

**PHÂN VÙNG ẢNH X QUANG VÚ
BẰNG PHƯƠNG PHÁP CHÊNH LỆCH MỨC XÁM ĐỒ
SEGMENTING MAMMOGRAM BY HISTOGRAM DIFFERENCE METHOD**

*Nguyễn Thái Hà, Nguyễn Đức Thuận, Phạm Mạnh Hùng, Đào Trang Linh, Trần Thanh Minh
Trường Đại học Bách khoa Hà Nội*

TÓM TẮT

Chụp X quang vú là một phương pháp thường được lựa chọn để phát hiện sớm căn bệnh ung thư vú. Thực tế, ảnh X quang vú được chia thành 4 vùng cơ bản: vùng nền, vùng mô mỡ, vùng mô tuyến và cơ ngực. Các khối u hay những điểm bất thường ở vú chủ yếu xuất hiện ở vùng mô tuyến. Trong quá trình xử lý ảnh X quang vú thì bước phân vùng ảnh là một bước xử lý quan trọng. Bài báo giới thiệu thuật toán phân vùng ảnh X quang vú bằng phương pháp chênh lệch mức xám đồ. Thuật toán dựa trên lược đồ xám của ảnh để tìm ra đường biên giữa các phần khác nhau của ảnh X quang vú. Thuật toán được thực hiện bằng Matlab 7.04, sử dụng cơ sở dữ liệu MINI-MIAS (là bộ CSDL chuẩn về ảnh X quang vú) đã đưa ra được kết quả phân vùng chính xác.

ABSTRACT

Mammography is currently the method of choice for early detection of breast cancer. In fact, a mammogram is divided into four segments: background, fat tissue, fibroglandular tissue and muscle. It is postulated that most cancers or abnormal regions arise in fibroglandular tissue. In the processing of mammogram, segmentation plays a very important role. This article presents the algorithm to segment a mammography by the histogram difference method. The algorithm bases on histogram to find the boundaries between components in the image. It was implemented by Matlab 7.04, use the MINI-MIAS database that gives good results.

I. LỜI MỞ ĐẦU

Căn bệnh ung thư vú là một loại ung thư phổ biến nhất ở phụ nữ. Thời gian phát hiện càng sớm thì càng tăng khả năng chữa khỏi bệnh. Chụp X quang vú là một cách phát hiện bệnh hiệu quả nhất. Ảnh X quang vú ở nhiều tư thế khác nhau được dùng để phát hiện, dò tìm ra những khối u, vùng bất thường và có thể chẩn đoán được trường hợp này là u lành tính hay u ác tính [1].

Tuy nhiên, để đọc được ảnh X quang vú cần bác sỹ có trình độ chuyên môn cao và cả kinh nghiệm chẩn đoán. Phát hiện sớm hay muộn và độ chính xác phụ thuộc rất nhiều vào trình độ bác sỹ. Giờ đây, ngành công nghệ thông tin phát triển đã can thiệp vào quá trình tái tạo ảnh và chẩn đoán, bác sỹ nhanh chóng chẩn đoán được chính xác kết quả của bệnh nhân giúp quá trình điều sau này hiệu quả hơn.

Trong quá trình xử lý ảnh X quang vú thì bước phân vùng ảnh là một bước xử lý quan trọng. Thực tế, ảnh X quang vú được chia thành 4 vùng cơ bản: vùng nền, vùng mô mỡ, vùng mô tuyến và cơ ngực. Các khối u hay những

điểm bất thường ở vú chủ yếu xuất hiện ở vùng mô tuyến [1]. Vì vậy, phân vùng là một bước tiền xử lý để lấy ra mô tuyến giúp giảm tính toán ở các vùng không cần thiết cho các bước sau. Độ chính xác ở bước phân vùng ảnh hưởng trực tiếp đến kết quả ở các bước sau đó.

Phân vùng ảnh X quang vú có nhiều phương pháp như phân vùng dựa theo miền đồng nhất, phân vùng theo mạng neuron hay phương pháp phân vùng theo ngưỡng [2], [3]... Vấn đề đặt ra là tìm được phương pháp nào hiệu quả mà thuật toán đơn giản, phương pháp phân vùng theo ngưỡng là một phương pháp đơn giản tuy nhiên làm thế nào để tìm ra được ngưỡng chính xác của các vùng. Phương pháp chênh lệch mức xám đồ là một phương pháp không cần bộ dữ liệu đã được huấn luyện, thuật toán đơn giản nhưng cho một kết quả chính xác.

II. CƠ SỞ LÝ THUYẾT

Phương pháp phân vùng ảnh bằng chênh lệch mức xám đồ là phương pháp phát triển lên từ phương pháp phân vùng dựa trên ngưỡng biên độ. Người ta coi lược đồ xám của một bức

ảnh hoàn chỉnh hay một bức ảnh con là một vector. Với ảnh mức xám 8 bit thì lược đồ xám có 256 giá trị, vì vậy vector có 256 chiều. Để so sánh thuận tiện giữa 2 vector, người ta tính tích vector vô hướng của vector $v(v_1, v_2, \dots, v_n)$ và $u(u_1, u_2, \dots, u_n)$ như sau:

$$\vec{v} \bullet \vec{u} \equiv \sum_{i=1}^n (v_i * u_i)$$

Trong đó, góc θ là góc giữa hai véc tơ \vec{u} và vecto \vec{v} được xác định bằng:

$$\cos(\theta) \equiv \frac{\vec{v} \bullet \vec{u}}{|\vec{v}| * |\vec{u}|}$$

Với $|\vec{u}| = \sqrt{\sum_{i=1}^n (u_i)^2}$; $|\vec{v}| = \sqrt{\sum_{i=1}^n (v_i)^2}$

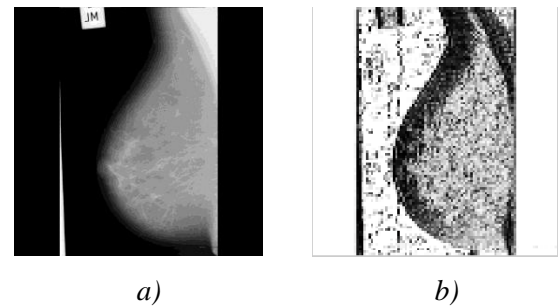
Để tính $\cos\theta$ của góc giữa 2 vector không khó, nó sẽ cung cấp một thông số về khoảng cách giữa hai vector. Nếu hai lược đồ xám này giống nhau thì $\cos\theta=1$, nếu trực giao thì $\cos\theta=0$. Từ đó ta sẽ tìm được cách phân loại các phân trên ảnh X quang vú dưới dạng chênh lệch giữa các cặp lược đồ xám.

Người ta cũng có thể sử dụng các phương pháp khác để so sánh các lược đồ xám ví dụ như tương quan thống kê. Một số các nhà nghiên cứu đã chứng minh được rằng các phân bố thống kê ví dụ như phân bố Poisson có thể được dùng để tính lược đồ xám [2,4,5]. Tuy nhiên, trong nghiên cứu này, ta chọn tích vô hướng vì chúng đơn giản và tính toán nhanh, ngoài ra luôn cho một giá trị từ 0 đến 1 trong tất cả trường hợp.

Đại lượng về độ chênh lệch giữa hai lược đồ xám của các ảnh con liên kế cho ta một phương thức rất tốt để phát hiện các thay đổi dần trong mật độ quang của ảnh X quang vú. Ranh giới giữa cơ ngực và nhu mô, giữa nhu mô và nền ảnh có thể bị nhiễu làm mờ đi. Nhưng so sánh giữa các lược đồ xám giúp cho ta một định vị chính xác ranh giới giữa cơ ngực, nhu mô, lớp mỡ dưới da và nền ảnh.

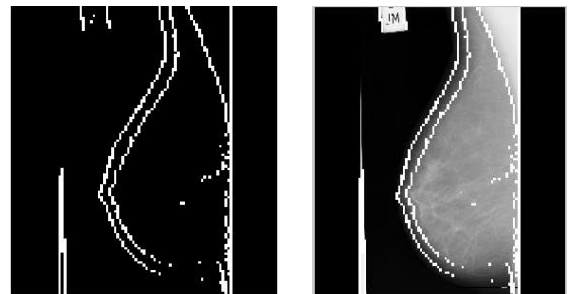
Để công việc này được thực hiện hiệu quả, ta xây dựng phương pháp lấy mẫu. Ảnh X quang ban đầu được chia thành các hàng và các cột, tạo thành các bức ảnh con. Các ảnh con này có thể chồng hoặc không chồng lên nhau. Kích thước ảnh con có thể thay đổi tùy thuộc vào độ

phân giải mong muốn của thuật toán. Để xác định chính xác ranh giới giữa các nhu mô, cơ ngực, lớp mỡ và nền ảnh, các ảnh con liên kế được sử dụng từ trái sang phải và từ trên xuống dưới. Sau quy trình lấy mẫu là tính tích vô hướng giữa các lược đồ xám của các cặp ảnh con. Tích vô hướng này được dùng để xây dựng một bức ảnh mới biểu thị các thay đổi về mật độ quang trên ảnh X quang ban đầu. Ảnh mới được giảm tỷ lệ xuống. Ví dụ, nếu ảnh ban đầu là 5000x5000 điểm ảnh và sử dụng các ảnh con kích thước 10x10 thì ảnh mới sẽ có kích thước là 499x499 điểm ảnh do cột cuối cùng không có ảnh con nào ở bên phải nó để so sánh.



Hình 1. a) Ảnh X quang ban đầu;
b) Kết quả chênh lệch lược đồ xám

Hình 2 minh họa tốt hơn mức độ hiệu quả của phương pháp chênh lệch mức xám độ. Nó biểu diễn kết quả của phương pháp này chỉ sử dụng lược đồ xám 4 giá trị thay vì lược đồ xám 256 giá trị như trên hình 1, hình 1 minh họa những thay đổi nhỏ trên ảnh đi từ trái sáng phải và trên xuống dưới, nhưng không đưa ra những thay đổi giữa cơ ngực, lớp mỡ và nền ảnh.

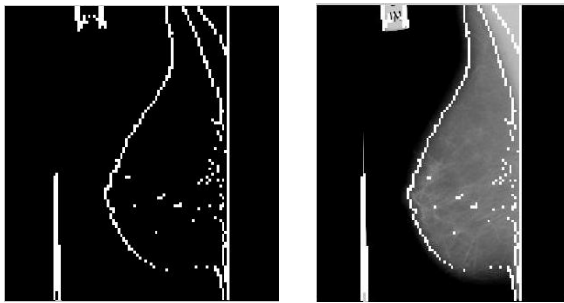


Hình 2. Kết quả chênh lệch lược đồ xám 4 giá trị và đưa chênh lệch lược đồ xám vào ảnh X quang vú

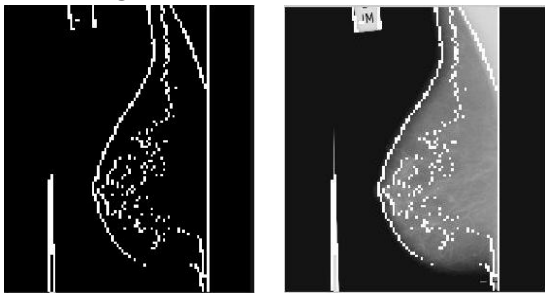
Trên ảnh có quá nhiều sự thay đổi không có ý nghĩa. Sử dụng lược đồ xám 4 giá trị có tác dụng làm nổi bật các thay đổi quan trọng nhất

trong cường độ giữa các ảnh con. Ranh giới phù hợp cho các thành phần trên ảnh được đánh dấu bằng đại lượng chênh lệch lược đồ xám.

Phương pháp chênh lệch lược đồ xám 4 giá trị có thể hoạt động tốt trong trường hợp này nhưng với các trường hợp khác, nó có thể không xác định chính xác ranh giới. Khi lược đồ xám được nén từ 256 giá trị xuống còn 4 giá trị, ranh giới mà sẽ được phát hiện, được đưa vào trong các khoảng cường độ sau: [0,63], [64, 127], [128, 191] và [192, 255]. Chúng hoạt động khá tốt trên ảnh này, đã phân ra được vùng cơ ngực, mô vú, mô mỡ và nền. Tuy nhiên không phải tất cả các ảnh X quang vú đều giống nhau.



Hình 3. Kết quả chênh lệch mức xám đồ và đưa chênh lệch mức xám đồ vào ảnh trên đã được giảm cường độ



Hình 4. Kết quả chênh lệch mức xám đồ và đưa chênh lệch mức xám đồ vào ảnh đã được tăng cường độ

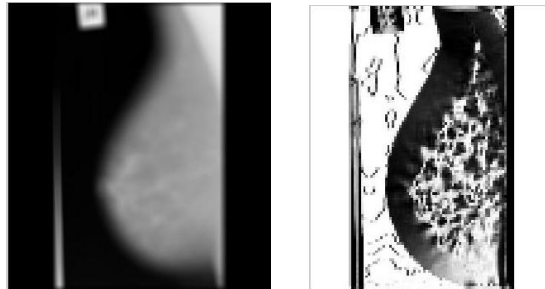
Ta thấy rằng kết quả chênh lệch lược đồ xám đối với ảnh giảm cường độ sáng chỉ cho ra ba vùng, vùng cơ ngực phân vùng không chính xác và có nhiễu, vùng mô mỡ không xác định được, ranh giới giữa ngực và nền ảnh xác định không chính xác. Với ảnh tăng cường độ cũng chỉ có thể phân vùng thành ba vùng, xuất hiện nhiều nhiễu ở vùng ngực. Trong ví dụ này, tất cả các ranh giới đều bị xác định sai và nhiều ở phần mô vú khiến không đưa ra được kết quả.

Như vậy, sử dụng lược đồ xám 4 giá trị không đủ để đưa ra được ranh giới giữa các thành phần ảnh. Bởi vậy, ta phát triển một phương pháp định vị các ranh giới này sử dụng phương pháp chênh lệch lược đồ xám 256 mức.

Tìm ranh giới giữa các thành phần

Kết quả của phương pháp chênh lệch lược đồ xám 256 giá trị tạo ra một ảnh mới trong đó cường độ điểm ảnh tương ứng với lượng cường độ thay đổi trong khu vực đó. Tuy nhiên, nhiều khu vực thay đổi thể hiện các đặc trưng tỷ lệ nhỏ trên ảnh không phải là các ranh giới quan trọng giữa các thành phần ảnh. Để đơn giản quy trình phân vùng, ta sử dụng bộ lọc trung bình 30x30.

Hình 5 biểu diễn ảnh X quang vú sau khi sử dụng bộ lọc trung bình 30x30 và kết quả chênh lệch mức xám đồ.

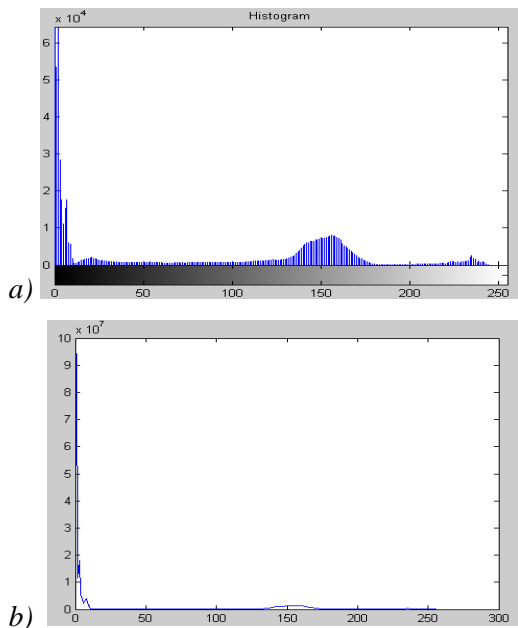


Hình 5. Ảnh X quang vú đã loại bỏ những tỷ lệ đặc trưng nhỏ và kết quả chênh lệch lược đồ xám đã được làm nhẵn

Hình 5 cho thấy các vùng thay đổi rõ ràng hơn trước và cũng phù hợp với các đường ranh giới quan trọng giữa các thành phần ảnh. Ta xác định các ranh giới giữa các thành phần bằng cách tạo ra một lược đồ xám trọng số của bức ảnh ban đầu với các trọng số là các giá trị chênh lệch lược đồ xám. Ta xây dựng lược đồ xám cho ảnh ban đầu nhưng thay vì thêm 1 mỗi khi gặp lại một điểm ảnh mang giá trị cường độ đó thì ra thêm vào giá trị chênh lệch lược đồ xám vào điểm ảnh đó trên ảnh. Tất cả giá trị này nằm trong khoảng từ 0 đến 1. Việc tạo ra lược đồ xám trọng số có ý nghĩa nhấn mạnh vào các cường độ xuất hiện trong các vùng thay đổi trên ảnh.

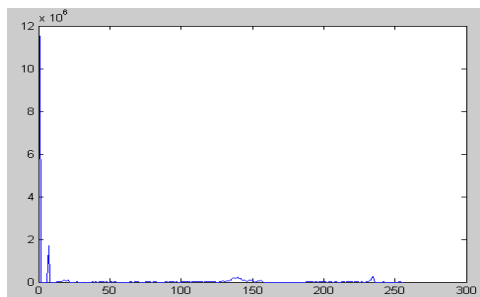
Sau đó, các giá trị cường độ trên lược đồ xám trọng số được chuẩn hóa để loại bỏ bất cứ độ lệch nào từ các vùng biên giới lớn hơn.

Sử dụng lược đồ xám trọng số trên hình 6b, các ranh giới mong muốn được xác định rõ ràng hơn.



Hình 6. a) Lược đồ xám ban đầu; b) Lược đồ xám trọng số

Hình 7 biểu diễn xấp xỉ đạo hàm bậc nhất đơn giản của lược đồ xám trọng số đã được lấy ngưỡng là 0. Cực đại cục bộ từ các vị trí này là các vị trí ranh giới giữa các thành phần.



Hình 7. Đạo hàm bậc nhất của lược đồ xám trọng số

Bằng các sử dụng lược đồ xám trọng số, ta đã định vị chính xác ranh giới cường độ phù hợp cho các thành phần ảnh và phân vùng ảnh ban đầu. Hơn nữa, ảnh đã được điều chỉnh tăng hoặc giảm cường độ sáng đã không được phân vùng chính xác bằng phương pháp chênh lệch lược đồ xám 4 giá trị, cũng đã được phân vùng đúng. Các bức ảnh này giống như ban đầu ngoại trừ việc các giá trị điểm ảnh đã thay đổi

cùng tăng hoặc giảm. Bởi vậy, lược đồ xám trọng số gần như giữ nguyên. Sự khác biệt duy nhất là chúng dịch chuyển sang bên trái hoặc phải tùy theo mức độ tăng hay giảm sáng.

Thêm bối cảnh không gian

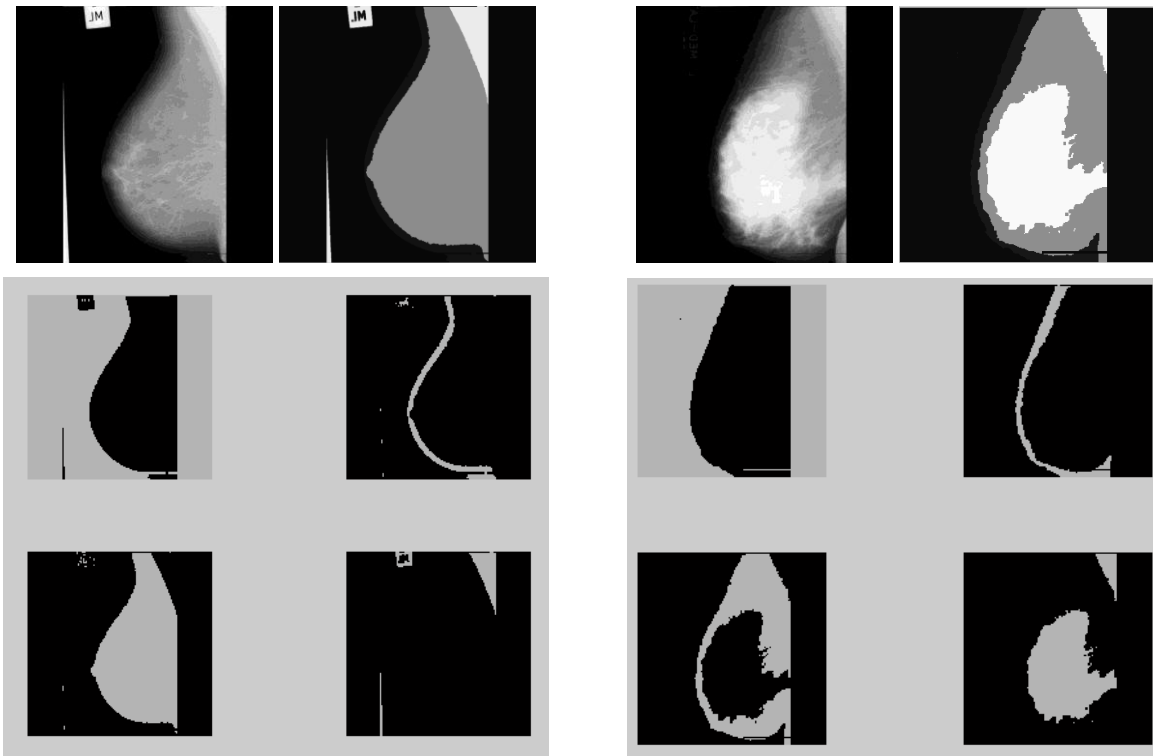
Phương pháp chênh lệch lược đồ xám giới thiệu ở phần trên tạo ra ranh giới giữa các thành phần trên ảnh X quang vú. Tuy nhiên, các ranh giới này có dạng cường độ điểm ảnh. Để phân vùng ảnh thành các phần riêng biệt, cường độ điểm ảnh phải được dùng với ảnh X quang. Từ đạo hàm xấp xỉ của lược đồ xám trọng số, ta xác định được 4 cực đại cục bộ tương trưng cho mức xám trung bình của 4 vùng cần chia lần lượt là vùng nền, vùng mô mỡ, vùng mô tuyến và cơ ngực. Để tìm được ngưỡng giá trị giữa các vùng, ta tìm cực tiểu giữa hai cực đại cục bộ. Giá trị đó chính là ngưỡng của từng vùng. Như vậy, dựa vào đạo hàm xấp xỉ của lược đồ xám trọng số, ta xác định ranh giới về cường độ giữa các vùng.

Sử dụng ranh giới về cường độ, ta phân ngưỡng ảnh X quang ban đầu. Một số điểm ảnh bị phân loại nhầm. Các điểm ảnh này được định vị về không gian trong một thành phần những giá trị cường độ của chúng lại nằm ngoài phạm vi phát hiện với thành phần đó. Để phân vùng phù hợp và phân tích chuẩn, chúng cần được thay đổi để phù hợp với vùng chúng thuộc về. Điều này dễ dàng được thực hiện bằng các thực hiện phép toán ăn mòn trên ảnh. Phép toán ăn mòn ảnh kiểm tra từng điểm ảnh trên ảnh và nếu một điểm ảnh nào đó khác với 4 trong 8 điểm ảnh lân cận của nó, nó được thay đổi để phù hợp với 4 điểm ảnh nào đó. Nói cách khác, nếu điểm ảnh nổi bật trên nền của nó, nó sẽ được thay đổi để phù hợp với nền. Việc này loại bỏ việc phân loại sai khỏi bức ảnh phân vùng.

Sau khi ảnh đã được phân vùng đúng, ta chia các vùng vào các ảnh riêng biệt. Điều này cho phép xử lý độc lập với các vùng.

III. KẾT QUẢ THỰC HIỆN

Thuật toán được thực hiện bằng Matlab 7.04 và được dùng với cơ sở dữ liệu MINI-MIAS, các ảnh có kích thước 1024x1024.



Hình 8. Kết quả phân vùng (ảnh được chia thành 4 vùng riêng biệt: cơ ngực, mô mỡ, mô tuyến và nền ảnh)

IV. KẾT LUẬN

Như vậy, với thuật toán đơn giản, phương pháp chênh lệch mức xám đồ được thực hiện theo các bước như sau:

- Ảnh được đưa xử lý qua bộ lọc trung bình để loại bỏ các đặc trưng nhỏ.
- Tạo ảnh từ các giá trị chênh lệch mức xám đồ.
- Lược đồ xám trọng số được tạo ra từ ảnh chênh lệch mức xám đồ.
- Tính đạo hàm xấp xỉ của lược đồ xám trọng số từ đó tìm ra ngưỡng của các vùng.

- Ảnh được phân vùng theo ngưỡng tìm được.

- Ăn mòn điểm ảnh đối với ảnh đã phân vùng nhiều lần để loại bỏ các điểm ảnh bị phân loại sai.

- Các thành phần khác nhau được chia thành các thành các ảnh khác nhau.

Phương pháp này đã dùng với rất nhiều ảnh X quang trong bộ cơ sở dữ liệu MIAS và đều ra kết quả khá chính xác. Hiện phần mềm đang tiếp tục được nghiên cứu và phát triển hoàn thiện phục vụ cho công tác hỗ trợ chẩn đoán ung thư vú tại bệnh viện và hỗ trợ nghiên cứu.

TÀI LIỆU THAM KHẢO

1. Bích Thủy, Hải Yến; Cẩm nang vú và bệnh ung thư vú; NXB Y học, 1997.
2. John Terry Sample; Computer Assisted Screening of Digital Mammogram Images; University of Southern Mississippi; 2003
3. Martin Masek; Hierarchical segmentation of mammograms based on pixel intensity; The University of Western Australia, 2004.
4. Aamir Mukhdoomi; On the problem of Segmentation Fibroglandular Tissue in Mammograms; Umea University, Sweden; 2006
5. Mohammad Sameti; Detection of soft tissue Abnormalities in mammographic Images for early diagnosis of breast cancer; University of Waterloo; 1994.

Địa chỉ liên hệ: Nguyễn Thái Hà - Tel: (04) 3868.2099, 0904.277.585, Email: thaiha-fet@mail.hut.edu.vn
 Khoa Điện tử Viễn thông, Trường Đại học Bách khoa Hà Nội
 Số 1, Đại Cồ Việt, Hà Nội