American Academic Scientific Research Journal for Engineering, Technology, and Sciences
ISSN (Print) 2313-4410, ISSN (Online) 2313-4402
http://asrjetsjournal.org/

Computational Dynamic Features Extraction from Anonymized Medical Images

Angela U. Makolo[a*], Opoola Sunday O[b]

[a,b]*Department of Computer Science, University of Ibadan, Nigeria*
Tel: +234 803 862 9948 & +234 703 049 1222
[a]*Email: aumakolo@gmail.com*
[b]*Email: opoolasolakunle@gmail.com*

Abstract

Images depict clearer meaning than written words and this is reason they are used in a variety of human endeavors, including but not limited to medicine. Medical image datasets are used in medical environment to diagnose and confirm medical disorders for which physical examination may not be sufficient. However, the medical profession's ethics of patient confidentiality policy creates barrier to availability of medical datasets for research; thus, this research work was able to solve the above stated barrier through anonymization of sensitive identity information. Furthermore, the Content Based Image Retrieval (CBIR) using texture as the content was developed to overcome the challenge of information overloading associated with data retrieval systems.

Images acquired from various imaging modalities and placed into Digital Imaging and Communications in Medicine (DICOM) formats were obtained from several hospitals in Nigeria. The database of these images was created and consequently anonymized, then a new anonymized database was created. On anonymized images, feature extraction was done using textures as content and the content was considered for the implementation of retrieval system.

The anonymized images were tested in DICOM view to see if all files were successfully anonymized; the result obtained was 100%. A texture retrieval test was performed, and the number of precisely returned search images using the Similarity Distance Measure formulae resulted in a significant reduction in image overload. Thus, this research work solved the problem of non-availability of datasets for researchers in medical imaging field by providing datasets that can be used without violating law of patient confidentiality by the interested parties. It also solves the problem of hackers obtaining useful information about patients' datasets. The CBIR using texture as content also enhances and solves the problem of information overloading.

Keywords: DICOM; Medical Image; CBIR; Datasets; Texture.

----

* Corresponding author.

1. Introduction

Digitization is one of very rapid area of Artificial Intelligence (AI) that involves various applications of computer techniques ranging from simple signal analogue-digital conversion to the highly powerful and complex computational approaches. AI is being implemented practically in every area of human endeavour with the ultimate objective of optimizing productivity, efficiency and accuracy [1]. For instance, Machine learning; an arm of Artificial Intelligence has gained wide popularity in all sector of the world for its efficiency and accuracy in providing solution to both computational and non-computational problems.

In contrast to the old systems for acquisition and communication of medical images with films, the communications and the analysis of images from existing manual methods were inefficient [1]. Such inefficient systems which include limited durability of film, film deformation and the likes affect the medium for their archiving. Therefore, secure, fast retrieving medical repository, transfer, storage and viewing of medical images was not plausible but necessitate the alternative techniques of digitization of medical images through the acquisition of medical images from diverse image modalities. These images are crucial tools for medical education, research and health environment in the process of diagnosing and treating patients [2]. Medical imaging is more informative than ordinary visual inspection of the damaged human body component. However, they can be difficult to interpret, needing the involvement of highly competent medical specialists (radiologists or pathologists). This strategy aids the medical professional in the diagnosing process and allows for more precise medical decisions. It resulted in more consistent and accurate interpretations of medical images by removing some human bias in the confirmation or reassurance that the proper diagnosis was reached.

The rightful application of Artificial Intelligence technique will definitely enhance the system performance and make it more efficient by using appropriate machine learning metrics trained with enormous quality datasets. The system required quality and sufficient datasets because the systems will be dealing with critical life decisions and such case data quality will always be a highest priority so as to ensure optimum safety performance. One of the major challenges in medical imaging research is the availability of sufficient datasets and constraint of ethical policy of the patient health information system and its usage in DICOM. Such ethical policy stated that “Any personal health information must be deleted from both the DICOM metadata and the images” [3].

The DICOM file format is built on the ideas of pixel depth, photometric interpretation, metadata and pixel data. The combination of header size and pixel data resulted in the DICOM file format. DICOM standard is very useful in integrating all modern imaging gadgets and ease image transmission for medical diagnostic and researching purposes [4]. Despite the standardization of DICOM, digital images for research still suffer some setbacks due to non-availability or no access to raw medical data as a result of patient privacy and confidentiality policy of patient health information. Hence, to overcome this setback, there is the need to employ anonymization technique that will ensure data privacy and confidentiality of image metadata without any degradation to the image quality [3]. Through this de-identification, it will be impossible to determine the identity of the owners.

The Picture Archiving and Communication Systems (PACS) is used to manage the images workflow from acquisition to display. There are numerous PACS workstations linked to the primary central PACS that serves as the image server; these nodes access the images in the PACS database archive. These workstations query the database for information about the content and the database retrieves the images from the storage backend the mechanism is simply bidirectional. Content-Based Image Retrieval (CBIR) techniques were utilized in the medical imaging sector to aid radiologists in obtaining images with the most similar visual features. Content Based Image Retrieval (CBIR) frameworks are crucial in the medical area because the technique returns only similar image(s) from a large dataset search, which allowing for faster and more accurate patient identification [5].

2. Review of Related Works

2.1 k-anonymity models

A variety of strategies based on various k-anonymity models have been developed to achieve k-anonymity. Some of the models are reviewed as follows;

i. Incognito Algorithm

Sweeny's Incognito Algorithm in 1998 established the set of all conceivable k-anonymous full-domain generalizations of relation T with an optional tuple suppression threshold [8]. For each iteration, the algorithm is divided into two halves, it begins by testing single-attribute of subsets of the quasi-identifier before repeating the process with bigger subsets of quasi-identifiers to check k-anonymity. Despite the fact that the algorithm detects all k-anonymous in full domain of generalizations and returns the optimal solution based on the various criteria, the technique employs a breadth-first search method that traverses the solution space slowly.

ii. Samarati's Algorithm

This program examines multiple layers of the Domain Generalization Hierarchy (DGH) for possible k-anonymous solutions. It makes use of binary search to locate a solution in less time. According to Samarati, outstanding solutions have the fewest generalizations in their final outcomes in a table [9]. As a result, the technique will examine the cross section and locate the lowest level at which, atleast one solution vector may be located. However, this approach has a problem because the likelihood of finding an optimal solution changes dramatically with k, MaxSup lattice size.

iii. Sweeney's Algorithm- Datafly

In this method, Anonymization is performed in demographic research by mechanically generalizing, substituting, adding and eliminating facts without losing details. Because the technique simply examines a limited number of nodes for k-anonymity, the results are produced quickly. The method is a practical implementation of a greedy strategy that constructs frequency lists and continually generalizes combinations with fewer than k occurrences.

iv. K-Anonymity Algorithm Based on Improved Clustering

Wantong Z. and his colleagues provide a new way of improving clustering and achieving k-anonymity in 2017 and their strategies optimize the clustering process by taking into account the total distribution of quasi-identifier groups in a multidimensional space [10]. The strategy works well with local optimum clustering because it minimizes intra-cluster distances while increasing inter-cluster distances. The experimental results indicate that the technique may successfully limit information loss while establishing equivalence classes. It also performs well with numerical and categorical attributes. However, the approach does not prioritize quasi-identifiers.

v. PrioPrivacy Algorithm

This is a universal greedy local recoding algorithm that uses the size of anonymous groups of records produced by applying generalisation rules as a heuristic rather than any utility metric [11]. Apply generalisation first, followed by suppression of quasi-identifiers, beginning with the least important quasi-identifiers and making a few modifications per row as possible. It prioritizes quasi-identifiers and reduces information loss in the anonymised dataset during the anonymization process.

2.2 Content Based Image Retrieval (CBIR) systems

The method of obtaining suitable images from a large collection of images by utilizing qualities (such as colour, texture and shape) that may be extracted automatically from the images themselves is known as content-based image retrieval (CBIR) [12]. The availability of massive amounts of graphic and multimedia content, as well as the introduction of the internet, underlines the need for theme access strategies that go beyond simple text-based queries or requests based on matching exact database criteria.

Going by the current technology, digital images are being made in ever-increasing volumes in the medical industry and are being used for various medical examination and therapy. The retrieval features can be simple or complex, but the extraction procedure must be largely automated. Many CBIR methodologies are taken from image processing and computer vision and it is regarded by some to be a subset of that field. It is distinct from these fields because it focuses on retrieving images with certain characteristics from a huge collection.

Previously, the retrieval strategy was to attach textual metadata to each image and then retrieve them using typical database query techniques. The lack of systematization in the annotation process degrades keyword-based image search performance. These systems' image processing algorithms extract feature vectors that represent image properties. In this manner, images comparable to those chosen by the user can be obtained [13]. This means that the user is shown images in the vicinity of the feature vector from which to choose.

In content-based image retrieval system, image distance measure is the most commonly used method for comparing two images. Image distance quantifies how two images are similar in terms of colour, texture, shape and so on. A distance of zero, for example, indicates an identical match with the query in terms of the dimensions examined. A score greater than zero, indicates varying degrees of similarity between the images.

The search results can then be sorted based on how close they are to the image in question.

3. Research Methodology

The research activity is divided into six primary processes as shown in figure 1, and each process is describe as follows:

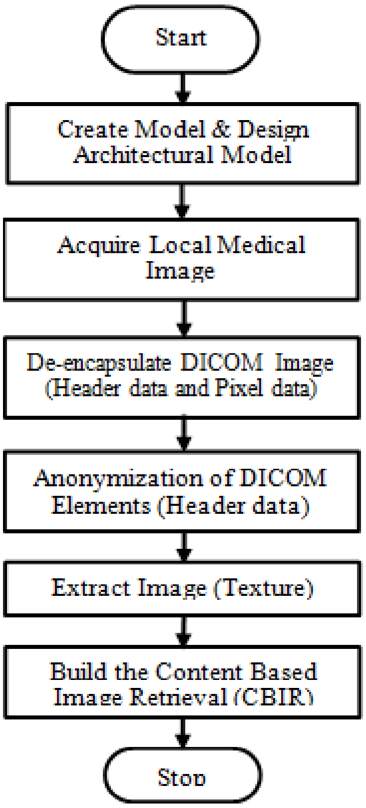


Figure 1: Flowchart depicting the research's Method and Structure.

a. Image Acquisition

Medical images of different image modalities from diverse sources were obtained from few selected hospitals and research centres in Nigeria.

Digitized hardcopy images, image file servers, isolated medical image databases and other image sources are examples of such image sources.

Regardless of where the images came from, they were collected, preprocessed and cleaned for inclusion in this research work as shown in figure 2.

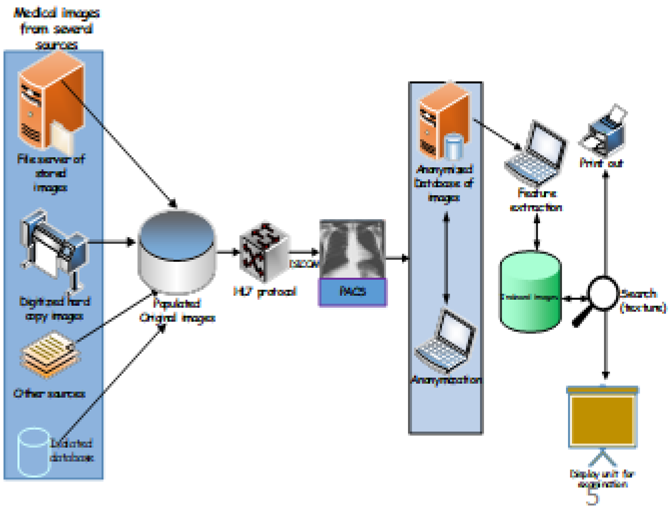


Figure 2: Architectural layout for the research.

b. Datasets and Feature Preprocessing (The input models)

There is a need for a common standard for storing and sharing information across imaging equipment and other systems because there are so many different techniques and equipment used to take biological images. As a result, all medical images gathered were subjected to preprocessing models to ensure that they met the DICOM standard. DICOM is the current standard, and it has enabled the deployment of (PACS). To ensure the convenience of PACS, all medical images would adhere to the HL7 protocol. The HL7 standard is intended for software developers and manufacturers of medical equipment with the sole goal of standardizing how information is exchanged, processed, and stored in medical units and institutions using a common format agreed upon by all parties involved [14].

c. Anonymization Model (Detection and Extraction of Features)

A DICOM file is made up of an image and a DICOM header file. The standard completely specifies the image header format, which includes information such as patient's name, date of birth, gender, phone number, address, scan date, imaging modality, imaging equipment vendor's name, image dimensions (rows, columns, slices), number of bits per pixel, frames, magnetic field strength, diagnosis, heart rate, referring physician and so on. At this stage, the selected features from the image header tags were anonymized through the anonymization engine and the result file was stored in database of anonymized images.

Anonymization Algorithm

STEP 1: Input dicom medical images

STEP 2: Construct the preprocessing features for the database

STEP 3: Detect and remove the sensitive identities/attributes information (identifiers)

STEP 4: Anonymize the identities of the image

STEP 5: The newly created images are saved in a database known as Anonymized Databases.

K-anonymity algorithm was used and implemented with python.

d. The Research Content Based Image Retrieval (CBIR) system

The key modules of the CBIR system in the research are as follows;

(i) The user interface
(ii) The feature extraction subsystem
(iii) The indexing subsystem.
(iv) The query processing subsystem.
(v) The feature matching subsystem
(vi) Relevance feedback module

(i) **User interface**: The user interface facilitates interaction between the user and the system, specifically when presenting a query to the database and viewing the retrieval results.

(ii) **Feature Extraction**: A subsystem for extraction of visual properties, for proper representation of images. Image processing and pattern analysis are mainly used in order to detect visual properties and compute their measures. The features are divided into two broad categories, the low-level features and high-level features. Low-level features are those that are calculated from pixel values while the high-level features involved semantics of the entire image.

(iii) **Indexing Subsystem:** The indexing subsystem facilitates the discovery of information about each image in the database. If the database is large enough, adequate index structures are required for careful preservation of the feature vectors. Points in a multidimensional feature space convey colour, texture, shape and other visual properties. Among the different multidimensional point indexing algorithms, the dynamic data structures, R-tree and K-d tree are widely applied in CBIR for feature vector indexing from a large database.

(iv) **Query Processing Subsystems**: It performs operations on the features collected from the query image that are required, such as (a portion of the image may be submitted as a query). To start a query, the user chooses which features or parameters are significant for a certain situation, where the user's question could be in terms of image texture attributes like coarseness, contrast and directionality.

(v) **Feature Matching Subsystem**: This module used similarity matching to compare each image in the database to the query image. Similarity-based image indexing is the process of re-ordering database images based on distance measures such as Euclidean distance, Mahalanobis distance and so on.

(vi) **Suitable Relevance Feedback:** The system began automated image analysis in order to offer information about the image without the need for time-consuming human intervention. Despite its appealing features, a fully automated system will not produce adequate outcomes because images with high degree of features similarity to the query may differ, the N-top most similar images would be retrieved.

e. Texture Feature Extraction

According to Tamura, texture qualities are classified into six categories: coarseness, contrast, directionality, line-likeness, regularity and roughness [15]. Only a few of the most relevant criteria, such as coarseness, contrast and directionality were considered for the similarity measurement in this study.

i. Contrast is defined as

$$Contrast = \frac{\sigma}{(\alpha_4)^n}$$

Where: n = 0.25 (recommended)

σ^2 = Variance

α_4 = kurtosis (degree of flatness)

$$\sigma^2 = (q - m)^2 pr(q/l)$$

$$\alpha_4 = \frac{1}{\sigma^4} \sum_{q=0}^{qmax} ((q - m)^4 pr(q/l))$$

ii. Coarseness is measure in term of pixel p(x,y) distribution within an image. The procedure is to compute (six) average for the windows of size k = 0, 1, 2, . . . 5 around the pixel. At each pixel compute the absolute differences at scale $E_k(x,y)$ between pairs of non-overlapping averages on opposite directions.

$E_{k,a}$ (P) = $|A_k^1 - A_k^2|$

$E_{k,b}$ (P) = $|A_k^3 - A_k^4|$

P(x,y) = { $E_{1,a}$, $E_{1,b}$, $E_{2,a}$, $E_{2,b}$ …}

E_k = max (E_1, E_2, E_3, . .).

The best pixel window size S_{best} is 2^k.

Then Coarseness is the average S_{best} over the entire image.

iii. Directionality: It considers the edge strength and directional angle. It is calculated using Prewitt's edge detector and pixelwise derivatives.

Edge strength = 0.5 ($|\Delta_x (x,y)| + |\Delta y (x,y)|$)

Directionality angle = $\text{Arctan}\left(\frac{\Delta x}{\Delta y} + \frac{\pi}{2}\right)$

Δx, Δy are the pixel differences in the x and y directions.

iv. Line likeness is defined as the average coincidence of edge directions that occur at pixels separated by 'd' along the direction of α.

v. Regularity is defined as:

(1 – r

$\sigma Coarseness + \sigma Contrast + \sigma Directionality + \sigma Linelikenes$

Roughness is defined as; Coarseness + Contrast

$$Roughness = \text{Coarseness} + \text{Contrast}$$

(f) **Similarity Measurement**

The extraction of features is an important stage in image retrieval and it is followed by similarity measurement, which is based on similarity evaluation; at the moment, distance and correlation are commonly employed to determine similarity. The narrower the distance between two feature points, the closer two images are to each other. The most common ways are Euclidean distance, Hausdorff distance, Manhattan distance, and EMD distance. The closely related matching output result was decided by the degree of similarity between the query image and relevant texture feature images, as shown in figure 3.

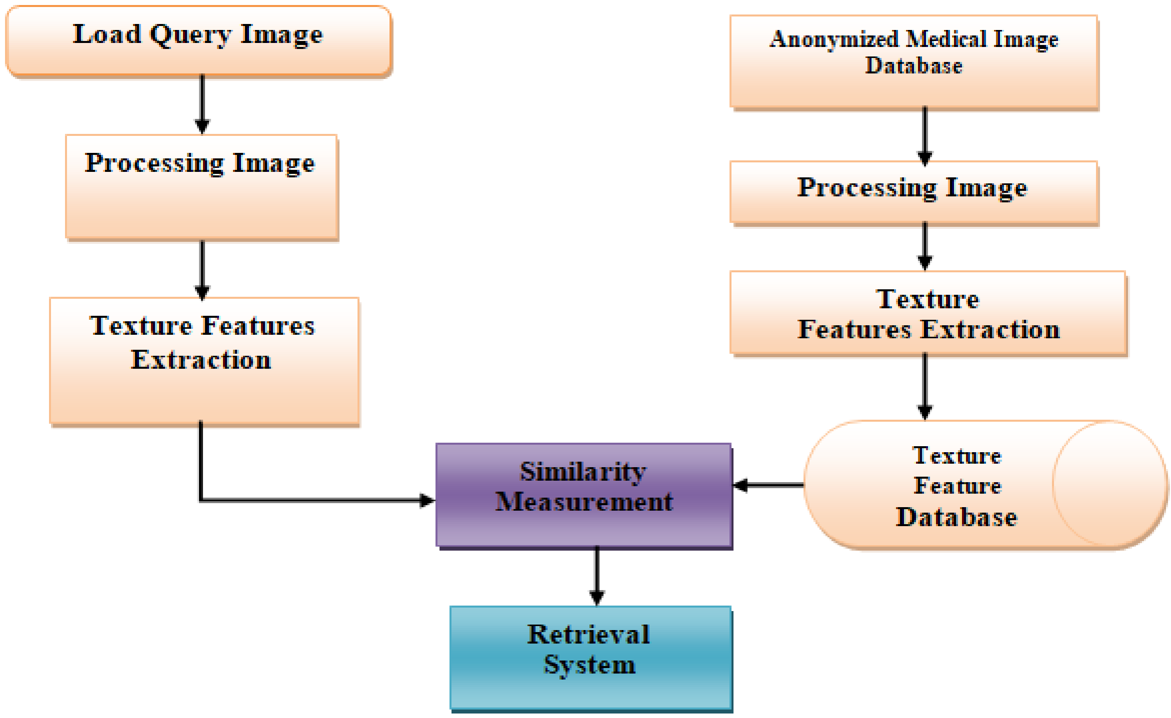


Figure 3: Similarity Measurement Model.

Adapted from [16]

Algorithm CBIR

STEP 1: Create a database with diverse medical images (import anonymized files).

STEP 2: For each image in the database, extract the Texture feature.

STEP 3: For Texture, create a composite feature vector.

STEP 4: The newly created images are saved in a separate database named Featured Databases.

STEP 5: Measure the distance between the query image feature vectors and database images feature vectors to compute the similarity ratio.

STEP 6: Sort the distance and find the issue's N-top most comparable images.

CBIR Algorithm for Query Matching

input Q; */Q is the query image,

find Fq; */ calculate the feature vector of query image Q, like step 1

for J=1: N */ N is the number of images in database Dd

compare Q; */ search for the most similar images in database Dy

*/ (compare number of components and feature vectors)

report similarities; */ report partially or completely similar cases.

end;

4. Implementation And Results

a. Implementation

The medical images obtained were saved in a folder titled Dataset (Dicom format), and following the captured of the raw medical images, a small amount of preprocessing was done on the dataset to verify that all of the medical images were in DICOM format (.dcm file extension), follow by anonymization and de-encapsulation processes.

i. Anonymization Process and Result

The images were anonymized with Sweeny K-anonymity algorithm implemented in python, both unique identifiers and quasi-identifiers were anonymized through generalization and suppression and all the dicom images saved in anonymized folder.

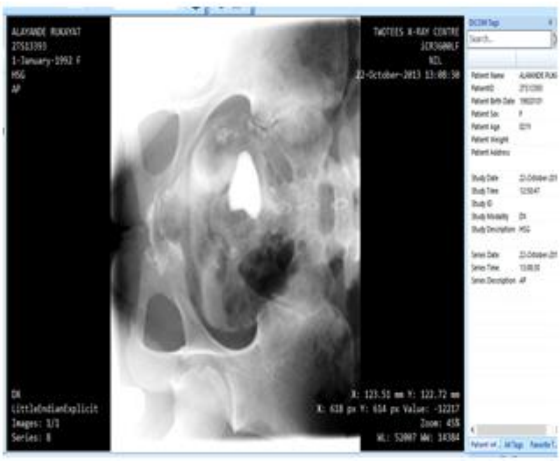
Quasi-identifiers: Let T (A1, . . ,An) be a table. A Quasi-identifier of T is a set of attributes {A1, . . , Aj} € (A1, . . ,An) whose release must be controlled.

The table 1. shows some of the attributes that anonymized

Figure 4a. shows DICOM image sample before anonymization and figure 4b.shows DICOM image sample after anonymization.

Table 1: Attributes Anonymization Details.

S/N	PATIENT IDENTIFIERS	K
1	Name	“ ”
2	Phone Number	“ ”
3	Address	“ ”
4	Sex	“ ”
5	Date of Birth	“ ”
6	Zip Code	“ ”
7	Modality	“DX ”
8	Referral Name	“ ”
9	Study Date	“ ”
10	Study Time	0.00.00
11	Series Time	0.00.00
12	Ailment	“ ”
,		“ ”
,		“ ”
An		“ ”



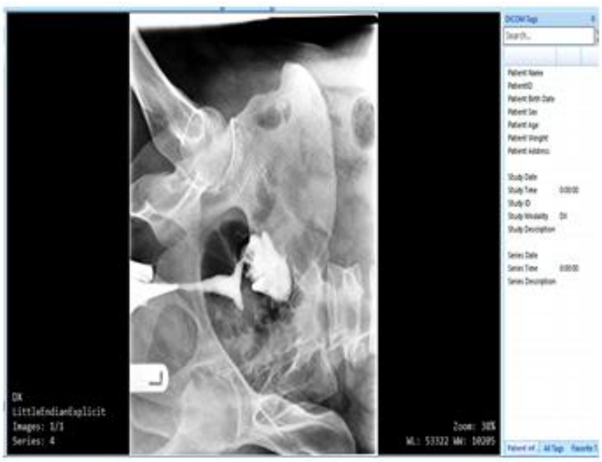


Figure 4a: DICOM medical image before anonymization. **Figure 4b:** DICOM medical image after Anonymization.

ii. **Feature Extraction Stage and Image Statistics**

Coarseness, Contrast and Roughness are three of the six qualities of Tamura texture features that were examined for feature extraction and to determine the computational properties of each extracted feature evaluated, the Tamura Algorithm was employed and implemented in Python. The statistical Texture characteristics findings from the implementation are shown in Table 2.

Table 2: Sample of statistical texture feature extracted properties.

S/n	Image Shape (H, W)	Coarseness	Contrast	Roughness
1	(1512, 1260)	9.928393	53.09822	63.02661
2	(2148, 1768)	9.76814	40.879402	50.647542
3	(1512, 1260)	9.443163	8.340268	17.783431
4	(1512, 1260)	11.193451	47.27779	58.471241
5	(1771, 1162)	8.938565	4.595516	13.534081
6	(1512, 1260)	9.554855	48.120229	57.675084
7	(1512, 1260)	9.173846	4.560938	13.734784
8	(1512, 1260)	10.281169	50.313627	60.594796
9	(2148, 1768)	8.064974	8.773325	16.838299
10	(1403, 1088)	10.380238	4.609936	14.990174

iii. Image Search and Retrieval of Image(s)

The search in this solution was designed to be beneficial for both new and experienced users. To conduct the search, the texture statistical properties are used.

This search is divided into three levels: level one, level two, and level three, and it aids in the classification of images by texture. When the number of images in the database grew, it was discovered that the time spent searching was insignificant in comparison to the time spent extracting and acquiring these images. The query technique employed for the searching was based on the level searching with the use single extracted feature and combination of extracted feature properties.

It was observed that searching with combination of two or more statistical properties outperform single searching and far better than context

searching. The query feedback is highly accurate and timely.

Table 3: Retrieval (levels 1, 2 and 3).

S/n	Image	Level 1 Searching (Using Contrast)	Level 1 Searching (Using Coarseness)	level 3 search (Using contrast + coarseness)	Irrelevant Retrievals (By Count)	Precision (%)	Recall (%)
1	Query1 (1512, 1260)	8	4	1	0	1	1

Since the combination of feature attributes provides no irrelevant data, the accuracy and recall at this level are considered to be 100 percent. The formulae below were used to calculate precision and recall.

Since the combination of feature attributes provides no irrelevant data, the accuracy and recall at this level are considered to be 100 percent. The formulae below were used to calculate precision and recall.

$$Precision = \frac{NRIR}{NIR} \quad \text{-------------- (i)}$$

Where;
NRIR = Number of Relevant Items Retrieved, and NIR = Number of Items Retrieved
For level one search,
The number of relevant records retrieved for a search = 1
The number of relevant records not retrieved = 0

Therefore;

$$Precision = \frac{1 * 100\%}{1 + 0}$$

Precision = 100%.

For recall,

$$Recall = \frac{NRIR}{NRT} \quad \text{--------------- (ii)}$$

Where NRIR = Number of relevant items retrieved, and
NRT= Number of relevant items

Therefore

$$Recall = \frac{NRIR}{NRT}$$

$$Recall = \frac{1 * 100\%}{1 + 0} = 100\%$$

The result of precision and recall being 100% means the search is effective and accurate. The F-measure was calculated using equation;

$$F.Measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad \text{------- (iii)}$$

Since the Precision is 100% and Recall is 100%
The F-Measure = 100%

5. Conclusion

This research work was able to solve the problem of non-availability of medical image datasets without violating medical ethics or medical information privacy policies through anonymization of sensitive information. It also provides high level of efficiency and accuracy in retrieving the most appropriate and relevant image with little delay. Since the query image is not examined with all of the images in the database, rather it only examined with those of similar statistical relevance thereby enhanced search performance. This approach addresses sensory and semantic gaps by improving the precision and speed with which images may be searched and retrieved in minimum time frame.

The architectural base of PACS has also been strengthened and improved through the inclusion of additional elements. Such additional elements including; anonymization engine and texture feature extraction tool. PACS's primary advantage is its capacity to deliver images and information to referring physicians in a timely and accurate manner. Researchers in this field can now use the data texture attributes of the implementation to do quick and precise analysis, as well as feature extracted statistical inference. The findings in this research work will definitely make it possible for any interested parties to have access to medical datasets without being violate private policy of medical patient's information.

6. Funding

This research was funded by National Institutes of Health, grant number U41HG006941; H3Africa Bioinformatics Network.

7. Conflicts of Interest

The authors declare no conflict of interest. The funders or the H3Africa Initiative had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the result.

References

[1]. RSNA (2019).Challenges Related to Artificial Intelligence Research in Medical Imaging and the Importance of Image Analysis Competitions. *Radiology: Artificial Intelligence* 2019; 1(1):e180031 https://doi.org/10.1148/ryai.2019180031

[2]. Burbridge, B. (2020). Dicom Image Anonymization and Transfer to Create a Diagnostic Radiology Teaching File. Int J Radiol Imaging Technol 6:065. doi.org/10.23937/2572-3235.1510065, Volume 6, ISSN: 2572-3235.

[3]. Martin, J. W. et al. (2020). Preparing Medical Imaging Data for Machine Learning. Radiology 2020; 295:4–15, https://doi.org/10.1148/radiol.2020192224

[4]. Trivedi, D.N. et al. (2019). *Dental Image Processing for Human Identification*. https://doi.org/10.1007/978-3-319-99471-0

[5]. Aishawariya, R.N., Sushmitha, N.S., Nalini, M.K. and Semester, B.E. (2019). Content Based Image Retrieval for the Medical Domain. International Journal of Engineering, Research &Technology (IJERT), ISSN: 2278-0181, Vol. 8, Pg 1345 – 1351

[6]. Feng L. et al, (2017). The Current Role of Image Compression Standards in Medical Imaging, www.mdpi.com/journal/information.

[7]. Anisha T. and Minu C. (2017). K-Anonymization Techniques for hiding multi-sensitive information. Scholars Journal of Engineering and Technology (SJET), Sch, J.Eng. Tech., 2017; 5(6):232-237.

[8]. Sweeney, L. (1998). Datafly: a system for providing anonymity in medical data and Database Security, XI: Status and Prospects, IFIP TC11 WG11.3 11th Int'l Conf. on Database Security, 356-381, 1998.

[9]. Samarati, P. and Sweeney, L. (1998). Generalizing data to provide anonymity when disclosing information. In Proc. of the 17th ACM SIGACTSIGMOD-SIGART. Symposium on Principles of Database Systems (PODS). page 188, Seattle, WA, June 1998.

[10]. Wantong, Z. et' al (2017). K-Anonymity Algorithm Based on Improved Clustering, Springer International Publishing.

[11]. Alexandros, B., Ioannis, M. and Mihai, L. (2019). PrioPrivacy: A local recoding K-Anonymity Tool for prioritized quasi-identifiers. Association for Computing Machinery, ACM ISBN 978-1-4503-6988-6/19/10.

[12]. Hirata, K. and Kato, T. (1992). Query by visual example – content-based image retrieval. In EDBT'92, Third International Conference on Extending Database Technology, 56-71

[13]. Xiang-Yang, Wan., Bei-Bei, Zhang and Hong-Ying Yang, (2014), Content-based image retrieval by integrating colour and texture Features. Multimed Tools Appl (2014) 68:545–569 DOI 10.1007/s11042-012-1055-7

[14]. HL7. (2008). Website of the Health Level 7 Inc., http://www.hl7.org. Health Level Seven, German usergrouphttps://doi.org/10.1007/978-3-319-99471-0_4

[15]. Priyanka S. (2021). An Overview of Feature Extraction Techniques in Content-Based Image Retrieval Systems. International Journal of Current Engineering and Technology. Vol. 11, No. 1 (Jan./Feb. 2021), P-ISSN 2347-5161.

[16]. Olaleke, et al. (2019). Appraisal of Content Based Image Retrieval (CBIR) Methods, Asian Journal of Research in Computer Science Creative ISBN: 2581-8260.