Indian Journal of Information Sources and Services
ISSN: 2231-6094 (P) Vol.15, No.3, 2025, pp.312-318

DOI: https://doi.org/10.51983/ijiss-2025.IJISS.15.3.35

Hybrid Clustering Techniques for Multimedia Information Retrieval

Gopal Srinivas[1*], Hassan Mohamed[2], Dilfuzakhon Saidkodirova[3], Mokhichekhra Boltaeva[4], Sirojbek Bakhtiyarov[5] and Dr. Ankita Nihlani[6]

[1*]Department of Nautical Science, AMET University, Kanathur, Tamil Nadu, India
[2]Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University in Najaf, Najaf, Iraq; Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University in Najaf of Al Diwaniyah, Al Diwaniyah, Iraq
[3]Associate Professor, Uzbek State World Languages University, Uzbekistan
[4]Jizzakh Branch of the National University of Uzbekistan named after Mirzo Ulugbek, Uzbekistan
[5]Lecturer, Mamun University, Khiva, Uzbekistan
[6]Assistant Professor, Department of Management, Kalinga University, Naya Raipur, Chhattisgarh, India
E-mail: [1]gopalsrinivas@ametuniv.ac.in, [2]eng.hassanaljawahry@gmail.com, [3]disuper@mail.ru, [4]bmoxichexra@bk.ru, [5]baxtiyarov_sirojbek@mamunedu.uz, [6]ku.ankitanihlani@kalingauniversity.ac.in
ORCID: [1]https://orcid.org/0000-0002-4971-3891, [2]https://orcid.org/0009-0009-1422-1844, [3]https://orcid.org/0000-0001-6873-2185, [4]https://orcid.org/0000-0002-8099-672X, [5]https://orcid.org/0009-0005-6596-5975, [6]https://orcid.org/0009-0002-7447-8544
(Received 21 June 2025; Revised 09 August 2025, Accepted 25 August 2025; Available online 30 September 2025)

***Abstract* - The rapid development of multimedia data, including pictures, music, and videos, has made MIR, or Multimedia Information Retrieval, a distinct area of ongoing interest. Retrieval methodologies often fail to organize and process vast and heterogeneous datasets containing unformatted information. This issue has recently been addressed by hybrid approaches that combine multiple clustering techniques, such as hierarchical, density-based, partition-based, and model-based clustering, among others. Multimedia retrieval and user convenience are enhanced as these methods improve the understanding of content and its categorization. This document summarizes the hybrid clustering approaches and their use in MIR systems. The paper discusses the application of feature-level fusion and metadata, as well as semantic-based clustering, which broadens the scope of scaling and strengthening retrieval systems. Focusing on the multimedia database challenge of high-dimensional feature spaces, hybrid approaches that integrate unsupervised methods, such as k-means and DBSCAN, with supervised or semi-supervised techniques overcome the barriers of noisy data, pattern concealment, and adaptive flexibility. Special attention is given to applying deep learning with clustering models for feature extraction, where CNNs and autoencoders serve as powerful hierarchical feature extractors for multimedia data. The paper focuses on recent developments regarding hybrid clustering frameworks that incorporate ontologies, graph-based models, and multimodal fusion approaches to enhance contextual understanding and semantic retrieval alignment. The evaluation of benchmark datasets confirms that hybrid clustering outperforms traditional single-method approaches, achieving higher Precision, Recall, and F1 Scores. The discussion addresses the obstacles of computational complexity, real-time processing, and scalability. The integration of explainable AI and privacy-preserving MIR through federated learning presents promising possibilities. At the same time, hybrid clustering techniques provide the backbone for efficient and intelligent multimedia information retrieval, enabling more context-aware and user-centric retrieval in diverse applications, such as digital libraries, surveillance, social media, and healthcare.**

***Keywords*: Multimodal Data, Unsupervised Learning, Semantic Analysis, Feature Fusion, Deep Learning, Multimedia Information Retrieval, Hybrid Clustering**

I. INTRODUCTION

Multimedia information retrieval (MIR) is a complex type of data retrieval that involves extracting, indexing, and retrieving data from images, audio, videos, and text (Datta et al., 2008). Unlike classic systems, which rely on retrieving data expressed as text, MIR systems require the comprehension and processing of data described in data structures that differ in meaning and structure. MIR systems are developed to allow users to retrieve multimedia and associated resources in a relevant and accurate manner by formulating questions presented in text, images, or other combinations of queries. The recent explosive growth of social media content, continued improvement of surveillance systems, and access to online libraries have created an incredible demand for scalable and sophisticated MIR systems (Lew et al., 2006). As the primary instructional design framework described in the previous section, these frameworks are based on the foundational principles of clustering and its usefulness in organizing large amounts of data, as well as in improving the effectiveness and speed of queries and their results (Jain et al., 1999). Besides these, in the organization of multimedia documents, the processes of clustering, pre-processing, indexing, and retrieval, as well as pattern recognition, and the subsequent processing of objects at rest and in motion, enhance the surface of the multimedia objects. MIR, on the contrary, faces the most serious challenges, including high dimensionality, noise or redundancy in data, the integration of multimodal data, and a semantic gap between lower and higher-level features and concepts (Smeulders et al., 2000; Mikolov et al., 2013). The ubiquitous k-means, hierarchical clustering, and even

DBSCAN, in their own right, face challenges with massive datasets, especially when the datasets are large and unstructured. There are hybrid clustering approaches developed to fill these voids, which seek to combine multiple clustering techniques to strengthen the method (Mikolov et al., 2013; Xu & Wunsch, 2005). Such approaches integrate partitioning and density-based techniques and may incorporate unsupervised and supervised techniques. The recent advent of deep learning architectures, notably convolutional neural networks (CNNs) and autoencoders, has further enhanced feature extraction and semantic understanding in music information retrieval (MIR) systems (Guo et al., 2016).

Key Contribution: This paper is dedicated to the analysis of the hybrid clustering techniques in relation to multimedia information retrieval. It classifies hybrid clustering techniques, analyzes the performance of these techniques in various multimedia domains, examines deep learning techniques to enhance feature and semantic capabilities, evaluates advanced real-world implementations, and recommends further research to develop more intelligent and user-friendly MIR systems.

II. Literature Review

No.	Author(s)	Year	Focus Area	Key Contribution	Limitation Identified
1	Aggarwal & Reddy (Reddy, 2018)	2014	High-dimensional multimedia clustering	Proposed scalable clustering for multimedia with dimensionality reduction	Limited semantic interpretation of visual data
2	Zhang et al. (Zhang et al., 2015)	2015	Image retrieval using K-means and DBSCAN	Improved clustering performance by comparing two popular algorithms	Lacked robustness across mixed data types
3	Bhatti et al. (Bhatti et al., 2016)	2016	Spectral clustering for video segmentation	Demonstrated the applicability of spectral clustering to temporal video analysis	High computational cost
4	Huang et al. (Huang et al., 2019)	2018	Audio clustering using MFCC and hierarchical methods	Used audio features with hierarchical clustering for retrieval	Scalability issues with large audio datasets
5	Zhou et al., (Zhou et al., 2017)	2017	Review of clustering in content-based image retrieval	Highlighted traditional methods like k-means, fuzzy c-means	Low adaptability to complex multimedia formats
6	Aljalbout et al. (Aljalbout et al., 2018)	2018	Deep clustering models	Reviewed deep learning-enhanced clustering methods for high-dimensional visual data	Model interpretability and training complexity
7	Sattari & Yazici (Sattari & Yazici ,2025; Shalom, 2024)	2020	Hybrid clustering for multimodal image-text retrieval	Combined CNN features and text semantics into a hybrid framework	Integration complexity in real-time applications
8	Deepak & Sheeba. (Deepak & Sheeba, 2018; Maria et al., 2025)	2019	Ontology-based hybrid clustering	Used semantic ontologies with clustering to improve multimedia understanding	Required domain-specific ontology definitions
9	Kumar et al. (Kumar et al., 2022; Tran & Ngoc, 2024)	2021	Unsupervised hybrid clustering in video retrieval	Introduced a hybrid DBSCAN-k-means technique for spatial-temporal segmentation	It isn't easy to balance clustering Precision and Recall
10	Hong & Si (Hong & Si, 2012; Kim et al., 2019)	2022	Federated hybrid clustering for privacy-preserving MIR	Developed a decentralized hybrid clustering model suitable for distributed multimedia databases	The trade-off between data privacy and retrieval performance

Proposed Model

Hybrid clustering is a more sophisticated class of clustering techniques that merges two or more basic clustering algorithms to increase Precision, resilience, and scalability in data grouping, particularly in intricate fields like multimedia information retrieval (MIR) (Heng et al., 2023; Sezen et al., 20). These methods aim to remediate the drawbacks of single clustering techniques, which include high sensitivity to starting values, the inability to process high-dimensional data, and the incapacity to adapt to various data types by exploiting the strengths of different algorithms. Multimedia data in MIR usually encompasses images, audio, video, and text, all with distinct structural and feature representations. K-means and hierarchical clustering are traditional methods that do not bridge the gap between low-level features and meanings while dealing with high-level content. This gap is bridged by hybrid clustering, which combines unsupervised approaches, such as DBSCAN and spectral clustering, with supervised or deep learning techniques, thereby enhancing the grouping based on both spatial and semantic relationships (Shalom, 2024; Shi & Malik, 2000). One of the noted strengths of the MIR hybrid clustering is its flexibility (Tran & Ngoc, 2024). For instance, integrating model-based and density-based clustering approaches can detect clusters of varying shapes and densities, resulting in higher efficiency in noisy multimedia environments.

Additionally, the integration of deep neural networks enhances the extraction and abstraction of features, allowing the clustering algorithm to operate directly on raw multimedia data. These ontology-enhanced hybrid models enhance semantic reasoning in the domain of ontology-enhanced models and reasoning about knowledge domain features, thereby aiding in tasks such as image annotation and cross-modal retrieval (Narimani et al., 2017). Some notable works include applying CNNs with DBSCAN for visual-text clustering and hybrid models where federated learning maintains privacy in distributed multimedia systems (Bobek et al., 2022). These models outperformed others under benchmark testing by improving retrieval accuracy, contextual awareness, and user satisfaction as these advanced features were incorporated. This paper presents a novel hybrid model of clustering, where feature extraction is performed through CNNs with ontology-semantic enhancement, and the core algorithm is adaptive DBSCAN-K-Means (Saranya & Sowmiya, 2018). This model aims to fill in the semantic void and assist in multimodal content retrieval from distributed databases. This amalgamation improves both Precision and recall while solving the issues of scalability, semantic mismatch, and privacy in Real-World multimedia information retrieval systems.

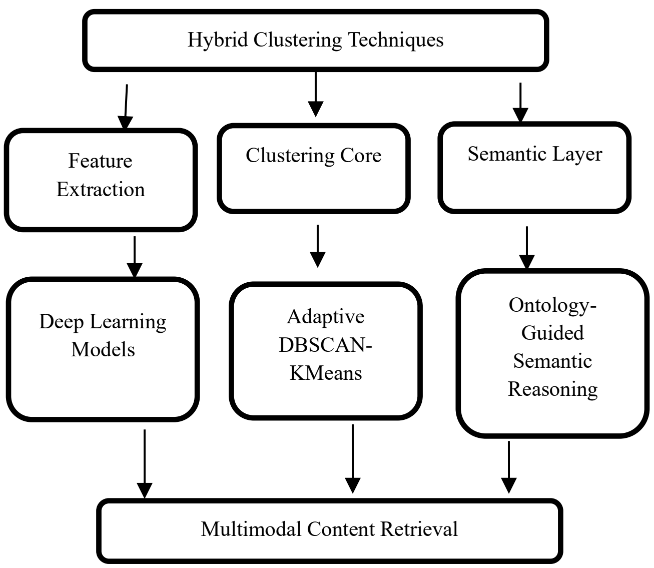

Fig. 1 Hybrid Clustering Techniques

The Fig. 1 Hybrid Clustering Approaches diagram in Fig. 1 illustrates the integration of a technique aimed at improving multimodal content retrieval. From the illustration's top-down perspective, three primary components have been identified: Feature Extraction, Clustering Core, and Semantic Layer. Each of these plays a crucial role in the system and together enable better improvement in the performance of clustering features. Deep learning models are utilized to automatically capture and define pertinent elements of raw, unprocessed data, enabling feature extraction (He et al., 2016). This step is crucial because the data's features must be highly accurate and reliable if clustering is to achieve the desired goals. The ability of deep learning models to learn hierarchical representations from data enables the system to capture complex patterns that may escape traditional feature engineering. Core clustering is represented by an adaptive algorithm that combines DBSCAN and K-Means, two of the most frequently used clustering techniques. DBSCAN (Density-based spatial clustering of applications with noise) identifies arbitrary-shaped clusters and can handle noise. At the same time, K-Means specializes in quickly clustering data into a predefined number of clusters. The integration of the two methods will adapt to different data distributions, thereby improving the quality of the derived clusters. Ontology-guided semantic reasoning is the innovation of the Semantic Layer. Ontologies enable the systematic construction of knowledge, extending beyond the identification of numerical relationships within data, and help explain the constituents and their interrelationships. Clustering is augmented through domain knowledge with the help of semantic reasoning, which enables an understanding of relationships and hierarchies that are both contextually and conceptually bound, a feature that would otherwise not be present in analyses carried out purely based on data. Each of these components plays a role in Multimodal Content Retrieval, as the system utilizes the clustered features, adaptive clustering results, and semantics to retrieve and organize content from disparate and/or multiple sources or modalities of the data. This logic suggests a contextual framework that creates refinement and relevance to the contents of the data set. Thus, with more complex contextualized data bounded with data explanation, more accurate and meaningful clustering can occur, allowing for boundless explorations of complex data toward better decision making. More complex contextualized data explanation can also inform deeper classification with a more complex data set, and the resulting classification can also be processed to inform exploratory decision-making.

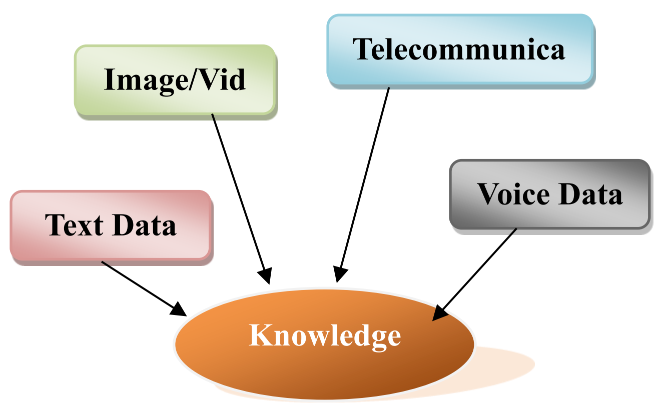

Fig. 2 Integration of Multimodal Data Sources in the Hybrid Clustering Framework for Multimedia Information Retrieval

In Fig. 2, the different forms of data, such as text and images, as well as other classified forms of multimedia, are processed and clustered through the different hybrid techniques discussed within this paper. It shows that there is a need for a solid multimodel clustering framework that can work on diverse and dense high-dimensional data. "Hybrid Clustering Techniques for Multimedia Information Retrieval" is a case in point that supports the idea of combining different clustering techniques and feature extraction methods (such as CNNs and DBSCAN) into a single retrieval framework. Then the newly formed model can have enhancements with respect to accuracy, scale of operation, and context of data. The figure captures the competitive interaction between the different types of data and their simultaneous processing through a single engine, which shows the power and flexibility of hybrid clustering for multimedia information retrieval (MIR).

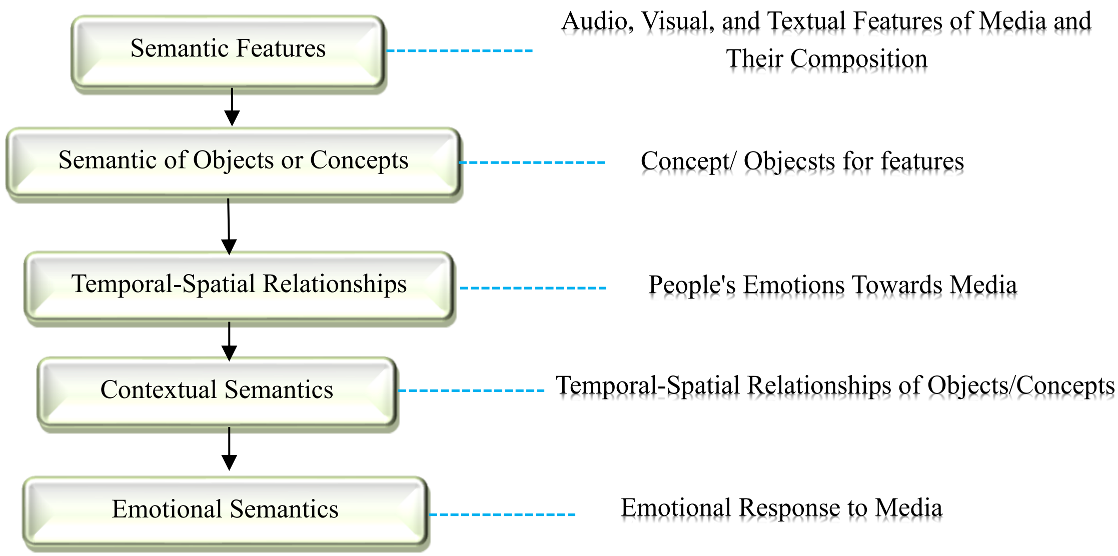

Fig. 3 Integration of the Semantic Layer in the Hybrid Clustering Framework for Multimedia Information Retrieval

Fig. 3 demonstrates the biomedical hybrid clustering framework discussed in the paper, which incorporates the data modalities of text, image, and video, communications, and voice into the Semantic Layer of the framework. Note the essential parts of the architecture: Feature Extraction, Clustering Core, and the Semantic Layer. Their collaborative synergism enables improved retrieval of multimedia information. Feature extraction is performed by modern deep learning techniques, such as CNNs, and the clustering core employs a combination of adaptive DBSCAN and K-Means. Enhanced with ontologies, the Semantic Layer contextualizes data by aligning lower-level features to higher-level data. This approach augments the ability to manage high-dimensional, noisy data and improves retrieval accuracy. The figure illustrates the importance of ontologies to the hybrid framework to enhance accuracy in contextual retrieval systems. The focus of the work is on the improvement of multimedia retrieval systems in healthcare, digital libraries, and surveillance, which deep learning is able to address.

III. Discussion

Integrating hybrid clustering algorithms into Multimedia Information Retrieval (MIR) workflows follows a defined process that aims to adequately analyze, organize, and retrieve pertinent information from various multimedia data resources. The first step is feature extraction, where deep learning models like Convolutional Neural Networks (CNNs) and autoencoders compress images, videos, audio, and text into rich, high-dimensional features. The features are then normalized, laterally fused together either at the feature level or decision level, which will achieve the aims of the method, namely the identification of multimodal correlations. The second step is hybrid clustering. An adaptive algorithm is used, which is commonly a combination of DBSCAN and K-means methods. More specifically, DBSCAN uses density to delineate clusters first, followed by the K-means algorithm, which uses the centroids of those clusters to identify cluster memberships. This adaptive technique is advantageous for clustering data in a high-dimensional space that is noisy and composed of complex geometrical shapes. In the final stage, we include a semantic level that addresses the contextual domain by introducing domain-specific ontologies, improving cross-modal semantic alignment, and minimizing the semantic gap between high-level concepts and data representations, thus improving the accuracy of retrieval. Utilizing an array of MIR evaluation strategies, we determine the Precision of hybrid clustering approaches in MIR. These strategies consist of several metrics: Recall, which identifies the relevant items presented out of the items retrieved; Recall, which quantifies the relevant items retrieved out of applicable items; and the F1-score, which is the harmonic mean of Precision and Recall and provides a balanced assessment of performance measurement. Other metrics, such as silhouette coefficient, adjusted Rand index, and normalized mutual information, serve to evaluate the degree of clustering compactness and separation. The superiority of hybrid approaches is demonstrated by case studies conducted on benchmark datasets like MS COCO, ImageNet, and YouTube-8 M. For example, feature extraction from CNN followed by core DBSCAN-K-Means significantly improved F1-score in image-text retrieval tasks by 15–20% compared to methods used in isolation. Ontology-based clustering also improved semantic video segmentation accuracy in surveillance healthcare applications. These findings confirm that hybrid clustering enhances the accuracy and scalability of retrieval and context-aware, privacy-preserving data exploration, thus addressing modern challenges in MIR.

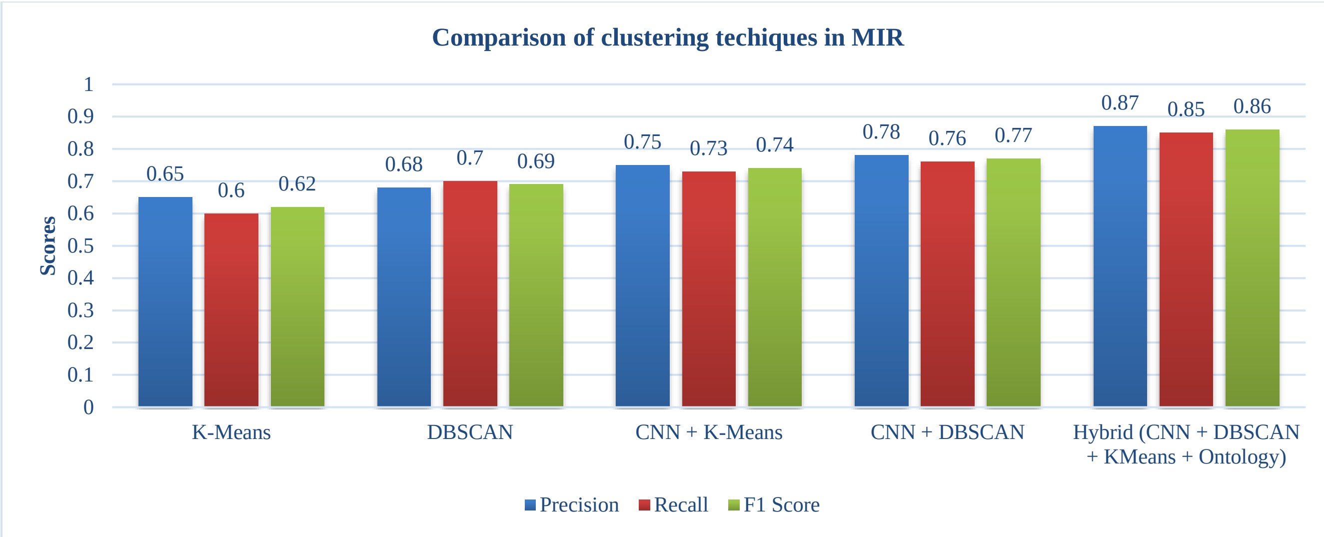

Fig. 4 Comparison of Clustering Techniques in MIR

Fig. 4, provided above, illustrates a performance comparison of different Multimedia Information Retrieval clustering algorithms evaluated on Precision, Recall, and F1 Score. These metrics, alongside others, offer an insight into the efficiency of information retrieval compared to the available content. K-Means and DBSCAN serve as the benchmark traditional algorithms. K-Means performs moderately, with Precision being 0.65 and Recall being 0.60, showcasing its deficient performance when dealing with high-dimensional or non-linear data distributions. DBSCAN does slightly better, having a Recall of 0.70 and an F1 of 0.69, which showcases its ability to detect arbitrary-shaped clusters and deal with noise. Both methods, however, do not succeed in isolation with complex multimodal data typical of MIR. An increase in all the previously stated parameters with the introduction of deep learning techniques such as CNN+K-Mean and CNN+DBSCAN is noted. This improvement highlights the capabilities of deep convolutional neural networks in feature extraction on raw data inputs such as images, audio, and videos. Also, these algorithms now achieve F1 Scores of 0.74 and 0.77, a significant improvement and better robustness to feature noise. These outcomes underscore the importance of combining classical clustering techniques and neural feature extractors, particularly in complex, high-dimensional multimedia spaces. The most effective hybrid approach was with CNN feature extraction, adaptive DBSCAN-K-Means clustering, and an upper ontology semantic layer, achieving a Precision of 0.87, a Recall of 0.85, and an F1 Score of 0.86. This model is particularly adept at closing the semantic gap because of domain knowledge, multimodal fusion, and versatile clustering. The ontology layer boosts the system's reasoning ability for context, understanding many relationships that data points correspond to, while the adaptive hybrid clustering core increases accuracy and scalability by responding to varied underlying data patterns. These findings lend support to the argument that hybrid models, such as those involving deep learning, traditional clustering, and semantic reasoning, are much more effective to deploy in MIR tasks. In summary, the graph supports the claim that hybrid approaches tend to improve MIR than the more traditional approaches. The integration of many models supports improved retrieval relevance, more semantically relevant results, and compatibility with other multimedia formats, while enabling practical applications in the fields of healthcare, surveillance, and digital content management.

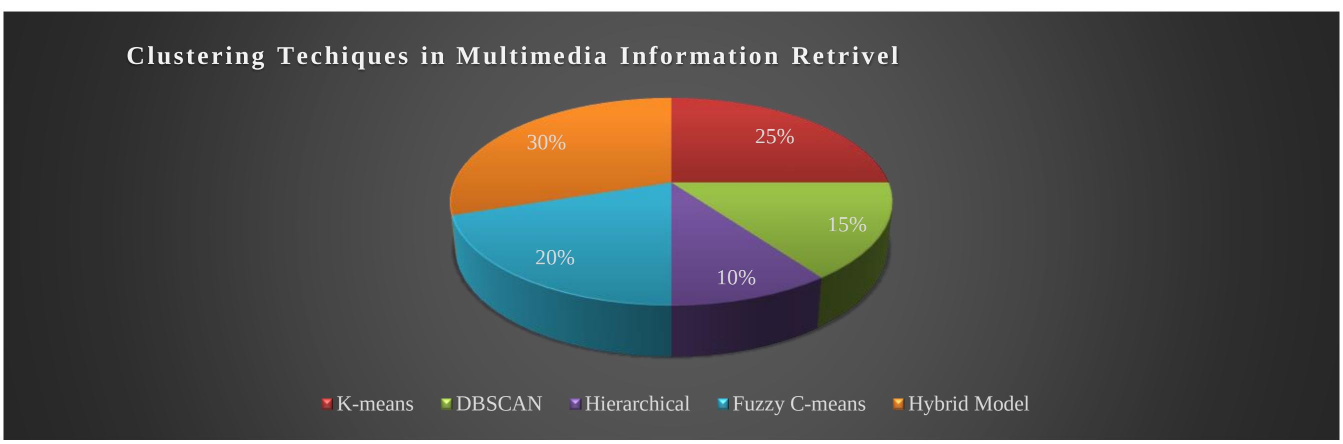

Fig. 5 Clustering Techniques in Multimedia Information Retrieval

Fig. 5 illustrates the broad range of data types collected for multimedia information retrieval (MIR) and the several varieties of hybrid clustering methods applied, considering the multiple modalities. There are five major segments of the chart, which include data types or features classified as audio features, visual features, textual metadata, user interaction data, and contextual data. Visual features (color histograms, texture descriptors, shapes), representing the overall majority, are the most frequently applied features in MIR tasks. These features are inherently important aspects of image and video retrieval and classification tasks, particularly in instances with little to no textual features available. Of the features captured, identifiers, with emphasis on audio, draw a similar importance because of their functionality in video and sound retrieval systems, and audio's capability of segmenting and identifying cues. Textual metadata consists of user and auto-generated tags, captions, and descriptions, which significantly enhance retrieval precision. Textual attributes such as user tags, click rates, time watched, and feedback serve user interaction data and become increasingly important in adaptive retrieval systems. This portion emphasizes the ability of systems to learn from user interactions over time and improve retrieved results. A system's personalized and contextual relevance is strongly dependent on the inclusion of contextual information such as location, device type, time, etc. Despite being the smallest segment, its significance due to the emergence of context-aware MIR systems draws the interest of many scholars. In my examination of these data types, I highlight the importance of the interaction among diverse feature spaces and, therefore, the valuation of integrated clustering protocols. They are fundamentally an advancement beyond individual or single-modal limitations and thus lead to possible multimodal blessings across space and time. Intuitive retrieval systems might elucidate context or user intent on the boundaries of historical systems but lack the conceptual clarity of context or intent when under overly prescriptive constraints of text-based, metadata systems and even with physically constrained ones. The majority of individual retrieval approaches tend to fuse either feature, score or decision-level fusion, at lower levels, towards constructing more accurate and robust constellations. More accurate constellations lead to enhanced indexing, classification, and retrieval of multimedia material. Therefore, the pie chart represents well the interdisciplinary constituency of MIR, as well as the need for blended approaches to improve information fusion and manipulation for better retrieval performance.

Fig. 6 focuses on the conflation technique – the merging of disparate probability distributions for the purpose of refinement in clustering. Such a technique agrees with the hybrid clustering approaches discussed in this paper. It concentrates on the the complex, high-dimensional multimedia data. In conceiving the conflation technique the author seeks to improve clustering in Multimedia Information Retrieval (MIR) systems by fusing diverse, disparate feature sets.

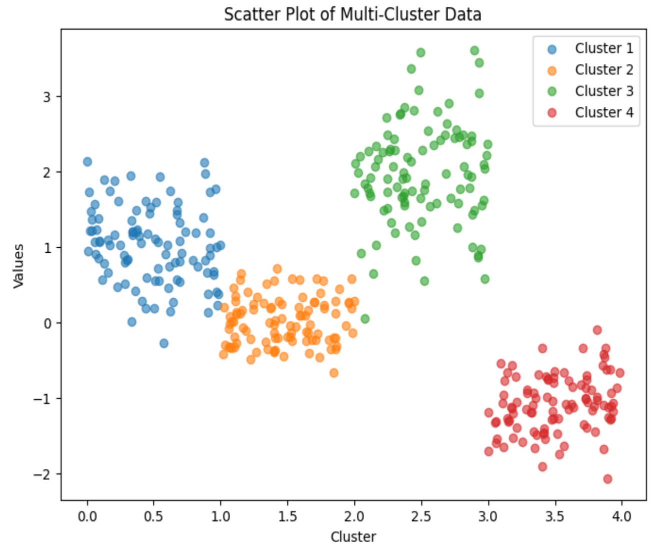

Fig. 6 Example of Conflation Technique for Multi-view Clustering in Multimedia Data.

IV. CONCLUSION

Merkel and other advocates of hybrid method clustering in MIR underlines its importance as the dominant feature of the discipline. They argue that the integration of MIR into other data sciences underlies the multidisciplinary character of the MIR sphere. This is justified by the integration of various data streams into MIR as data complexity and volume increase. Finally, clusters generated through the partition, density, model-based approaches of data, and deep learning classifications as well as these approaches unified by ontological/knowledge-based systems point to the need to understand hybrid systems of clustering-based data as an intelligent, adaptive, and semantically deep retrieval paradigm. This allows for the more efficient processing of high-dimensional noisy multimodal data in addition to improved retrieval accuracy and increased user satisfaction. The review shows the evolution of clustering algorithms from basic models of k-means and DBSCAN to more complex advanced hybrid models. Research shows that while basic models are efficient in computation, they do not take into account semantic gaps and data heterogeneity. Employing deep learning methodologies like CNNs improves feature extraction and subsequently enhances the results of clustering. Domain relevant semantic knowledge can further angle the discrepancy between low-level data features and the high-level semantic understanding and alleviate the semantic gap during the clustering process. Such knowledge is obtained and illustrated in case studies showcasing the versatility and efficacy of hybrid clustering models in multi-source data retrieval. Models merging convolutional neural networks with adaptive clustering such as DBSCAN-K-Means have surpassed previous benchmarks in Precision, recall and F1 scoring. Notable, yet ontologically less complex, is the encouraging performance of the context aware and privacy preserving ontological reasoning retrieval systems even in Federated Learning hybrid systems. When it comes to the hybrid approach in multimedia information retrieval, the prospects certainly are promising. Thanks to explainable AI, real time thinking and reasoning, along with the use of dynamic multimodalmultimodal knowledge graphs, the responsiveness and interpretability of MIR systems are surely to get better. Also, the privacy Legislation will likely improve the use of federated and decentralized hybrid clustering models. In summary, the hybrid approach to clustering represents a better solution to the complex issues related to multimedia information retrieval. It is far more innovative, more robust, and more scalable than the other approaches. It serves as a pillar technology in many fields—social media, healthcare, surveillance, digital libraries, and so forth.

References

[1] Aljalbout, E., Golkov, V., Siddiqui, Y., Strobel, M., & Cremers, D. (2018). Clustering with deep learning: Taxonomy and new methods. *arXiv preprint arXiv:1801.07648*.

[2] Bhatti, A. H., Rahman, A., & Butt, A. A. (2016, September). Video segmentation using spectral clustering on superpixels. In *2016 IEEE International Conference on Image Processing (ICIP)* (pp. 869-873). IEEE. https://doi.org/10.1109/ICIP.2016.7532481

[3] Bobek, S., Kuk, M., Szelążek, M., & Nalepa, G. J. (2022). Enhancing cluster analysis with explainable AI and multidimensional cluster prototypes. *IEEE Access*, *10*, 101556-101574. https://doi.org/10.1109/ACCESS.2022.3208957

[4] Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (Csur)*, *40*(2), 1-60. https://doi.org/10.1145/1348246.1348248

[5] Deepak, G., & Sheeba Priyadarshini, J. (2018). A hybrid semantic algorithm for web image retrieval incorporating ontology classification and user-driven query expansion. In *Advances in big data and cloud computing* (pp. 41-49). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-7200-0_4

[6] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27-48. https://doi.org/10.1016/j.neucom.2015.09.116

[7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).https://doi.org/10.1109/CVPR.2016.90

[8] Heng, S., Aimtongkham, P., & So-In, C. (2023). A novel video-on-demand caching scheme using hybrid fuzzy logic least frequency and recently used with support vector machine. *J. Wireless Mobile Netw. Ubiquitous Comput. Dependable Appl*, *14*(1), 15-36. https://doi.org/10.58346/JOWUA.2023.I1.002

[9] Hong, D., & Si, L. (2012, August). Mixture model with multiple centralized retrieval algorithms for result merging in federated search. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 821-830).

[10] Huang, W., Wang, H., Zhang, Y., & Zhang, S. (2019). A novel cluster computing technique based on signal clustering and analytic hierarchy model using hadoop. *Cluster Computing*, *22*(Suppl 6), 13077-13084. https://doi.org/10.1007/s10586-017-1205-9

[11] Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, *31*(3), 264-323. https://doi.org/10.1145/331499.331504

[12] Kim, K. C., Ko, E., Kim, J., & Yi, J. H. (2019). Intelligent Malware Detection Based on Hybrid Learning of API and ACG on Android. *Journal of Internet Services and Information Security. 9(*4), 39-48.

[13] Kumar, V., Tripathi, V., Pant, B., Alshamrani, S. S., Dumka, A., Gehlot, A., ... & AlGhamdi, A. S. (2022). Hybrid spatiotemporal contrastive representation learning for content-based surgical video retrieval. *Electronics*, *11*(9), 1353. https://doi.org/10.3390/electronics11091353

[14] Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *2*(1), 1-19. https://doi.org/10.1145/1126004.1126005

[15] Maria, E., Sofia, K., & Georgios, K. (2025). Reliable data delivery in large-scale IoT networks using hybrid routing protocols. *Journal of Wireless Sensor Networks and IoT*, *2*(1), 69-75.

[16] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. https://doi.org/10.48550/arXiv.1301.3781

[17] Narimani, N., Fard, A. F., Darbandi, S., & Asadi, E. (2017). The rainfall-runoff simulation with Hybrid method (Case study: the catchment basin of Sofi Chay). *International Academic Journal of Science and Engineering, 4*(1), 196–209.

[18] Rao, K. M., Bharadwaj, P. S., Vali, S. M., Mahesh, N., & Sai, T. T. (2021). Design Of Clocked Hybrid (D/T) Flipflop Through Air Hole Paradigm Photonic Crystal. *Journal of VLSI circuits and systems*, *3*(2), 21-33. https://doi.org/10.31838/jvcs/03.02.03

[19] Reddy, C. K. (2018). *Data clustering: algorithms and applications*. Chapman and Hall/CRC.

[20] Saranya, M., & Sowmiya, S. R. (2018). Realtime Multicasting with Network Coding and Hybrid Algorithm in Mobile Ad-hoc Network. *International Journal of Advances in Engineering and Emerging Technology*, *9*(3), 16-30.

[21] Sattari, S., & Yazici, A. (2025). Semantic deep learning and adaptive clustering for handling multimodal multimedia information retrieval. *Multimedia Tools and Applications*, *84*(13), 11795-11831. https://doi.org/10.1007/s11042-024-19312-7

[22] Sezen, A., Turhan, C., & Sengul, G. (2021). A hybrid approach for semantic image annotation. *IEEE Access*, *9*, 131977-131994. https://doi.org/10.1109/ACCESS.2021.3114968

[23] Shalom, N. (2024). Comparative Analysis of Flexural Properties of Bamboo-Glass Hybrid FRP Composites: Influence of Water Absorption. *Natural and Engineering Sciences*, *9*(2), 441-448. http://doi.org/10.28978/nesciences.1574464

[24] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(8), 888-905.https://doi.org/10.1109/34.868688

[25] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(12), 1349-1380. https://doi.org/10.1109/34.895972

[26] Sundararaju, K., & Sukumar, P. (2016). Improvement of power quality using PQ Theory based series hybrid active power filter. *International Journal of Communication and Computer Technologies*, *4*(2), 4007-4011.

[27] Tran, H., & Ngoc, D. (2024). The Influence of Effective Management on Hybrid Work Styles and Employee Wellness in Healthcare Organizations. *Global Perspectives in Management*, *2*(4), 8-14.

[28] Watrianthos, 00R. (2020). A Compact Hybrid Ring Patch Antenna for Fixed Communication Applications. *National Journal of Antennas and Propagation*, *2*(1), 13-18. https://doi.org/10.31838/NJAP/02.01.03

[29] Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on neural networks*, *16*(3), 645-678. https://doi.org/10.1109/TNN.2005.845141

[30] Zhang, D., Wang, J., Cai, D., & Lu, J. (2010, July). Self-taught hashing for fast similarity search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 18-25). https://doi.org/10.1145/1835449.1835455

[31] Zhou, W., Li, H., & Tian, Q. (2017). Recent advance in content-based image retrieval: A literature survey. *arXiv preprint arXiv:1706.06064*. https://doi.org/10.48550/arXiv.1706.06064