

Enhanced Visual Analytics Technique for Content-Based Medical Image Retrieval

S. Abinaya¹ and T. Rajasenbagam²

¹School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India, s.abinaya@vit.ac.in ²Department of Computer Science and Engineering, Government College of Technology, Coimbatore, India, trajasenbagam@gct.ac.in

*Corresponding Author: S. Abinaya; Email: s.abinaya@vit.ac.in

ABSTRACT- Content-based image retrieval (CBIR) is a method for searching that finds related images in a medical database. Furthermore, a clinical adaptation of CBIR is hampered in part by a contextual gap that is the disparity among the person characterization of the picture and the framework characterization of the image. This technique makes it tough for the user to validate the fetched images that are similar to the query image in addition to that it only fetches the images of top-ranked and ignores the low-ranking ones. Visual Analytics for Medical Image Retrieval is a novel procedure for medicinal CBIR proposed in this research (VAMIR). By integrating human and machine analysis, Visual Analytics provides the potential to address the abovementioned significant challenges. The texture properties are retrieved using the shape features extraction and Gray Level Cooccurrence Matrix (GLCM) is performed by contour-based shape descriptor. Using the Euclidean distance correlation metric, related medical pictures will be fetched by distinguishing the query image's attribute vector with the database images' respective attribute vectors. A vector of multiple features outperforms a vector of a single feature in terms of quality. The VAMIR implementation demonstrates that the search outcome for the user is acquired with 90% of recall and precision.

Keywords: Image retrieval, Lung CT images, Visual Analytics for Medical Image Retrieval (VAMIR), Content-based image retrieval.



The action of collecting knowledge resources appropriate to a demand for knowledge from a repository of knowledge resources is known as information retrieval (IR). Entire content or further forms of content-based indexing could be used for searches. Information retrieval is the ability to retrieve relevant in a document and also look for metadata to characterize facts and datasets of pictures, sounds, or texts. While a person inserts a query into the model, the information retrieval procedure starts. Queries, like seek terms in web crawlers, be explicit representations of knowledge requests. The query cannot peculiarly identify a specific object in a group when it appears to retrieve information. Instead, numerous items might correspond to the query, with varying degrees of significance.

A computational framework for finding and obtaining images through a huge set of visual pictures is meant as a system of image retrieval. Digital images in the medical industry expand

by thousands daily, posing some s retrieval, storage, and interpretation issues. In recent times, retrieving stored images from these huge archives for utilization in the clinical appliance, teaching, and research has been a key focus of medical informatics research [1]. CBIR has been employed for this rationale.

CBIR is a strategy for searching employed to find images that are related toward the query image [2] which is meant as content-based visual information retrieval (CBVIR) and query by image content (QBIC). The phrase "content" refers to shapes, colors, and textures. One of the key issues in the medical industry in current times has been the retrieval of images. CBIR, furthermore recognized as Content-based Medical Image Retrieval (CBMIR), is employed for this purpose [3]. The primary issue in CBIR is the discrepancy, which is the disparity between the system and human descriptions of an image. Several models are integrated through CBIR, however, it overlooks the images of lowranked that are related to the query image. The CBIR model employs machine learning methods for feature selection and extraction to minimize the semantic gap.





Research Article | Volume 10, Issue 2 | Pages 93-99 | e-ISSN: 2347-470X

As illustrated within Fig.1, a user provides enters the system of retrieval framework in the form of an image. Then, for each of the query images and the database images, image features are extracted. The retrieval technique uses correlation measures to determine the correlation between the database images and query images [4]. The value of the similarity metric is ordered in descending format. Only the images with a high-rank value are taken into account which is highly correlated to the query image.

2. RELATED WORK

In the literature, Ashnil Kumar et al., (2015) the image is provided by the user as an input to the model. Images are used to extract features, which are then used in the selection process. The Pearson correlation coefficient approach is used in Automatic Feature Selection. The manually Selection procedure employs domain expertise and experience. The correlation between the database and query mages is computed after feature selection. This method lets users include features that are relevant to them where the user is unaware of the retrieval algorithm.

Weidong Cai et al., (2000) provided a prototype concept for a dynamical positron emission tomography database system of content-based retrieval of the functional image. TTAC is a preprocessing input in a functional image analysis engine. Using the Knowledge-based cluster assessment approach, the image database engine extracts characteristics from the TTAC curve. The Euclidian distance approach is employed in a dynamic image archive framework to compare the attributes of database and query images. Converting from PTAC to TTAC

might consume time but it can readily hold a bigger quantity of patient data.

Mahmudur Rahman et al., (2007) created a system comprised of both unsupervised and supervised learning approaches to minimize the conceptual gap and improve retrieval results. For extracting features, principal component analysis is performed. For classification and pre-filtering support vector machine (SVM) and fuzzy c-mean clustering methods are utilized. As a similarity comparison method, the statistical distance measure computes the similarity between database and query images. RF technology is utilized to do re-ranking, parameter updates, and query shifting all at the same time. Recall and precision results show that once the portion of categorized feedback records available is sparse, the SVMbased RF method performs poorly.

Gwenole Quellec et al., (2010) proposed a system for retrieving partial medical cases with contextual information This approach is used on datasets from mammography screening (DDSM) and diabetic retinopathy follow-up (DRD). Mages are defined in this context as digital material. For the generation of decision trees, the C4.5 classification method is employed. The signatures' distances and image signature are determined using Kullback–Leibler divergence and wavelet coefficient distribution. Clustering with a Fuzzy c-mean approach is used to cluster the signature. An AdaBoost method is employed to improve the accuracy of predicting by combining weak classifiers. Precision is utilized to compute the retrieval result's accuracy. Boosting does not result in a statistically meaningful gain in accuracy over the full dataset.

3. PROPOSED SYSTEM



Figure 2: VAMIR FRAMEWORK



Research Article | Volume 10, Issue 2 | Pages 93-99 | e-ISSN: 2347-470X

To address the semantic gap, incorporate visual analytics into the Medical CBIR approach known as Visual Analytics for Medical Image Retrieval (VAMIR). During a channeled search among the huge attribute set employing the query image as a reference position, VAMIR would allow more efficient and effective retrieval and inference of current medical images. VAMIR will also be able to assess 'missed' images from other approaches.

The system depicted in *Figure 2* includes a CBIR engine, representations for the query image, images obtained, in addition to VA modules that enable the person to strain and choose images and their characteristics for investigation. In our approach, no single CBIR engine is needed, and alternative engines can be employed for various databases or scenarios. The VAMIR modules in the process provide an intuitive graphical strategy for analyzing CBIR engine retrieval outcomes. Principles for visualization and search interfaces recommend that various differing opinions of data that are complex, explorations to comprehend it, the demonstration of auxiliary data, in addition to the potential to converse by the visualization to modify the data provided otherwise to tweak the subspace can facilitate appropriate human perception of visualized information.

Following parameters and our ultimate goal of employing VA to investigate the extracted features, VAMIR is created. The prerequisites comprise the subsequent:

1) Supplying alternative perspectives on the attributes gap.

2) Facilitating automated attribute choice depending on person input to colonize all of the points of view.

3) Avoiding the depiction of numerous huge images by extrapolating the separate images at every view.

4) Enabling direct image examination by integrating each conceptualization into its related image.

5) Permitting the person to filter the visualization

6) Facilitating the person to filter otherwise transform the seek-out gap and fine-tune the query with dynamism.

3.1 VAMIR visualization design

Our visualization approach is based on the idea that just a fraction of the figure characteristics are necessary for a person to comprehend the obtained outputs related towards the query, that is, a fragment of the attributes has top importance for a similar image. [5]. Based on the query, the user, or the concluding retrieval appliance, the subset may transform. Hence, by clustering images according to their correlation using the subset of features, the VAMIR visualization lets the user notice significant patterns that are not visible in a ranking. The nodes of usage (points) in the visualization of VAMIR serve as abstractions of images. If the variance between the feature subset of two nodes is minimal, then merge them. Interacting using visualization lets users peek in on specific areas of relevance and study the relationship between population dynamics throughout the connected feature clusters [6]. Users may subsequently use domain expertise about the significance of image attributes, such as tumor volume, to proactively change and optimize the query or investigate other subsets of features.

3.2 VAMIR analytics interactions

3.2.1 Image Visualization

Whereas the VAMIR process gives additional imminent keen on image correlation, a user must still inspect the real figure to determine the complete qualities of an instance (e.g., stages of infection). To determine whether an image obtained by feature space exploration meets the supplied search purpose, a connection between both the conceptual representation and the real image is provided [8].

3.2.2 Node Selection

The proposed visualization is built on solitary feature pairs, allowing persons to investigate the significance of certain attributes to image correlation. To permit the user to contrast images utilizing various groupings of characteristics, a global selection scheme [9] is employed. It connects the nodes from every sector as well as the super-Centre. When a user chooses nodes inside one area, they simultaneously choose the matching nodes in every subsequent sector and the super-Centre. By selecting numerous nodes in a single sector, the user may determine to suppose imagery be grouped close to any form in entire attribute pairs (relatively highly correlated) else merely part of it (an outlier).

3.2.3 Dynamic Querying

If an initial query was an outlier in comparison, the person might discover it beneficial to narrow it by employing an image in the dataset where the node is near to such query node inside a specific region. VAMIR subsequently receives the dataset's correlation to an incoming query from the CBIR framework. This correlated data is utilized to immediately revive the display while maintaining the attribute couplings [10]. After that, the modules depicting the query and the visuals obtained have been modified.

4. DATASETS USED

The image repository of the Lung Image Database Consortium (LIDC-IDRI) comprises diagnostic and screening thoracic computed tomography (CT) images through highlighted interpreted tumors. This records assortment includes 1018 instances and is produced in association with 7 academia and 8 imaging medical firms. Every context holds imagery commencing a medical thoracic CT scan plus an XML document with the outcomes of a 2 stage image analysis process completed via 4 experienced thoracic radiologists. During the first completely blinded stage, every radiologist separately assessed every CT scan moreover classified laceration into 1 of 3 groups ("non-nodule > or =3 mm", "nodule <3 mm", and "nodule > or =3 mm"). Before reaching a final decision, every radiologist independently reviewed the markings, and also the anonymous markings of the 3 additional radiologists, within the phase of unblended-read.



Table 1: Database description

COLLECTION STATSTICS	COUNTS AND DEPICTION
Modalities	CR(computed radiography) DX(digital radiography) CT(computed tomography)
Count of patients	1,010
Count of radiologists	4
Nodule size comparison	nodule>=3, nodule<3, non-nodule
Count of Images	244,527
Image Size (GB)	124

5. IMPLEMENTATION

5.1 Feature extraction

The feature gives relevant information about the images. Feature extraction is the most critical step because it signifies the move from graphical data description to non-pictorial data description. The resulting representation can be used as an input to further processing image processing. Shape and texture are the two most commonly employed attributes in medical image retrieval. Texture attributes provide information on the spatial configuration of color, a particular part of an image, and intensity levels in an image.

5.1.1 Texture feature extraction

As shown in Fig.3 Texture is described as an abstraction made up of pixels that are connected and a collection of pixels. It is classified into two approaches as Structural Approach, Statistical Approach [11]. Here, a statistical approach is used for retrieving texture characteristics.

5.1.2 Gray Level Co-occurrence matrix (GLCM)

Table.1, it is the tabulation of how often the different combination of pixel grey levels co-occurs in an image. It is the foremost traditional second-order statistical approach for analyzing texture. Second-ordered notes relationship between groups of two pixels in the original image.

Algorithm Steps

Step 1: Convert the image in color into an image in grayscale. Step 2: Grayscale image is transformed into matrix format

based on the gray level in the images.

Step 3: Construct the GLCM Matrix by counting the co-occurrence of pixels in the original image matrix

Step 4: Find Transpose of a GLCM matrix by interchanging the row and columns of the matrix.

Step 5: Add the original matrix to its transpose.

Step 6: Normalize the GLCM Matrix by using the Eq.(1)

$$P_{ij} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}$$
(1)

5.1.3 Shape feature extraction

Research Article | Volume 10, Issue 2 | Pages 93-99 | e-ISSN: 2347-470X

The form of an item, as seen in *Figure 4*, alludes to its intrinsic structure. The border, region, moment, and so on may all be used to depict the form. These visualizations could be employed to match shapes, recognize objects, and measure shapes. Contour-based shape descriptor used for Shape features extraction. Different kinds of shape attributes are constructed for object identification, and their accuracy in retrieving comparable shapes from a database is measured. Accurate form descriptors must be able to successfully discover comparable shapes from the database for excellent retrieval. A shape descriptor must be able to get images for a wide range of forms rather than just a few, and it should be application agnostic. Here, a contour-based shape descriptor is used for Shape features extraction.

5.1.4 Contour-based shape descriptor

Boundary information will be retrieved employing contourbased approaches. For extraction of features, information such as circularity, area, minimum length, maximum length, and so on are determined via the boundary parameters. Because simple shape characteristics may only distinguish between objects with considerable variances, they are frequently coupled with various shape attributes for an effective retrieval procedure. (*Figure 5*)

ance in	UY												4		- 6
Ranger 201	340 . 040	Column vecto Matrix Cell Army	1 (41)	nes suit for treat	ing unings	nav an		-	W. mart						
8.170		Sector Inc.	83	_	200	10.01	_		-						
languite 10					_										_
4		5		E		6	н	1	1		4	M	N	0	_
							-								
mages	35M	Contract	Conelation	Variance D	М	Similing	Sam Ver 3	iun Erbopy B	intropy	Off is 0	#Entropy 3	(C D)	M2 N	ACC .	
DATASET STELpp	13(2)	6,295	0.854	6206	1.859	3.8620	27,2022	13496	12255	8,2470	0.942	-1578	1,840	0.0549	
DATASET SSLipg	1516	0.277	0.2568	6.905	1923	33075	17837	1387	2,2134	0,346	0.5285	-1582	0.3048	0.0800	
DATASET MILIPS	1627	0.208	93417	14112	1905	3,4529	22,4967	1128	14507	0.1925	0,3696	-1004	5.8979	0.0620	
CATAGET SIM JPS	1.508	0.1977	0.9681	6.3677	1942	3,7543	263067	12001	1,892	0.1800	0.4057	-1889	0.5250	0.054	
54149E7 \$15,99	150	-0.553	0.9753	7.4349	1,678	39522	3.738	1,346	1,988	8,1738	64125	4.93	1,805	0.0608	
DATASET JE jpg	1:36	0,1130	0,857	25.0608	1366	11.1852	140.9075	22530	3,2768	0.1022	03474	-6.7894	2362	0.1598	
DATASET JUL PS	\$122	0.0889	0.9921	28.3760	0.05%	9.5700	114/0122	2260	1340	0.0800	0.2910	-1634	0.9960	0.152	
DATASET 211.jpg	82371	0,080	0.9857	21,6730	0.520	7,8528	\$7,7990	2255	1,290	1166	0.465	-5.7454	0.9906	0,2103	
DATASET 22.09	12487	0,1790	0.982	20,4950	1,532	7,5284	\$2,2699	133	1,000	0.1562	0.4301	-0.762	6,9905	0,200	
DATAGET (20) pg	1235	0,1798	6,870	21,298	1,639	7.852	85,3412	2,6%2	1173	0.1536	0.4042	4,7625	0.3699	0.2103	
DATAGET CLA.pg	5388	\$1378	0.9907	252822	LIKE	8.4674	201.2922	1,9683	3.0042	5,1230	0.3748	-1.787	63063	0.2100	
DATASET (25 jpg	12314	0.2990	1,946	2584ET	12.9	7.6145	85,6695	21586	3,3200	0,1727	64711	-67571	1.896	0,205	
DATASET 223.00	1308	0.1852	2 0.0749	TAID	1942	4,0058	30,5989	13290	20802	01710	0.3742	-6784	1942	0.0627	
DATASET 202,00	1505	0.1368	0,9749	1,000	1258	43545	8110	1,398	2,558	84110	0.3534	-67,89	0.824	0.0765	
DATASET 223 pg	1404	9,1822	0,5743	7,7,288	0,802	41190	30,985	1,116	2082	0.1700	0,7508	-8,7341	0.054	0.060	
DATASET JUN jpg	15154	6 0.1940	0,0719	1320	1949	3,3589	29.3529	13125	21702	0.1806	0.3731	-475.8	0,9481	0.0014	
04149ET-23.pg	1453	0.1503	0.9769	7.4943	1953	4,1488	30,3450	1367	21154	0.1400	0.3405	-63434	0.9580	0.1107	
DATASET 31.00	1528	0.185	0.8749	1637	1962	4,302.5	30,5989	13280	21832	0,1710	0.3742	-1.7384	1943	0.007	
DATASET USU pg	1,505	0.1366	0.5749	1,000	1208	4,0545	81,1713	1,398	2,0506	01348	0.394	-4728	0.9524	0.0768	
DATASET 20 pg	6404	0.1822	0.9043	7,7,288	1,812	4,1190	30,985	13296	23982	01705	0.2506	-67340	0.3540	0.060	
CATAGET 204,09	1354	0,294	0,9725	7,1281	1949	33589	29,1520	13123	23782	0.2896	6,3731	-173	0,940	0.0614	
BATAST 25 JPS	14571	0.150	0,09	7.4843	1,63	4148	30.5450	1,3967	21154	0.1400	03455	-17434	0.9580	01107	
POTSTT DE Les	1144	6144	6 909	77 2176	A Real	1136	110-4412	24736	195	5184	6.2579	1.7942	8 8872	0.2128	

Figure 3: Texture features extraction

	VEI										Acres
Rang	e 40	• 1 Colu	nn vectors h	241 1 100	nuie for treatie	g wingeda	ikate :			¥.	
ete lo	#1		Seaty	1					- 1	Selector +	
10.00	10	. 1870	no pra			2000	ALC: NO.			10.00	
data	X										
	1.	0	0	1	14	- 6	18	1	1		
-				stap	1	_		_		-	
1.000	100 0000		0	0.00	and a		renndet.	11700	Toundhes		
1000	206300	-	1	1	62 42	257		1100	-		
C STORE	301.5000	- 1	1	4		34		11084			
1 1000	30 100	100	w.	1548	778,7719	475.967	99.251	26.02	0.01		
1.988	111.500	1	1	1	2	340		11284	H.		
1.500	131-5000	2	2	2	2.5000	340,5000	2836	1.508	314		
1525	1.865	1000+00	0.0020	- 2	6.0220	1803	2	1.958	6.2		
1.5300	26.522	43	48	2429	252-4201	4353254	1111Be-53	554220	0.02		
6.588	40.500	294	40	1079	135.5284	484,7074	644.2253	36.9620	630		
7.500	845,500	52	36	115	112,7913	467.6348	131,9947	121005	0.05		
1.500	20,500	45	45	.2328	252,5906	427,9759	10490e-01	54408	0.02	4	
1,5000	252,5800	6	5	35	485	395,2000	26.7270	3,982	0.44	9	
300	20,500	1	1	1	11	29	1	11294	м		
1,5000	201.5000				14,7900	26,250	344	2,258	17		
1,200	394,5000	2		-	0120	26.3750	-	21955	9.82	1	
1.000	101 5000	- 1		1	24	344 0000	24764	1 0.000	110		
T Game	201500	1	1		10.000	711	-104	1136	34		
1.500	34 500	1	1		14,7500	291,7526	100	2.768	17		
1900	344,5000	3			251250	346,3750	11.6712	1185	0.32		
15880	23.500	1	1	1	34	13		11284	М	1	
7 9000	341500			2	15 5000	144 1000	7876	1.959	214		

Figure 4: Shape features extraction



Research Article | Volume 10, Issue 2 | Pages 93-99 | e-ISSN: 2347-470X

5.2 Feature selection

It signifies the task of deciding on a division of the retrieved characteristics. It will be known as an attribute or inconsistent selection [12]. The precision of the outcomes may be affected by the features used. By deleting unnecessary and redundant features, it picks the finest division of the attributes among the real retrieved attribute.

Regression: It is the statistical method for modeling the relationship between the variables. It starts with no independent variables. It adds the variable one at a time as long as a significant reduction in the error sum of squares (SSE) can be achieved. In the simple regression model, split the features into a dependent variable and independent variables. The dependent variable is the single variable being predicted by the regression model [13]. The Independent variable is the explanatory variable used to predict the dependent variable.

Dependent variable: Correlation

Independent variables: Inverse Difference Moment, Information Measure of Correlation II, Contrast, Variance, Energy, Sum Entropy, Entropy, Difference Entropy, Sum Average, Sum Variance, Difference Variance, Information Measure of Correlation I, Maximal Correlation Coefficient.

4453 (K.C.M.												4 0
10 FB 25A 300	1884	100					02			the C	anywind	q a
👻 📑 200 -												
There I a make set and												
14 🖂 🛃 🛔 e C. e Unice e Theorem e Decimi	Finance + 1941)	A loarshipperd + BUT										•
anext failur	4	Silver still ?					n minter b	- 25.04				8.
and Cl		T pakes										
ange with factor in	*	NACHCH III -										
manatik () manatarkestof tokowi ()		1 mpri	1 DME	1 381e	4 Difference	5 Cortest	i Loogy	4	1		-14	-
manifered in		E DATAITURAN	072.0	036.0	24020	SET 1	11111					
numper C	1.	PRINCTERATOR D	-13882	8445.8	10232.0	10021	1051					
metal.	0	PRACTEMENT E	A385	12843	MMCD -	1012.0	1007					
tan.matek		PRAETEATAT A	8885	10813	1040	1262.8	1001					
Sinuya milatarbo		MARCHEATER R	0805-	107.13	153.0	1302	1886					
Scool II		PERCINATAT &	10753-	11819	1010	0023	11W					
analysis and	5 <u> </u>	PRINTEATAD C	#835-	1080.0	0.000.0	10011	12460					
he daire -	11	BRUILLARANCE R	8853	80118	5343	1812.8	MILL					
Bre. Lyfalleis		T DATAGETUDINE	083	19810	3099A0	0821	ONGE					
fan belten. er	7	PRISTINATAD RE	6363	HELS .	0040	872.8	10.01					
of All decempts	A	BRADTEATAT H	1005	8118	9/12	0023	1881					
- tought	8	DE CLITERTAR II -	2015-	107.18	2550	0821	MILE					
		PR.ILTHATAT LE	467.3-	81519	SALE	LINE	1001					
Table of the local division of the local div		WATELTEATANT AN	8015-	10/018	XEED	8286	NRL .					
Concesso fraction		PRESS AND	0215-	10710	3051.0	LUE 8	1001					
		AN THIAT AN	853	0.011	2071.0	04018	1881					
		WESTINGTON CO	4953	10813	2043	10218	4005					
		621				1						
		Conversion Name										
		A >>	\$1 mm cards									

Figure 5: Feature selection

Select a file						1
Ubraries	 Documents + MATLAB + query ima 	pes				
New folde	ler -				¥• 1	1 0
Fevorites	Documents library				Amongo by Fold	
E RecentFlaces	ON IFEG image	GED PIES image 44.3 KB	Q385.jeg PEG image 43.3 KB	Q403,jeg IPEG image 41.6 ktl		
Documents	Q411(P) STEE-mape	Q411.jpg JFE5 image 44.9 KB	Q471.jpg PES image db.4 Kil	Q535,jpg IPEG image 43.3 12		
 Public Documents Music 	(C) (542/09 (FED intrage 52.0 XII	CONTRACT CONTRACT				
Videos						
Computer						
Network						
Eile na	ume (1X3.jpg				All Files (".")	
					Open 💌 Car	escel

Figure 6: Query image selection

5.3 Query image selection and feature extraction

As exposed in *Figure 6 and* 7, the user picks the query image from the test image dataset. The Query image attributes are obtained using GLCM Matrix which is used in dataset feature extraction.

5.4 Similarity matching technique

This is certainly solitary among the most important concerns in CBIR. It has to be chosen to determine by what means similar one vector is to some other. Finding effective correlation metrics across images depending upon characteristic sets is a difficult problem. This was a technique for evaluating the difference/correlation among database and query images [14]. The distance equations vector can be used to calculate similarity.

Euclidean distance: As shown in Figure 8 and 9, as shown in Equation (2), the correlation-based result is obtained by calculating the Euclidean distance amid the image in the database and the query image.

$$d(q, \boldsymbol{r}_{i}) = \sqrt{\sum_{j=1}^{n} (f_{j}(q) - f_{j}(\boldsymbol{r}_{i}))^{2}}$$
(2)

- q represents an image of query,
- *r_i* represents an image of reference *i* from the database of reference,
- f_i represents a j^{th} feature,
- *n* represents feature spaces' dimensionality.



Figure 7: Query feature extraction

Can and a state	• Res 0									(2)8	4 9 -		C inch	ter		
- 10.04			•													
+ 10 23 +	+ C + Uses + 78	entrati + Decum	em + 1827	All + Inscalle												
uner folde:				2 http://www.comedia.com												
C Name				withing it detends it detended it												
A stateling	e0.			III 16M ort												
Address at							1.014						14			
Thread Tetu	Follow.m.			1.0.00	1.000	10.000	81100	2.00						-44	11	
A second				1 0.201	1.000	4100	0.100	0.0100								
* perdana				1 0.000	2.4454	0.000	5.046	1 1446								
is detuning				2 Pa. 10		2.004										
				8 0.000	0.000	0.0704	1.040	0.000								
III otractesime	e.n.et			2 0100	mager.	1000	2000	a setter								
Contraction of the			_	1 acres	1000	1000	0.0401	0.000								
in which as not			1		0.000	41100		0.007								
in intellety me				1 1 1 1 1	0.0011	0.0077	0.0716	0.132								
ill similarity1.ma	£			IN RECT.	0.0005	0.001	0.056	0.5408								
its (brided war				in Arran	0.0040	0.000	in some	0.1078								
served and the first			~	and an other	0.0046	01007	1000	0.4947								
fulgers .				11.0.104	0.4750	0.0411	0.000	10.540								
and a	linke	Mar Mar		14 2 200	8.6272	0.635	1.1.1.4	8 342								
America	14hJ money			15 0.367	8.0702	40.00	1179	0.7079								
dimenoid	245d out			10.01128	Antes .	0.0407	0.0277	0.752								
0Gorethie	343-2 tota			17 8 57 5	0.0071	100	1.084	0.7070								
dittanoit 2023aaa					10000	1000	Committee of the local division of the local	-								
strugglit	and page of		-	_				_	_					_		
edinugel] addrug	2426 table 3426 table		General Testin Second Testing and " Solid Protocology and " Solid Protocology and "													

Figure 8: Distance calculation



Research Article | Volume 10, Issue 2 | Pages 93-99 | e-ISSN: 2347-470X

Image: Control of the state of the	P
Bartistic District Only District Onl	
Description Control (Control (Contro) (Control (Control (Contro) (Control (Control (Contr	
Image: Second	
Image: Solution of	
Card 1980 O Participation Paricipation Participation	0
Image:	() ×
Bit Mark Control Contro Control Control <t< td=""><td></td></t<>	
Internation	
Instantion Instantion Instantion 1 Statisticity Links Instantion 5 Statisticity Links Instantion 5 Statisticity Links Instantion 1 Statisticity Links Instantion 1 Statisticity Links Instantion 1 Statisticity Links Instantisticity Links Links Links	
Image: Section of the sectio	
Internet 2 Sch41110.pg 6635 Internet 1 Sch41110.pg 6635 Internet 1 Sch41110.pg 6635 Internet 5 Sch41110.pg 6635 Internet 5 Sch41110.pg 6635 Internet 5 Sch41110.pg 6635 Internet 5 Sch41110.pg 6636 Internet 5 Sch41110.pg 6666 Internet 5 Sch41110.pg	1
Queue p StatAffred, py 8005 StatAffred, py 8005 StatAffred, py 8006 StatAffred, py 8006 StatAffred, py 8006	
information information information information information infor	
Bit Marking American In 5 Statil 2005 pyr 6.000 Bit Marking American 5 Statil 2005 pyr 6.000 Bit Marking American 7 Statil 2005 pyr 6.000 Bit Marking American 8 Statil 2005 pyr 6.000 Bit Marking American 9 Statil 2005 pyr 6.000	
and state and	
Bandwind P DatATION DataG generation 1 DataTion Syst 0.000 generation 1 DataTion Syst 0.000 generation 1 DataTion Syst 0.000 means and 10 ⁻¹ O 1 DataTion Syst 0.000 means and 10 ⁻¹ O 1 DataTion Syst 0.000	
Ø monol Juliu Ø monononononono Juliu <td)< td=""><td></td></td)<>	
Image: Section galax 10 DatAssT1465.ppg 0.0000 finalma.mmt (VAT-fin) 11 DatAssT1464.ppg 0.0000	
finalmes.mat (UAT-Sie) 11 DATASETVASA.pg 0.0961	
@ 12 DaTaSFT/92 m/ 0.1007	
NUMBER OF A DATACTUM IN DITE	
Name - Value Min Max IS United Housing County -	
distance 140/2 toble 14 united realing 4000	
di distaveli 1406 eri	
In drawership Idazimine In university (100 gr UL24)	
17 UNIASE1552.99 ULDU	
a demogrative and a model of the second seco	-
entimated by Statis	۲
<pre>>> load('dist.mat')</pre>	
sum 142d double 1.078233794 (> 1000('TIMBIPS.mst')	

Figure 9: Image sorting based on distance

HOME	RLOTS	4095	VARIABLE		VEW							10 L S	2 6 0	Search Dec	umentation	P
🔮 🛃 Open	· Roves	Columns	# #	1 12 1	anspose											
New from A Print Selection +	- 1	1	Insert Dele	* <u>1</u> S	it •											
VARIABLE																
💠 🔶 💽 🔁 🍶 🖡	C: Huses	Thermodii >	Documents	 MATL 	AB > ImageRetri	A/Hiskory										•
Current Folder				۲	🕑 Editor - quer					1	/ariables - selch	ng				•
D Name					sum 🗏 distance248 🗶 estimage248 🗶 sekkung 🗶											
1 queylm.m					142-7 1444											
dist.mat					1	2	2	4	5	6	2		0	10	11	12
distsummet					images	MC	DiffVar	DiffEntropy	Contrast	Entropy	Variance			20		
Emiliar and	ng			- 10	1 DATASED	-0.5779	0.2470	0.5342	0.2905	2.2205	6.8306					
Ormat					2 DATASET	-0.5882	0.2449	0.5255	0.2772	2,2034	6,9659					
selfeature.mat				- 1	3 DATASET	-0.6634	0.1925	0.3698	0.2068	1.6537	5.6222					
🔢 similarity.mat					4 DATASED	-0.6689	0.3903	0.4037	0.1977	1.8682	6.5677					
🖶 similarity1.mat					5 DATASET	-0.6805	0.1738	0.4123	0.1920	1,9885	7,4349					
Untitled.mat					6 DATASED	-0.7994	0.3022	0.3474	0.1139	3,2768	35,0808					
and dotance.alox					7 DATASET	-0.8384	0.0833	0,2910	0.0889	3.3490	28.3750					
El distancellator					8 DATASET	-0.7464	0.3536	0.4535	0.1887	3.3390	21.8720					
i etimages243.ds	DK .				9 DATASED	-0.7603	0.1583	0,4301	0.1790	3,2000	20,4850					
🔄 sekchungulisk				*	10 DATASED	-0.7625	0.1536	0.4242	0.1738	3,1731	21,2593					
selfeature.mat (NAT-9	le)			^	11 DATASED	-0.7907	0.3230	0.3749	0.1370	3.0161	25.1922					
Washington					12 DATASED	-0.7371	0.1727	0.4711	0.1990	3,3200	20,8483					
					13 DATASED	-0.7064	0.1710	0.3742	0.1852	2,0832	7,6337					
Name +	Value	Min	Max		14 DATASED	-0.7289	0.1743	0.3534	0.1855	2,0536	7,3777					
🔚 selclung	142x7 toble	1			15 DATASED	-0.7343	0.1703	0.3508	0.1822	2,0912	7.7288					
					16 DATASED	-0.7125	0.1906	0.3731	0.1549	2,0702	7.3251					
					17 DATASED	-0.7424	0,3400	0.3405	0.1503	21154	7,4843					
					Command Wine >> load(>> load(ft >>	kw 'extractedim 'selfeature.	age.mat') mat')									

Figure 10: Visual analytics approach

5.4 Visual analytics approach

As shown in *Figure 10 and 11*, this technique permits the user to add attributes that are deemed important [15]. The person is unrequired to be aware of the algorithm utilized for retrieval. This allows persons to discover visuals by investigating the complete dataset by identifying the certain visuals overlooked by the correlation matching method.



Figure 11: Relevant images to the query image

5.6 Performance measures

CBIR models are evaluated based on effectiveness and efficiency. Efficiency is determined by the rate of restoration and efficacy employing the increased significance of the restoration. Frequently employed CBIR As shown in Figure 12, performance measures are obtained by precision and recall [17] by Equations (3) & (4). Precision is described as the proportion of images retrieved that are most relevant to the overall number of database images [16]. The recall is calculated by dividing the count of relevant images retrieved by the overall numeral of appropriate database images.

$$precision = \frac{Number of relevant images retrieved}{Number of images retrived}$$
(3)

$$\operatorname{Re} call = \frac{Number \ of \ relevant \ images \ retrieved}{Number \ of \ relevant \ images \ in \ the \ dataset} \tag{4}$$

Z-Score Normalization

As shown in *Figure 13*, Z-Score normalization means rescaling the feature X by deducting the average μ and dividing via its standard deviation σ as given in *Equations (5)*,

$$z - score = \frac{(X - \mu)}{\sigma}$$
(5)

Total the free view free view for the form of the second of the

2 9 0 2 2 0 0 a



Figure 12: Z-Score Normalization



Figure 13: Precision and Recall values



Research Article | Volume 10, Issue 2| Pages 93-99 | e-ISSN: 2347-470X

6. CONCLUSION & FUTURE ENHANCEMENT

In this work, a novel framework is depicted for medical CBIR known as Visual Analytics for Medical Image Retrieval (VAMIR). This approach offers the opportunity to solve the above complex problems by combining human and automated analysis. VAMIR's capabilities reduce semantic gaps by allowing users who may not search algorithm experts to use the search process to explore the attribute space and discover images considered to be similar. Users could indeed find images that could have been missed otherwise significantly lowered rankings through techniques that employ the ranked listing by reviewing the complete dataset.

The visual quantitative method is better suited for multimodality imaging and volumetric. This framework was only applicable to one clinical group. Our future research will focus on under what circumstances VAMIR could be utilized in various clinical groups of a user by incorporating domainspecific interactions and visualizations.

REFERENCES

[1] Ashnil Kumar, Member, IEEE, Falk Nette, Karsten Klein, Michael Fulham "A Visual Analytics Approach using the Exploration of Multi-Dimensional Feature Spaces for Content-based Medical Image Retrieval," IEEE Journal of Biomedical and Health Informatics, vol. 8, no. 6, pp. 2168-2194, 2015.

[2] WeidongCai, (David) Dagan Feng, Senior Member, "Content-based retrieval of dynamic PET functional images."in IEEE TRANSACTIONS on information technology, vol. 4, no. 2, June 2000.

[3] J. G. Dy, C. E. Brodley, A. Kak, L. S. Broderick, and A. M. Aisen, "Unsupervised feature selection applied to content-based retrieval of lung images," IEEE T Pattern Anal, vol. 25, no. 3, pp. 373 – 378, 2003.

[4] M. M. Rahman, P. Bhattacharya, and B. C. Desai, "A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback," IEEE T InfTechnol B, vol. 11, no. 1, pp. 58–69, 2007.

[5] G. Quellec, M. Lamard, L. Bekri, G. Cazuguel, C. Roux, and B. Cochener, "Medical case retrieval from a committee of decision trees," IEEE T InfTechnol B, vol. 14, no. 5, pp. 1227–1235, 2010.

[6] S. F. da Silva, M. X. Ribeiro, J. d. E. S. Batista Neto, C. Traina-Jr, and A. J. M. Traina, "Improving the ranking quality of medical image retrieval using a genetic feature selection method," Decis Support Syst, vol. 51, no. 4, pp. 810–820, 2011.

[7] Yu Cao, Shawn Steffey, Jianbiao He, Degui Xiao, Cui Tao, Ping Chen and Henning Muller, "Medical Image Retrieval: A Multimodal Approach," pp. 125-136, April 07, 2015.

[8] B.Jyothi, Y.MadhaveeLathaP.G.Krishna,"An Effective Multiple Visual Features for Content Based Medical Image Retrieval," IEEE Conference,, pp. 1227–1235, 2015.

[9] Jyothi, B., Y. MadhaveeLatha, and PG Krishna Mohan. "Multidimensional feature space for an effective content based medical image retrieval." Advance Computing Conference (IACC), 2015 IEEE International. IEEE, 2015.

[10] Satish, B., and K. P. Supreethi. "Content based medical image retrieval using relevance feedback Bayesian network." Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2017 International Conference on. IEEE, 2017.

[11] Dubey, Shiv Ram, Satish Kumar Singh, and Rajat Kumar Singh. "Local wavelet pattern: A new feature descriptor for image retrieval in medical CT databases." IEEE Transactions on Image Processing 24.12 (2015): 5892-5903.

[12] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases," Comput Vis Image Und, vol. 75, no. 1-2, pp. 111–132, 1999.

[13] W. Jiang, G. Er, Q. Dai, and J. Gu, "Similarity-based online feature selection in content-based image retrieval," IEEE T Image Process, vol. 15, no. 3, pp. 702–712, 2006.

[14] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges," ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 2, no. 1, pp. 1–19, 2006.

[15] T. Deserno, M. Guld, B. Plodowski, K. Spitzer, B. Wein, H. Schubert, H. Ney, and T. Seidl, "Extended query refinement for medical image retrieval," J Digit Imaging, vol. 21, no. 3, pp. 280–289, 2008.

[16] X. Xu, D.-J. Lee, S. Antani, and L. Long, "A spine x-ray image retrieval system using partial shape matching," IEEE T InfTechnol B, vol. 12, no. 1, pp. 100–108, . 2008.

[17] Abinaya, S., & Devi, M. K. (2021). Enhancing top-N recommendation using stacked autoencoder in context-aware recommender system. Neural Processing Letters, 53(3), 1865-1888.

[18] Smita S. Gajare, Dipti Y. Sakhare (2016), Analysis of Intra-body Communication Measurement in Biomedical Applications. IJEER 4(2), 53-56. DOI: 10.37391/IJEER.040203.

[19] Avadhesh Kumar Dixit, Rakesh Kumar Yadav and Ramapati Mishra (2021), Contrast Enhancement of Colour Images by Optimized Fuzzy Intensification. IJEER 9(4), 143-149. DOI: 10.37391/IJEER.090408.



© 2022 by S.Abinaya and T.Rajasenbagam. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY)

license (http://creativecommons.org/licenses/by/4.0/).