimized while dimensionality is reduced and high-dimensional data information is retained [17]. The whole training process does not depend on manual labels. It only requires the error between the output and the input to be minimized, and the network is forced to learn the most essential and key features in the data under the effect of the bottleneck layer. Xie et al. [18] carried out pioneering work in 2016 and proposed deep embedded clustering (DEC), which jointly optimized feature learning and clustering assignment for the first time. Later, many researchers also proposed various improved algorithms based on autoencoders, such as the method combined with k-means proposed by Yang et al. [19] and the unsupervised deep clustering method based on a convolutional residual autoencoder proposed by Guo [20]. Autoencoder-based deep clustering has evolved from a two-stage process of first pretraining the AE and then clustering to end-to-end joint optimization. Compared with traditional clustering, it has further changed into a method for learning data representation, so efficiency has been greatly improved. However, current autoencoder-based deep clustering methods are not suitable for all situations and may have limitations such as large fluctuations in algorithm results and high resource requirements.

Figure 1: Schematic Structure of an Autoencoder



3.2 Generative-model-based Deep Clustering

Generative-model-based deep clustering is a kind of clustering method that uses generative models to learn the rules behind data and naturally group similar data. Traditional clustering methods often do not perform ideally in processing high-dimensional complex data such as images and text. The core idea of this kind of deep algorithm is to use a variational autoencoder (VAE) or a generative adversarial network (GAN) [21] as the core, map data into a hidden space, generate new data from this space, and then let researchers design the mathematical distribution of the space. Finally, the model is trained and optimized at the same time. The main advantages of this kind of algorithm are as follows. First, it can capture more complex data distributions. Traditional clustering relies on Euclidean distance, while generative models are good at learning complex nonlinear distributions. For example, Mashfiquel et al. [22] studied and analyzed VAE-based generative models and improved clustering performance. Second, features and clustering promote each other. For example, Yan et al. [23] proposed a clustering framework based on DDPMs and achieved the goal of learning clustering-friendly latent features. However, the limitations of this clustering method are that training is difficult, the computational cost is high, and it is easy to fall into suboptimal solutions. The study of DGVAC by Ruan et al. [24] was intended to reduce training difficulty and ensure stability. Different from the method of calculating distance measures in traditional clustering, generative-model-based deep clustering identifies the hidden structure of data in a low-dimensional latent space through representation learning. This is the core driving force for the transformation of traditional clustering from a shallow paradigm that depends on preset distances to a deep paradigm that can independently mine structures.

3.3 Contrastive-learning-based Deep Clustering

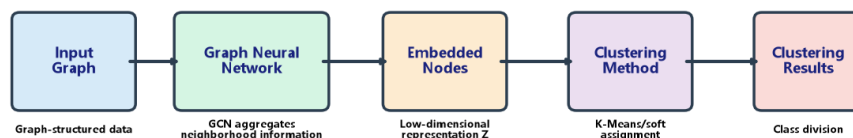
Contrastive learning is a kind of unsupervised learning method that aims to enable a model to learn the essential features of data by learning to distinguish similar and different data pairs [25]. Its core strategy is to construct positive and negative sample pairs through data augmentation [26], pull positive sample pairs, namely samples with similar semantics, closer in the latent space, and push negative sample pairs, namely samples with different semantics, farther away in the space. In this way, contrastive loss is used to learn discriminative and invariant feature representations. Relatively classic methods of contrastive learning include SimCLR [27], a simple framework for contrastive learning of visual representations; SwAV [28], which uses a contrastive method without computing pairwise comparisons and saves memory; and BYOL [29], a new method for self-supervised image representation learning. In recent years, there have been relatively few reviews of contrastive-learning-based deep clustering, but research in this area has made many key advances. Zhang et al. [30] proposed supporting clustering with contrastive learning (SCCL), which uses contrastive learning to promote a deep clustering effect with better separation. Li et al. [31] proposed contrastive clustering (CC), which conducts contrastive learning at both the instance level and the cluster level under a unified

framework. Xing [32] introduced the idea of self-paced learning, proposed an instance-level contrastive clustering algorithm based on self-paced learning, and further proposed a graph-structure-based instance-cluster-level contrastive clustering algorithm. This extends from the instance level to the cluster level and captures the potential relationships between data for feature learning and cluster assignment. The advantage of contrastive-learning-based deep clustering is that it has stronger feature discrimination and can better extract high-level semantics [25]. Compared with traditional clustering, its training efficiency is also higher. This kind of deep clustering algorithm focuses on learning data representation, is showing great potential, and has become a mainstream research direction. However, it currently depends highly on data augmentation strategies [33], and false negative samples may be introduced easily because of the randomness in selecting negative samples, resulting in low clustering efficiency [34]. These limitations need to be overcome by future research.

3.4 Graph-neural-network-based Deep Clustering

Traditional deep clustering methods usually regard samples as independent individuals and use only their attribute features for representation learning, while ignoring the topological dependence relationships between samples. In reality, data are often presented in graph structures, and the rich edge information between nodes contains strong cluster priors. Therefore, graph-neural-network-based deep clustering plays an important role. It is an algorithm that uses neural network models for deep clustering, can effectively process complex data, and can capture nonlinear node attributes [35]. Its structure is shown in Figure 2. Its core idea is to use a graph neural network (GNN) [36] to aggregate the feature information of nodes, so that the learned node representations can integrate graph-structure context and better capture the structures and relationships between graphs. With the deepening of research, neural networks have been continuously improved. Convolutional neural networks have formed a complete hierarchical structure [37]. Deep convolutional neural networks [38] have been optimized by methods such as local perception and pooling sampling. Kipf et al. [39] proposed a graph convolutional network (GCN), which improves classification performance by efficiently combining node features and graph topology. Fan et al. [40] made the first attempt to apply deep learning to attributed multi-view clustering and proposed a novel task-guided clustering framework, while jointly optimizing clustering label assignment and embedding representation. Graph-neural-network-based deep clustering adapts to graph-structure information with complex distributions and can integrate node attribute features and graph topology through end-to-end learning. This core transformation from designing measures to learning representations has also caused new challenges such as over-smoothing, high learning costs, and dependence on input quality.

Figure 2: Deep Clustering Structure Based on Graph Neural Networks



3.5 Comparative Analysis of Representative Algorithms

There are many schools of deep clustering algorithms, and the performance and efficiency of different algorithms also differ for different types of data. Therefore, this subsection conducts a comparative review of the performance and applicable objects of deep clustering algorithms based on existing literature and experimental data from research, and summarizes the performance of each algorithm in practical applications.

Evaluation results on standard visual data sets show that different architectures have significant data preferences. Xie et al. proposed the autoencoder-based DEC method, namely deep embedded clustering [18]. They evaluated it on a group of text data sets and two image data sets and compared it with k-means, SEC, and other algorithms. According to the experimental data of Xie et al., DEC achieved a good clustering effect on the structurally simple MNIST data set and showed a clustering accuracy of 84.30%. At the same time, DEC showed excellent performance and strong robustness in hyperparameter tuning, and it had the advantage of linear complexity in the number of data points, making it suitable for complex data sets. The contrastive-based PICA algorithm, or partition confidence maximization, was proposed by Huang et al. [41]. It enables

clusters to map one-to-one to real classes by maximizing the global partition confidence of the clustering solution. PICA has significantly higher NMI than other algorithms on complex visual tasks such as ImageNet-10 and has high ACC on STL-10, proving that PICA has a high level and excellent ability in image clustering. Chowdhury et al. proposed the graph-neural-network-based DMVC algorithm [42], which aims to learn consistent and complementary representations across views to improve clustering performance. This algorithm is suitable for complex image and video data and also plays an important role in multi-source heterogeneous data [43]. In terms of performance and efficiency, AE-based algorithms do not need to construct positive and negative sample pairs, so their training cost is relatively low and they are suitable for data sets with relatively simple structures. Contrastive-based algorithms shape the feature space through contrastive learning of positive and negative samples, so their training time is often several times that of the former. However, they have significant advantages in complex visual tasks and are more suitable for data sets with complex semantics and rich categories. GNN-based methods for graph data are almost specifically used for graph-structured data and multimodal graph structures and can also be extended through design, but their performance is limited by the quality of the graph structure, such as edge sparsity and over-smoothing.

4. Challenges and Prospects

4.1 Core Challenges

Although clustering algorithms have been widely applied in different fields, both traditional clustering algorithms and deep clustering algorithms still face different challenges. For traditional clustering algorithms, their core challenges are mainly reflected in the following aspects. First, they have strong assumptions about data distribution and are sensitive to hyperparameters. Taking the traditional k-means algorithm as an example, the number of clusters K needs to be specified in advance [4], but in practical applications, the true number of clusters is often unknown and may change dynamically. Misestimating K will lead to suboptimal or even misleading clustering results. In addition, k-means is sensitive to the selection of initial clustering centers and is easy to fall into a local optimum [44]. Second, it is difficult to process noisy data. For example, the sensitivity of hierarchical-clustering-based algorithms changes according to the linkage criterion used, and noise may hinder clustering from merging at the correct scale [45], thus affecting clustering performance. Although deep clustering algorithms use the strong nonlinear mapping ability of deep neural networks and alleviate the difficulty of traditional methods in high-dimensional data to a certain extent, they also face challenges. First, the coupling of representation learning and clustering objectives is one of the core difficulties. Early methods, such as DEC, adopted a multi-stage strategy of feature learning first and clustering later, and feature extraction and clustering assignment were independent of each other, so it was difficult to guarantee the optimal effect of representation learning. Second, deep clustering often faces complex high-dimensional data. For example, in contrastive clustering, data have complex semantic information, which brings great challenges to the selection of positive and negative sample pairs [46].

4.2 Future Directions

Facing the future, the development of clustering algorithms shows new research directions. For both traditional clustering and deep clustering, computational efficiency and stability should be key directions of concern. Most clustering methods will eventually be transformed into combinatorial optimization problems, and when large-scale complex data are involved, computational efficiency becomes very important. At the same time, a stable clustering algorithm should also be able to produce stable solutions when encountering data perturbations [5]. For deep clustering, exploratory research in the direction of prior knowledge is advancing. For example, Lu et al. [47] studied the innovation of prior knowledge and analyzed the performance of methods under different prior conditions. The diversification of prior knowledge may become a new research direction for deep clustering. Second, a unified clustering framework across data types remains to be explored. At present, clustering methods for images, text, and mathematical data each form their own systems, and there is a lack of a general clustering architecture that can uniformly process multi-source heterogeneous data. With the development of research, various clustering algorithms have gradually penetrated into many industry fields, including oceanography [48], architectural engineering [49], biology and medicine [16, 50, 51], community detection [52], and others, and they have become effective tools for processing complex tasks.

5. Conclusion

Clustering algorithms are very important unsupervised learning methods in modern times and play a significant role in data partitioning. This paper sorts out the evolutionary context of clustering algorithms, shows the key transition from designing distance measures in traditional clustering algorithms to learning data representations in deep clustering algorithms, and analyzes the challenges and future research directions faced by clustering algorithms. At present, classic clustering has laid the foundation of multiple paradigms, while deep clustering has broken through the bottleneck of traditional methods on high-dimensional complex data through representation learning. Contrastive learning has become a frontier research topic in current deep clustering, but challenges such as interpretability and generalization still need to be solved and broken through. In the future, clustering algorithms will be combined with large models, further adapt to dynamic data streams, introduce richer prior knowledge, and bring great advantages to various industry fields.

References

- [1] Singh, J., & Singh, D. (2024). A comprehensive review of clustering techniques in artificial intelligence for knowledge discovery: Taxonomy, challenges, applications and future prospects. *Advanced Engineering Informatics*, 62, Article 102799. <https://doi.org/10.1016/j.aei.2024.102799>
- [2] Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., ... & Lin, C. T. (2017). A review of clustering techniques and developments. *Neurocomputing*, 267, 664-681. <https://doi.org/10.1016/j.neucom.2017.06.053>
- [3] Wang, Q., Wang, C., Feng, Z. Y., & Ye, J. F. (2012). Review of the k-means clustering algorithm. *Electronic Design Engineering*, 20(7), 21-24. <https://doi.org/10.14022/j.cnki.dzsjgc.2012.07.034>
- [4] Yang, J. C., & Zhao, C. (2019). Review of the k-means clustering algorithm. *Computer Engineering and Applications*, 55(23), 7-14, 63. <https://doi.org/10.3778/j.issn.1002-8331.1908-0347>
- [5] Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8), 651-666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- [6] Celebi, M. E., et al. (2013). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*, 40(1), 200-210. <https://doi.org/10.1016/j.eswa.2012.07.021>
- [7] Wu, S. H., Cheng, Y., Zheng, Y. N., & Pan, Y. T. (2011). Review of the k-means algorithm. *New Technology of Library and Information Service*, (5), 28-35.
- [8] Bhuyan, R., & Borah, S. (2023). A survey of some density-based clustering techniques. *arXiv*. <https://doi.org/10.48550/arXiv.2306.09256>
- [9] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [10] Deng, L. P. (2019). Review of classic clustering algorithms. *Public Communication of Science & Technology*, 11(5), 108-110. <https://doi.org/10.16607/j.cnki.1674-6708.2019.05.053>
- [11] Watteau, T., Bonnefoy, A., Illouz-Laurent, S., Jusseau, J., & Iovleff, S. (2024). Advanced graph clustering methods: A comprehensive and in-depth analysis. (2024). *arXiv*. <https://doi.org/10.48550/arXiv.2407.09055>
- [12] Sun, J. G., Liu, J., & Zhao, L. Y. (2008). Clustering algorithms research. *Journal of Software*, 19(1), 48-61. <https://doi.org/10.3724/SP.J.1001.2008.00048>
- [13] Wang, C., Li, Y., & Xie, N. (2026). Beyond theory: A comprehensive survey of clustering algorithm applications in wireless and media communications. *Digital Communications and Networks*. <https://doi.org/10.1016/j.dcan.2026.06.004>
- [14] Dong, Y. F., Deng, Y. H., Dong, Y., & Wang, Y. C. (2022). A review of clustering based on deep learning. *Journal of Computer Applications*, 42(4), 1021-1028. <https://doi.org/10.11772/j.issn.1001-9081.2021071275>

- [15] Ji, Q., Sun, Y. F., Hu, Y. L., & Yin, B. C. (2021). Review of deep clustering algorithms. *Journal of Beijing University of Technology*, 47(8), 912-924. <https://doi.org/10.11936/bjutxb2021010013>
- [16] Zhou, S., Xu, H., Zheng, Z., Chen, J., Li, Z., Bu, J., Wu, J., Wang, X., Zhu, W., & Ester, M. (2025). A comprehensive survey on deep clustering: Taxonomy, challenges, and future directions. *ACM Computing Surveys*, 57(3), Article 69, 1-38. <https://doi.org/10.1145/3689036>
- [17] Tao, W. B., Qian, Y. R., Zhang, Y. Y., Ma, H. Z., Leng, H. Y., & Ma, M. N. (2022). Review of deep clustering algorithms based on autoencoders. *Computer Engineering and Applications*, 58(18), 16-25.
- [18] Xie, J., Girshick, R., & Farhadi, A. (2016). Unsupervised deep embedding for clustering analysis. In *Proceedings of the International Conference on Machine Learning* (pp. 478-487).
- [19] Yang, X., Deng, C., Zheng, F., et al. (2019). Deep spectral clustering using dual autoencoder network. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4066-4075). Long Beach, CA, United States.
- [20] Guo, M. (2024). Research on deep clustering methods based on convolutional residual autoencoders [Master's thesis, Lanzhou University of Finance and Economics].
- [21] Zhou, K., Huang, X., Song, Q., Chen, R., & Hu, X. (2022). Auto-GNN: Neural architecture search of graph neural networks. *Frontiers in Big Data*, 5, 1029307. <https://doi.org/10.48550/arXiv.1909.03184>
- [22] Chowdhury, M. H., Hirose, Y., Marsland, S., & Yao, Y. (2025). Mixtures of posterior and prior variational autoencoders for representation learning and cluster analysis in latent space. *Neurocomputing*, 657, Article 131524.
- [23] Yan, J., Liu, J., & Zhang, Z.-Y. (2025). ClusterDDPM: An EM clustering framework with denoising diffusion probabilistic models. *Information Sciences*, 720, Article 122518.
- [24] Ruan, T., Yu, N., Zhao, S., Chen, F., & Huang, Y. (2026). DGVAC: Disentangled Gaussian-vMF alignment for deep clustering. *Neural Networks*, 203, Article 109114.
- [25] An, J. X., Liu, Y., & Yang, L. W. (2025). Review of deep clustering based on contrastive learning. *Microelectronics & Computer*, 42(7), 1-10.
- [26] Dosovitskiy, A., Springenberg, J. T., Riedmiller, M., & Brox, T. (2014). Discriminative unsupervised feature learning with convolutional neural networks. *Advances in neural information processing systems*, 27.
- [27] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In *International conference on machine learning* (pp. 1597-1607). PmLR.
- [28] Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., & Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. *Advances in neural information processing systems*, 33, 9912-9924.
- [29] Grill, J. B., Strub, F., Altché, F., Tallec, C., Richemond, P., Buchatskaya, E., ... & Valko, M. (2020). Bootstrap your own latent-a new approach to self-supervised learning. *Advances in neural information processing systems*, 33, 21271-21284.
- [30] Zhang, D., Nan, F., Wei, X., Li, S. W., Zhu, H., McKeown, K., ... & Xiang, B. (2021, June). Supporting clustering with contrastive learning. In *Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 5419-5430).
- [31] Li, Y., Hu, P., Liu, Z., Peng, D., Zhou, J. T., & Peng, X. (2021). Contrastive clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 8547-8555). AAAI Press.
- [32] Xing, R. W. (2024). Research on deep clustering algorithms based on contrastive learning [Master's thesis, Shanxi University of Finance and Economics].
- [33] Li, H. J. (2024). Research on deep clustering algorithms based on contrastive learning [Master's thesis, Guilin University of Technology].

- [34] Ju, H., Guo, J., Ding, W., Pedrycz, W., Cheng, X., & Yang, X. (2025). FDGC: Fuzzy deep clustering with dual-granularity contrastive learning. Knowledge-Based Systems. <https://doi.org/10.1016/j.knosys.2025.114401>
- [35] Berahmand, K., Li, Y. F., & Xu, Y. (2023). DAC-HPP: Deep attributed clustering with high-order proximity preserve. *Neural Computing and Applications*, 35, 24493-24511.
- [36] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2008). The graph neural network model. *IEEE transactions on neural networks*, 20(1), 61-80.
- [37] Huang, X. Y., Zhang, C. H., Zhao, W., & Wang, J. K. (2024). Review of deep clustering research. *Journal of Sichuan University of Science & Engineering (Natural Science Edition)*, 37(4), 47-57.
- [38] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [39] Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv. <https://arxiv.org/abs/1609.02907>
- [40] Fan, S., Lu, E., Wang, X., Lin, K., Shi, C., & Wang, B. (2020). One2multi graph autoencoder for multi-view graph clustering. In *Proceedings of the Web Conference 2020* (pp. 3070-3076). Association for Computing Machinery.
- [41] Huang, J., Gong, S., & Zhu, X. (2020). Deep semantic clustering by partition confidence maximisation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 8849-8858).
- [42] Chowdhury, A. R., Gupta, A., & Das, S. (2025). Deep multi-view clustering: A comprehensive survey of the contemporary techniques. *Information Fusion*, 119, Article 103012. <https://doi.org/10.1016/j.inffus.2025.103012>
- [43] Jin, X., Chen, M., Wang, Y., & Li, S. (2026). CDMVC: Consensus-driven multiview clustering via adaptive hierarchical matrix factorization with dynamic view fusion. *IEEE Internet of Things Journal*, 13(8), 16480-16495. <https://doi.org/10.1109/JIOT.2026.3659704>
- [44] Wang, S., Liu, C., & Xing, S. J. (2022). Review of the k-means clustering algorithm. *Journal of East China Jiaotong University*, 39(5), 119-126.
- [45] Wani, A. A. (2024). Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions. *PeerJ Computer Science*, 10, e2286.
- [46] Li, Y., Cai, J., Yang, H., Yang, Y., Wang, J., & Hao, J. (2026). A survey of contrastive clustering research: algorithms, applications and challenges. *Complex & Intelligent Systems*. <https://doi.org/10.1007/s40747-026-02413-1>
- [47] Lu, Y., Li, H., Li, Y., Lin, Y., & Peng, X. (2024). A survey on deep clustering: from the prior perspective. *Vicinearth*, 1(1), 4. <https://doi.org/10.1007/s44336-024-00001-w>
- [48] Li, J. T. (2019). Research on density clustering algorithms and their application in ocean mesoscale phenomena [Master's thesis, Shandong University of Science and Technology].
- [49] Huang, Y. Y., Tursun, W., Tu, R. X., Zhang, X., & Chen, X. K. (2025). Research on traditional block building classification and style renovation strategies based on clustering algorithms: A case study of Laozheng Street in Dongshi Town. *Huazhong Architecture*, 43(4), 120–123.
- [50] Han, Q. H., Ma, Q., Liu, M., & Xu, J. (2021). Damage diagnosis of spatial grid structures based on natural-frequency clustering analysis under temperature changes. *Journal of East China Jiaotong University*, 38(4), 8-17.
- [51] Nawaz, M., Mehmood, Z., Nazir, T., Naqvi, R. A., Rehman, A., Iqbal, M., & Saba, T. (2022). Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microscopy research and technique*, 85(1), 339-351.

- [52] Wang, R., Wang, P., Li, L., & Liu, P. (2023, May). Community detection with graph convolutional auto-encoder and deep clustering. In 2023 26th international conference on computer supported cooperative work in design (CSCWD) (pp. 1589-1594). IEEE.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).