Research Article

Analysis of the Role of Design-Driven Innovation in the Interaction Design of Image Indexing Software under the Background of the Internet of Things

Li Rong

Hunan Geely Automobile College, Xiangtan, Hunan 411100, China

Correspondence should be addressed to Li Rong; lirong@st.bitbu.edu.cn

Received 10 June 2022; Revised 14 July 2022; Accepted 2 August 2022; Published 9 September 2022

Copyright © 2022 Li Rong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

More and more medical images are being produced quickly by modern imaging devices that are connected to the Internet of Things (IoT). For the retrieval of necessary images from a conventional big volume database, the continuous scanning of all image data in an IoT network appears to be ineffective and computationally expensive. Effective image retrieval from a big image database necessitates the use of an effective and scalable image indexing approach. Existing image indexing methods have significant drawbacks, including lower efficiency, constrained scalability, increased computing demands, and longer processing times. To index the medical images obtained by IoT sensors, we suggested a novel, effective Content-Based Cascaded Gabor Wavelet Algorithm (CBCGWA) in this study. Further in this study, the medical images from computed tomography (CT) are employed for medical indexing Totally, the dataset included 168 individually annotated square patches in a subset of 115 high-resolution CT (High-Resolution CT) slices used in this study. An adaptive median filter is used for preprocessing the CT medical images once they are acquired from IoT sensor nodes connected to medical imaging systems. The Gaussian Adaptive Attention Network is then used to cluster together images with comparable attributes (GAAN). The suggested method is used to index the images after image clustering. The cloud database is where the indexed images are eventually maintained. The comparison of the recommended indexing approach to the existing indexing strategies revealed that it is better in terms of processing time, power, and indexing efficiency.

1. Introduction

In the digital era, with the rising availability and use of the Internet, wireless networks, intelligent device makers, and dynamic sensing and processing technologies, IoT technology achieves fast growth, notably in smart healthcare development. The IoT is regarded as one of the most essential trends, drivers, and enablers for corporate transformation. In the current technology era, enhancing the efficiency of healthcare and biomedical systems is the hardest issue. IoT devices play a key role in healthcare applications notably in detecting and diagnosing different types of illnesses in smart cities (Chen et al. [1]).

Data from the IoT devices may be created quickly, in large volumes, and with a wide variety of data kinds. Traditional database solutions are not a better fit because of the sheer volume of data and the unique properties they possess. Various concepts guide the design of innovative management systems for IoT data. A variety of data management systems are built on these various concepts like middleware-based IoT focused on data and sources, efficient storage of data, indexing solutions, and Schematic support solutions for IoT data. Research in image retrieval focuses mostly on indexing the images captured by IoT sensors and extracting the relevant image from a large database. It is possible to look up previous diagnoses made on images with the same lesions using indexed medical images. In cases when the diagnosis is difficult, or when the physician is dealing with external factors like inexperience, uncertainty, fatigue, etc. that might lead to incorrect diagnoses indexed images are significant assistance for physicians (Sundararajan et al. [2]).
The way medical images are organized and managed is being revolutionized by new technologies. The medical field may benefit from new approaches in clinical data analytics and administration, introduced by Internet-based technologies, increased computing power, and quick, ubiquitous digital communications. Big Data in the medical field is being used to build innovative Medical Decision Support techniques, which aim to manage, convert, and portray it in a new understandable form. So, doctors may use a smart tool to help them make decisions based on easier real-time data analysis.

Because of rapidly growing data in the IoT environment, it is becoming more difficult to integrate, index, and manage time-series data from various sources for optimal data storage and/or obtain relevant information in real-time from them (Diene et al. [3]). Two factors need to be looked at for improvement to obtain successful medical image retrieval from large volumes of images. Reducing the size of the image feature vectors and refining the data indexing are two of the elements. Both of these problems are very difficult to solve using traditional indexing approaches. For the mining of big image data to succeed, the solution must be flexible and accessible to the end-user. As one of the most often utilized strategies for quick data access against a huge dataset, indexing is an important part of this process. Enabling effective data search and discovery are the motive of data indexing systems. Several operators may be used with image indexing, such as range and spatial searches as well as trajectory searches. These fundamental operations must be implemented efficiently so that the “Big Value” of a large image dataset may be quickly examined. Many indexing strategies have thus been devised and extensively utilized in many database management frameworks, scientific models, and the Spatial Data Infrastructure (Nashipudimath and Shinde [4]). But the wide variety of data properties makes it difficult to appropriately index a data item. Conventional indexing methods have several issues like scalability, time-efficiency, and dynamic features of data, which are depicted in Figure 1. So there is a need for an efficient IoT image indexing approach in the medical field.

In this study, we proposed a novel indexing scheme named Content-Based Cascaded Gabor Wavelet Algorithm (CBCGWA) for indexing the medical images acquired using IoT sensors. This is the order in which the remaining parts of the paper are organized. Section 2 includes a problem statement and associated literature. Section 3 outlines the materials and methodologies involved in the project. Section 4 is a summary of the study’s results and conclusions. The fifth section is the conclusion of the study.

2. Literature Survey

By assessing the query performance of each query factor, (Huang and Chang [5]) developed an adaptive approach for determining the optimum multi-attribute composite index. To find the most efficient combination index for each query, they used an index structure that considers all potential sequences. Using a novel efficient strategy, (Khettabi et al. [6]) found a way to make data indexing easier while also improving the quality and speed of similarity query searches in the IoT. A method known as DBSCAN (“density-based spatial clustering of applications with noise”) was employed at the clustering phase to split the obtained data into clusters so that parallel indexes could be constructed with little overlap. Then, the B3CF-tree structure, “binary tree based on containers at the cloud-clusters fog computing level,” was used to index the data in each cluster.

There is a new approach for building R-trees in parallel using Apache Spark depending on the IoT Zetta platform described by (Limkar and Jha [7]). When it comes to Apache Spark, the primary use is to index real-time geographic data in R-tree and its variations which makes the spatial range queries get real-time results. Employing feature transformation and selection techniques and (Nashipudimath et al. [8]) applied a probabilistic feature patterns (PFP) strategy to efficiently integrate data and feature latent semantic analysis (F-LSA) method to index the various heterogeneous cluster data sources.

Hierarchical multidimensional hybrid indexing (HMDH) was developed by (Zhu et al. [9]) to solve the issues underlying management and to increase query performance. They used a recursive neighborhood search technique and a fuzzy clustering model to classify the data before optimizing the resulting HMDH. For indexing huge IoT data, a novel and efficient indexing structure called BCCF-tree was presented by Benrazek et al. [10]. The k-means clustering technique is used to split space recursively into nonoverlapping subspaces, which enhances the search quality.

The time and energy-efficient multidimensional data indexing technique were introduced by (Wan et al. [11]) to handle range queries. Furthermore, they presented data indexing techniques that utilize hierarchical indexing structures, utilizing binary space partitioning (BSP), like the tree, k-means clustering, quad-tree, and Voronoi-based algorithms, to enable more efficient routing with reduced delay. Spark, an extension of Apache Spark and SparkSQL, that supports spatial data types, geometrical operations, and
indexes, is described in depth by (Yu et al. [12]). Each Spark data partition may benefit from having a local spatial index built on top of it.

A radix trie indexing (RTI) methodology depending on semantic visual indexing is presented by (Krishnaraj et al. [13]) to retrieve images from cloud systems. Interactive optimization models are used to find the descriptor space with both semantic and visual features. As a further step, this joint space model is integrated into an RTI model to identify an efficient search solution for large-scale datasets. Distributed access pattern R-tree (DAPR-tree) is a unique indexing technique for geographical data retrieval in the computing environment proposed by Xia et al. [14]. The indexing penalty matrix depending on the quantity of data, load, and topology is provided in this novel indexing method.

A unique CPU+GPU co-processing approach for multidimensional indexing was introduced. It allows the CPU to browse the inside nodes of hierarchical tree structures, while the GPU scans the leaf nodes linearly utilizing a larger number of processing elements in the hybrid tree. Using MapReduce, Alkathiri et al. [15] proposed a platform for analyzing and indexing huge volumes of multispectral raster data, based on the Apache Hadoop environment. The input components of Hadoop's distributed file system may be indexed using Abdullahi et al.'s [16] partitioned B+ Tree data structure. An effective query processing using Hadoop MapReduce was achieved by doing this.

Double-bit quantization (DBQ) is a new method for increasing retrieval accuracy by allocating additional bit weights to each dimension, proposed by (Xie et al. [17]). A double-bit index hashing (DBIH) is also a new method for indexing binary codes generated by DBQ, and the measurement of weighted distance for these binary codes is proposed to improve the ranking of search results generated by DBIH. (Liu et al. [18]) explored numerous IoT search situations and proposed a common representation model for sensor data recordings. To allow historical and real-time searches for spatiotemporal observation data, they presented a unique R tree indexing structure.

Odontogenic keratocysts are seen in several disorders, according to (Mody and Bhosreddy [19]). On the teeth of a 12-year-old girl, several odontogenic keratocysts were discovered. The study did not find any other abnormalities that may suggest a disease. According to Garg [20], fine-grained data was utilized to find individual departures from the norm. From a theoretical and ethical aspect, Digital Twins in engineering were employed to explore these growing data-driven health care approaches. Digital methods were utilized to link physical things and to represent their state constantly. Moral differences may be found through analyzing data structures and the interpretations made of them. Digital twins are examined in terms of their ethical and sociological ramifications. The importance of data in healthcare has increased. This technology has the potential to be a social equalizer by supplying excellent equalizing improvement strategies. According to Ahmed and Ali [21], allergy rhinitis will be a worldwide epidemic. Medications that are both Chinese and Western are used in Taiwan to treat patients. Allergic rhinitis was the most common cause of respiratory illness in traditional Chinese medicine. For allergic rhinitis in Taiwan, traditional Chinese medicine is compared to western medical treatment. As mentioned by Shahbaz and Afzal [22], the use of high-dose-rate (HDR) brachotherapy eliminates radiation, enables outpatient treatment, and reduces diagnostic time. A single-stepping source may be able to increase dosage dispersion by altering delay at each dwell point. HDR brachotherapy treatments must be performed accurately because of the smaller processing intervals, which make it unable to perform error checks. Li [23] provided treatment and technologies for residential sewage to improve the rural environment. Organic and physicochemical pesticides were found in soil samples from vegetable fields in Nigeria’s Zamfara State by Salihu and Zayyanu [24]. Testing procedures and results were evaluated using GC-MS and QuEChERS.

Battur and Jagadisha [25] employed a novel idea known as Content-based image retrieval (CBIR) is predicted to be helpful for many types of medical images with distinct imaging modalities, anatomical regions with variable orientations, and biological systems. So, based on similarity matching, the support vector machine (SVM) classifier may be advantageous for grouping prediction of query and database photos. To categorize and retrieve biomedical images using the SVM classifier method, the suggested SVM-MIR was created. For analysis, the SVM-MIR-based categorization takes into account a variety of medical picture categories.

Pinapatruni and Chigarpalle [26] utilized a strong adversarial image reconstruction learning framework for the retrieval of medical images. Before reconstructing the input medical image from the encoded features, the adversarial image reconstruction network (AIR-Net) is suggested to first encode the input medical image into a collection of features. These encoded characteristics provide latent representation for reliable input image reconstruction. Therefore, to retrieve medical images, these encoded attributes are utilized in the index matching and retrieval module.

(Rao and Prasad [27]) introduce an effective image retrieval system for medical applications based on the brand-new Brownian motion weighting deep learning neural network (BMWDLNN) classifier and Canny steerable texture filter (CSTF) feature descriptor. Images’ noise is initially reduced using a Modified Kuan Filter (MKF). The Gaussian Linear Contrast Stretching Model (GLCSM) technique is then used to improve the picture contrast. The CSTF technique is then used to extract the picture characteristics. After using the Mean Correlation Coefficient Component Analysis (MCCCCA) technique to lessen the dimensionality of the retrieved features, the BMWDLNN classifier is then used. Table 1 compares existing approaches, as well as their drawbacks and findings.

2.1. Problem Statement. Whenever a medical image is generated, IoT sensors automatically gather it and send it immediately and constantly to the required apps for
processing and storage. Models for stream data compression and subsequent indexing have been developed in some previous studies. The traditional indexing strategies are computationally less efficient. There is a pressing need to develop an effective indexing method for IoT data, which is generated by a range of sensors and is constantly changing in terms of variety and velocity.

3. Proposed Work

The focus of this study is to develop an efficient IoT image indexing technique for faster and more accurate query responses to assist physicians. The adaptive median filter is used to preprocess medical images obtained from IoT sensor nodes coupled to imaging equipment. GAAN is then used to aggregate related characteristics into clusters. The CBCGWA is applied to index the images once they have been clustered. Finally, all of the indexed images are uploaded to a cloud database. The framework of the discussed work is presented in Figure 2.

3.1. Dataset

The dataset included 168 individually annotated square patches in a subset of 115 high-resolution CT (High-Resolution CT) slices. Figure 3 displays 168 61 × 61 pixel emphysema-CT patches of three different categories. The patches were annotated in never smokers, healthy smokers, and smokers in areas of the leading pattern (Figure 4). There are four different types of emphysema-CT slices, each with a file size of 0.18 and a total number of series files of 46.48.

3.2. Medical Image Acquisition by IoT Sensors

In this study, complementary metal-oxide-semiconductor (CMOS) sensors are employed for capturing patient images from the CT (Computed Tomography) scanning machine [1]. These sensor nodes are attached to a CT imaging system in an IoT healthcare environment. Whenever the images are generated, they are acquired by these sensors and sent for further processing and storage.

3.3. Image Preprocessing using Adaptive Median Filter (AMF)

AMF is employed in this process to preprocess the medical images acquired using IoT sensors (Figure 5). AMF preserves the sharpness and edges of the images while removing the noise. The median value of the pixels in the immediate vicinity is used to replace each pixel in this strategy. This filter uses a mask with a window size of 3 × 3. This is one of the finest conventional filters for removing speckle noises. The AMF preserves the image’s tiny structure and edges. The steps in image preprocessing using AMF are explained as follows.

(i) Initially the pixels of the input image is mapped into the input matrix “I” with m rows and n columns.

| Table 1: Comparison of existing algorithms. |
|---|---|---|
| S.No | Existing work | Findings | Drawbacks |
| 1 | Huang and Chang [5] | By developing an adaptive approach for determining the image index, the optimum multiattribute composite index is developed | A composite index is altered when any one of its attributes is updated. Large entries frequently make up composite indexes |
| 2 | Khettabi et al. [6] | The acquired data were divided into clusters during the clustering phase using DBSCAN (density-based spatial clustering of applications with noise), allowing for the creation of parallel indexes with minimal overlap | The DBSCAN method does not work with clusters of different densities. It fails while dealing with high-dimensional data |
| 3 | Limkar and Jha [7] | The new method for sequentially creating R-trees with Apache spark. The usage of the IoT zetta platform depends on how real-time data is indexed in R-tree and its variants, enabling real-time responses to geographical range queries | It is quite expensive and lacks real-time data processing |
| 4 | Nashipudimath et al. [8] | Probabilistic feature patterns | It is quite expensive and lacks real-time data processing |
| 5 | Zhu et al. [9] | Hierarchical multidimensional hybrid indexing | The clustering in image makes more complex n images |
| 6 | Benrazek et al. [10] | Fuzzy clustering model | The segmentation of image is not precise |
| 7 | Wan et al. [11] | Voronoi-based algorithm | It is very energy consumption |
| 8 | Yu et al. [12] | Geospark-R tree | It more complex in indexing large data sets |
| 9 | Krishnaraj et al. [13] | Radix tree indexing (RTI) | It lacks in data processing |
| 10 | Xia et al. [14] | Distributed access pattern R-tree (DAPR-tree) | It has numerous tree nodes |
| 11 | Alkathiri et al. [15] | Multispectral raster data | It is more complexity and cost expensive |
| 12 | Abdullahi et al. [16] | B + tree data structure | The drawback is difficulty of traversing the keys sequentially |
| 13 | Xie et al. [17] | Double-bit quantization | This method is more complexity during clustering |
| 14 | Liu et al. [18] | Common representation model | It has low level features that are not able to describe |
Efficient Image retrieval

Medical images

Query Image

Indexed image database

Efficient Image retrieval

Smart Assistance to physicians

Figure 2: Framework of the proposed work.

Figure 3: Layout of Image Retrieval from Indexed images.

Figure 4: Sample images.
3.4. Image Clustering using Gaussian Adaptive Attention Network. Clustering is a fundamental task in several applications, which is defined as dividing images into groups based on the similarity of their images. Image clustering relies hea\v{v}yly on the evaluation of similarity or difference between images. Differences or similarities are determined by how data is represented and how distance functions are used. Here, we utilized GAAN for image clustering. The proposed GAAN is composed of three units namely the feature learning unit, feature fusion unit, and cluster assignment unit. Figure 6 shows the main layout of GAAN.

For clustering, GAAN receives the previously processed images as input. As a first step, an encoding function is used to encode the features of the input images. Images’ features are reduced to a low-dimensional feature space in GAAN’s feature learning unit. This unit carries out the primary feature learning based on estimating the Gaussian metric of data in feature space. In contrast to conventional deep clustering algorithms, Gaussian attention weights prevent the degraded metric structure of data. The Gaussian attention weight is calculated for each feature of the input image. The significant features are learned depending on this weight. As a result, attention weights will be kept just for features that are near to the most important features and those that are away from significant features will be deleted. The feature fusion unit combines features from several categories into a single feature vector.

Assume the image feature vector $I$ with "b" number of significant features ($i^1, i^2, \ldots$) according to equation (1).

$$L = [i^1, i^2, \ldots, i^b].$$

(1)

The softmax function is defined by equation (2).

$$W = \text{Softmax}(\frac{\text{sigmoid}(\text{FC}(I))}{P}),$$

(2)

where $P$ defines the calibration factor. An effective way of keeping feature scores from being too near to one is to use the combined effects of the sigmoid function and calibration factor.

Fused feature representation is represented by equation (3).

$$l_f = \sum_b E_b i^b,$$  \hspace{1cm} (3)

where $E_b = \text{mean}(W)$, $l_f$ means the fused feature vector.

Soft clustering assignment is obtained by feeding the fused feature vector into the clustering layer. The clustering unit comprises two fully linked layers and one softmax layer. In the softmax layer, a soft cluster membership matrix is produced that reflects the precise cluster assignment of each data point to a particular cluster. The specified loss is then used as a guide for network training.

The optimization objective to steer feature distribution learning can be expressed by equation (4).

$$\text{loss}_{\text{dist}} = \min_{\theta_b} \max_{\theta_f} \sum_{b=2}^{B} W_{b-1} \log S_b (I^1) + W_{B-b} \log (1 - S_b (I^b)).$$  \hspace{1cm} (4)

The fusion loss is defined by equation (5).

$$\text{loss}_{\text{fusion}} = \|G^* - G^f\|^2 C,$$  \hspace{1cm} (5)

where $G^*$ is estimated using the fused features along with Gaussian kernel and $G^f = \sum E_b G^b$.

Recent developments often employ Kullback-Leibler (KL) divergence-dependent loss to guide the clustering process. It focuses on data points with a high degree of
certainty for its operation. If the marginal samples are ignored, then the cluster compactness is not guaranteed. A novel clustering loss depending on Cauchy–Schwarz divergence is introduced here to address this problem. The loss of clustering stimulates the separation of clusters and the compactness of clusters.

The generalization of the probability distribution function (PDF) of Cauchy–Schwarz (CS) divergence is provided by equation (6).

\[
S_{cs} = -\log \left( \frac{1}{G} \sum_{i=1}^{g} \sum_{y=x}^{G} \frac{W_{t_i} \left( t_y \left( l \right) \right)}{W_{t_{y-x}} \left( t_y \left( l \right) \right) W_{t_{y+x}} \left( t_y \left( l \right) \right)} \right),
\]

where \( G \) denotes the total features distributions, \( t_x \) and \( t_y \) represent the PDFs of the clusters \( C_x \) and \( C_y \).

Clusters would be well-spaced and compact if there was a big divergence. It is possible to use data-driven methods to modify equation (6) and create equation (7).

\[
S_{cs} = \frac{1}{G} \sum_{i=1}^{g} \sum_{y=x}^{G} \frac{\alpha_x \beta_y G \alpha_y}{\alpha_x \beta_x G \alpha_x \beta_y},
\]

where \( G \) means feature metric depending on Gaussian kernel, \( \alpha \) denotes the column of the cluster assignment matrix.

\[
K = w^{-1/2} \mu \left( 0, ..., 0 \right) \sum_{\rho} \rho^{\alpha - \rho},
\]

where \( \rho \) denotes the coordinate vector, \( \mu \) means the Gaussian parameters.

Output space induced by Softmax activation is used to ensure the correctness of the output to prevent a degraded clustering partition. Equation (9) integrates this geometric structure into CS divergence.

\[
S_{gb} = \frac{1}{G} \sum_{i=1}^{g} \sum_{y=x}^{G} \beta_{x}^{\rho} G \beta_{y},
\]

where \( \beta_x \) and \( \beta_y \) denote the \( x \)th and \( y \)th column of the matrix with components \( \beta_{xy} \). It is possible to compactly center the cluster assignment vectors. In a two-dimensional observation space, the clusters are orthogonal to each other. This cluster arrangement is defined by equation (10).

\[
S_{reg} = F(Z^T Z),
\]

where \( F(.) \) denotes the sum of the upper triangular elements of the cluster. The overall clustering loss is defined using equation (11).

\[
\text{loss}_{\text{cluster}} = S_{cs} + S_{gb} + S_{reg},
\]

The optimization function is conducted using equation (12).

\[
\theta^* = \min_{\theta} \left( \text{loss}_{\text{dist}} + \text{loss}_{\text{fusion}} - \gamma \text{loss}_{\text{cluster}} \right),
\]

where \( \theta^* \) means the optimized parameters of GAAN for reducing the fusion and clustering loss. Using the optimized parameters, images with similar features are assigned to respective clusters. The clustered images are generated as the output of GAAN.

3.5. Image Indexing using Content-Based Cascaded Gabor Wavelet Algorithm (CBCGWA).

Image Indexing is defined as the provision of a tag or indicator to images. Here we employed CBCGWA for image indexing. In the first step, Gabor wavelets are used to perform a wavelet transform of the input images. Afterward, the wavelet coefficients are quantized into four different levels. For each scale, auto correlograms of quantized coefficients are produced in the horizontal and vertical directions for each scale. The wavelet coefficients are quantized into four levels based on their scale. Noise is removed when coefficients near 0 are tiny. The index vector is finally built using the autocorrelation coefficients of the autocorrelation matrix. The process of indexing using CBCGWA is explained below.

Gabor function according to CBCGWA is a Gaussian modulated function defined by a complex sinusoid. Dilatation and rotation of the generating function yield Gabor wavelets.

\[
\phi_{nm}(i, j) = z^{-\sigma_0} \exp \left\{ -\frac{1}{2} \left( \frac{(a - o)^2}{\sigma_a^2} + \frac{(b - b)^2}{\sigma_b^2} \right) \right\},
\]

where \( a = m/k, n \in \{0, \ldots, k-1\}, \) and \( m \in \{0, \ldots, s-1\} \) define image orientation and scale, and \( k \) and \( s \) are the number of desired orientations and scales, respectively. Gabor function is defined using the equation as follows:

\[
\phi(i, j) = \frac{1}{2 \sigma_a \sigma_b} \exp \left\{ -\frac{1}{2} \left( \frac{(a - o)^2}{\sigma_a^2} + \frac{(b - b)^2}{\sigma_b^2} \right) \right\},
\]

where \( \sigma_a \) and \( \sigma_b \) are parameters of CBCGWA.

This technique is guaranteed to collect the most amount of information with the least amount of duplication. The parameters \( z, \sigma_o, \) and \( \sigma_b \) of the Gabor function of CBCGWA are determined using equations (15)–(17) respectively.

\[
Z = \left( \frac{O_i}{O_h} \right) ^{(1/D-1)},
\]

\[
\sigma_o = \frac{(z - 1)O_i}{(z + 1) \sqrt{2 \ln 2}},
\]

\[
\sigma_b = \frac{\tan \left( \frac{\pi}{2g} \right) \left( \frac{O_i - (\sigma_o/O_h) 2 \ln 2}{(2 \ln 2) 2^{\sigma_o^2/O_h^2}} \right)}{\sqrt{2 \ln 2 - \left( (2 \ln 2) 2^{\sigma_o^2/O_h^2} \right)}},
\]

where \( O_i \) and \( O_h \) mean the lower and upper bound of the designing frequency band, respectively.

To calculate the wavelet coefficients, CBCGWA initially uses Gabor wavelets to do so. To achieve high performance, the quantization thresholds used to discretize the coefficients were found experimentally. Sinusoid oscillations in the Gabor wavelets are reduced in magnitude by truncating the negative components. In the end, equation (18) is used to calculate the quantized coefficients’ autocorrelogram in the direction normal to the Gabor wavelet orientation.
Figure 7: Comparative assessment of various indexing techniques based on efficiency.

![Image Description](image-url)

**Equation 18:**

\[
\alpha_{nm}(x, g) = \begin{cases} 
E_{nm}(i'_{g}, j'_{g}) = x; \\
E_{nm}(i'_{g}, j) = x, \text{or} \ E_{nm}(i, j'_{g}) = x; \\
2 \times |(i, j)|E_{nm}(i, j) = x,
\end{cases}
\]

where \( E_{nm} \) denotes the matrix composed of quantized wavelet coefficients estimated by \( \varphi_{nm} \), \( \theta_{g} \) depicts the autocorrelogram’s distance parameter, and \( i'_{g} \) and \( j'_{g} \) are is given by equation (19).

\[
\left[ i'_{g}, j'_{g} \right]^{p} = \begin{bmatrix} \sin \theta & \cos \theta \\ -\cos \theta & \sin \theta \end{bmatrix} \begin{bmatrix} i' \\ j' \end{bmatrix} + \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix}.
\]

4. Results and Discussion

This section deals with the evaluation of the efficiency of the proposed innovative image indexing algorithm. The efficacy of the suggested technique CBCGW A was analyzed using the performance indicators such as indexing efficiency, execution time, energy consumption, and memory utilization for image indexing. The indexing performance of CBCGW A was compared to the existing indexing techniques like Voronoi diagram (VD), R Tree-Apache Spark, B + Tree, and Geospark-R Tree.

4.1. Indexing Efficiency. The indexing efficiency is defined as how appropriately the approach generates the index for images acquired using IoT sensors. Figure 7 shows the comparative assessment of various approaches in image indexing based on efficiency. The image indexing efficiency rate of the CBCGW A was greater than that of existing approaches namely VD, R Tree-Apache Spark, B + Tree, and Geospark-R Tree. The VD received 60 percent, R Tree-Apache Spark received 63 percent, B + Tree received 70 percent, Geospark-R Tree received 76 percent the proposed model received 85 percent which is very high than other approaches. This indicates that the proposed technique efficiently assigns indexes for IoT images.

4.2. Execution Time. All images captured by IoT sensors must be indexed before they can be processed. The time taken to execute this indexing process is defined as execution time. The difference between the start and finish times necessary to create the image indexes is used to determine this value. It is measured in seconds. Figure 8 shows the comparative assessment of various approaches in image indexing based on execution time. The execution time for image indexing using CBCGW A was lower than that of existing approaches namely VD, R Tree-Apache Spark, B + Tree, and Geospark-R Tree. The VD scored 54 percent, R Tree-Apache Spark scored 46 percent, B + Tree scored 75 percent, Geospark-R Tree received 60 percent the proposed model scored 34 percent which is very low than other approaches. This indicates that the proposed technique is a time-efficient image indexing approach.

4.3. Energy Consumption. An indexing process’ energy usage may be used to estimate overall energy consumption for
the proposed technique. Figure 9 shows the comparative assessment of various approaches in image indexing based on energy consumption. The energy consumption for image indexing using CBCGW A was lower than that of existing approaches namely VD, R Tree-Apache Spark, B+ Tree, and Geospark-R Tree. The VD has achieved 35 percent, R Tree-Apache Spark has achieved 46 percent, B+ Tree has achieved 55 percent, Geospark-R Tree has achieved 44 percent the proposed model achieved 30 percent which is very less than other approaches. This indicates that the proposed technique is an energy-efficient image indexing approach.

4.4. Memory Utilization. Memory utilization defines the amount of memory utilized for image indexing. If the amount of memory used by the process is unusually high despite accurate indexing, then the memory utilization must be optimized for image indexing. Figure 10 shows the comparative assessment of various approaches in image indexing based on memory utilization. The VD received 44 percent, R Tree-Apache Spark received 54 percent, B+ Tree received 56 percent, Geospark-R Tree received 65 percent the proposed model received 35 percent which is very high than other approaches. The memory usage for image indexing using CBCGW A was lower than that of existing approaches namely VD, R Tree-Apache Spark, B+ Tree, and Geospark-R Tree.

5. Discussion

The IoT brings together a variety of technologies which include analytics on real-time data, machine learning, ubiquitous sensors, and embedded systems, to create a world of interconnected experiences. Interoperability will be ensured by systems that allow for flexible schema support because of the wide range of data quantities. To make good use of enormous amounts of data, an appropriate indexing strategy is a need. A database’s speed may be improved via indexing, which reduces the number of disc accesses necessary for each query. You may easily identify and retrieve database data by using this data structure method.

Big data indexing is used to retrieve information from large, complicated datasets stored in scalable, distributed cloud storage. It is impossible to do a manual investigation of this kind of data. Optimizing data query activities would be made easier with a high-throughput indexing strategy [28]. In this study, we propose CBCGW A-based image indexing and evaluate its performance in comparison to other approaches like VB, R Tree-Apache Spark, B+ Tree, and Geospark-R Tree that are already available. It is more difficult to store data in a B+ tree since the leaf and nonleaf nodes have varying sizes [17]. Most of the techniques examined only two-dimensional geographical data and all were based on the MapReduce framework is a drawback to using R-Tree indexes in parallel creation of geospatial data. This MapReduce framework has several drawbacks, including the fact that it only works with static data, takes up a lot of disc space and time, and has a high latency and fault tolerance of the whole system as a whole [7]. Scalability is a problem with traditional indexing techniques [14]. Because of its scalability and time savings, the indexing method we’ve presented is ideal for large datasets. The CBCGW A indexing strategy for IoT-produced images has high indexing efficiency, less energy consumption, and memory efficiency.

6. Conclusion

The most significant aspect to consider when processing images from IoT sensors is how to generate an index more quickly while maintaining a higher level of index quality. For the effective indexing of images acquired from IoT sensor nodes, we suggested the CBCGW A in this paper. We utilized GAAN to cluster the images, which helps the indexing
process dynamically. We examined the effectiveness of CBCGWA in comparison to established techniques such as VB, R Tree-Apache Spark, B+ Tree, and Geospark-R Tree. The outcomes showed that CBCGWA is significantly quicker than existing image indexing approaches. An essential step in Information Retrieval (IR) systems is indexing. Since it is the preliminary step in IR and helps in effective image information retrieval, it constitutes the essential functionality of the IR process. The suggested approach takes less time to execute than the other techniques. The amount of energy used is important for one strategy. The suggested approach consumes less energy. The suggested approach only consumes a small memory space overall and it gives accurate image indexing for medical image applications. Hence, this technique efficiently assists the image retrieval process. Providing real-time security to IoT-produced data might expand the scope of this project. In this method, unnecessary images are filtered out and the component weights are changed in a direct mix of closeness coordinating. Typically, this retrieval method is beneficial as a front end for sizable medical image databases where an inquiry may be made in various images for teaching, preparing, and investigating objectives.

Data Availability
The data used to support the findings of this study are available from the author upon request.

Conflicts of Interest
The author declares that there are no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

References
[1] H. Chen, S. Khan, B. Kou, S. Nazir, W. Liu, and A. Hussain, “A Smart Machine Learning Model for the Detection of Brain Hemorrhage Diagnosis Based Internet of Things in Smart Cities,” Complexity, vol. 2020, Article ID 3047869, 10 pages, 2020.
[2] S. K. Sundararajan, B. Sankaragomathi, and D. S. Priya, “Deep belief CNN feature representation based content based image retrieval for medical images,” Journal of Medical Systems, vol. 43, no. 6, pp. 174–179, 2019.
[3] B. Di`ene, J. J. Rodrigues, O. Diallo, E. H. M. Ndoye, and V. V. Korotaev, “Data management techniques for internet of things,” Mechanical Systems and Signal Processing, vol. 138, Article ID 106564, 2020.
[4] M. M. Nashipudimath and S. K. Shinde, “Indexing in big data,” in Computing, Communication and Signal Processing, pp. 133–142, Springer, Singapore, 2019.
[5] C. Y. Huang and Y. J. Chang, “An adaptively multi-attribute index framework for big IoT data,” Computers & Geosciences, vol. 155, Article ID 104841, 2021.
[6] K. Khettabi, Z. Kouahla, B. Farou, H. Seridi, and M. A. Ferrag, “Clustering and parallel indexing of big IoT data in the fog-cloud computing level,” Transactions on Emerging Telecommunications Technologies, vol. 33, no. 7, p. e4484, 2022.
[7] S. V. Limkar and R. K. Jha, “A novel method for parallel indexing of real time geospatial big data generated by IoT devices,” Future Generation Computer Systems, vol. 97, pp. 433–452, 2019.
[8] M. M. Nashipudimath, S. K. Shinde, and J. Jain, “An efficient integration and indexing method based on feature patterns and semantic analysis for big data,” Array, vol. 7, Article ID 100033, 2020.
[9] L. Zhu, X. Su, and X. Tai, “A high-dimensional indexing model for multi-source remote sensing big data,” Remote Sensing, vol. 13, no. 7, p. 1314, 2021.
[10] A. E. Benraezek, Z. Kouahla, B. Farou, M. A. Ferrag, H. Seridi, and M. Kurulay, “An efficient indexing for Internet of Things massive data based on cloud-fog computing,” Transactions on emerging telecommunications technologies, vol. 31, no. 3, p. e3868, 2020.
[11] S. Yan, W. Zhao, T. Wang, Z. Gu, Q. H. Abbasi, and K. K. R. Choo, “Multi-dimensional data indexing and range query processing via Voronoi diagram for internet of things,” Future Generation Computer Systems, vol. 91, pp. 382–391, 2019.
[12] J. Yu, Z. Zhang, and M. Sarwat, “Spatial data management in Apache spark: the geospark perspective and beyond,” Geo-Informatica, vol. 23, no. 1, pp. 37–78, 2019.
[13] N. Krishnaraj, M. Elhoseny, E. L. Lydia, K. Shankar, and O. ALDabbas, “An efficient radix tri-based semantic visual indexing model for large-scale image retrieval in cloud environment,” Software: Practice and Experience, vol. 51, no. 3, pp. 489–502, 2021.
[14] J. Xia, S. Huang, S. Zhang et al., “DAPR-tree: a distributed spatial data indexing scheme with data access patterns to support Digital Earth initiatives,” International Journal of Digital Earth, vol. 13, no. 12, pp. 1656–1671, 2020.
[15] M. Alkatiri, A. Jhummarwala, and M. B. Potdar, “Multi-dimensional geospatial data mining in a distributed environment using MapReduce,” Journal of Big Data, vol. 6, no. 1, pp. 1–34, 2019.
[16] A. U. Abdullahi, R. Ahmad, and N. M. Zakaria, “Big data analytics: partitioned B+-Tree-Based indexing in MapReduce,” in Machine Learning and Data Mining for Emerging Trend in Cyber Dynamics, pp. 217–239, Springer, Cham, Switzerland, 2021.
[17] H. Xie, Z. Mao, Y. Zhang, H. Deng, C. Yan, and Z. Chen, “Double-bit quantization and index hashing for nearest neighbor search,” IEEE Transactions on Multimedia, vol. 21, no. 5, pp. 1248–1260, 2019.
[18] M. Liu, D. Li, Q. Chen, J. Zhou, K. Meng, and S. Zhang, “Sensor information retrieval from internet of things: representation and indexing,” IEEE Access, vol. 6, pp. 36509–36521, 2018.
[19] A. R. Moody and A. R. Bhosreddy, “Multiple odontogenic keratocysts: a case report,” Annals of Dentistry, vol. 54, no. 1-2, pp. 41–43, 1995.
[20] H. Garg, “Digital twin technology: revolutionary to improve personalized healthcare,” Science Progress and Research (SPR), vol. 1, p. 1, 2020.
[21] B. Ahmed and A. Ali, “Usage of traditional Chinese medicine, western medicine and integrated Chinese-western medicine for the treatment of allergic rhinitis,” Science Progress and Research, vol. 1, pp. 1–9, 2020.
[22] A. Shahabaz and M. Afzal, “Implementation of high dose rate brachytherapy in cancer treatment,” Science Progress and Research (SPR), vol. 1, no. 3, pp. 77–106, 2021.
[23] Z. Li, “Treatment and Technology of Domestic Sewage for Improvement of Rural Environment in China: Jiangsu: A Research,” Science Progress and Research (SPR), vol. 2, no. 1, 2022.

[24] S. O. Salihu and I. Zayyanu, “Assessment of Physicochemical parameters and Organochlorine pesticide residues in selected vegetable farmlands soil in Zamfara State, Nigeria,” Science Progress and Research (SPR), vol. 2, p. 2, 2022.

[25] R. Battur and N. Jagadisha, A Supervised Learning Approach for Classification of Medical Image Retrieval, Research Square, Durham, NC, USA, 2022.

[26] R. Pinapatruni and S. B. Chigarapalle, “Adversarial image reconstruction learning framework for medical image retrieval,” Signal, Image and Video Processing, vol. 16, no. 5, pp. 1197–1204, 2022.

[27] R. V. Rao and T. J. C. Prasad, “An efficient content-based medical image retrieval based on a new Canny steerable texture filter and Brownian motion weighted deep learning neural network,” The Visual Computer, pp. 1-17, 2022.

[28] M. Rashad, S. Nooh, I. Afifi, and M. Abdelfatah, Effective of Modern Techniques on Content-Based Medical Image Retrieval: A Survey, 2022.