

디지털 유방영상에서 멀티영상 기반의 컴퓨터 보조 진단에 관한 연구

최형식¹, 조용호¹, 조백환¹, 문우경², 임정기², 김인영¹, 김선일¹

¹한양대학교 의용생체공학과

²서울대학교 의과대학 진단방사선과

(Received January 22, 2007. Accepted January 30, 2007)

A Study on the Multi-View Based Computer Aided Diagnosis in Digital Mammography

Hyung-Sik Choi¹, Yong-Ho Cho¹, Baek-Hwan Cho¹, Woo-Kyoung Moon², Jung-Gi Im², In-Young Kim¹, Sun-I. Kim¹

¹Department of Biomedical Engineering, Hanyang University

²Department of Diagnostic Radiology, College of Medicine, Seoul National University

Abstract

For the past decade, the full-field digital mammography has been widely used for early diagnosis of breast cancer, and computer aided diagnosis has been developed to assist physicians as a second opinion. In this study, we try to predict the breast cancer using both mediolateral oblique(MLO) view and craniocaudal(CC) view together. A skilled radiologist selected 35 pairs of ROIs from both MLO view and CC view of digital mammogram. We extracted textural features using Spatial Grey Level Dependence matrix from each mammogram and evaluated the generalization performance of the classifier using Support Vector Machine. We compared the multi-view based classifier to single-view based classifier that is built from each mammogram view. The results represent that the multi-view based computer aided diagnosis in digital mammogram could improve the diagnostic performance and have good possibility for clinical use to assist physicians as a second opinion.

Key words : digital mammogram, textural features, support vector machine, multi-view analysis

1. 서론

서구식 생활문화의 유입으로 인해 우리나라의 유방암이나 유방질환은 점점 증가하고 발생연령도 낮아지고 있다. 따라서, 이러한 유방암에 있어서 조기진단은 아무리 강조해도 지나침이 없다. 유방암의 조기진단을 위해서 유방촬영(mammography)을 비롯한 유방초음파촬영술, 그리고 유방자기공명촬영술 등의 유방영상진단이 그 근간을 이루고 있다. 그 중 유방촬영술의 결과로 만들어진 유방영상(mammogram)이 유방의 미세석회화(microcalcification)나 유방 종괴(mass) 등의 유방 병변을 검출하는데 가장 많이 사용된다 [1, 2].

석회화(calcification)는 유방 병변에 형성되어 있는 조그마한 칼슘 침전물인데, 대부분 양성 과정으로써, 활동적인 유즙분비

에 의한 석회화 소견과 악성 병변에서 보이는 괴사성 세포의 침전물에 의한 것이며, 유방영상에서 이러한 석회화 소견은 밝은 하얀색 점들처럼 보인다. 유방 병소의 양성 및 악성을 판단하기 위해서는 석회화 소견의 분석이 매우 중요하며, 석회화 소견의 형태, 크기, 음영, 석회화 군집내의 개수 등이 양성 및 악성을 판단하는 중요한 파라미터이다. 그러나 유방영상에서 석회화 소견의 크기는 매우 작고 때로는 유방조직과 겹쳐있어서 눈으로 식별하기 힘든 경우가 있다 [3].

이런 문제점을 해결하기 위하여 최근 컴퓨터를 이용한 방법이 많이 도입되고 있다. 컴퓨터 보조 검출(computer aided detection, CAD) 기술을 이용하여 의심되는 병변 영역을 먼저 검출한 후, 다양한 알고리즘의 컴퓨터 보조 진단(computer aided diagnosis, CADx) 방법을 적용하여 해당 영역의 영상 특성을 분석하여 병변 영역의 양성 및 악성의 정도를 자동으로 예측하고 있다 [4-6].

일반적으로 유방영상을 이용한 CADx 시스템에서는 크게 두 종류의 특징점이 이용되는데, 형태학적 특징점과 비형태학적 특징점이 그것이다. 형태학적 특징점은 병변의 크기나 모양 등과 같이 형

이 논문은 원자력중장기 연구개발사업 지원에 의한 결과입니다

Corresponding Author : 김인영

서울특별시 성동구 행당동 17 한양대학교 의용생체공학과

Tel : +82-02-2291-1713 / Fax : +82-02-2296-5943

E-mail : iykim@hanyang.ac.kr

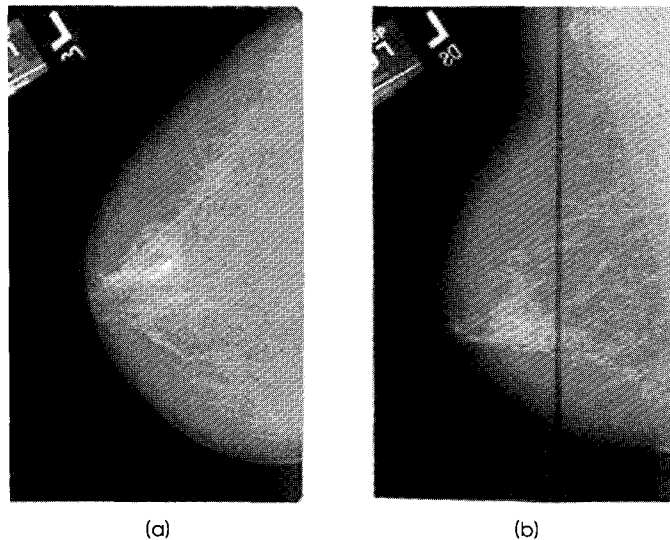


그림 1. 유방암 진단용 유방영상에서 두 영상. 왼쪽 그림은 상하측 촬영영상을 나타내고, 오른쪽 그림은 내외방향사위 촬영영상을 나타내고 있다.
Fig. 1. Two views in screening mammography. The left image shows a craniocaudal (CC) view and the right shows a mediolateral oblique (MLO) view.

태학적 관점에서 양성과 악성 영상을 분류할 수 있는 것을 의미한다. 그러나, 방사선과 전문의는 유방영상을 이용한 진단시 이러한 형태학적 특징점 뿐 만 아니라 비형태학적 특징점도 이용하며, 그 중에서 영상에 있어서 질감을 나타내는 텍스처 특징점들이 유방영상을 판별하는데 자주 이용되어 왔다 [7, 8].

또한, 방사선과 전문의는 유방영상에서 영상의 전반적인 조건을 분석하기 위하여 그림 1과 같이 일반적으로 양측의 내외방향사위 촬영(mediolateral oblique view, MLO view)과 상하촬영(craniocaudal view, CC view)을 근접시켜 비교, 분석하는 방법을 사용한다 [1]. 예를 들어, MLO view에서 의심되는 병변이 발견되면 CC view에서도 같은 병변이 발견되는지 확인하여, 해당 병변이 실제로 존재하는지 여부를 판별한다. Thurfjell 등은 하나의 영상에서만 병변이 발견될 때 보다, 두 영상 모두에서 병변이 발견될 때 오진률이 더욱 낮아진다고 보고하였다 [9-11].

그러나, 유방영상에 있어서 현재까지의 많은 CAD 및 CADx 기법들은 이러한 두 영상 비교의 장점을 이용하기 어려웠다. 유방 자체가 유연하기 때문에 각 view의 영상을 촬영할 때 변형이 일어나기 쉽고, 이로 인해, 컴퓨터를 이용한 영상처리에 많은 어려움이 따랐기 때문이다. Huo 등은 MLO view, CC view 및 특정 지점 압축 촬영(spot compression view)의 세 영상을 이용하여 유방 종괴의 악성 여부를 판별하였다 [12]. 즉, 각 영상에 대한 분류기(classifier)를 따로 생성한 후, 특정환자의 세가지 유방영상을 모두 적용하였을 때, 각 분류기의 결과값을 평균하여 최종 판단을 내리도록 하였다. 그렇지만, 각각의 영상에서 먼저 결과를 도출한 뒤에, 그 결과만을 가지고 종합적인 판단을 한다는 것은 오류를 낳을 가능성이 크다 할 수 있다.

따라서, 본 연구에서는 MLO view와 CC view에서 텍스처 특징점을 추출한 후, 모든 특징점을 동시에 적용하여 Support Vector

Machine(SVM)을 이용한 기계 학습 방법으로 유방영상의 컴퓨터 보조 진단의 진단 성능을 향상시키는 연구를 수행하였다.

II. 재료 및 방법

A. 데이터 수집

서울대학교병원을 내원한 31명의 여성환자로부터 디지털 유방영상촬영기(GE Senography 2000D)를 이용하여 유방영상 데이터를 수집하였으며, 영상의 공간해상도는 $100\mu\text{m} \times 100\mu\text{m}$ 이고, 화소당 12 bit로 저장하였으나, 계산의 효율성과 결과의 향상을 위해 경험적으로 12 bit 영상을 7 bit로 다운 샘플링 하였다. 숙련된 방사선과 전문의가 각 환자의 유방영상에서 1군데 이상의 군집된 미세석회화가 있는 관심영역(region of interest, ROI)을 추출하였으며, 동일한 영역에 해당하는 MLO view와 CC view의 35쌍(총 70개의 관심영역)을 연구에 포함하였으며, 양성(benign) 16건과 악성(malignant) 19건으로 구성되어 있다. 관심영역의 크기는 45×59 pixel에서 318×368 pixel까지 이다.

B. SGLD 매트릭스를 이용한 텍스처 특징점 추출

Spatial Grey Level Dependence(SGLD) 매트릭스는 픽셀의 그레이 레벨(grey level)들간의 조합이 얼마나 자주 일어나는지를 나타낸다 [13]. 이는 거리 d 와 방향 θ 에 떨어진 두 픽셀의 그레이 레벨의 빈도수를 나타내는 매트릭스이며, 그림 2는 거리 1과 방향 0° 일 때의 SGLD 매트릭스를 나타내고 있다. 이 SGLD 매트릭스로부터 Contrast, Correlation, Energy, Homogeneity, Inverse Difference Moment, Sum of Entropy, Difference Entropy, Information Measures of Correlation 등 8개의 특징점을 추출하였으며, 다음과 같이 계산하였다.

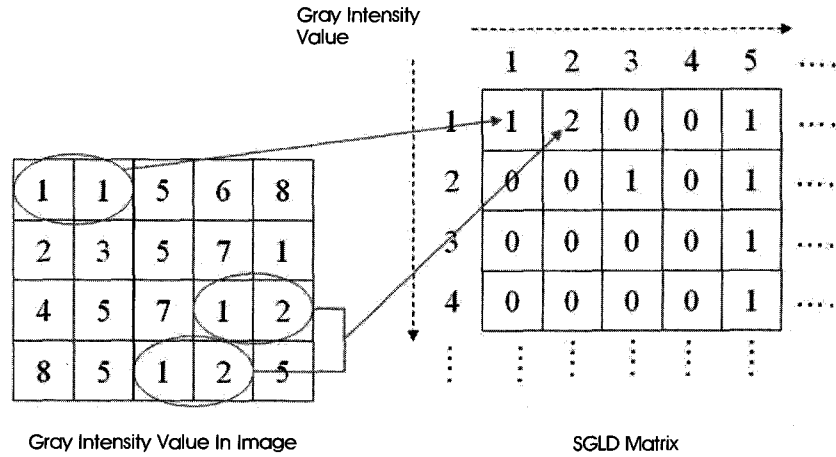


그림 2. 거리 1과 방향 0 일 때, SGLD 매트릭스를 계산하는 실례. 그레이레벨 (1,1)쌍은 관심영역에서 1번 존재하고, (1,2)쌍은 2번 존재한다.
 Fig. 2. An example of SGLD matrix for distance of 1 and direction of 0 . There are one grey level pair of (1,1) and two pairs of (1,2).

Notation

$P(i,j)$ 는 SGLD 매트릭스에서 (i,j) 번째 원소를 나타내며, $p(i,j)$ 는 SGLD내의 모든 원소값들의 합으로 다시 나누어 정규화한 SGLD 매트릭스에서의 (i,j) 번째 원소를 나타낸다. 이를 기반으로 다음과 같은 값들을 계산할 수 있다.

$$p_x(i) = \sum_{j=1}^{Ng} P(i,j), \text{ where } Ng \text{ is the number of grey levels} \quad (1)$$

$$p_y(j) = \sum_{i=1}^{Ng} p(i,j) \quad (2)$$

$$p_{x+y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j), \text{ } k = 2,3,\dots,2Ng, \text{ where } i + j = k \quad (3)$$

$$p_{x-y}(k) = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j), \text{ } k = 0,1,\dots,Ng-1, \text{ where } |i - j| = k \quad (4)$$

$$HXY = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j) \log(p(i,j)) \quad (5)$$

$$HXY1 = - \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j) \log\{p_x(i)p_y(j)\} \quad (6)$$

Definition

위에 정의된 Notation 값들로부터 아래의 특징점들을 계산하였다.

$$\text{Contrast} : f_1 = \sum_{n=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i,j) \right\}, \text{ where } |i - j| = n \quad (7)$$

$$\text{Correlation} : f_2 = \frac{\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

$$\text{Energy (Angular Second Moment)} : f_3 = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \{p(i,j)\}^2 \quad (9)$$

$$\text{Homogeneity} : f_4 = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{p(i,j)}{1 + |i - j|} \quad (10)$$

$$\text{Inverse Difference Moment} : f_5 = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{1}{1 + (i - j)^2} p(i,j) \quad (11)$$

$$\text{Sum of Entropy} : f_6 = - \sum_{i=2}^{2Ng} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad (12)$$

$$\text{Difference of Entropy} : f_7 = - \sum_{i=0}^{Ng-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (13)$$

$$\text{Information Measures of Correlation} : f_8 = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (14)$$

이러한 SGLD 매트릭스는 각 pixel 쌍의 거리와 방향에 따라 하나씩 생성할 수 있다. 즉, pixel간 거리를 1~4까지 계산하고, 각도를 0°, 45°, 90°, 135°의 4방향에 해당하는 모든 매트릭스를 계산한다면 총 16개의 매트릭스가 생성된다. 따라서, 각각의 매트릭스에 8개의 특징점을 구할 수 있기 때문에 총 128개의 특징점이 각 유방영상의 ROI에서 추출된다. 그러나, 이는 데이터 레코드의 개수에 비해 매우 많기 때문에, 특징점 개수의 감소를 필요로 한다. 이를 위해 본 연구에서는 경험적인 방법에 의해서, 0°와 90°방향의 매트릭스의 각 원소값을 평균화하여, 하나의 매트릭스로 대체한다. 45°와 135°방향의 매트릭스도 마찬가지로 Half averaging한다. 따라서, 최종적으로 각각의 영상에 64개의 특징점을 구할 수 있다.

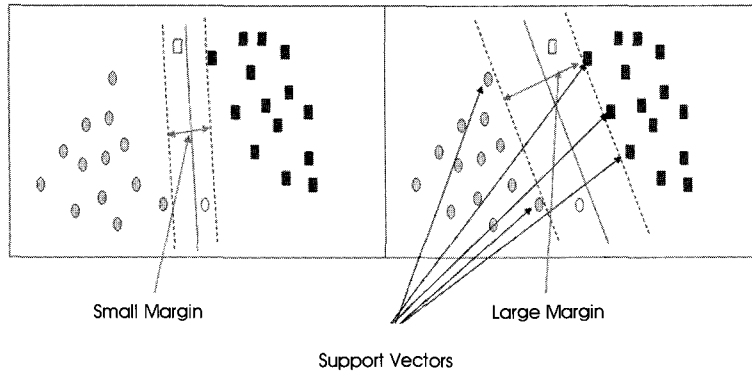


그림 3. Margin의 크기에 따른 분류기. SVM은 두 그룹을 구분 짓는 초월평면중에서 가장 큰 Margin을 갖는 것을 찾으며, 초월평면에서 가장 가까운 거리에 있는 데이터 샘플을 Support Vector라고 한다. 왼쪽 그림의 초월평면, 즉 분류기는 테스트 데이터(하얀색 사각형과 원)를 정확하게 분류하지 못하는 반면, 오른쪽 그림의 초월평면은 Margin이 크기 때문에 정확하게 분류한다.

Fig. 3. Classifiers that have smaller margin and larger margin. SVMs seek to find maximal margin hyperplane that separates the two groups like the right figure. The closest data samples from the hyperplane are support vectors. Although the classifier in the left figure can not discriminate the test data (white square and circle), that in the right figure exactly classify the test data, since the right one has larger margin.

C. Support Vector Machine(SVM)

SVM은 분류 및 예측 기법에 사용되는 기계 학습법의 일종이다 [14-16]. 그림 3에서 나타낸 바와 같이, SVM은 두 그룹을 분류하는데 있어서, 두 그룹을 구분시키는 초월평면(hyperplane)의 계수 값(weight)을 계산하는데, 마진(margin)을 최대화 하도록 한다. 마진은 초월평면과 직교방향으로 가장 가까운 데이터 샘플과의 거리를 의미한다. 마진이 큰 초월평면을 찾는 이유는, 기계학습에 의해 두 그룹을 구분 짓는 분류기, 즉 초월평면을 계산한 후에 새로운 데이터인 테스트 데이터로 분류할 때, 보다 더 정확한 예측을 할 수 있다는 가정 때문이다. 그림 3에서 마진이 작은 왼쪽 그림의 분류기가 흰색의 테스트 데이터를 잘못 분류한 반면, 마진이 큰 오른쪽 그림의 분류기는 정확하게 판별한 것을 확인할 수 있다. 여기서, 초월평면과 가장 가까운 데이터 샘플들을 support vector라고 하는데, SVM은 이 support vector를 찾아서, 그것들만을 가지고 분류기를 모델링한다.

양성(positive)인지 음성(negative)인지를 미리 알고 있는 학습 데이터 세트 $(x_i, y_i), i=1, \dots, n, y_i \in R^d$ 를 가정 하였을 때, 선형분리 가능한 경우에는 $F(x) = (w \cdot x) + b$ 인 초월평면을 구할 수 있다. 이때 w 는 가중치 벡터이고, b 는 바이어스이며, (\cdot) 는 벡터내적을 의미하며, 두 그룹을 구분 지을 수 있는 초월평면은 $(w \cdot x + b) * y \geq 1$ 을 만족한다.

마진은 $|F(x)|/\|w\|$ 로 계산할 수 있는데, 마진을 최대화하는 초월평면을 구하는 것은 다음과 같은 최적화 문제가 된다..

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|w\|^2 \\ &\text{subject to } y_i(w \cdot x_i + b) \geq 1, \quad i = 1, \dots, n \end{aligned} \tag{15}$$

이 최적화 문제를 풀기 위해 Karush-Kuhn-Tucker(KKT) 이론

을 이용하여, 다음과 같이 dual problem으로 전환하여 계산한다.

$$\begin{aligned} &\text{maximize } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \langle x_i \cdot x_j \rangle \\ &\text{subject to } \alpha_i \geq 0, \quad i = 1, \dots, n \\ &\sum_{i=1}^n \alpha_i y_i = 0 \end{aligned} \tag{16}$$

여기서, α 는 Lagrange multiplier로써, 모든 학습데이터 마다 하나씩의 양수값을 갖게 되는데, support vector 외에 다른 데이터는 모두 0의 값을 갖게 된다. 그리하여, 최종적으로 다음 식과 같이 초월평면, 즉 분류기의 식을 구할 수 있게 되는데, 일반적으로 테스트 데이터를 이 식에 적용하였을 때, 양수값이 나오면 양성, 음수값이 나오면 음성으로 판별한다.

$$F(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i \langle x \cdot x_i \rangle + b \right) \tag{17}$$

위와 같은 SVM의 학습방법은 초월평면에서 가장 가까운 데이터인 support vector를 찾을 때 엄격한 규칙을 적용 시키기 때문에 hard margin 방법이라 하는데, 이는 학습데이터가 선형분리 가능한 경우를 가정한 학습법이다. 그러나 만약 학습데이터가 선형분리 불가능한 경우에는, support vector를 찾는 규칙을 어느 정도 완화하는데, 즉 학습시의 에러를 어느 정도 허용하도록 한다. 이를 soft margin 방법이라 하는데, 다음과 같은 최적화 문제로 해를 구한다.

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ &\text{subject to } y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad i = 1, \dots, n \\ &\xi_i \geq 0, \quad i = 1, \dots, n \end{aligned} \tag{18}$$

표 1. 실제 진단과 예측된 결과를 판단하는 일반적인 결정 매트릭스

Table 1. A general decision matrix that contains information on actual diagnosis and prediction output.

Actual diagnosis	Predicted result	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Hard margin 방법에 비하여 C 값을 포함한 여러 부분이 추가된 것을 확인 할 수 있다. 따라서, 위 식을 이용하여 최적화를 한다면, 마진의 크기뿐만 아니라 여러 부분도 동시에 최소화 하도록 하기 때문에 계산상으로 매우 복잡해질 수 있다. 그러므로, hard margin 방법에서와 마찬가지로 dual problem으로 변경하면 다음과 같이 쉽게 계산할 수 있다.

$$\begin{aligned}
 & \text{maximize } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j \langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle \\
 & \text{subject to } 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \\
 & \sum_{i=1}^n \alpha_i y_i = 0
 \end{aligned} \tag{19}$$

위 식을 hard margin 방법과 비교하였을 때, 값의 범위만 파라미터 C로 한정시키고 나머지 부분은 완전히 동일함을 확인할 수 있다. 여기서 파라미터 C가 무한대로 크면, hard margin 방법과 동일한 효과를 나타낸다. 따라서, C값이 작을수록 유연한 규칙을 적용하여 좀 더 많은 support vector를 찾을 수 있게 된다.

위와 같은 선형 커널(linear kernel)을 이용하는 방법 외에도, 학습데이터가 선형분리 불가능한 경우에는 비선형 커널(nonlinear kernel)을 이용하기도 한다. 비선형 커널은 원 데이터들을 선형분리 가능한 고차원으로 매핑시키는 역할을 한다고 할 수 있다. 식 (19)에서 선형 커널인 $\langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle$ 부분을 커널의 일반적인 형태인 $K(\mathbf{x}_i, \mathbf{x}_j)$ 로 전환할 수 있는데, 가장 널리 사용되는 비선형 커널인 RBF(radial basis function) 커널은 다음과 같은 식으로 표현할 수 있으며, 여기서 γ 는 데이터간 유사도를 결정하는 파라미터이다.

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \tag{20}$$

이러한 SVM은 1990년대 후반 및 2000년대 초반까지 널리 사용되었던 인공신경망(artificial neural network)에 비하여 수많은 영역에서 높은 성능을 나타낼 수 있다고 보고되어 왔다. 또한 인공신경망에서는 학습률과 은닉층의 개수, 각 은닉층에 사용된 뉴런의 개수 등 사용자가 조절해야 할 많은 파라미터가 있는 반면에, SVM에서는 선형 커널의 경우 마진을 조절하는 파라미터 C와 비선형 커널인 RBF에서는 γ 값 하나만을 조절하면 되므로, 사용자에게 많은 편의성을 제공한다고 할 수 있다.

D. 실험절차

하나의 영상을 이용한 유방영상의 판독과 MLO view와 CC

view를 모두 이용한 판독 결과를 비교하기 위하여 기저 실험에서는 MLO view와 CC view를 개별적인 데이터 샘플로 여겨서, 총 64개의 특징점으로 이루어진 70개의 데이터로 실험을 하였고, 본 실험에서는 MLO view와 CC view의 특징점을 모두 합쳐서, 총 128개의 특징점으로 이루어진 35개의 데이터 샘플로 실험하였다. 두 실험에 대한 예측성능을 객관적으로 비교하기 위하여, 예측정확도(accuracy), 평균오차율(balanced error rate), 민감도(sensitivity) 및 특이도(specificity)를 비교하였다.

표 1은 실제 진단과 예측된 결과를 비교할 때 사용되는 결정 매트릭스를 나타내고 있다. 여기서 계산할 수 있는 TP, FP, TN, FN을 이용하여 예측정확도, 평균오차율, 민감도 및 특이도를 다음과 같이 계산할 수 있다.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{21}$$

$$\text{Balanced Error Rate} = \frac{1}{2} \left(\frac{FN}{TP + FN} + \frac{FP}{FP + TN} \right) \tag{22}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{23}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{24}$$

SVM을 이용하여 학습 및 테스트를 수행하였고, 데이터 개수가 비교적 많지 않기 때문에, Leave One Out Cross Validation (LOOCV) 방법으로 평가하였다. LOOCV 방법은 전체 데이터에서 하나만을 제외한 나머지 데이터를 이용하여 학습한 후, 제외된 하나의 데이터로 테스트하는 방법인데, 이러한 과정을 전체 데이터 개수만큼 반복한다. 이는 전체 데이터를 임의로 학습 데이터와 테스트 데이터로 구분하여 평가하는 hold out 방법에서 발생할 수 있는 bias를 최소화 할 수 있는 장점이 있는 반면, 대단위 데이터에서는 많은 계산 시간이 소요된다는 단점이 있다.

일반적으로 SVM을 학습할 때, 학습데이터가 선형분리 가능한 경우는 매우 이상적인 경우라 할 수 있다. 따라서, 본 연구에서 사용된 데이터가 선형 분리 가능하다는 보장이 없기 때문에, soft margin 방법과 비선형 커널인 RBF 커널을 사용하여 학습한 후 그 결과를 비교하였다. 이때, SVM의 soft margin을 제어하는 C값을 1에서 1000까지 변경시켜 가면서 학습하였다. 또한 비선형 커널인 RBF 커널을 사용할 때에는, 데이터간 유사도에 대한 파라미터

표 2. 단일 영상 기반 방법의 실험 결과.

Table 2. Experimental results using single-view based method.

데이터 개수	특징점 개수	커널	C ¹	gamma ²	예측정확도	평균오차율	민감도	특이도
70	64	Linear	1	-	0.643	0.351	0.579	0.719
			10	-	0.8	0.199	0.789	0.813
			100	-	0.743	0.254	0.711	0.781
			1000	-	0.771	0.223	0.711	0.844
		RBF	1	0.1	0.657	0.343	0.658	0.656
			10	100	0.714	0.283	0.684	0.75
			100	100	0.729	0.275	0.763	0.688
			1000	1	0.686	0.319	0.737	0.625

¹C: A user parameter that controls allowable errors in learning SVM.

²gamma: A user parameter that controls similarity between data in RBF kernel.

표 3. 멀티 영상 기반 방법의 실험 결과

Table 3. Experimental results using multi-view based method.

데이터 개수	특징점 개수	커널	C ¹	gamma ²	예측정확도	평균오차율	민감도	특이도
35	128	Linear	1	-	0.714	0.288	0.737	0.688
			10	-	0.743	0.257	0.737	0.75
			100	-	0.829	0.173	0.842	0.813
			1000	-	0.771	0.235	0.842	0.688
		RBF	1	0.01	0.628	0.372	0.632	0.625
			10	100	0.714	0.278	0.632	0.813
			100	10	0.771	0.230	0.789	0.75
			1000	1000	0.714	0.288	0.737	0.688

¹C: A user parameter that controls allowable errors in learning SVM.

²gamma: A user parameter that controls similarity between data in RBF kernel.

인 gamma값을 10⁻⁵에서 10⁵까지 변화시키면서, 가장 성능이 뛰어난 분류기를 찾도록 하였다. SVM의 학습은 전 세계적으로 가장 널리 알려진 SVM 학습 라이브러리인 LIBSVM을 이용하여, python 코드로 리눅스 기반에서 수행하였다 [17].

III. 결과 및 고찰

표 2는 각각의 영상을 개별적으로 하나의 데이터 샘플로 취급한 단일 영상 기반의 방법에 대한 실험 결과를 나타내고 있다. 선형 커널로 학습하였을 때의 예측 정확도가 0.8로 가장 높았고, 각 그룹에 대한 오차율의 평균이 0.199로 가장 낮음을 알 수 있다. 이때의 민감도와 특이도는 각각 0.789와 0.813을 나타내었다.

표 3은 MLO view와 CC view의 각 특징점을 한꺼번에 이용한 멀티 영상 기반 방법에 대한 결과를 나타내고 있다. 비선형 커널인 RBF 커널보다, 선형 커널에서 가장 높은 결과를 나타내었고, 이때의 예측정확도가 0.829, 평균오차율은 0.173이었다. 단일 영상의 경우보다 더 좋은 성능을 나타내었으며, 이때의 민감도가 0.842이고, 특이도는 0.813이었다.

실험 결과, 전체적으로 RBF 커널을 사용했을 때 보다, 선형 커

널을 사용하였을 때 더 나은 성능을 나타내는 것을 확인하였다. 이는 각 특징점들이 각각 적절한 가중치를 두고 선형적으로 더하면, 비선형적인 결합보다 두 그룹을 더 잘 분류할 수 있다는 의미이다.

뿐만 아니라, 멀티 영상 기법이 단일 영상기법의 경우보다 예측 정확도, 평균오차율 및 민감도 측면에서 우수한 성능을 나타내었다. 이는 MLO view나 CC view 각각의 영상만을 가지고 유방영상을 분석하는 것 보다, 동시에 두 개의 영상을 하나의 데이터로 묶어 분석하는 것이 더 정확한 결과를 나타낼 수 있다는 것을 의미한다. 즉, 각각의 영상을 따로 분석할 경우에 동일한 환자의 유방에 대한 유방영상에서 서로 다른 결과를 얻을 수 있기 때문이다. 본 연구는 각각의 영상을 따로 분석한 후에, 그 결과값들만을 가지고 종합적으로 분석했던 기존의 연구와 차이가 있다고 할 수 있으며, 향후 더 많은 데이터로 검증할 수 있는 가능성을 열어두고 있다.

본 연구에서 사용된 특징점은 각각의 영상에서 64개였고, 두 영상을 동시에 이용하였을 때에는 128개였다. 이는 사용한 데이터 개수에 비해 여전히 많은 특징점을 가지고 있다고 여겨지며, 이를 위해 특징점 선택(feature selection) 방법의 적용을 고려해 볼 수 있다. 특징점 선택 방법은 모든 특징점들 중에서 특정 특징점 집합을 선택하고, 필요 없는 특징점들은 제거하는 기계 학습 과정의 일

부분이다. 특징점들이 많으면 많을수록 필요한 계산량이 기하급수적으로 늘어나며, 기계 학습의 결과에도 나쁜 영향을 미칠 수 있다. 이는 "차원의 저주(curse of dimensionality)"에 빠지는 것을 피하기 위해 여러 가지 특징점 선택 기법들이 사용되어 왔다 [18]. 이러한 특징점 선택 기법은 데이터의 시각화를 용이하게 할 수 있고, 데이터를 보다 직관적으로 이해할 수 있도록 해준다. 뿐만 아니라, 측정에 걸리는 시간이나 저장장치의 용량도 줄여 줄 수 있으며, 기계 학습시에 필요한 계산량을 줄여 줄 수 있다는 장점과 함께, 기계 학습의 성능 향상도 꾀할 수 있다. 그러나, 이러한 특징점 선택시 실제로 유용한 특징점들이 제거될 수도 있다는 단점이 있기 때문에, 적절한 사용이 요구된다. 유방영상에서 적용한 특징점 선택 기법의 일종으로 Hall 등은 집단과의 상관관계가 높고, 특징점들간의 상관관계는 낮은 특징점들이 분류 기법에 있어서 좋은 성능을 나타낸다는 가정하에, 예측 결과와 관련이 없고 불필요한 특징점을 제거하는 것이 효과적이라고 발표하였다 [19]. 이러한 가정은 완전한 상관관계를 갖는 특징점들은 부가적인 정보를 갖지 않기 때문이다. 따라서, 향후에는 MLO view와 CC view에서 각각 추출된 특징점의 상관관계를 이용한 연구를 할 필요가 있다고 판단된다.

본 연구에서는 SGLD 매트릭스를 이용하여 텍스처 특징점을 추출한 후, 기계 학습방법의 일종인 SVM을 이용하여 예측률을 측정하였다. 그러나, 방사선과 전문의가 유방영상을 판독할 때에는 텍스처 뿐만 아니라 병변의 형태학적인 소견도 중요하게 여긴다. 그러므로, 추가 연구에서는 형태학적인 특징점들도 동시에 이용하면, 보다 더 높은 예측률을 기대할 수 있으리라 판단된다.

참고문헌

- [1] D. B. Kopans, *Breast Imaging*: Lippincott Williams & Wilkins, 1997.
- [2] S. H. Heywang-Kobrunner, *Diagnostic Breast Imaging*: Thieme, 1997.
- [3] E. Alberdi, R. Lee, and P. Taylor, "Radiologists' description and interpretation of mammographic microcalcifications: A knowledge elicitation study for computerized decision support," presented at IWDM, Toronto, Canada, 2000.
- [4] R. M. Nishikawa, M. L. Giger, K. Doi, C. J. Vyboryny, and R. A. Schmidt, "Computer aided detection of clustered microcalcifications on digital mammograms," *Medical and Biological Engineering & Computing*, vol. 33, pp. 174-178, 1995.
- [5] W. Veldkamp, N. Karssemeijer, J. Otten, and J. Hendriks, "Automated classification of clustered microcalcifications into malignant and benign types," *Medical Physics*, vol. 27, pp. 2600-2608, 2000.
- [6] M. Kallergi, "Computer aided diagnosis of mammographic microcalcification clusters," *Medical Physics*, vol. 31, pp. 314-326, 2004.
- [7] A. Petrosian, H. P. Chan, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis," *Phys. Med. Biol.*, vol. 39, pp. 2273-2288, 1994.
- [8] H. P. Chan, Sahiner B., W. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces," *Medical Physics*, vol. 25, pp. 2007-2019, 1998.
- [9] E. Thurfjell, "Mammography screening: One versus two views and independent double reading," *Acta Radiologica*, vol. 35, pp. 340-344, 1994.
- [10] E. Thurfjell, A. Taube, and L. Tabar, "One- versus two-view mammography screening: A prospective population-based study," *Acta Radiologica*, vol. 35, pp. 340-344, 1994.
- [11] A. Tiedeu, C. Daul, P. Graebbling, and D. Wolf, "Correspondence between microcalcification projections on two mammographic views acquired with digital systems," *Comp. Med. Imag. and Graph*, vol. 28, pp. 151-158, 2005.
- [12] Z. Huo, M. L. Giger, D. E. Wolverton, W. Zhong, and S. Cumming, "Computerized analysis of mammographic parenchymal patterns for breast cancer risk assessment: Feature selection," *Medical Physics*, vol. 27, pp. 4-12, 2000.
- [13] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans on Systems, Man, Cybernetics*, vol. SMC-3, pp. 610-621, 1973.
- [14] V. Vapnik, *The nature of statistical learning theory*, New York: Springer, 1995.
- [15] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.
- [16] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*, Cambridge, UK: Cambridge University Press, 2000.
- [17] C.C. Chang and C.J. Lin, "LIBSVM : a library for support vector machines," Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> 2001.
- [18] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J Mach Learn Res*, vol. 3, pp. 1157-1182, 2003.
- [19] M. Hall, "Correlation-based feature selection for discrete and numeric class machine learning," presented at International Conference on Machine Learning, Stanford University, 2000.