# Journal of Digital Imaging

# A Mammography Database and View System for African-American Patients

Shani Ross,<sup>1,2</sup> O'tega Ejofodomi,<sup>2</sup> Ahmed Jendoubi,<sup>2</sup> Lisa Kinnard,<sup>3</sup> Mohamed Chouika,<sup>2</sup> Ben Lo,<sup>3</sup> Paul Wang,<sup>4</sup> and Jianchao Zeng<sup>2</sup>

We have digitized mammography films of African-American patients treated in the Howard University Hospital Radiology Department and have developed a database using these images. Two hundred and sixty cases totaling more than 5,000 images have been scanned with a high resolution Kodak LS85 laser scanner. The database system and web-based search engine were developed using MySQL and PHP. The database has been evaluated by medical professionals, and the experimental results obtained so far are promising with high image quality and fast access time. We have also developed an image viewing system, D-Viewer, to display these digitized mammograms. This viewer is coded in Microsoft Visual C# and is intended to help medical professionals view and retrieve large data sets in near real time. Finally, we are currently developing an image content-based retrieval function for the database system to provide improved search capability for the medical professionals.

KEY WORDS: African-American women, breast cancer, database, mammography, mammogram, digitization, image viewer, Howard University

## INTRODUCTION

T he American Cancer Society estimates that in the United States alone, approximately 40,000 women die each year from breast cancer and that over 200,000 new breast cancer cases are diagnosed each year.<sup>1,2</sup> Typically, there are four main types of breast cancer: ductal carcinoma in situ (DCIS) where the cancer is confined within the ducts of the breast, lobular carcinoma in situ (LCIS) where the cancer is confined within the lobules or glands of the breast, invasive ductal carcinoma (IDC), and invasive lobular carcinoma (ILC).<sup>1,3</sup> IDC and ILC refer to the type of breast cancer where the tumor has spread from the ducts or lobules it originated from, respectively, into the surrounding tissue of the breast. Other less common breast cancers include medullary carcinoma, mucinous carcinoma, Paget's disease of the nipple, Phyllodes Tumor, and tubular carcinoma.<sup>3</sup>

Breast cancer is grouped into stages which indicate the invasiveness of the disease. There are four stages—I, II, III, IV—defined by the American Joint Committee on Cancer<sup>1,3</sup> based on a combination of tumor size, lymph node involvement, and presence or absence of distant metastasis. There is also a more general classification: early/local stage where tumor is confined to the breast, late/regional stage where cancer has spread to the surrounding tissue or nearby lymph nodes, and advanced/distant stage where cancer has spread to other organs beside the breast.<sup>1,4</sup>

There has been a decline in breast cancer mortality rates of about 2.3% over the last decade<sup>1,2</sup> due to improved screening techniques leading to earlier detection, increased awareness,

Online publication 29 March 2007 doi: 10.1007/s10278-007-9019-6

<sup>&</sup>lt;sup>1</sup>From the Biomedical Engineering, University of Michigan, Ann Arbor, MI, USA.

<sup>&</sup>lt;sup>2</sup>From the Electrical and Computer Engineering, Howard University, Washington, DC, USA.

<sup>&</sup>lt;sup>3</sup>From the Radiology, Georgetown University, Washington, DC, USA.

<sup>&</sup>lt;sup>4</sup>From the Radiology, Howard University Hospital, Washington, DC, USA.

Correspondence to: Mohamed Chouika, 2300 Sixth St., NW, Rm 1022, Washington DC 20059; tel: 202-806-4816; fax: 202-332-3152; e-mail: mchouikha@howard.edu

Copyright  $\bigcirc$  2007 by Society for Imaging Informatics in Medicine

and improved treatments. Improved screening techniques include regular self breast examination, breast physical examination by a doctor, and regular mammogram examinations. In a mammogram, a radiologist looks for any unusual signs. These include asymmetry between the two breasts, any irregular areas of increased density, or any areas of skin thickening all of which can indicate whether there is a mass, or lump, in the breast tissue.<sup>1,3</sup> As this mass can either be cancerous or benign, further tests such as a biopsy of the suspicious area are done to differentiate between the two. In addition, doctors look for calcifications which are tiny calcium deposits that indicate changes within the breast possibly pointing to cancer. Microcalcifications especially are usually associated with cancer.

Despite improved screening techniques, this decline has been much smaller in African-American women compared to Caucasian-American women even with increased participation by African-Americans in routine mammography screening since 1980.<sup>5</sup> Furthermore, statistics have shown that although Caucasian-American women have an overall higher incident rate of breast cancer than African-American women, African-American women have a slightly higher incident rate of breast cancer before the age of 35 and are more likely to die of breast cancer at any age.<sup>1,6</sup> Also, in comparison to Caucasian women, African-American women are less likely to be diagnosed with tumors of diameter less than or equal to 2.0 cm and in the early stage of breast cancer, but more likely to be diagnosed with tumors of diameters greater than 2.0 cm and in later and advanced stages of the cancer.<sup>1,6,7</sup> Some of this disparity could be explained by an overall lower socioeconomic status, unequal access to medical care, and additional illnesses,<sup>8-10</sup> which results in a much smaller percentage of African-American women getting early and regular mammograms compared to Caucasian women. Hence, they are usually not diagnosed with cancer until at a much later date when the cancer is at an advanced stage, which increases their mortality risk.

Some studies have also reported differences in breast cancer histology in African-American women and Caucasian women, although they noted that these results are not conclusive. Some of these studies found that African-American women had a greater chance of developing ductal tumors, whereas Caucasian women had a greater chance of developing more lobular tumors.<sup>11,12</sup> This could be significant in the difference in outlook for the two racial groups, as lobular carcinoma is usually less aggressive. Other studies reported that African-American women had a greater frequency of medullar carcinoma than Caucasian women.<sup>5,7</sup> Medullary carcinoma is similar to ductal carcinoma but has a clearer distinction between the cancerous and normal cells.

Perhaps the most statistically significant differences seen between the two racial groups are the steroid receptor status of the tumor and the tumor proliferation rates. These differences are evident even when comparisons were adjusted for age, stage, treatment and screening opportunities, and risk factors. Typically African-American women had more estrogen- and progesterone-receptor negative tumors. Such tumors are generally unaffected by changing levels of these hormones and are thus less responsive to antihormonal treatments. African-American women also tended to have more poorly defined tumors, increased cell growth, and marked tumor necrosis. All these factors contribute to more aggressive tumors for African-American women.<sup>5,7,11,12</sup>

The epidemiological and tumor biology differences seen in African-American women compared to Caucasian women, leading to higher mortality rates, stress the particular importance in trying to fully understand why these disparities arise. One possible step towards addressing these issues is to have a larger number of African-American breast cancer cases available for study. Currently, there are no breast cancer databases primarily developed for African-American patients, although general breast cancer database systems do exist.<sup>13,14</sup> Hence, the main objective of this research is to develop the first major African-American Breast Cancer Mammography database for the use of training radiologists and other personnel involved in the detection and treatment of breast cancer. This could thus help in further understanding the nature of the breast cancer in African-American women. Part of this study also includes developing a view system that will allow radiologists to retrieve and view the digital imaging and communications (DICOM)-formatted images in the database in near real time.

This research is divided into three main parts. The first part is the collection of actual data for the database, which includes the obtaining, scanning, and organization of the mammography films. The second part is the development of the database and web-search engine using database management software and webpage development programming languages. The third part is the development of the viewing software.

#### METHODS

This study is in collaboration between the Howard University Electrical and Computer Engineering Department, the Howard University Hospital Radiology Department, and the Georgetown University Radiology Department. The database and viewing software development are being done through the Howard University Electrical Engineering Department with the films being supplied by the Howard University Hospital Radiology Department. Consulting support is provided by the Georgetown University Radiology Department. Before conducting the study, approval was obtained from the Institutional Review Board. The primary investigators, who worked directly with the films, took the Human Participants Protection Education for Research Teams online certification course<sup>15</sup> to meet the National Institute of Health (NIH) requirements on patient privacy issues. The investigators also attended seminars required by Howard University Hospital on the Health Insurance Portability and Accountability Act (HIPAA) of 1996.

#### Digitization of the Mammography Films

#### Film Collection

Currently, mammography films from 260 patients from the Howard University Radiology Department have been digitized and entered into the database. All of these patients were diagnosed with breast cancer between 1994 and 2004 and were between the ages of 24 and 88. The patients' histology includes masses of different sizes with and without poorly defined borders, microcalcifications, clusters of calcifications, adenopathy, heterogeneous density, and solid nodules. Furthermore, the majority of these patients had undergone more than one mammography examination. For

these patients, the earlier exams typically were regular screening mammograms, which showed normal and benign images. Normal images were mammograms in which nothing suspicious was seen, and benign images were mammograms in which something suspicious was seen initially but with further examination found to be nonmalignant. The cancer appeared and was diagnosed in the later exam dates. Thus, we were able to obtain a wide variety of films showing normal, benign, and cancerous images. Table 1 gives an overview of the patient characteristics of the database.

Each examination date, on average, contained four films corresponding to the four typical views taken of the breasts during a mammogram. These views are: left mediolateral oblique view (LMLO), right mediolateral oblique view (RMLO), left craniocaudal view (LCC), and right craniocaudal view (RCC) (Fig. 1). The films generally came in two sizes:  $8 \times 10$  and  $10 \times 12$  in.

Majority of these mammograms had been done at Howard University Hospital using the MGX 2000 mammography machine from Instrumentarium Corporation, Imaging Division, Tuusula, Finland. The machine parameters include: a film focus distance of 26 in., AGFA Mamoray HDR-C high dynamic range film 24 cm  $\times$  32 cm, exposure of 75–130 mAs (depending on patient), tube voltage of 27–30 KeV, and a filter of 0.025 mm molybdenum, or 0.025 mm rhodium, or 0.5 mm aluminum. Some of the older mammograms had been done at the patient's previous hospital and the technical parameters of the machine used are unknown.

## Scanning and Technical Procedure

The scanner used is the Kodak LS85 Laser Digitizer, specifically designed for high resolution teleradiology. Films can be converted into high resolution digital images of up to 5,120 pixels by 6,140 lines with 12 bits per pixel. Each pixel is assigned an optical digital value equal to  $1,000\times$ optical density and has a spot size of 50 µm. The digitizer has a scan rate of approximately 75 lines per second. It was purchased along with accompanying support software from RADinfo Systems, Dulles, Virginia. This accompanying software supports the digitization and burning of the images and includes: the Inter2000 Film Digitizer TM

Patient	Age	Year Cancer Diagnosed	Histology	Digitization Date
MK833AI543	24	1999	Spiculated mass and speculated density in R. breast	6/10/2004
WM980RI062	26	2001	Abnormal, dilated ducts and microcalcifications	6/14/2004
BR815AA561	32	1999	Cluster of calcifications	1/22/2004
TN989IA352	39	2001	Malignant calcifications in L. breast, extensive ductal carcinoma in situ	6/2/2004
HS963AH074	45	1999	Large medial calcifications in L. breast	4/6/2004
BM8690I885	46	1998	Microcalcifications in L. Breast	6/19/2004
ME7290M197	55	2001	Malignant calcifications and solid mass	5/19/2004
AJ337DU655	56	1997	Solid mass with poorly defined borders in L. breast	7/7/2004, 7/8/2004
BJ567RA821	58	2001	Cluster of calcifications in L. Breast	1/30/2004
GI759HZ353	59	1998	Microcalcifications and lateral calcifications in L. breast	6/21/2004
DM795AA866	62	1999	Spiculated mass in R breast	3/25/2004
BM169AE858	62	1996	Density with obscured borders in L. breast	6/20/2004
MA935CN074	63	1999	Metastative adenopathy in L axillary	6/10/2004
BD759A0805	65	2000	Breasts heterogeneously dense, highly suspicious mass with speculated borders and associated microcalcifications	2/24/2004, 3/1/2004
SS913TY027	66	2000	R. breast ductal carcinoma in situ	6/1/2004
DM710AA630	70	1998	Large mass and microcalcifications in R. breast	6/22/2004
GN21600404	70	1997	Macro and micro calcifications in R breast	6/22/2004
KC717EL219	71	2000	R. breast carcinoma—irregular nodal density; Previous L. breast carcinoma (now irradiated)—spiculated mass with microcalcifications at the peripheral	5/28/2004
CG362AE888	71	1998	Large pulp mass solid with lobulated borders, small satellite lesions and large abnormal axillary node in L. breast	6/21/2004
BL316YO065	72	1999	Cluster of calcifications	1/23/2004
DS833AA858	73	1997	Mass and poorly defined post border in R. breast	6/21/2004
BM065RA569	74	1996	Malignant appearing calcifications in R. breast	6/19/2004
FD103LE602	76	1998	Irregular asymmetric density with associated microcalcifications in L breast	7/15/2004
JF393AR669	77	2001	Two spiculated masses, poorly defined	5/28/2004
MD782IE169	82	1998	Spiculated mass in L breast	6/24/2004
HK862AA996	83	2000	Solid and irregular polyp mass	5/17/2004

Table 1. Representative Sample of the Patient Characteristics of the Database

(I2000FD), RSVS Viewer  $^{\text{TM}},$  and CD Power-PACS  $^{\text{TM}}.$ 

The I2000FD<sup>TM</sup> is the interface for the film digitizer. This is where patient information is

entered, the digitizer is accessed for scanning, and where any necessary changes can be made to the film before the digital image is saved. Before scanning, patient information required by



Fig 1. A typical mammography view: a LMLO, b RMLO, c LCC, and d RCC.  $^{\rm 13}$ 

I2000FD<sup>™</sup> is entered. This information includes: the patient name written in a code to preserve patient's privacy rights, birth date, exam date, the ascension number used to keep track of each series, and the time and date of digitization of each film. This information is the essential information required by the program to stamp and classify each resulting digitized image. Next, the user specifies the resolution, bit depth, window level settings, and size of the film to be scanned. In this study, films are digitized at a resolution of 512 pixels/in. and a bit depth of 12, both of which are maximum values of the scanner, to give the highest image quality. The scan size chosen is either  $8 \times 10$  or  $10 \times 12$  in., depending on the size of the film to be scanned. Once the settings have been entered, the film is loaded into the Kodak scanner and digitized. Furthermore, the software allows us to cover the patient's name and other identifying information so that this is not seen in the digitized image in the software. The images are then saved and copies of them are transferred to the viewing software (RSVS TM) and burning software (CD PowerPACS TM). The sizes of these saved images are generally 40 MB for the  $8 \times 10$  films and 60 MB for the  $10 \times 12$  films. The RSVS<sup>TM</sup> is a viewing software for DICOM-formatted images, which allow us to further view and verify the digitized images. The CD PowerPACS<sup>™</sup> software allows us to save and transfer the images to CD-R storage media. This program also saves a copy of the RSVS viewer on each CD.

#### Development of the Database

The Howard University Mammography Database<sup>16</sup> was developed using PHP and MySQL. MySQL is a simple relational database management system that allows users to add, access, and process data stored in a database.<sup>17</sup> The database contains relevant information such as encrypted patient name, patient ID, date of birth, type of scanner used, number of series in a case, exam date, total image count in a case, malignancy, and available image views (RMLO, LCC, etc.). New information can also be added when necessary. Our database is linked to a web-based search engine powered by PHP-recursive acronym for "PHP: Hypertext Preprocessor,<sup>18</sup> which is a general-purpose open source scripting language that can be embedded in hypertext markup

language (HTML) for web-page development and which can run directly on the server, unlike other languages such as java script. Oueries on the database are carried out using PHP. In addition, digitized images and patient information are stored in the database using PHP. Information in the database is first entered into Microsoft EXCEL, and then from there, is imported into the MySQL tables via PHP. This process makes entering new information into the database less tedious than directly entering it into MySQL from the Microsoft Windows command line. Furthermore, a login web page for the web-based search engine is also developed using PHP to limit use of this system to authorized users only. After being logged on to the system, a user can then specify search criteria via the web page. The query is then searched through the MySOL database, and the search results are displayed on the web page. The results are displayed as a list of case names with links to the actual images.

#### Development of the D-Viewer

The D-viewer is a viewing system developed for visualizing the digitized mammograms. By using this system, radiologists would not only be able to perform various image manipulations, such as circling the area of the cancer and making notes on the image, but could also save these markings as needed. The view system uses graphical user interface (GUI) for easy use and is to be DICOM compliant. The D-viewer system is developed with Microsoft Visual C#, an objectoriented programming language from Microsoft that is easy for manipulating graphics and creating window forms. Currently, the system accepts the following image formats: Windows bitmap (bmp), joint photographic experts group (jpg), tagged image file format (tif/tiff), graphics interchange format (gif), and portable network graphics (png).

#### RESULTS

Mammography films from 260 patients, each having approximately 20 to 40 images, have been digitized using the Kodak LS85 Laser Scanner. This has resulted in over 5,000 digitized images that have been stored into our database. Many patients had typically 20 to 40 images because they did several regular screening mammograms before the mammography exam in which the cancer was diagnosed. Hence, on average for each patient, one could be looking at 4-10 different exam dates each of which would contain the four films representing the four typical views (LMLO, RMLO, LCC, RCC). Sometimes, for a particular examination, extra images were taken, depending on if the radiologist was interested in a particular region. All these patients are African-American women from Howard University Hospital who have been diagnosed with cancer between the years 1994 and 2004. In addition, we can compare normal images and images of benign and malignant lesions of the same patient. The image quality is excellent, as the films were digitized at the highest resolution possible by the Digitizer, which was at 512 pixels per inch at a bit depth of 12. Furthermore, there were no observable distortions or artifacts due to the digitization process, although there is some concern that some noise may have been picked up due to possible dirt spots on the films. These images were also approved by the radiologist consultant associated with this study. The digitized image size is either 40 or 60 MB, depending on the size of the mammogram film (8  $\times$  10 or 10  $\times$  12 in., respectively).

A database and a web-based search engine have been developed using MySQL and PHP. Patient information and the corresponding images can be retrieved when queries are made by selecting various criteria. A user is first required to login to access the database (Fig. 2). User accounts will be set up for general users just wanting to learn more about breast cancer and to peruse the database and for users requesting more functionality from the database, such as being able to download the images.

After a user logs in, he or she is directed to the web-search page where he or she can select his or her search criteria (Fig. 3). Currently, the search criteria available include whether the image is normal, benign, or cancer, the year the patient was born, the year the exam was done, total number of examinations for a patient, total number of images for a patient, and the mammography view. For example, a user may want to see all the cancer images of patients who were diagnosed with cancer during 1999–2000 and were born during 1961–1970. There is also a reset button on this page that clears all the search fields. After the search criteria has been selected and submitted, the database returns a list of all patient cases corresponding to that search. The list of cases contains the code name associated with that patient, each series and corresponding exam date of that series, the digitization date for each series, the total number of images in each series, and all available image views done. Figure 4 shows a sample results page of a search.

Finally, Figure 5 shows a snapshot of the D-viewer system with the LCC and RCC views of one of our digitized mammograms. The system can accept bitmap, tiff, jpg, gif, and png image formats. The D-viewer is currently under development.

#### DISCUSSION

Our research attempts to provide the first major breast cancer database for African-American women. This would allow radiologists and other applicable parties to study various types of African-American women breast cancer cases. Howard University Hospital was chosen not only

1.0.1		1994 B	and the second	1	[II] /71	20 51
4= Back • ⇒ • 🕤	C C C Search	Avorite	s (Friedia	3	43° 38	12 · 12
Address 🕘 http://l-lat	imer.eng.howard.edu/bo	x4/recent_log	in/login.html			
msn <sup>N</sup> ? •	*	Search	- PHighligh	E A	Options	Pop-ups Blocked (60)
Welcome to	the hubc onli	ne	-			
Welcome to	the hubc onli	ne webuser				
Welcome to Usemame: Password:	the hubc onli	webuser	r			
Welcome to Usemame: Password: login	the hubc onli	webuser				

Fig 2. Login page of the web-based search engine linked to the breast cancer database.

File Edit View	Favorites Tools He	Þ							
⇔Back • ⇒ -	C 1 G Searc	h 🔂 Favorites 🖣	Moda 🎯 🔂 - 🖾	) 🖸 - 🖃					
Address 🗃 http://	l-latimer.eng.howard.edu/	bcx4/formp_e.php						· 260	o Li
msn <sup>N</sup> ! •		Search •	Highlight AOptions	Pop-ups Blocks	ed (60) 🔹 🖂 Hotmail 👙	Messenger	2 My MSN		
Welcome	to the hubc on	line							
							submit res	et [	
Volume Type: case name:	Cose002	🕫 Benign	I ⊂ Cancer	⊢ Any			outrint 100		
Density:	<b>D</b> 1	<b>C</b> 2	□3	Γ4	₩ Anv				
Scanner	DBA	Lumisve	□ Howtek	□ LS85	Any				
Pathology:	Malignant	□ Benign	E Benign Wit	hout Callback	Any Any				
Assessment	Π1	Γ2	Γ3	Γ4	E s	I⊽ Ar	W.		
Subtlety:	<b>□</b> 1	E 2	□ 3	□ 4	E 3	I Ar	y		
Lesion Type:	( Mass	C Calcificatio	n						
[	N	lass Options			c	alcification (	ptions		
[ ·		19 m		Т	vper(@and @or)		Distribution: ( @ a	nd ©orì	
Shape:(	Cand Cor)	Ma	rgin: (Cand @or)	F Punc	tate	E	Clustered		
T Round		Circums	cnbed	E Amo	phous	E	Linear		_
D Oval		I Microlo	Microlobulated		Pleomorphic		Segmental		
Lobulated		Dbscur	C Obscured		Round_And_Regular		Regional		_
🗖 Inegular	🗂 Irregular		ned	E Luce	Lucent Center		Diffusely Scattered		
Architectua	Architectural_Distortion		D Spiculated		Fine Linear Branching		Any		
Tubular	Tubular		Any Any		Skin Calcification				
Lymph Not	lo			E Vasc	ılar				
Asymmetri	c_Breast_Tissue			E Coar	:*				_
Focal_Asy	mmetric_Density	3. 2		E Large	Rodlike	-	-		
Any Any				E Roun	d				
				Egg S	ihell				_
							-		

## Fig 3. Web-based search page.

for its accessibility, but more importantly, because it services a large African-American community in the Washington DC, Maryland, and Virginia area, thus, providing a large number of African-American women mammography cases. This research is also significant in an effort to close the gap of diversity in breast cancer mortality.

However, this research is not just about digitizing a large number of mammograms. The database is also being developed as an online

		100	00.0	Access	dam			
nsn": •		-	Jo Search •	#Highaght	Aroptions	X Pop-ups Blocked (60)	• Hotmail	2 Messenge
Welcon	ne to the	hubc onl	ine					
Authenticate	đ							
Volume_No	Volume_type	case_name						
benign_07	benign	case1599						
benign 04	benign	case0280						
benign 06	benign	case3226						
benign 13	benign	case3460						
benign 01	benign	case3091						
cancer 02	cancer	case0022						
cancer 02	cancer	case0025						
cancer 01	cancer	case3025						
cancer 05	cancer	case0223						
cancer 07	capcer	case1233						
currin_o/		THEFT						
four search (	Criteria was:							
Volume type						benign cancer		
Density						any		
Patholom						any		
Assessment	8					any		
Subtlety						uny		
Mass Shepe						eny		
Mass Margin						any		

Fig 4. Results page of search.



Fig 5. D-viewer user interface.

interactive teaching/training tool geared towards both the radiology students and others interested in learning more about breast cancer. For example, the site would contain general information on breast cancer and more detailed information. Medical students could be asked to perform tasks such as determining, from the image alone, the search criteria required to produce that particular image. This would require the students to be able to accurately identify specific characteristics of that image. Radiologists could also use this database to help their diagnosis process of difficult new cases by searching for similar cases or images.

Our database differs from current major databases such as the one developed by the University of South Florida<sup>14</sup> in that our patient data set consists of African-American women only. Furthermore, we intend our database to be an interactive training tool where the users can look at images and guess the search criteria that they would have had to put in to obtain that image. We also plan to extend the functionality of the database by making it an image-based retrieval system in which features in the image can be used to define the search criteria.

#### CONCLUSIONS

We have digitized more than 5,000 mammography images of 260 African-American women patients and developed a database and viewing system for these images. All patients in this study have already been diagnosed with breast cancer. The digitized images include normal, benign, and malignant mammograms, which were taken between the years 1994 and 2004. The normal and benign mammograms of a patient were from routine screenings done before breast cancer was diagnosed. All the mammography images are searchable in the database system via a web-based search engine. They can be viewed in the viewing system and easily manipulated during the viewing process.

Further research will be performed to improve the database and the viewing system. First of all, we will seek input from radiologists when they test our database and viewing system. They will be asked to identify key information from the images such as the presence or absence of malignant cancer, location of cancer, type of cancer (mass or calcification), assessment, subtlety, and density of the breast. Locations of the cancers, if present, on the digitized mammograms, will be marked and saved. The database will also be updated with these added data.

New mammograms will also be collected and digitized to increase the size of our database. The D-viewer system will be further enhanced to improve its functionalities based upon input from the radiologists after testing the system. We will also develop an interactive web-based learning/ training/testing system, which can provide detailed information on breast cancer, training sessions for breast cancer diagnosis, and test platforms. To improve the quality and versatility of our database, we plan to develop an image-based retrieval database, whereby, we can use images instead of words to determine our search criteria. For example, if we circle a region of interest on an image such as a mass, we could query the database to return all images and corresponding information that have a similar region of interest.

This database will be made accessible to radiologists and students via announcements on the Howard Medical School website and on the Howard University Electrical and Computer Engineering Department website. No fees will be charged for basic database access, but fees will be charged for any advanced functions such as image processing or content-based retrieval operations.

#### ACKNOWLEDGEMENTS

We would like to thank the radiologists and other medical professionals in the Department of Radiology at Howard University Hospital who made significant contributions to this research. This research was supported by the US Army MRMC Breast Cancer Research Partnership Program (No. DAMD17-01-1-0267). Paul Wang was also supported by the following grants: NIH NCRR No. 2G12RR003048, NIH No. 5U54CA091431, US Army MRMC No. W81XWH-05-1-0291 and No. DAMA17-00-1-0291.

#### REFERENCES

1. American Cancer Society: Breast Cancer Facts and Figures 2005–2006. Atlanta, GA: American Cancer Society, Inc., 2005

2. American Cancer Society: Cancer Facts and Figures 2006. Atlanta, GA: American Cancer Society, Inc., 2006

3. Abeloff MD, Wolff AC, Wood WC, McCormick B, Weber BL: Cancer of the Breast. In: Clinical Oncology, Philadelphia, PA: Elsevier, 2004

4. Young Jr, JL, Roffers SD, Ries LAG, Fritz AG, Hurlbut A, Eds.: SEER Summary Staging Manual—2001: Codes and Coding Instructions, Bethesda, MD: National Cancer Institute, 2001 5. Rovak-Schaler R, Rose DP: Mammography screening and breast cancer biology in African American women—a review. Cancer Detec Prev 26:180–191, 2002

6. SEER Cancer Statistics Review, 1975-2002, National Cancer Institute. Available at http://seer.cancer.gov/csr/1975\_2002/. Accessed 31 May 2006

7. Elledge RM, Clark GM, Chamness GC, Osborne CK: Tumor biologic factors and breast cancer prognosis among white, hispanic, and black women in the United States. J Natl Cancer Inst 86:705–712, 1994

8. Bach PB, Schrag D, Brawley OW, Galaznik A, Yakren S, Begg C: Survival of blacks and whites after a cancer diagnosis. JAMA 287:2106–2113, 2002

9. Josyln SA: Racial differences in treatment and survival from early stage breast carcinoma. Cancer 95:1759–1766, 2002

10. Shavers VI, Brown ML: Racial and ethnic disparities in the receipt of cancer treatment. J Natl Cancer Inst 95:334–357, 2002

11. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, Dolan N, Paskett E, McTiernan A, Hubbell FA, Adams-Campbell LL, Prentice R: Ethnicity and breast cancer factors influencing difference in incidence and outcome. J Natl Cancer Inst 97:439–448, 2005

12. Porter PL, Lund MJ, Lin MG, Yuan X, Liff JM, Flagg EW, Coates RJ, Eley JW: Racial differences in the expression of cell cycle-regulatory proteins in breast carcinoma: study of young African American and white women in Atlanta, Georgia. Cancer 100:2533–2542

13. Digital Database for Screening Mammography. Available at http://figment.csee.usf.edu/Mammography/DDSM/ thumbnails/normals/normal\_01/case0130/A-0130-1.html. Accessed 31 May 2006

14. University of South Florida Digital Mammography Home Page. Available at http://figment.csee.usf.edu/Mammography/ Database.html. Accessed 31 May 2006

15. Human Participant Protections Education for Research Teams, National Cancer Institute. Available at http://cme.cancer. gov/clinicaltrials/learning/humanparticipant-protections.asp. Accessed 31 May 2006

16. Ross S, Ejofodomi O, Jendoubi A, Chouikha M, Lo B, Wang P, Zeng J: A mammography database and view system for the African American patients. In: Emerging Technologies and Applications for Imagery Pattern Recognition, Los Alamitos, CA: IEEE Computer Society, 2005

17. MySQL Reference Manual. Available at http://dev.mysql. com/doc/refman/5.0/en/. Accessed 31 May 2006

18. PHP: PHP Manual - Manual. Available at http:// www.phpbuilder.com/manual/. Accessed 31 May 2006