

유방 종양 세포 조직 영상의 분류

Classification of Breast Tumor Cell Tissue Section Images

황해길, 최현주, 윤혜경, 남상희, 최흥국

Hae-Gil Hwang, Hyun-Ju Choi, Hye-Kyoung Yoon, Sang-Hee Nam, Heung-Kook Choi

요약

본 논문은 유방질환 중에서 유관(duct)에 발생하는 유방 종양을 Benign, DCIS(ductal carcinoma in situ), NOS(invasive ductal carcinoma)로 분류하기 위해 3가지 분류기(classifier)를 생성한 후, 비교·분석하였다. 분류기 생성에서 가장 중요한 단계인 특징 추출 단계에서 세포핵의 기하학적 특징을 나타내는 형태학적 특징을 추출하여 분류기를 생성하고, 염색질 패턴의 내부적 변화를 나타내는 질감 특징을 추출하여 2가지 배율(100/400배)에서 2개의 분류기를 생성하였다. 400배 배율의 유방 질환 영상에서 세포핵을 추출하여, 핵의 형태학적 특징값인 핵의 면적, 둘레, 가로·세로(장·단축)의 길이, 원형성의 비율을 구한 후, 이 특징값들을 조합하여 판별분석에 의해 분류기를 생성하고, 분류 정확도를 검증하였다. 100배 배율과 400배 배율의 유방 질환 영상에서 1, 2, 3, 4단계(level)의 wavelet 변환을 적용한 후, 분할된 서브밴드에서 GLCM(Gray Level Co-occurrence Matrix)을 이용하여 질감 특징(Entropy, Energy, Contrast, Homogeneity)을 추출하고, 이 특징값들을 조합하여 판별 분석에 의해 분류기를 생성한 후, 분류 정확도를 검증하였다. 이 세 분류기를 비교·분석하였을 때, 현미경 100배 배율의 영상을 3단계 wavelet 변환을 적용하고 질감 특징을 추출하여 생성한 분류기가 다른 두 분류기보다 유방 질환 Benign, DCIS, NOS를 분류하는데 더 나은 결과를 보였다.

ABSTRACT

In this paper, we propose three classification algorithms to classify breast tumors that occur in duct into Benign, DCIS(ductal carcinoma in situ), NOS(invasive ductal carcinoma). The general approach for a creating classifier is composed of 2 steps: feature extraction and classification. Above all, feature extraction for a good classifier is very significance, because the classification performance depends on the extracted features. Therefore, in the feature extraction step, we extracted morphology features describing the shape, size of nuclei and texture features. The internal structures of the tumor are reflected from wavelet transformed images with 10× and 40× magnification. Particularly, to find the correlation between correct classification rates and wavelet depths, we applied 1, 2, 3 and 4-level wavelet transforms to the images and extracted texture features from the transformed images. The morphology features used are area, perimeter, width of X axis, width of Y axis and circularity. The texture features used are entropy, energy, contrast and homogeneity. In the classification step, we created three classifiers from each of extracted features using discriminant analysis. The first classifier was made by morphology features. The second and the third classifiers were made by texture features of wavelet transformed images with 10× and 40× magnification. Finally, we analyzed and compared the correct classification rate of the three classifiers. In this study, we found that the best classifier was made by texture features of 3-level wavelet transformed images.

Key words: breast tumor cell tissue section images, morphology feature, texture feature, wavelet transform, classifier

I. 서론

유방암은 서구여성에게 있어서 가장 빈번하게 발생하는

악성 종양이며, 최근 우리 나라 여성에게도 위암에 이어 두번째로 많은 악성 종양이다. 매년 여성에게 발생하는 위암이나 자궁암은 발생률이 동일하거나 감소하는 추세

이나 유방암은 매년 발생률이 급격히 증가하고 있는 추세이다. 이에 따라 유방암에 대한 연구도 활발히 이루어지고 있는데, 유방암 진단 방법으로는 유방촬영술(Mammography), 초음파검사(Ultrasonography), 유선조영술(Galactography), 전산화단층촬영(CT), 자기공명영상(MRI) 등의 방사선학적 방법과 세포조직학적 방법이 있다. 특히 현미경을 통해 획득한 세포조직영상은 질병의 유무나 암의 진행 정도를 파악하기 위한 매우 중요한 요소이다[1][2].

유방 질환에서 세포 조직 영상의 분석을 위한 병리진단에서는 암종세포의 세포형태 정도와 핵의 다형성, 유사분열 정도를 기준으로 하여 나누는 방법이 용이하고 재현성이 높다[1]. 그리고 세포핵의 크기, 방향, 세포핵 사이의 거리 등에 대한 규칙성과 불규칙성은 암의 진행 정도를 판단하는데 있어 중요한 요인이므로, 그 특성을 분석하고 분류하기 위해서는 질감 분석이 필요하다[3]. 질감분석 방법 중 텍스처 정보에 기반한 방법에는 Gray Level의 변환 정도에 따른 Co-occurrence Matrix(GLCM)를 이용한 영상의 통계적 텍스처 분석 방법[4], wavelet 변환에 기반한 텍스처 특징을 분석하는 방법[5] 등이 제안되고 있다.

세포조직영상의 분석은 대부분 병리전문가의 육안에 의해 주관적으로 이루어지고 있으므로, 이를 보다 객관적이고 높은 재현성을 가지도록 하기 위해서는 눈에 보이는 시각적 판단 기준을 수치적으로 표현하는 특징의 정량적 해석과 전문가의 판단 기준을 가장 잘 나타낼 수 있는 유의한 특징값 추출이 필수적이다.

본 논문은 유방 질환 중에서 유관(duct)에 발생하는 유방종양을 Benign(양성종양), DCIS(ductal carcinoma in situ 관상피 내암종), NOS(invasive ductal carcinoma 침윤성 관상피)로 분류하기 위해 각 그룹별 특성을 잘 나타내는 특징을 추출하고, 추출한 특징들에 대한 유의성 평가로 특징들간의 상관관계까지 고려한 특징 벡터를 구성하여 분류기를 생성함으로써 보다 객관적이고 재현성이 높은 분류 방법을 개발하는데 목적이 있다.

따라서 분류하고자 하는 각 그룹(Benign, DCIS, NOS)에서 현미경 100배 배율(10×)와 400배 배율(40×)의 유방 종양 세포 조직 영상을 각각 20개씩, 총 60개(20×3)의 영상을 획득하였다. 100배 배율의 유방 종양의 세포 조직 영상과 400배 배율의 유방 세포 조직 영상을 비교해 보면 100배 배율 영상에서는 세포 하나 하나 보다는 유관 구조를 중심으로 영상을 볼 수 있고, 400배 배율 영상에서는 100배 배율과는 반대로 세포 하나 하나를 중심으로 영상을 볼 수 있다. 따라서 400배 배율의 영상에서 세포핵을 추출하여 형태학적 특징값을 추출하여 분류기를 생성하고, 같은 배율에서 질감특징값을 비교하기 위해 wavelet 변환을 적용하여 특징값을 추출하여 분류기를 생성하여 두 분류기를 비교 하였다. 그리고, 세포의 크기 조가 잘 나타내는 100배 배율의 영상을 wavelet 변환을 적용, 특징값을 추출하여 분류기를 생성하여 400배 배율

에서 같은 방법으로 생성된 분류기와 비교 하였다.

400배 배율의 유방 종양의 세포 조직 영상에서 20개씩, 총 60개 영상에서 세포핵을 추출하여, 핵의 형태학적 특징값인 핵의 면적, 둘레, 가로·세로(상·단축)의 길이를 구하고 핵의 원형성의 비율을 계산하였다. 30개의 영상은 training data set으로 나머지 30개는 test data set으로 사용하여 분류 정확도를 검증하였다. 100배 배율과 400배 배율의 유방 종양 세포 조직 영상을 1, 2, 3, 4 단계의 wavelet 변환을 적용한 후, 분할된 서브밴드에서 GLCM을 이용하여 질감 특징(Entropy, Energy, Contrast, Homogeneity)을 추출하였다. 이 특징값들을 조합하여 개별 분석에 의해 분류기(classifier)를 생성한 후, 총 60개의 영상 중에서 30개는 training data set으로 나머지 30개는 test data set으로 사용하여, 분류 정확도를 검증하였다.

그림 1은 분류 시스템 전체 구성도이다.

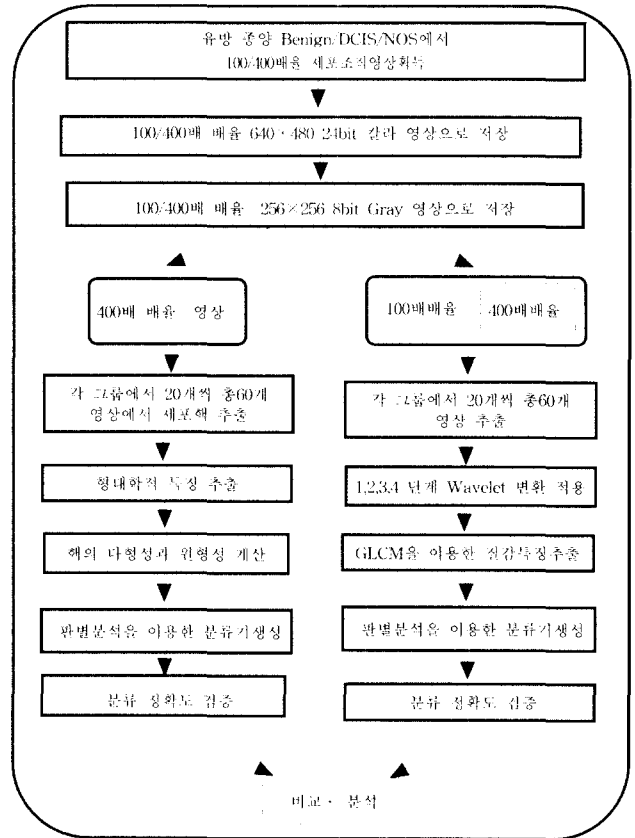


그림 1. 분류 시스템 전체 구성도

Fig. 1. The entire structure diagram of the classification system

II. 재료와 이미지 획득

본 연구에 사용된 Breast 세포조직영상들은 1999년과 2000년 유방 질환 환자의 슬라이드로 인제대학교 백병원 해부병리학교실에서 제공 받았다. Hematoxylin and eosin 염색된 양성 종양(Benign) 슬라이드와 DCIS 슬라이드,

NOS 슬라이드를 CCD 칼라 비디오 카메라(sony)가 연결되어 있는 Olympus 광학 현미경을 통해 병리 전문가가 흥미 있어 하는 영역을 100배 배율과 400배 배율에서 640×480 크기의 칼라 이미지로 획득하였으며, 256×256 크기의 8bit Gray 영상으로 저장하였다.

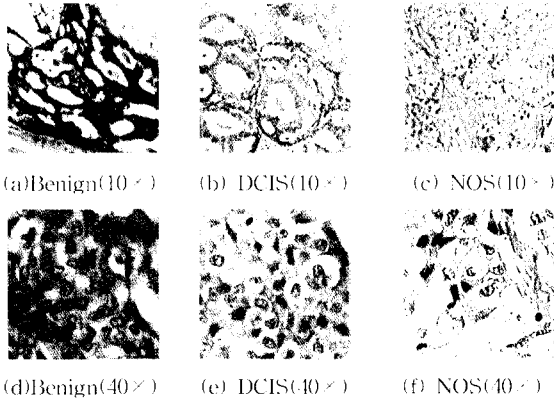


그림 2. 유방 중앙 세포 조직 영상

Fig. 2. Breast tumor cell tissue section images

III. 형태학적 특징

1. 세포핵 추출을 위한 분할

영상 분할(image segmentation)은 입력된 영상을 영상의 구성요소나 물체들의 집합으로 분리하는 과정이다. 영역 분할 방법은 경계선 기반 방법(boundary based method)과 영역 기반 방법(region based method)으로 크게 구분할 수 있다. 경계선 기반 방법은 각 화소의 위치에서 그래디언트를 사용하여 에지를 검출한 후에 그래프 탐색, 동적 프로그래밍, 하프 변환 등의 방법을 사용하여 인접한 에지를 연결하는 방법이고, 영역 기반 방법은 영상 내의 히스토그램을 사용한 임계치 방법, 신경망을 이용한 화소 분류 방법, 영역 내의 유사성 및 연결성을 이용하여 영역을 확장시켜 가는 영역 성장(region growing) 기법 등이 있다[6][7][8]. 경계선 기반 방법은 세포 영상에서 세포의 위치, 모양과 크기, 질감(texture)이 어떠한지 알 수 있으나, 세포 영상의 많은 artifact나 핵 안의 작은 점이나 핵소체(nucleolus), 세포의 그룹에 대해 많은 에지(edge)가 생성되어 영역 분할이 어렵다[9]. 본 논문에서는 각 화소들의 특성이 동일하다고 분류되는 경우 같은 영역으로 나누는 영역 기반 분할 방법을 사용하여 세포핵 영역을 추출하였다.

ROI 획득을 위해 저장된 RGB 칼라 영상을 삼색 스펙트럼에서의 빛의 분포를 감안하여 아래의 식(1)을 사용하여 8bit Gray Level 영상으로 전환하였다[10].

$$Gray = Red \times 0.35 + Green \times 0.58 + Blue \times 0.07 \quad (1)$$

히스토그램을 이용하여 영역의 분리를 위해 그림3과 같이 적절한 임계치(threshold)(120~165)를 구하여 설정하였다[8].

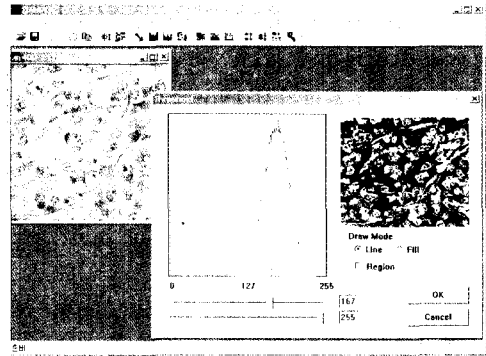


그림 3. Threshold 적용 과정

Fig. 3. The process of thresholding operation

임계치를 설정한 후 영상에서 관심영역 외의 부분 중에 같이 분할된 같은 명도 값을 가지는 부분의 제거를 위해, 팽창(dilation), 침식(erosion), 새끼(opening), 채움(closing)등의 형태연산자를 이용하여 분할하고, 세포핵 내에 구멍을 제거하기 위해 Fill hole를 반복적으로 수행하였다[11]. 다음 이 영상을 8 connectivity를 바탕으로 한 labeling을 적용한 후 각각의 객체에 대하여 형태학적 특징을 추출하였다. 추출한 형태학적 특징은 Area object의 면적, Perimeter object의 둘레, Xwidth object의 가로 길이, Ywidth object의 세로길이 이다.[6][12]

2. 세포핵 형태학적 특징 추출

유방 질환 중에서 유관에 발생하는 유방중앙을 Benign, DCIS, NOS으로 나누는 중요한 요소 중에 하나가 세포핵의 형태학적 특징이다. Benign에서 DCIS, NOS 진행 될 수록 상피세포의 핵(nucleus)이 커지며, 핵 모양이 다형성(pleomorphism)을 나타낸다.

Benign, DCIS, NOS의 세포핵 크기와 다형성의 비교를 위해, 세포핵의 형태학적 특징값인 핵의 면적, 둘레, 가로·세로(장·단축)의 길이를 구하고, 가로·세로와 둘레·넓이와의 상관관계 및 세포핵의 다형성을 아래의 식을 이용하여 계산하여 보았다[13][14].

* Morphologic Features

- 특성1

$$F_{px} = \frac{X_{width}}{P_r} \quad F_{py} = \frac{Y_{width}}{P_r}$$

- 특성2

$$F_{ax} = \frac{X_{width}}{Area} \quad F_{ay} = \frac{Y_{width}}{Area}$$

특성3

$$F_{x\pi} = \frac{x\pi}{P_r} \quad F_{y\pi} = \frac{y\pi}{P_r}$$

특성4

$$F_{\text{원}} = \frac{x^2}{4} \pi \quad F_{\text{원}} = \frac{y^2}{4} \pi$$

세포핵의 원형성(circularity)을 알아 보기 위해 원의 경 우 직을 유도하여 보면, 원의 둘레(perimeter)와 원의 면 적(area)은 각각 $Perimeter = 2\pi r$, $Area = \pi r^2$ 이다. 원 둘 레 를 세 우 하 면 $Perimeter^2 = (2\pi r)^2 = 4\pi^2 r^2 = 4\pi(\pi r^2)$ 이다. 원면적은 πr^2 이므로 위 식은 $Perimeter^2 = 4\pi Area$ 가 된다. 그러므로, $Perimeter^2 / (4\pi Area) = 1$ (2)이 된다. 만일 세포핵이 원형에서 멀 어질수록 면적에 비해 둘레가 더 많이 커지므로, 식(2)의 값이 크면 할수록 원형성은 나빠짐을 알 수 있다.

IV. Wavelet 변환을 적용한 질감 특징

1. Wavelet 변환

Wavelet 변환은 최근 들어 신호 처리나 패턴 인식 분야 에서 가장 많이 쓰이고 있는 변환 방법 중 하나이다. Wavelet 변환 방법은 공간 영역과 주파수 영역에서 고 지역적 특성을 잘 나타낸다는 특성과 다중해상도 분석 (Multiresolution Analysis)이 가능하므로, 서로 다른 해상 도로 영상을 분할하여 개별적으로 부영상을 분석할 수 있다는 특징으로 인해 질감 정보를 얻고자 할 때 많이 쓰이고 있다[15]. 이 특성은 wavelet이라고 불리는 기저 함수(base function)를 이동(translating)하고 확장 (scaling)함으로써 저대역 부밴드와 고대역 부밴드로 주파 수 영역에 따른 다해상도를 갖게 된다[16]. Wavelet 기저 함수는 식(3)과 같은데, 여기서 a 는 스케일 변수(scale variable)이고, b 는 이동 변수(translation variable)이다.

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{2}} \Psi\left(\frac{x-b}{a}\right) \quad (3)$$

2차원 영상의 wavelet 변환은 그림 3의 대역 분할 부분 에서 나타난 것과 같이 필터 뱅크 구조를 구성하여 수평 방향과 수직 방향으로 1차원 wavelet 변환을 반복적으로 수행해서 얻어진다.

아래 그림 4에서 표시된 LL, LH, HL, HH은 분해된 대 역을 나타낸다. L과 H는 각각 저주파 대역 통과 필터 (Low Pass Filter h_0)와 고주파 대역 통과 필터(High Pass Filter h_1)을 이용하여 필터링함을 의미한다. 분해 과정에서는 먼저 입력 영상을 수평 방향에 대해서 각각 저주파 대역 통과필터링과 고주파대역 통과필터링 한 후 부표본화(sub sampling)하여 저주파 대역과 고주파 대역 으로 분해한다. 그리고, 분해된 각각의 저주파 대역과 고 주파 대역에 대하여 다시 수직 방향으로 저주파 대역 통 과 필터링과 고주파 대역 통과 필터링한 후에 부표본화 하여 분해 대역 LL, LH, HL, HH을 얻게 되며, 이를 반

복적으로 적용하여 다단계 영상을 얻게 된다.

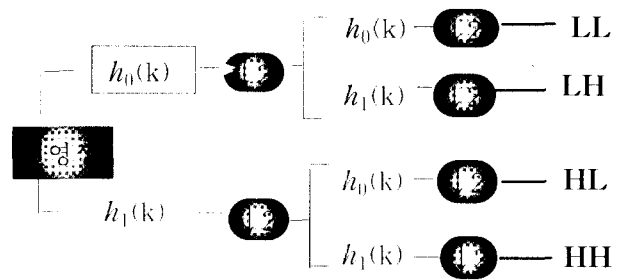


그림 4. Wavelet 변환의 다단계 분할 과정

Fig. 4. Process of the two dimensional multi level wavelet decomposition

그림 5는 2-채널 필터로 2단계 wavelet 변환한 후 이미 지의 분할된 영역이다. LL_2 는 해상도가 반으로 줄어든 고 원 영상의 블러링한 영상을 얻게되고, LH_1, LH_2 는 원 영상수평성분의 에지탐색이 적용된 영상을 얻게 되고, HL_1, HL_2 는 수직성분이, HH_1, HH_2 는 대각선 성분의 에지탐색이 적용된 영상을 얻게 된다.

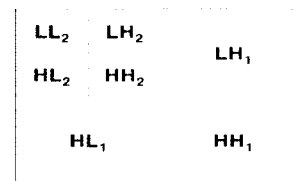
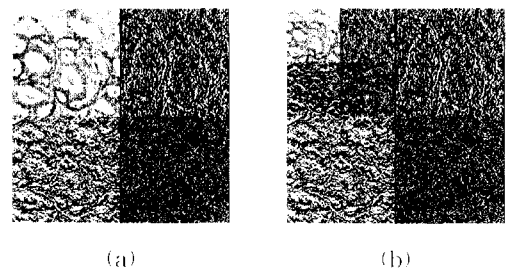


그림 5. 2단계 wavelet 분할영역에서의 주파수 성분의 위치
Fig. 5. Location of frequency bands in a sub band of 2 level wavelet transformed image

그림 6은 100배 배율의 유방 중앙 세포 조직 영상을 각 단계별 wavelet 변환이 적용된 결과들을 보여주어 있다.

Benign, DCIS, NOS의 해상도 256×256 8bit 그레이 영 상에서 각각 20개씩 총 60개의 영상에 대해 1, 2, 3, 4 단 계의 wavelet 변환을 적용하고, 분할 된 시브 영역에서 LL 영역을 제외한 저역영역 LH_1, HL_1, HH_1 총 3개 영 역(1단계), $LH_1, LH_2, HL_1, HL_2, HH_1, HH_2$ 총 6개 영



(a)

(b)

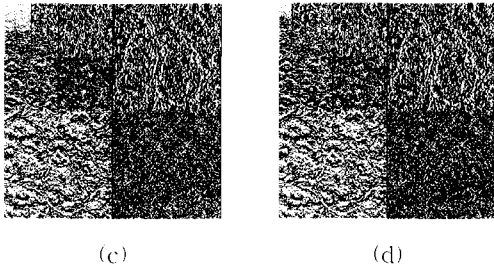


그림 6. 단계(level)별 wavelet 변환 영상의 예
(a) 1 단계 (b) 2 단계 (c) 3 단계 (d) 4 단계

Fig. 6. Examples of multi level wavelet transformed images
(a) 1 level (b) 2 level (c) 3 level (d) 4 level

역(2단계), $LH_1, LH_2, LH_3, HL_1, HL_2, HL_3, HH_1, HH_2, HH_3$ -총 9개영역(3단계), $LH_1, LH_2, LH_3, LH_4, HL_1, HL_2, HL_3, HL_4, HH_1, HH_2, HH_3, HH_4$ -총 12개 영역(4단계)에서 질감 특징값(Entropy, Energy, Contrast, Homogeneity)을 추출한다.

2. GLCM을 이용한 질감 특징 추출

영상에서의 질감 분석은 영상이 가지고 있는 질감의 상호 대비성, 질감의 방향성, 질감 부위의 규칙성을 분석하여 객관적 수치로 나타내는 것이다. 질감 특징 추출 방법은 세포 조직 영상이 거의 불규칙적인 패턴을 보이므로 통계학적 방법을 사용하였다. Joint Probability Density 함수 $f(i, j, d, \theta)$ 를 통해 Gray level의 Co-occurrence Matrix(GLCM)을 생성한 후, 두 화소간의 거리 d값은 1에 대해, 각 θ 은 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 에 대해 계산하였다. 원 영상의 256 gray level에 대해 모든 두 화소 간의 밝기값의 분포를 계산 할 경우, 크기도 큰 뿐만 아니라 계산량도 너무 많이 요구되므로, 16 gray level로 축소하여 16×16 GLCM을 생성하여 질감 특징을 추출하였다 [7]. Entropy와 Energy는 텍스처 패턴의 영향을 많이 받는 특징이고, Contrast와 Homogeneity는 그레이 레벨의 차이(different)에 영향을 많이 받는 특징이다. 추출한 질감 특징의 식은 아래와 같다[17].

- Entropy = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (PM) \log (Px(i)Py(j))$
- Energy = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (PM)^2$
- Contrast = $\sum_{n=0}^{N-1} n^2 (\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} PM)$ (단, $|i-j|=n$)
- Homogeneity = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1+n}$ (단, $|i-j|=n$)

V. 실험 결과 및 분석

1. 형태학적 특징

1.1 실험 결과 및 분석

현미경 400배 배율의 유방 종양 세포 조직 영상을 각 그룹(Benign, DCIS, NOS)에서 20개씩, 총 60개의 영상을 추출하였다. 하나의 영상에서 10개의 세포핵을 추출하여, 핵의 형태학적 특징값인 핵의 면적, 둘레, 가로·세로(상·단축)의 길이를 구하고 핵의 원형성의 비율을 계산하였다. 하나의 영상에서 10개의 세포핵을 추출해서 10개의 평균을 구한 후, 그 영상을 대표하는 값으로 사용하였다. 60개의 영상에서 추출한 세포핵의 형태학적 특징값은 아래와 같다.

표 1. Benign, DCIS, NOS 세포핵 크기
Table 1. The nucleus size of Benign, DCIS, NOS

	면적 (Area)	둘레 (Perimeter)	가로길이 (Xwidth)	세로길이 (Ywidth)
Benign	304.035	67.450	20.690	20.585
DCIS	329.840	72.875	21.595	22.680
NOS	469.010	89.205	25.630	27.775

표 2. 형태학적 특성값

Table 2. The value by morphology features

영상 특성	Benign	DCIS	NOS
F_{px}	0.30722	0.30377	0.29301
F_{py}	0.30688	0.31837	0.31999
F_{ax}	0.70463	0.66780	0.55277
F_{ay}	0.70443	0.70281	0.61061
$F_{x\pi}$	0.97185	0.95388	0.92006
$F_{y\pi}$	0.96359	0.99999	1.00477
F_{xA}	1.14448	1.13454	1.09529
F_{yA}	1.13928	1.25555	1.34582

표 3. 세포핵의 원형성

Table 3. Nuclear circularity

	Benign	DCIS	NOS
비율	1.20428	1.29999	1.42319

Benign, DCIS, NOS의 세포핵 크기 비교에 대한 결과를 보면 면적, 둘레, 가로·세로축의 길이 모두 Benign < DCIS < NOS 순으로 나타남으로 세포핵의 크기는 Benign, DCIS, NOS을 분류하는데 있어서 유의한 특징값이라 할 수 있다. 그리고, 가로·세로의 길이를 보면 Benign보다는 DCIS가 DCIS보다는 NOS가 더 세포핵의 다형성을 나타낸다.

세포핵의 다형성을 알아보기 위해 계산한 형태학적 특

성값의 식에 의한 결과값은 Benign, DCIS, NOS의 차이를 거의 보이지 않는다. 세포핵의 가로길이와 상관 관계에 있는 특징값들은 Benign>DCIS>NOS 순으로 나타나고 세로길이와 상관 관계를 가지는 특징값들은 Benign<DCIS<NOS 순으로 나타낸다.

세포핵의 원형성의 비율도 Benign<DCIS<NOS 순으로 나타남으로 Benign, DCIS, NOS을 분류하는데 있어서 유의한 특성값이라 할 수 있다. Benign이 가장 좋은 원형성을 보였다.

1.2 형태학적 특징 유의성 검증

추출한 형태학적 특징들에 대한 유의성을 평가하기 위해 통계 패키지인 SAS을 이용하여 판별 분석에 의해 분류기를 생성한 후, 분류 정확도를 검증하였다. 판별분석은 집단간의 차이를 의미있게 설명해 줄 수 있는 독립변수들을 찾아내고, 이들의 선형결합으로 판별식을 만들어 내는 것으로, 추정된 판별식이 집단간의 차이를 잘 반영하는지에 대한 검증을 할 수 있다[18].

총 60개의 영상 중에서 30개는 training data set으로 나머지 30개는 test data set으로 사용하여, 13개의 형태학적 특징값을 모두 고려하여 분류기를 생성한 후, 분류 정확도를 검증하였다. 첫번째 열은 병리전분가에 의한 분류를 말하며, 첫번째 행은 컴퓨터에 의한 분류를 뜻한다. 즉 병리전분가가 분류한 영상을 컴퓨터가 몇%나 같이 분류했는가에 대한 결과값은 표 4와 같다 (전문가가 Benign으로 분류한 20개의 영상을 컴퓨터가 Benign으로 분류한 것은 18개, DCIS로 1개를 잘못 분류했고, NOS로 1개를 잘못 분류하여 90%의 정확도를 보였다)

표 4. 형태학적 특징에 의한 분류 결과
Table 4. The result of the classification by the morphological features

	Benign	DCIS	NOS	CCR
Benign	18	1	1	90%
DCIS	5	14	1	70%
NOS	0	0	20	100%
Total	23	15	22	86.7%

13개의 형태학적 특징값이 Benign, DCIS, NOS을 분류하는데 있어서 유의한 특성값인지 알아 보기 위해 분산 분석(ANOVA)을 한 후, F-통계량(F-value)과 유의확률값(P-value)으로 판단하였다. 유의확률값을 0.001로 하였을 때는 세포핵의 면적, 둘레, 가로·세로(장·단축)의 길이와 특성2의 F_{ax} , F_{ay} 와 특성4의 F_{ya} , 원형성-8개의 특징값이 유의한 특성값이었다. 표 5는 8개의 형태학적 특징값에 따른 분류의 결과이다.

표 4와 표 5를 비교하여 볼 때, 86.7%로 동일한 결과를 보인다. 그러나 13개의 형태학적 특징값을 모두 사용하여 분류기를 생성하였을 경우 Benign에서 더 나은 결과를

보이며, 8개의 형태학적 특징값을 사용하여 분류기를 생성하였을 경우 DCIS에서 더 나은 결과를 보이고 있다.

표 5. 8개의 형태학적 특징에 따른 분류 결과
Table 5. The classification result of 8 morphological features

	Benign	DCIS	NOS	CCR
Benign	15	5	0	75%
DCIS	2	17	1	85%
NOS	0	0	20	100%
Total	17	22	21	86.7%

2. Wavelet 변환을 적용한 질감 특징

2.1 실험 결과 및 분석, 유의성 평가

위와 동일한 방법으로 추출한 질감 특징들에 대해 유의성을 평가하였다. 총 60개의 영상 중에서 30개는 training data set으로 나머지 30개는 test data set으로 사용하여, 분류 정확도를 검증하였다.

표 6. Wavelet 변환의 단계별에 기반한 질감 특징에 의한 분류 결과

Table 6. The results of classifications by texture features for different wavelet transform levels

	Benign	DCIS	NOS	CCR		Benign	DCIS	NOS	CCR
Benign	15	1	4	75%	Benign	12	8	0	60%
DCIS	3	12	5	60%	DCIS	2	18	0	90%
NOS	3	1	16	80%	NOS	0	2	18	90%
Total	21	14	25	71.7%	Total	14	28	18	80.0%

(a) 1-level (10x)

(b) 1-level (40x)

	Benign	DCIS	NOS	CCR		Benign	DCIS	NOS	CCR
Benign	15	5	0	75%	Benign	19	1	0	95%
DCIS	1	19	0	95%	DCIS	3	16	1	80%
NOS	0	0	20	100%	NOS	1	1	18	90%
Total	16	24	20	90.0%	Total	23	18	19	88.3%

(c) 2-level (10x)

(d) 2-level (40x)

	Benign	DCIS	NOS	CCR		Benign	DCIS	NOS	CCR
Benign	18	2	0	90%	Benign	18	1	1	90%
DCIS	0	20	0	100%	DCIS	2	18	0	90%
NOS	0	0	20	100%	NOS	0	1	19	95%
Total	18	22	20	96.7%	Total	20	20	20	91.7%

(e) 3-level (10x)

(f) 3-level (40x)

	Benign	DCIS	NOS	CCR		Benign	DCIS	NOS	CCR
Benign	18	1	1	90%	Benign	18	1	1	90%
DCIS	0	20	0	100%	DCIS	2	17	1	85%
NOS	0	0	20	100%	NOS	0	0	20	100%
Total	18	21	21	97.7%	Total	20	18	22	91.7%

(g) 4-level (10x)

(h) 4-level (40x)

표 6에서 wavelet 변환 단계에 따른 분류 결과를 살펴 보면, 1-level wavelet 변환과 2-level wavelet 변환을 적용한 결과 보다는 3-level wavelet 변환을 적용한 결과가 훨씬 좋은 분류를 보였고, 3-level wavelet 변환과 4-level wavelet 변환을 적용한 결과는 비슷한 %의 분류 결과를 보였다. 따라서 Benign, DCIS, NOS를 분류하려면 최소 3-level 이상의 wavelet 변환을 적용시켜야 함을 알 수 있었다.

100배 배율의 유방 중앙의 세포 조직 영상과 400배 배율의 유방 세포 조직 영상을 비교 해 보면 100배 배율의 세포 조직 영상이 96.7%로 400배 배율의 세포 조직 영상 91.7%보다 더 나은 분류 결과를 보임을 알 수 있다. 앞에서 말한 것과 같이 이는 현미경 100배 배율 슬라이드에서는 세포 하나 하나 보다는 유관 구조를 중심으로 영상을 볼 수 있고, 400배 배율의 조직 영상에서는 100배 배율과는 반대로 세포 하나 하나를 중심으로 영상을 보게 되므로, 400배 배율의 영상을 wavelet 변환을 적용하면 암의 Grading을 분석하는데 좋은 결과를 보임을 알 수 있었고[19], 100배 배율의 영상을 wavelet 변환을 적용하면 Benign, DCIS, NOS를 분류하는데 좋은 결과를 보임을 알 수 있었다.

3. 생성된 분류기 비교 · 분석

Benign, DCIS, NOS를 분류하기 위한 같은 현미경 400배 배율의 세포핵의 형태학적 특징으로 생성된 분류기와 3단계 wavelet 변환을 적용 후 질감 특징으로 생성된 분류기를 비교하면 86.7%와 91.7%로 3단계 wavelet 변환을 적용 후 질감 특징으로 생성된 분류기가 좋은 결과를 나타냈다. 같은 방법으로 3단계 wavelet 변환을 적용 후 질감 특징으로 생성된 분류기를 비교하면 100배 배율의 분류기는 96.7%를, 400배 배율의 분류기는 91.7%로 100배 배율의 분류기가 더 좋은 결과를 보였다.

Benign, DCIS, NOS를 분류하는데는 형태학적 특징으로 생성된 분류기보다 wavelet 변환을 적용 후 질감특징으로 생성된 분류기가 더 유용하고, 현미경 400배 배율에서 생성한 분류기 보다 100배 배율의 세포 조직 영상을 3단계 wavelet 변환을 적용 후 질감 특징으로 생성한 분류기가 더 유용함을 알 수 있었다.

VI. 결론 및 향후 과제

유방 질환 중에서 유관에 발생하는 유방종양을 Benign, DCIS, NOS로 분류하기 위해 현미경 100배 배율과 400배 배율의 유방 중앙 세포 조직 영상 각 그룹(Benign, DCIS, NOS)에서 20개씩, 총 60개의 영상에서 실험하였다.

400배 배율의 영상에서 세포핵을 추출하여 형태학적 특징값을 추출하여 분류기를 생성하고, 같은 배율에서 질감 특징값을 비교하기 위해 wavelet 변환을 적용하여 특징

값을 추출하여 분류기를 생성하여 두 분류기를 비교 하였다. 그리고, 100배 배율의 영상을 wavelet 변환을 적용, 특징값을 추출하여 분류기를 생성하여 400배 배율에서 같은 방법으로 생성된 분류기와 비교 하였다.

같은 400배 배율에서는 3단계 wavelet 변환을 적용하여 추출한 질감 특징값으로 생성한 분류기가 91.7%로 형태학적 특징값으로 생성한 분류기 86.7%로 더 좋은 결과를 보였다. 3단계 wavelet 변환을 적용하여 추출한 질감 특징값으로 생성한 두가지 분류기 중에 100배 배율의 분류기 96.7%로 같은 방법의 400배 배율 분류기 91.7%보다 좋은 결과를 보였다. 따라서 Benign, DCIS, NOS를 분류하려면 최소 3단계 이상의 wavelet 변환을 적용하여 추출한 질감 특징값으로 생성된 분류기를 이용하며, 400배 배율의 세포 조직 영상 보다는 100배 배율의 세포 조직 영상에서 분류기를 생성하는 것이 더욱 나은 결과를 얻을 수 있다.

향후 과제로는 형태학적 특징으로 분류기의 분류 정확도를 높이는 방법에 대한 연구와 wavelet 변환을 적용한 분류기에서 최적의 특징값들을 선택하여 더 좋은 분류 결과를 보이는 분류기를 생성하는 것이다. 그리고 이 두가지 특징을 조합하여 최적의 분류기를 생성하는 것도 연구되어야 할 것이다. Gray 영상뿐만 아니라 RGB 컬러 영상에서 적용 가능한 분류기를 생성하여 병리 진단의 보조도구로 유용하게 사용될 수 있는 연구가 필요하다.

접수일자 : 2001. 9. 20

수정완료 : 2001. 10. 24

본 논문은 2000년도 인재대학교 학술연구조성비 지원에 의하여 연구되었음 (This work was supported by the 2000 Inje University reserch grant).

참고문헌

- [1] 대한병리학회, "병리학(PATHOLOGY)," 고문사, 2000
- [2] 김동석, 이수정, "유방질환의 진단병리(Diagnostic Pathology of the Breast)", 아카데미아, 1998
- [3] Heung-Kook Choi, Janos Vasko, Ewert Bengtsson, Torsten Jarkrans, Per-Uno Malmstrom, Kenneth Wester, and Christer Busch, "Grading of Transitional Cell Bladder Carcinoma by Texture analysis of Histological Section," *Analytical Cellular Pathology* Vol.6, pp.327-343, 1994
- [4] H. Tamura, S. Mori and T. Yamawaki, "Textures corresponding to visual perception," *IEEE TransSyst. Man Cybern.* Vol. SMC-8 No. 6, pp.460-473, 1978
- [5] 서상환, 이연숙, 김상균, 김홍식, "웨이브릿 변환에 기반한 색상과 질감 특징을 이용한 이미지 검색 시스템," 1999년도 한국정보과학회학술발표논문집 Vol. 26. No. 2, pp.30-32, 1999

[6] 최현주, 최민권, 최홍국, 김상균, 최항목, 박세명, "칼라유방암조직영상에서 질감 특징과 신경회로망을 이용한 양성세포핵과 음성세포핵의 자동 분할," 정보과학회 학술발표논문집, Vol. 26, No. 2, pp.422-424, 1999

[7] 최예찬, 김호영, 김선아, 김백집 "세포영상에서의 텍스처 특징을 이용한 영역 분할," 1999년도 한국정보과학회 봄 학술발표논문집 Vol. 26, No. 1, pp.576-578, 1999

[8] 김재환, 하진영, 김백집, 김호성, "자강경부암 세포 영상 분할을 위한 Threshoding 기법," 한국정보과학회 학술발표논문집, Vol. 26, No. 2, pp.419-421, 1999

[9] 은성경, 박찬모, 박화순, 윤소영, 조민선, 조수연, 김성숙, "영상치료를 이용한 자강경부 세포질의 자동탐색 방법에 관한 연구," 대한세포병리학회지, Vol. 5, No. 1, pp.15-22, 1994

[10] 최홍국, "ML 분류를 사용한 유방암 항체 조직 영상 분할," 멀티미디어논문지, pp108-115, Vol.4, No.2 2001

[11] Scott E Umbugh, Ph.D., "Computer Vision and Image Processing," Prentice Hall International, Inc, pp.79-148

[12] Earl Gose, Richard Johnsonbaugh, and Steve Jost, "Pattern Recognition and Image Analysis," Prentice Hall, pp. 372-379, 1996

[13] 황해길, 최현주, 이병일, 최홍국, "폐암변이 분석을 위한 영상 특성 추출," 한국정보과학회 가을 학술발표논문집 Vol. 27, No. 3, pp.473-475, 2000

[14] 황해길, 최현주, 윤혜경, 최홍국, "유방질환 진단을 위한 유방 중앙 세포핵의 형태학적 특성 비교 분석," 2001년도 한국멀티미디어 춘계학술발표논문집, Vol. 4, No. 1, pp.91-94, 2001

[15] 최현주, 이병일, 이인숙, 최홍국, "Wavelet 변환에 기반한 암세포 조직 영상의 질감 분석," 2000년도 한국 신호처리 · 시스템학회 하계 학술발표논문집, Vol. 1, No 2, pp.53-56, 2000

[16] 김준복, 정원용, "강사 투영을 사용한 컬러 이미지 스케일링," 2000년 한국 신호처리 · 시스템학회 추계 학술발표논문집, Vol. 1, No 1, pp.305-308, 2000

[17] Robert M. Haralick, K. Shanmugam, and Its' Hak Dinstein, "Textural Features for Image Classification," *IEEE Trans. On System, Man, and Cybernetics*, Vol. SMC 3, No. 6, pp. 610-624, 1973

[18] Richard A. Johnson, Dean W. Wichern, "Applied Multivariate Statistical Analysis," Prentice Hall Inc. pp.629-725, 1998

[19] Hun-Ju Choi, Heung-Kook Choi, "Classification of Bladder Carcinoma Cell Tissue Sections by Image Analysis," 제 5회 한국 집단 의료영상처리 워크샵 발표 논문집, 2001



황해길(Hae Gil Hwang)
 准會員
 2000년 인제대학교 전산학과
 2000년~현재 인제대학교 전산학과
 대학원 석사과정
 관심분야 : 영상처리 및 분석



최현주(Hyun Ju Choi)
 准會員
 1995년 인제대학교 전산학과
 2001년 인제대학교 전산학과
 (이학석사)
 2001년~현재 인제대학교 전산학과
 대학원 박사과정
 관심분야 : 의료영상처리 및 분석,
 패턴인식



윤혜경(Hye Kyoung Yoon)
 1982년 부산대학교 의과대학 졸업
 1990년 부산대학교 대학원 의학박사
 1993년 미국 플로리다대학
 부속병원연수
 1999년 미국 콜리플랜드 칼리닉
 해부병리과 연수
 1991년~현재 부산 백병원
 해부병리학 교수

전문분야 : 부인과 병리, 유방병리, 유방세침흡인병리



남상희(Sang-Hee Nam)
 1978년 영남대 물리학과
 1980년 영남대 물리학과 방사선물리학
 (이학석사)
 1988년 영남대 물리학과 방사선물리학
 (이학박사)
 1992년~현재 인제대학교 의용공학과
 교수

관심분야 : 아날로그/디지털 영상신호 분석,
 디지털 방사선 검출기



최홍국(Heung-Kook Choi)

正會員

1988년 Linköping University,
Computer Engineering
Linköping, Sweden (공학사)

1990년 Linköping University,
Computer Engineering
Linköping, Sweden (공학석사)

1996년 UppsalaUniversity, Computerized Image
Analysis Uppsala, Sweden (공학박사)

1997년~현재 인제대학교 정보컴퓨터공학부 조교수

관심분야 : 멀티미디어, 컴퓨터그래픽스,
영상처리 및 분석
