



Centric Multi-Feature-Based Retrieval Approach for Fast Handling Large Medical Image Datasets

Najjar Mohammed^{*}, Bouyghf Hamid^{*}, Nahid Mohammed^{*}

LSIB Laboratory Faculty of Sciences and Techniques Mohammedia, University Hassan II of Casablanca, Mohammedia 20650, Morocco

Corresponding Author Email: mohammed.najjar-etu@etu.univh2c.ma

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300524>

ABSTRACT

Received: 26 March 2025

Revised: 4 May 2025

Accepted: 22 May 2025

Available online: 31 May 2025

Keywords:

multi-dimensional medical images, information retrieval, feature vector database, multi-features indexing, computer vision, vectorization

Querying and filtering multidimensional biomedical images in large datasets is one of the most difficult and time-consuming tasks. In this paper, a new approach called Centric Multi-Feature-Based Medical Image Retrieval from volumetric biomedical image datasets is presented to improve computer vision operations for effective medical diagnosis and research purposes. In today's digital era, we have witnessed unexpected advances in medical imaging technology as well as an ever-increasing number of digital healthcare image datasets, which also negatively impact feature detection, matching, filtering, and computational time. Motivated by the performance efficiency of feature-based algorithms such as 3D SIFT, which enable robust feature extraction, our approach allows the processing and extraction of features from different types of image modalities stored in highly heterogeneous decentralised datasets to feed and serve computer vision and machine learning models in real-time. Experimental evaluations of 500 volumetric MRA brain scans on public datasets such as LIDC-IDRI and OASIS show that our approach achieves 100% query fidelity compared to conventional pairwise methods, while reducing the overall processing time by 55% and accelerating the query latency by 850 times. The experimental results demonstrate the effectiveness of our approach and its potential as a pre-trained input to improve feature vector-based queries.

1. INTRODUCTION

Imaging techniques are a fundamental part of modern medicine; they are used to medically image the human body from the inside and outside. The first technique that comes to mind is X-ray, also called computed radiography (CR). While X-rays are the oldest and still the most widely used method, many different scanners, such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET), generate large amounts of data [1].

Regardless of the imaging modality, medical images can be digitally stored, accessed, and shared through an integrated system known as PACS (Picture Archiving and Communication System). This system includes software capable of aggregating data from various imaging technologies and supplementing it with descriptive metadata—such as the physician's name, patient ID, and diagnostic reports—providing a structured and efficient method for annotating, organizing, and retrieving image-related information for both clinical and research applications [2]. For instance, magnetic resonance imaging (MRI) files often contain detailed parameters like pulse sequence timing, flip angles, and acquisition counts [3].

Furthermore, modern PACS solutions are often integrated with Radiology Information Systems (RIS), enabling the management of complex, multidimensional image data. These systems utilize the standardized DICOM (Digital Imaging and

Communications in Medicine) protocol to store and transmit both the raw image files and all relevant metadata concerning image acquisition procedures [3]. However, despite the advanced infrastructure, the images stored in PACS are not inherently useful for clinical interpretation or decision-making unless properly structured and annotated. For example, considering a repository of over 900,000 images lacking metadata or organization, without contextual information, extracting meaningful insights would require manually reviewing the entire dataset, which is highly impractical [4]. Since recent biomedical image datasets [5] contain multi-dimensional images (3D+time), analysis by computer vision and biomedical image processing algorithms may take longer. This dimensionality of medical images is the real performance issue, as multidimensional computational complexity increases with each additional dimension since the processing of image data must be performed accurately with all available computational power. Therefore, the biomedical image data generated daily requires algorithms that can efficiently process and analyze this data to produce meaningful insights.

Many retrieval algorithms and architectures have been adapted for 3D medical image retrieval, e.g. Deep Convolutional Neural Networks (CNNs) [6], Deep Similarity Learning for multimodal images [7], the algorithm developed [8] or the registration algorithm based on Compressive Sensing and Local Scale-Invariant Feature Transform (SIFT) [9]. However, a major limitation of the pairwise method, even

when based on these algorithms, is the high computational cost incurred by exhaustive comparisons that scale poorly with increasing dataset size. Traditional retrieval pipelines often calculate similarity scores between a query and each entry in the dataset, resulting in significant delays that make such methods impractical for use in real-time or on a large clinical scale. Attempts to reduce these costs through downsampling, segmentation or subsampling [10-12] often result in a loss of accuracy and diagnostic precision.

Consequently, there remains a need for a retrieval framework that avoids the computational inefficiencies of pairwise comparisons, eliminates the need for data reduction techniques, and scales effectively with the growing volume of 3D medical image data. Our approach addresses these challenges through a centralised, multi-feature-based pipeline that supports efficient, accurate and timely retrieval without compromising data quality.

1.1 Problem formulation

Despite the advancements in medical processing, efficient retrieval of relevant medical images from large-scale datasets remains a significant challenge due to the high computational cost associated with traditional methods. Conventional approaches rely on direct image-to-image comparisons, which are inefficient for handling large volumetric medical datasets. Additionally, existing methods often lack an optimized, memory-efficient strategy for managing and indexing extracted features, leading to scalability limitations and may not fully utilize the potential of 3D and 4D data, leading to suboptimal feature extraction.

1.2 Major contributions

This study presents an effective retrieval framework based on 3D SIFT feature extraction [13] in combination with FAISS [14], which was chosen over other deep learning techniques and key-point methods such as HOG, HDFS [15] and SuperPoint [16] due to its feature descriptor based on a 64-dimensional descriptor, this descriptor provides higher efficiency than classical 128-dimensional and also for 3D medical image search, 3D SIFT is a GPU-optimised implementation that achieves about 7x speed improvement while maintaining robustness in MRI scans, making 3D SIFT both computationally efficient and robust to noise and intensity changes in 3D volumes resulting in similar image matching performance with a much larger memory footprint [17] and less time-consuming when dealing with large medical datasets to enhance both the storage and indexing of volumetric medical images. By replacing traditional image-to-image comparisons with a more efficient feature-to-feature matching strategy, the framework substantially reduces computational demands while preserving high retrieval. Tailored for real-time diagnostic use, the architecture supports high-performance querying and is well-suited for large-scale clinical and research imaging environments.

By tackling the key challenges associated with big medical data, the proposed framework delivers a robust and scalable solution for efficient image retrieval. Leveraging ranked 3D SIFT features for optimized storage and indexing, it significantly improves data organization and search precision, all while minimizing processing overhead. Its feature-driven architecture, combined with a high-capacity database, ensures fast and reliable access to relevant medical imagery, making it

highly applicable to real-time clinical workflows, biomedical research, and secure data-intensive applications.

The remainder of this paper is organized as follows: Section II reviews related work and current strategies for improving medical image storage and retrieval. Section III details the design and components of the proposed framework. Section IV presents the experimental evaluation of the system, analyzing its performance in terms of retrieval speed, precision, and scalability. Section V concludes the paper by outlining limitations and suggesting directions for future enhancements.

2. RELATED WORK

To effectively exploit valuable information contained in medical images, various retrieval techniques have been developed to identify image similarities using large datasets, where an image serves as the query input. These techniques include text-based image retrieval [18], content-based image retrieval (CBIR) [19], and hybrid approaches combining both [20]. In text-based systems, experts manually annotate images with descriptive metadata—such as keywords, labels, or tags—stored in a database to enable keyword-based searches. In contrast, CBIR focuses on analyzing the visual content of the image itself, rather than relying on external annotations [21-25].

CBIR systems extract and index visual features, which are typically divided into global features (e.g., color, texture, shape) and local features (e.g., keypoints like corners or edges). Global features represent the overall characteristics of the image and are useful for tasks such as classification and general image retrieval, while local features are more suitable for detailed object detection and recognition. Feature extraction is central to CBIR, as it enables the system to compute descriptors for both the query and dataset images, facilitating efficient matching based on visual similarity. Compared to single-feature approaches, multi-feature CBIR methods significantly improve retrieval accuracy and robustness.

Unlike TBIR which is a very tedious and expensive approach, The Content-based image retrieval (CBIR) is a computer vision image search technique that uses mostly low-level features such as color, texture, and shape to explore images from databases. Indeed, it saved countless lives over the years along with analyzing traditional 2D medical images (rows and columns) which have focused on Chest X-ray, such as plain X-rays and mammograms using Convolutional Neural Networks (CNNs) models like ResNet and VGG applied to tasks such as image classification, segmentation, and anomaly detection [26] or combined with Recurrent Neural Networks (RNNs) architecture CNN-LSTM [27] that achieved unprecedented performance in medical imaging-based diagnosis particularly Long Short-Term Memory (LSTM) networks. However, these approaches may have many limitations especially when processing medical images with additional dimensions, such as a three-dimensional (3D) volume or 3D volume changes over time (4 dimensions), or much more if the images are multidimensional or n-dimensional stored in large image databases.

Multidimensional imaging provides flexibility to perform functions for traditional two-dimensional filtering in scientific applications. The first multimodal imaging technique used in clinical settings and modern hospitals combines positron emission tomography and computed tomography (PET-CT),

which produce images of the body and its functions that enable better cancer diagnosis and effective treatment, among other benefits [28].

To improve accuracy and increase the computational complexity related to these additional dimensions of medical images, significant new algorithms are identified for analysis by invariant keypoint techniques such as 3D SIFT, a powerful approach that performs a variety of imaging tasks, including classification [29], registration [30], without the need for training procedures or data. Our work is not intended to replace the 3D SIFT, but to complement it and use it to build an alternative based on the results of the recent study describing a novel implementation of the 3D SIFT algorithm to enhance keypoints and features extraction based on GPU [31]. Also, as the results prove it uses three-dimensional volumetric data and real objects to detect keypoints and extract a robust description of their contents to support a variety of applications such as video processing [32] and 3D object recognition [33].

The reviewed studies establish a solid groundwork for enhancing the storage and retrieval of medical images. Building on these advancements, the proposed framework in this research is designed to tackle key challenges related to efficiency and scalability in managing medical images data stored in large datasets.

3. MATERIALS AND METHODS

The proposed method presents a lightweight, feature-centric biomedical storage and retrieval system aimed at improving the efficiency of feature extraction algorithms such as 3D SIFT for large-scale medical image queries. Designed to enhance both speed and accuracy, the framework also addresses the challenge of data scarcity in biomedical imaging by supporting scalable processing without requiring massive training datasets. It integrates three core components - feature extraction, indexing, and real-time matching- while optimizing overall processing time as a key performance indicator. Additionally, the system leverages features from both 2D and 3D deep learning models, enabling accurate and cost-efficient retrieval, making it highly suitable for advanced medical imaging and research applications.

Most machine learning and computer vision workflows explored in this study rely on a core computational framework for processing image queries against reference datasets, summarized by the following equation:

Processing Time=Fr Extraction Time + Comparison Time

Feature extraction time: the period needed to generate meaningful descriptors from medical images using techniques like SIFT, CNN-based embeddings, or hybrid methods.

Comparison time: the duration required to match these extracted features against those in the dataset, typically using similarity metrics such as Euclidean distance, cosine similarity, or domain-specific medical measures.

The proposed Centric Feature-Based Retrieval Approach introduces a two-stage pipeline—Indexing (offline) and Query (online/real-time)—to optimize medical image retrieval. By extracting and storing features only once during the indexing stage, the system eliminates the need for repeated dataset processing at query time. This design follows the principle of Separation of Concerns, enabling real-time performance by

decoupling expensive computations from the user-facing query phase. Unlike traditional full pairwise methods that perform full pairwise comparisons at every query, the Centric strategy drastically reduces processing overhead while maintaining high retrieval accuracy. It is particularly well-suited for large-scale medical datasets, offering scalable, storage-efficient, and fast retrieval performance ideal for real-time clinical applications.

The main difference between the two approaches is explained in Figure 1.

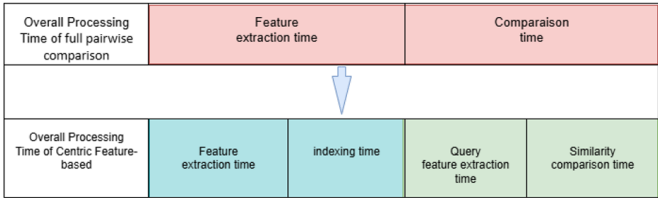


Figure 1. Separation of concerns of centric vs full pairwise approach

This enhanced formulation ensures that the system not only extracts and compares features efficiently but also structures and stores extracted features in an optimized manner, facilitating rapid and scalable real-time retrieval. By incorporating indexing mechanisms, such as KD-trees, FAISS, or Using vector database, the proposed framework significantly improves query performance and scalability for large-scale medical imaging datasets as shown in Figure 2.

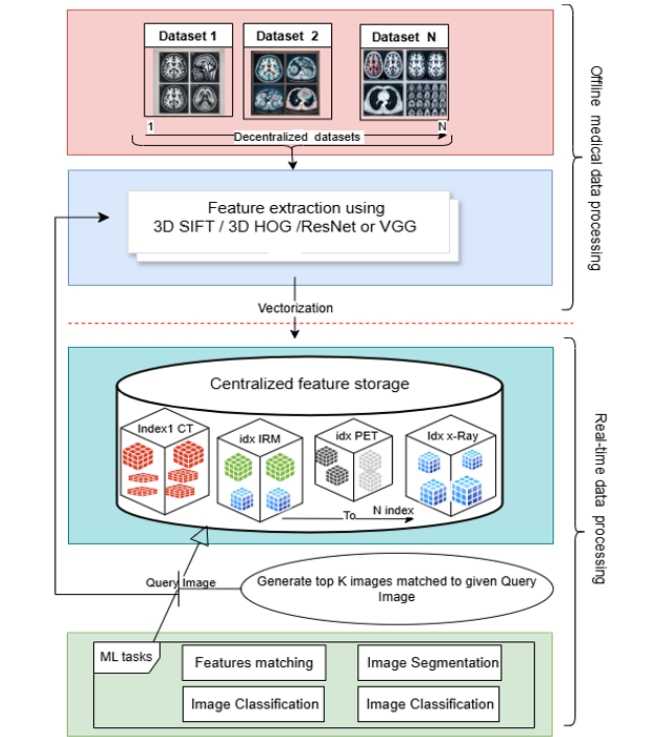


Figure 2. Bloc diagram for building an end-to-end medical image based on multi-features approach

The diagram above shows a typical workflow for indexing medical image features extracted from many mid-range and high-end biomedical datasets. Our process consists of two main stages line-dashed which are The Offline and Real-time data processing to speed up processing historically and

archived medical data stored in original datasets.

The Offline processing stage consists of two blocks, the first one is crawling and importing images from one or more datasets (e.g., ultrasound, MRI, CT, and radiography) provided to Feature extraction algorithms (the 3D SIFT to extract keypoint descriptors) from the corpus of parsed images.

To go through the scarcity of biomedical datasets, medical images from TCIA dataset (e.g., from ultrasound, MRI, CT, and radiography), are provided to the 3D SIFT extraction method. As shown in the diagram, multiple public datasets, e.g., OASIS [34], ADNI [35], and HCP [36], can be used simultaneously for feature extraction to feed our feature database. As mentioned earlier, TCIA is a service that provides a large archive of medical cancer images that can be publicly downloaded. The data are organized as “collections”; typically, patient images are organized by a common disease (e.g., lung cancer), image modality, or image type (MRI, CT, digital histopathology). DICOM is the primary file format used by TCIA for radiologic imaging. Supporting data for images such as patient outcomes, treatment details, genomics, and expert analysis are also provided when available.

The second block addresses feature extraction using 3D SIFT and vectorization, a technique for converting raw data into numerical features that can be processed to produce what is known as a feature vector. 3D SIFT -Rank is designed for feature extraction from 3D volumes. The main format is Nifti, but (.hdr,.nii.gz) or raw data (IEEE 32-bit float, Little Endian) are also supported.

The third block of the process is indexing. As the name suggests, indexing data refers to the process of creating a structured and organized reference system to efficiently retrieve and access specific information from a data set. Simply put, indexing helps to improve the speed and efficiency of data queries by allowing systems to quickly find and retrieve the desired information without having to search through the entire data set.

This improvement is achieved through a Java-based multi-feature wrapper that plays a central role in managing the flow between raw medical image ingestion, feature extraction and vector database indexing. This wrapper performs several automated tasks, such as error handling by automatically filtering out unsupported or corrupted medical images based on format and size constraints, as well as batch crawling, which allows the system to process medical volumes in optimized batches and what is more important is feature aggregation, where multiple feature vectors are encapsulated into a unified metadata object per image. These aggregated feature vectors are then passed in bulk to the FAISS-based indexing engine, allowing for a fast build-up of the similarity search index. This modular approach not only improves fault tolerance and scalability but also enables parallelism and pre-fetching mechanisms that significantly reduce latency and indexing time.

This modular and scalable structure allows for dynamic index updates each time a new image is processed, enabling real-time retrieval tasks such as classification, object recognition and clustering with significantly reduced latency, as illustrated in Figure 2.

4. MORE EXPERIMENTAL EVALUATION

4.1 Experimental setup

Our approach is both a cloud-based platform service and can

be run as a standalone solution on a simple machine. It is designed to run in a server cluster to distribute the load across multiple nodes. However, in our case, we can run it with only one node, as setting up a single node is easy to implement and manage, making it a good choice for development or test environments where high availability and scalability are not critical; moreover, our application is about small applications with low data volume, where a single node can effectively handle the workload. The operating system where feature extraction is performed on is a Mac with 10-core CPU, 16-core GPU, 8 GB of unified memory, 512 GB of SSD storage¹, and 16-core neural engine.

4.2 Feature vector generation and indexing

In most cases, biomedical 3D images feed private lab datasets and they rely on famous 3D algorithms such as 3D-SIFT for analysis and research. However, to prove the performance of the approach, we decided to conduct studies based on other extraction algorithms to prove the effectiveness of this intermediate multi-feature extraction system (see Fig. 1). Besides 3D-SIFT, whose generated data is represented as a matrix of n vectors with M elements, three other algorithms are also used to extract 2D and 3D medical images from the original datasets over many iterations, as well as 3D-SIFT support algorithms (HOG) feature extraction from both formats, DICOM and NIFTI(.nii).

5. RESULTS

5.1 Evaluating centric feature query-based stage

To rigorously evaluate our approach to medical image retrieval and similarity assessment, we conducted a series of experiments using benchmark datasets commonly employed in clinical research. These included the LIDC-IDRI dataset, consisting of 1609 2D thoracic CT scans from 214 patients for nodule detection, and two well-established 3D brain MRI datasets -OASIS and ²- which contain tumor annotations across multiple institutions. Additionally, the TCIA brain metastasis dataset was used, comprising 560 multimodal MRI volumes from 412 patients, annotated with over 5000 brain metastases, and available in both DICOM and NIFTI formats. These diverse, multimodal, and publicly available resources ensured clinical relevance and scientific reproducibility for our evaluation.

Our experimental study focused specifically on 3D volumetric medical data, performing controlled tests using both the 3D SIFT and 3D HOG feature extraction algorithms. The datasets were scaled progressively, ranging from as few as 1 image to as many as 500 NIFTI volumes, totaling approximately 93 GB in raw (.nii) format, or 15.9 GB when compressed (.nii.gz). For each configuration, we compared the performance of two retrieval strategies: the Centric Feature Query-Based approach, which separates offline indexing from real-time querying, and the traditional method, which performs full pairwise comparisons at query time. Across all configurations, we evaluated the feature extraction time, indexing overhead, and similarity search latency, under conditions designed to reflect realistic clinical use cases as shown in the Tables 1-12.

Our demonstration was primarily conducted on 3D medical imaging datasets, where the Centric Feature-Based Retrieval

Approach consistently exhibited real-time query performance, even at the largest tested scales. Using 3D SIFT, the total processing time—including a one-time offline indexing step—was approximately 4807.67 seconds, with the query stage executing in under 10 seconds. Similarly, when applying 3D HOG, the system achieved even faster results: 3026.59 seconds for offline processing and ~5.10 seconds per query. In both cases, the traditional full pairwise method required significantly more time—over 10,721 seconds for SIFT and 7,322 seconds for HOG—underscoring the computational burden and poor scalability of full pairwise comparison strategies Tables 7-12.

Table 1. Traditional full pairwise comparisons at query time using ResNet50 model

Total Images	Data Size (KB)	Feature Extraction Time(s)	Comparison Time(s)	Overall Total Processing Time(≈s)
1	16	3.52	0.7343	4
10	350	6.22	8.2823	14
50	1351	20.61	38.4243	59
100	2447	39.12	80.1948	119
150	2539	60.70	125.163	185
200	3215	77.27	160.073	237
250	4679	97.99	197.2865	295
300	6604	112.95	243.1207	356
350	7802	137.41	276.9598	414
500	9113	192.48	397.6731	590
1000	24371	386.60	786.8337	1173
2000	43622	777.87	1661.0856	2438
2500	52224	980.08	2078.7321	3058

Table 2. Indexing stage of centric feature-based extraction using ResNet50 model

Total Images	Data Size (KB)	Indexing Stage(s)		
		Feature Extraction Time	Indexing Time	Overall Processing Time
1	0.01	0.01	0	0.01
10	0.06	0.08	0.000099	0.08
50	13.53	0.63	0.000085	0.64
100	12.82	18.8	0.000089	18.81
150	13.45	21.13	0.000098	21.13
200	13.88	20.35	0.000061	20.35
250	13.48	19.72	0.000063	19.74
300	13.45	20.93	0.00008	20.94
350	12.44	21.65	0.000051	21.66
500	12.83	19	0.000077	19.01
1000	13.18	19.43	0.000056	19.44
2000	34.33	20.92	0.000055	20.93
2500	80.39	20.35	0.000081	20.41

Table 3. Query stage of centric feature-based extraction using ResNet50 model

Total Images	Data Size (KB)	Query Stage(s)		Overall Total Processing Time
		Query Feature Extraction	Similarity Comparison Time	
1	0.01	0.0052	0.0001	0.0053
10	0.06	0.0051	0.0001	0.0051
50	13.53	0.0053	0.000073	0.0053
100	12.82	0.4135	0.000079	0.4136
150	13.45	0.0043	0.000095	0.0044
200	13.88	0.4636	0.000105	0.4636
250	13.48	0.4407	0.000056	0.4408
300	13.45	0.456	0.000062	0.4561

350	12.44	0.0042	0.000043	0.0042
500	12.83	0.0046	0.000045	0.0047
1000	13.18	0.3247	0.00006	0.3248
2000	34.33	0.3139	0.000075	0.314
2500	80.39	0.4382	0.000054	0.4382

Table 4. Traditional full pairwise comparisons at query time using VGG16 model

Total Images	Data Size (KB)	Feature Extraction Time(s)	Comparison Time(s)	Overall Total Processing Time(≈s)
1	16	0.4682	0.4683	1.205
10	350	6.8388	6.8407	15.0307
50	1351	32.2695	32.2763	65.9343
100	2447	87.8673	87.8820	176.4271
150	2539	141.0949	141.1187	283.1777
200	3215	191.9745	191.9995	385.0010
250	4679	246.9657	247.0018	494.6738
300	6604	293.5808	293.6212	587.8835
350	7802	348.2412	348.2912	697.1526
500	9113	489.1017	489.1657	979.6022
1000	24371	1004.2016	1004.3512	2009.2659
2000	43622	1988.1951	1988.4769	3977.3399
2500	52224	2499.1334	2499.4762	4999.3054

Table 5. Indexing stage of centric feature-based extraction using VGG16 model

Total Images	Data Size (KB)	Indexing Stage(s)		
		Feature Extraction Time	Indexing Time	Overall Processing Time
1	0.01	0.01	0.000099	0.01
10	0.06	0.06	0.000071	0.06
50	13.53	0.54	0.000077	0.55
100	12.82	0.24	0.000050	0.24
150	13.45	0.26	0.000077	0.26
200	13.88	0.19	0.000122	0.19
250	13.48	0.26	0.000056	0.27
300	13.45	0.30	0.000055	0.0043
350	12.44	0.26	0.000063	0.26
500	12.83	0.29	0.000089	0.29
1000	13.18	0.3	0.00007	0.31
2000	34.33	0.24	0.000067	0.25
2500	80.39	0.3	0.000078	0.32

Table 6. Query stage of centric feature-based extraction using VGG16 model

Total Images	Data Size (KB)	Query Stage(s)		Overall Total Processing Time
		Query Feature Extraction	Similarity Comparison Time	
1	0.01	0.0041	0.000054	0.0042
10	0.06	0.0042	0.000048	0.0042
50	13.53	0.0047	0.000062	0.0047
100	12.82	0.0037	0.000041	0.0038
150	13.45	0.0053	0.000061	0.0053
200	13.88	0.0039	0.000046	0.0039
250	13.48	0.0071	0.000047	0.0071
300	13.45	0.0043	0.000051	0.0043
350	12.44	0.0042	0.000045	0.0042
500	12.83	0.0044	0.000051	0.0045
1000	13.18	0.0049	0.00006	0.005
2000	34.33	0.0075	0.000063	0.0075
2500	80.39	0.0061	0.000068	0.0061

Table 7. Traditional full pairwise comparisons at query time using 3D SIFT

Total Images	Data Size (KB)	Feature Extraction Time(s)	Comparison Time(s)	Overall Total Processing Time(≈s)
1	10.9	5.44	5.85	11.29
5	56.7	42.92	45.35	88.27
10	115	111.36	141.36	252.72
50	604	575.24	699.34	1274.58
70	860	775.75	933.46	1709.21
100	1023	1146.85	1212.68	2359.52
130	1065	1488.18	1580.34	3068.52
150	1093	1703.74	1831.17	3534.91
200	2063	2163.62	2436.14	4599.75
300	3087	3096.75	3396.96	6493.71
500	6043	5081.41	5640.18	10721.59

Table 8. Indexing stage centric feature-based extraction using 3D SIFT

Total Images	Data Size (KB)	Indexing Stage(s)		Overall Processing Time
		Feature Extraction Time	Indexing Time	
1	10.9	10.59	0.007696	10.6
5	56.7	55.99	0.000123	55.99
10	115	100.9	0.000092	100.9
50	604	511.73	0.000073	511.73
70	860	715.44	0.000086	715.44
100	1023	968.08	0.000074	968.08
130	1065	1218.12	0.007834	1218.13
150	1093	1406.3	0.000175	1406.3
200	2063	1922.84	0.00039	1922.84
300	3087	2840.71	0.001226	2840.71
500	6043	4807.64	0.019482	4807.67

Table 9. Query stage of centric feature-based extraction using 3D SIFT

Total Images	Data Size (KB)	Query Stage		Overall Total Processing Time
		Query Feature Extraction	Similarity Comparison Time	
1	10.9	8.145	0.004036	8.149
5	56.7	8.7178	0.0001	8.7179
10	115	9.2392	0.000086	9.2393
50	604	9.1528	0.000091	9.1529
70	860	10.1524	0.000056	10.1525
100	1023	12.59	0.000215	12.5902
130	1065	8.4019	0.00432	8.4062
150	1093	8.4324	0.000076	8.4324
200	2063	9.5973	0.000121	9.5975
300	3087	9.6777	0.000409	9.6781
500	6043	9.9755	0.01	9.9855

Table 10. Traditional full pairwise comparisons at query time using 3D HOG

Total Images	Data Size (KB)	Feature Extraction Time(s)	Comparison Time(s)	Overall Total Processing Time(≈s)
1	10.9	5.93	6.39	19.57
5	56.7	32.8	45.59	85.11
10	115	66.03	89.78	162.3
50	604	334.3	466.78	807.46

70	860	470.05	470.05	1108.59
100	1023	667.85	758.4	1432.89
130	1065	867.19	976.73	1850.76
150	1093	1008.21	1008.21	2157
200	2063	1317.51	1482.97	2806.96
300	3087	2019.53	2233.84	4262.98
500	6043	3416.02	3899.84	7322.72

Table 11. Indexing stage of centric feature-based extraction using 3D HOG

Total Images Per (Folder)	Data Size (KB)	Indexing Stage		Overall Processing Time(s)
		Feature Extraction Time(s)	Indexing Time(s)	
1	10.9	4.00	0.000067	4.26
5	56.7	22.87	0.000155	30.55
10	115	43.06	0.000085	60.61
50	604	237.43	0.000248	255.69
70	860	335.74	0.000281	442.15
100	1023	452.42	0.000384	603.68
130	1065	795.37	0.01666	879.38
150	1093	929.38	0.000782	1039.32
200	2063	1246.9	0.001293	1387.1
300	3087	1865.68	0.001883	2105.8
500	6043	2338.87	0.020366	3026.59

Table 12. Query stage of centric feature-based extraction using 3D HOG

Total Images	Data Size (KB)	Query Stage(s)		Overall Total Processing Time
		Query Feature Extraction	Similarity Comparison Time	
1	10.9	4.9475	0.000104	4.9476
5	56.7	4.6932	0.00006	4.6933
10	115	4.2852	0.000112	4.2853
50	604	4.6594	0.000145	4.6596
70	860	4.6394	0.000305	4.6397
100	1023	4.8789	0.000287	4.8792
130	1065	5.5825	0.007281	5.5898
150	1093	5.5712	0.000445	5.5716
200	2063	6.0327	0.000643	6.0333
300	3087	5.4867	0.0008	5.4875
500	6043	5.0964	0.001186	5.0976

Table 13. Computing accuracy and consistency of traditional full pairwise strategy based on 3D SIFT

Total Images	Traditional Feature-Based Extraction	
	Image Label	Score (L2 Distance)
500	IXI050-Guys-to-1-MRA.nii.gz	0.0000
	IXI050-Guys-to-2-MRA.nii.gz	0.0000
	IXI050-Guys-to-3-MRA.nii.gz	0.0000
	IXI050-Guys-to-311-MRA.nii.gz	1.9126
	IXI050-Guys-to-309-MRA.nii.gz	1.9126

Table 14. Computing accuracy and consistency of centric similarity search strategy based on 3D SIFT

Total Images	Centric Feature-Based	
	Image Label	Score (Cosine)
500	IXI050-Guys-to-1-MRA.nii.gz	1.0000
	IXI050-Guys-to-2-MRA.nii.gz	1.0000
	IXI050-Guys-to-3-MRA.nii.gz	1.0000
	IXI050-Guys-to-311-MRA.nii.gz	0.9905
	IXI050-Guys-to-309-MRA.nii.gz	0.9905

Table 15. Computing accuracy and consistency 3D HOG traditional full pairwise strategy

Total Images	Traditional Feature-Based Extraction	
	Image Label	Score (Cosine)
500	IXI050-Guys-to-1-MRA.nii.gz	0.0000
	IXI050-Guys-to-2-MRA.nii.gz	0.0000
	IXI050-Guys-to-3-MRA.nii.gz	0.0000
	IXI050-Guys-to-175-MRA.nii.gz	9.4248
	IXI050-Guys-to-176-MRA.nii.gz	9.4248

Table 16. Computing accuracy and consistency of centric similarity search strategy based on 3D HOG

Total Images	Centric Feature-Based	
	Image Label	Score (Cosine)
500	IXI050-Guys-to-1-MRA.nii.gz	0.0000
	IXI050-Guys-to-2-MRA.nii.gz	0.0000
	IXI050-Guys-to-3-MRA.nii.gz	0.0000
	IXI050-Guys-to-175-MRA.nii.gz	88.8264
	IXI050-Guys-to-176-MRA.nii.gz	88.8264

A key advantage of the Centric architecture is the explicit separation between indexing and querying. The indexing phase, which is executed offline, is remarkably efficient (e.g. ~ 0.02 seconds), so that only the lightweight query phase needs to be executed in real-time. This decoupled design ensures low latency and scalable performance, making the method ideal for high-throughput clinical environments. After validating our approach for large 3D volumes, we extended our evaluation to 2D image datasets using deep learning-based feature extraction models (VGG16 and ResNet50) Tables 1-6 to demonstrate the generality and robustness of our framework for different modalities and architectures.

Having analysed the computational advantages of each approach, we now focus on another crucial aspect: the accuracy and consistency of the search results presented in the Tables 13-16.

5.2 Evaluating efficiency of feature search queries

To evaluate the accuracy and retrieval consistency of the two approaches, we conducted a comparative assessment on the same datasets of high-resolution 3D MRA brain scans. Feature extraction was performed using 3D SIFT and 3D HOG descriptors as shown in Tables 9 and 10, chosen for their robustness in preserving spatial and structural features. The Traditional method employed full pairwise comparisons with cosine similarity, while the Centric approach relied on approximate nearest neighbor (ANN) search using L2 distance for 3D HOG and cosine similarity for 3D SIFT. Despite differences in similarity metrics and numerical score scales, both methods retrieved the exact same Top 5 most similar images for each query. This validates that the Centric Feature-based approach maintains 100% retrieval fidelity when equivalent feature representations and comparable metrics are used.

While score values varied due to the nature of the distance measures—L2 distance in Centric queries producing lower-is-better values (e.g., 0.0000 to ~ 1.91), and cosine similarity in Traditional comparisons yielding higher-is-better results (e.g., 1.0000 to ~ 0.9905)—the relative ranking of retrieved results remained identical. This ranking consistency is crucial for content-based medical image retrieval, where diagnostic

decisions may depend on precise match ordering. Consequently, the Centric approach offers not only superior computational efficiency but also retrieval accuracy equivalent to full pairwise methods, making it highly suitable for real-time clinical applications.

To consolidate the insights gained from the results presented above, we now provide a comprehensive summary of the experimental studies conducted across both 2D (ResNet50 and VGG16) and 3D (3D SIFT and 3D HOG) feature extraction models across a wide range of dataset sizes to assess scalability, efficiency, and retrieval performance focusing 3D algorithms like 3D SIFT for Summarizing experimental studies and compared two strategies: a traditional approach, which performs real-time pairwise comparisons for every query, and a Centric Approach, which decouples the pipeline into offline indexing and fast real-time querying stages. Among all models and settings, the 3D SIFT-based evaluation on a large 500 files of 6.5 GB volumetric brain MRA dataset highlighted the most significant performance distinctions between the two paradigms illustrated on Tables 13-16. Using 3D SIFT features, the Traditional approach required over 10,700 seconds per query (5081.41s for feature extraction + 5640.18s for comparison), making it impractical for real-time applications. In contrast, the Centric pipeline processed the same dataset in 4820.26 seconds total, with the majority (4807.64s) spent during a one-time offline indexing phase. At query time, only 12.59 seconds were needed for feature extraction, and FAISS similarity search executed in just 0.0002 seconds, enabling truly responsive queries. This design as shown in Figure 3 yields more than 55% reduction in total processing time, but more importantly, an 850 \times speedup in online execution—demonstrating that separating concerns between indexing and querying not only optimizes resource use but transforms the retrieval system into a scalable, real-time solution for medical imaging applications.

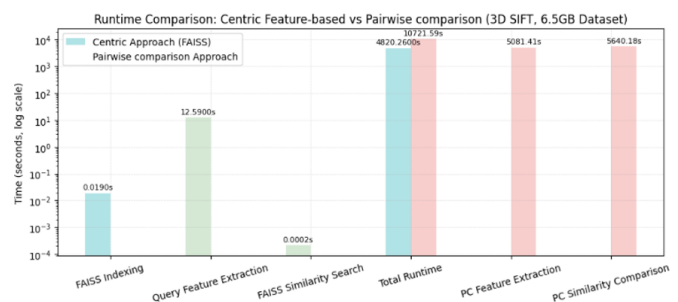


Figure 3. 3D SIFT accuracy and retrieval consistency comparison of traditional full pairwise and centric similarity search strategies

Experimental results highlight the effectiveness and practicality of the proposed Centric feature-based method. Its core strength lies in the architectural separation between offline indexing and online querying, which allows for rapid, real-time retrieval without the need to reprocess the entire dataset for each query. This design ensures seamless responsiveness, making the system highly suitable for real-time applications, including clinical and machine learning environments. The indexing process, performed only once and offline, enables the system to handle queries almost instantly, even on large-scale datasets.

In addition to its performance benefits, the method is both flexible and resource-efficient. It does not require prior image

segmentation [10] and runs effectively on standard computing hardware, avoiding reliance on costly GPU setups

. The indexing structure supports both visual and textual queries and can easily accommodate various feature extraction algorithms, whether replacing 3D SIFT with HOG or integrating GPU-accelerated methods for more demanding applications. Offline processing tasks can also be scheduled during low-demand periods, further optimizing system efficiency.

Table 17. Requirements and resources for experimental study

Total Memory	Distributed Across Multiple Nodes	A Single Pool of RAM
Operating System	Linux	MacOS
RAM Per Node	12.7 GB	32 GB
CPU Cores	8	16-core
GPU Cores	-	16-core
Hard Disk	107.7 GB	512 GB SSD
Entire server RAM	Configurable per node	Entire server RAM
Flexibility	Distributed workload	Single workload
Complexity	More complex setup	Simpler setup

Importantly, the Centric approach also offers a significant advantage in terms of data privacy and security. Because it stores only extracted features rather than raw medical images, it minimizes the risk of exposing sensitive patient information. This feature-centric design reduces the need to handle or share original data directly, thereby enhancing compliance with privacy regulations and contributing to more secure medical data management. Altogether, the Centric approach presents a scalable, secure, and adaptable solution for managing and retrieving large volumes of medical imaging data in support of precision medicine and biomedical research.

6. CONCLUSIONS

In this article, we first provided an overview of medical image modalities stored as multidimensional datasets and highlighted the major challenges in analyzing and extracting diagnostic insights from large, heterogeneous archives. These challenges include high computational costs and limited accessibility for clinical practitioners. To overcome them, we proposed the Centric Approach, which separates one-time feature indexing from real-time retrieval. This concept provides a scalable and efficient solution for 3D medical image retrieval that achieves the same precision as traditional pairwise methods while drastically reducing processing time - ideal for real-time clinical applications and large medical archives.

Our method also simplifies tedious manual data entry and retrospective analyses by enabling automatic, batch processing of historical data. In addition, the system is extensible to support 2D feature extraction and hybrid multimodal queries. Although the experimental evaluations in this article were performed with curated datasets, the architecture is designed to scale to large clinical repositories and integrate seamlessly with PACS environments, enabling direct interoperability with radiologists' workflows. This opens the door for intelligent queries, automated tagging and fast similarity-based retrieval in the clinical environment.

Future work will explore the extension of this system to 4D

modalities and closer coupling with real-time diagnostic decision aids.

ACKNOWLEDGMENTS

This work and the results shown here are supported by Matthew Toews based in whole or in part on data from the TCGA Research Network: <http://cancergenome.nih.gov/>.

REFERENCES

- [1] Tsoumpas, C., Gaitanis, A. (2013). Modeling and simulation of 4D PET-CT and PET-MR images. *PET Clinics*, 8(1): 95-110. <https://doi.org/10.1016/j.cpet.2012.10.003>
- [2] Samaan, S.S. (2016). Picture archiving and communication system design and implementation. *Nahrain University College of Engineering Journal*, 19(1): 124-136.
- [3] Aiello, M., Esposito, G., Pagliari, G., Borrelli, P., Brancato, V., Salvatore, M. (2021). How does DICOM support big data management? Investigating its use in medical imaging community. *Insights into Imaging*, 12(1): 164. <https://doi.org/10.1186/s13244-021-01081-8>
- [4] Amirrajab, S., Al Khalil, Y., Lorenz, C., Weese, J., Pluim, J., Breeuwer, M. (2022). Label-informed cardiac magnetic resonance image synthesis through conditional generative adversarial networks. *Computerized Medical Imaging and Graphics*, 101: 102123. <https://doi.org/10.1016/j.compmedimag.2022.102123>
- [5] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., et al. (2013). The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26: 1045-1057. <https://doi.org/10.1007/s10278-013-9622-7>
- [6] Yu, H., Yang, L.T., Zhang, Q., Armstrong, D., Deen, M.J. (2021). Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*, 444: 92-110. <https://doi.org/10.1016/j.neucom.2020.04.157>
- [7] Cheng, X., Zhang, L., Zheng, Y. (2018). Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(3): 248-252. <https://doi.org/10.1080/21681163.2015.1135299>
- [8] Kumar, A., Kim, J., Wen, L., Fulham, M., Feng, D. (2014). A graph-based approach for the retrieval of multi-modality medical images. *Medical Image Analysis*, 18(2): 330-342. <https://doi.org/10.1016/j.media.2013.11.003>
- [9] Sa, Y. (2015). Medical image registration algorithm based on compressive sensing and scale-invariant feature transform. In *2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, Nanchang, China, pp. 547-551. <https://doi.org/10.1109/ICICTA.2015.140>
- [10] Said, Y., Alsheikhy, A.A., Shawly, T., Lahza, H. (2023). Medical images segmentation for lung cancer diagnosis based on deep learning architectures. *Diagnostics*, 13(3): 546. <https://doi.org/10.3390/diagnostics13030546>
- [11] Shaul, R., David, I., Shitrit, O., Raviv, T. R. (2020). Subsampled brain MRI reconstruction by generative adversarial neural networks. *Medical Image Analysis*, 65:

101747. <https://doi.org/10.1016/j.media.2020.101747>
- [12] Díaz García, J., Brunet Crosa, P., Navazo Álvaro, I., Vázquez Alcocer, P.P. (2017). Downsampling methods for medical datasets. In Proceedings of the International conferences Computer Graphics, Visualization, Computer Vision and Image Processing 2017 and Big Data Analytics, Data Mining and Computational Intelligence 2017: Lisbon, Portugal, pp. 12-20.
- [13] Toews, M., Wells, W. (2009). SIFT-Rank: Ordinal description for invariant feature correspondence. In IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, pp. 172-177, <https://doi.org/10.1109/CVPR.2009.5206849>
- [14] Yang, Q., Ji, H., Xu, Z., Li, Y., Wang, P., et al. (2023). Ultra-fast and accurate electron ionization mass spectrum matching for compound identification with million-scale in-silico library. *Nature Communications*, 14(1): 3722.
- [15] Dhulavvagol, P.M., Gadagkar, A., Ateeth, K.J., Hegade, G., Poonia, R., Totad, S.G. (2023). Optimised image storage and retrieval on Hadoop. In ITM Web of Conferences, Gujarat, India, p. 03001. <https://doi.org/10.1051/itmconf/20235303001>
- [16] DeTone, D., Malisiewicz, T., Rabinovich, A. (2018). Superpoint: Self-supervised interest point detection and description. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, pp. 224-236. <https://doi.org/10.1109/CVPRW.2018.00060>
- [17] Xu, Y., He, X., Xu, G., Qi, G., Yu, K., et al. (2023). A medical image segmentation method based on multi-dimensional statistical. *Multimodal Brain Image Fusion: Methods, Evaluations, and Applications*, 16648714: 78.
- [18] Unar, S., Wang, X., Zhang, C., Wang, C. (2019). Detected text-based image retrieval approach for textual images. *IET Image Processing*, 13(3): 515-521. <https://doi.org/10.1049/iet-ipr.2018.5277>
- [19] Liu, S., Hadi, N., Liu, S., Pujol, S., Kikinis, R., et al. (2015). Content-based retrieval of brain diffusion magnetic resonance image. In *Multimodal Retrieval in the Medical Domain: First International Workshop, MRMD 2015, Vienna, Austria*, pp. 54-60. https://doi.org/10.1007/978-3-319-24471-6_5
- [20] Mohamadi, H., Shahbahrami, A., Akbari, J. (2013). Image retrieval using the combination of text-based and content-based algorithms. *Journal of AI and Data Mining*, 1(1): 27-34.
- [21] Saminathan, K. (2023). Content based medical image retrieval using deep learning algorithms. *Journal of Data Acquisition and Processing*, 38: 3868.
- [22] Kobayashi, K., Hataya, R., Kurose, Y., Miyake, M., Takahashi, M., et al. (2021). Decomposing normal and abnormal features of medical images for content-based image retrieval of glioma imaging. *Medical Image Analysis*, 74: 102227. <https://doi.org/10.1016/j.media.2021.102227>
- [23] Praveena, H.D., Guptha, N.S., Kazemzadeh, A., Parameshachari, B.D., Hemalatha, K.L. (2022). Effective CBMIR system using hybrid features-based independent condensed nearest neighbor model. *Journal of Healthcare Engineering*, 2022(1): 3297316. <https://doi.org/10.1155/2022/3297316>
- [24] Anjomshoe, S., Omeiza, D., Jiang, L. (2021). Context-based image explanations for deep neural networks. *Image and Vision Computing*, 116: 104310. <https://doi.org/10.1016/j.imavis.2021.104310>
- [25] Srinivas, M., Naidu, R.R., Sastry, C.S., Mohan, C.K. (2015). Content based medical image retrieval using dictionary learning. *Neurocomputing*, 168: 880-895. <https://doi.org/10.1016/j.neucom.2015.05.036>
- [26] Kolarik, M., Sarnovsky, M., Paralic, J., Babic, F. (2023). Explainability of deep learning models in medical video analysis: A survey. *PeerJ Computer Science*, 9: e1253. <https://doi.org/10.7717/peerj-cs.1253>
- [27] Islam, M.Z., Islam, M.M., Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*, 20: 100412. <https://doi.org/10.1016/j.imu.2020.100412>
- [28] Monnier, F. (2018). *Imagerie multimodale et quantitative en TEP/IRM* (Doctoral dissertation, Brest).
- [29] Toews, M., Wachinger, C., Estepar, R.S.J., Wells III, W.M. (2015). A feature-based approach to big data analysis of medical images. In *International Conference on Information Processing in Medical Imaging*, Springer, Cham, pp. 339-350. https://doi.org/10.1007/978-3-319-19992-4_26
- [30] Machado, I., Toews, M., Luo, J., Unadkat, P., Essayed, W., et al. (2018). Non-rigid registration of 3D ultrasound for neurosurgery using automatic feature detection and matching. *International Journal of Computer Assisted Radiology and Surgery*, 13: 1525-1538. <https://doi.org/10.1007/s11548-018-1786-7>
- [31] Carlier, J.B., Chauvin, L., Luo, J., Wells III, W.M., Machado, I., Harmouche, R., Toews, M. (2021). GPU optimization of the 3D scale-invariant feature transform algorithm and a novel BRIEF-inspired 3D fast descriptor. *arXiv preprint arXiv:2112.10258*. <https://doi.org/10.48550/arXiv.2112.10258>
- [32] Scovanner, P., Ali, S., Shah, M. (2007). A 3-dimensional sift descriptor and its application to action recognition. In *Proceedings of the 15th ACM International Conference on Multimedia*, Germany, pp. 357-360. <https://doi.org/10.1145/1291233.1291311>
- [33] Rani, S., Lakhwani, K., Kumar, S. (2022). Three dimensional objects recognition & pattern recognition technique; related challenges: A review. *Multimedia Tools and Applications*, 81(12): 17303-17346. <https://doi.org/10.1007/s11042-022-12412-2>
- [34] Popuri, K., Ma, D., Wang, L., Beg, M.F. (2020). Using machine learning to quantify structural MRI neurodegeneration patterns of Alzheimer's disease into dementia score: Independent validation on 8,834 images from ADNI, AIBL, OASIS, and MIRIAD databases. *Human Brain Mapping*, 41(14): 4127-4147. <https://doi.org/10.1002/hbm.25115>
- [35] Weber, C.J., Carrillo, M.C., Jagust, W., Jack Jr, C.R., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W. (2021). The worldwide Alzheimer's disease neuroimaging initiative: ADNI-3 updates and global perspectives. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 7(1): e12226. <https://doi.org/10.1002/trc2.12226>
- [36] Zeng, M., Yao, B., Wang, Z.J., Shen, Y., Li, F., et al. (2019). CATIRI: An efficient method for content-and-text based image retrieval. *Journal of Computer Science and Technology*, 34: 287-304. <https://doi.org/10.1007/s11390-019-1911-2>