

# 흉부 CT 영상에서 다중 뷰 영상과 텍스처 분석을 통한 고형 성분이 작은 폐 간유리음영 결절 분류

이선영<sup>†</sup>, 정주립<sup>\*\*</sup>, 이한상<sup>\*\*\*</sup>, 홍헬렌<sup>\*\*\*\*</sup>

## Classification of Ground-Glass Opacity Nodules with Small Solid Components using Multiview Images and Texture Analysis in Chest CT Images

Seon Young Lee<sup>†</sup>, Julip Jung<sup>\*\*</sup>, Han Sang Lee<sup>\*\*\*</sup>, Helen Hong<sup>\*\*\*\*</sup>

### ABSTRACT

Ground-glass opacity nodules(GGNs) in chest CT images are associated with lung cancer, and have a different malignant rate depending on existence of solid component in the nodules. In this paper, we propose a method to classify pure GGNs and part-solid GGNs using multiview images and texture analysis in pulmonary GGNs with solid components of 5mm or smaller. We extracted 1521 features from the GGNs segmented from the chest CT images and classified the GGNs using a SVM classification model with selected features that classify pure GGNs and part-solid GGNs through a feature selection method. Our method showed 85% accuracy using the SVM classifier with the top 10 features selected in the multiview images.

**Key words:** Pulmonary Nodule Classification, Ground-glass Opacity Nodule, Texture Analysis, Machine Learning, Support Vector Machine

### 1. 서 론

간유리음영 결절(ground-glass opacity nodule; GGN)은 흉부 컴퓨터단층촬영(computed tomography; CT) 영상에서 폐 실질보다 상대적으로 밝기 값이 높고 흐릿한 영역으로 나타나며 결절 내부에 고형 성분(solid component)을 포함하지 않고 간유

리음영 성분(ground-glass opacity; GGO)만을 가지는 순수 간유리음영 결절(pure GGN)과 결절 내부에 고형 성분과 간유리음영 성분을 모두 포함하는 혼합 간유리음영 결절(part-solid GGN)로 분류된다[1]. 혼합 간유리음영 결절은 순수 간유리음영 결절보다 악성도가 높은 것으로 알려져 있고, 폐암의 조기발견 과도 관련이 있기 때문에 간유리음영 결절을 분류하

\* Corresponding Author : Helen Hong, Address: (01797) 621, Hwarang-ro, Nowon-gu, Seoul, Republic of Korea, TEL : +82-2-970-5756, FAX : +82-2-970-5981, E-mail : hlhong@swu.ac.kr

Receipt date : Mar. 13, 2017, Revision date : June 20, 2017  
Approval date : June 21, 2017

<sup>†</sup> Dept. of Software Convergence, Seoul Women's University (E-mail : seonyounglee@gmail.com)

<sup>\*\*</sup> Dept. of Software Convergence, Seoul Women's University (E-mail : jjulip@swu.ac.kr)

<sup>\*\*\*</sup> School of Electrical Engineering, KAIST (E-mail : hansanglee@kaist.ac.kr)

<sup>\*\*\*\*</sup> Dept. of Software Convergence, Seoul Women's University

\* This research was supported by a research grant from Seoul Women's University(2017) and the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW(2016-0-00022) supervised by the IITP(Institute for Information & communications Technology Promotion) (2016-0-00022). We would like to thank to Prof. Chang Min Park and Jong Hyuk Lee at Seoul National University Hospital to provide Chest CT dataset.

는 것이 중요하다[2-3]. 그러나 고형 성분의 크기가 작은 혼합 간유리음영 결절의 경우 순수 간유리음영 결절과 유사하게 보이기 때문에 두 결절을 분류하는 것이 어렵다. Fig. 1은 흉부 CT 영상에서 나타나는 폐 결절 유형별 영상으로 Fig. 1(a)의 고형 결절에 비해 Fig. 1(b)의 순수 간유리음영 결절은 간유리음영 성분이 상대적으로 어두운 밝기값을 보이고, Fig. 1(c)의 혼합 간유리음영 결절은 흐릿하게 나타나는 간유리음영 성분과 내부에 밝게 나타나는 고형 성분을 함께 가지고 있다. Fig. 1(d)의 혼합 간유리음영 결절은 내부에 작은 고형 성분을 포함하고 있으나 그 크기가 작아 Fig. 1(b)의 순수 간유리음영 결절과 유사한 밝기값 특성을 보인다.

2D 단면 상에서 결절의 텍스처 정보를 이용한 연구는 폐 결절에 대한 연구와 폐 간유리음영 결절에 대한 연구로 나누어 볼 수 있다. 폐 결절의 특성을 분석하고 분류한 기존 연구는 다음과 같다. Shewaye 등은 폐 결절의 기하학적 특징과 텍스처 특징을 이용하여 악성 폐 결절과 양성 폐 결절을 분류하였다[4]. Kumar 등은 인공 신경망의 종류 중 하나인 오토인코더(autoencoder)로부터 추출된 딥 피쳐(deep feature)를 이용하여 악성 폐 결절과 양성 폐 결절을 분류하였다[5]. Dilger 등은 폐 결절에서 추출된 결절의 형태 정보와 텍스처 특징, 폐 결절 주변의 폐 실질에서 추출된 경계 정보와 텍스처 특징 그리고 폐 내 결절의 위치, 폐엽 등의 정보를 이용하여 악성 폐 결절과 양성 폐 결절을 분류하였다[6]. Kamiya 등은 폐 결절 영역의 히스토그램으로부터 얻어진 텍스처 특징인 침도, 왜도를 이용하여 악성 폐 결절과 양성 폐 결절을 분류하였다[7].

순수/혼합 간유리음영 결절의 특성과 악성도 간 관계를 분석하고 분류한 기존 연구는 다음과 같다.

Chae 등은 간유리음영 결절 영역의 밝기값 평균, 표준편차와 침도, 왜도, 엔트로피, 백분위 CT 밝기값 등의 텍스처 특징, 부피, 지름, 표면적 등의 체적 특징 그리고 구형도, 경도 등 형태학적 특징과 명암도 동시발생 행렬 기반의 텍스처 특징을 이용하여 간유리음영 결절의 특징을 분석하였다[8]. Mattonen 등은 간유리음영 결절의 텍스처 특징을 이용하여 폐암 발생 가능성에 대해 분석하였다[9]. Son 등은 간유리음영 결절의 크기, 부피, 백분위 CT 밝기값, 엔트로피, 균일성 등 결절의 형태 정보와 텍스처 특징을 이용하여 악성 병변으로 진단된 간유리음영 결절의 특성을 분석하였다[10,11]. Lee 등은 간유리음영 결절의 밝기값 평균, 표준편차, 침도, 왜도, 엔트로피 그리고 명암도 동시발생 행렬을 이용한 텍스처 특징을 이용하여 순수 간유리음영 결절과 혼합 간유리음영 결절을 분류하였다[12]. 대부분의 폐 결절 연구에서는 텍스처 특징 정보와 형상 정보를 이용해 악성 폐 결절과 양성 폐 결절을 분류하였다. 간유리음영 결절 연구의 경우 대부분 특징 분석만 수행하였으며 혼합 간유리음영 결절 내부의 고형 성분이 5mm 미만으로 작은 경우에 대해서는 거의 다루고 있지 않다. 또한, 고형 성분의 크기가 작은 결절의 경우 2D 대표 단면을 사용하더라도 고형 성분을 표현하는 정보량이 적어 순수 간유리음영 결절과의 분류 정확도가 낮을 수 있다는 한계점이 있다.

본 논문에서는 5mm 이하의 작은 고형 성분을 가지는 혼합 간유리음영 결절과 순수 간유리음영 결절을 분류하기 위해 다중 뷰 영상을 생성하여 여러 2D 단면의 정보를 고려한 텍스처 분석과 주요 특징 선별을 통해 분류하는 방법을 제안한다. 다중 뷰 영상을 이용하여 하나의 대표 단면만 고려할 때 보다 결절의 모양, 위치 등의 정보를 다양하게 고려하고, 특징 선

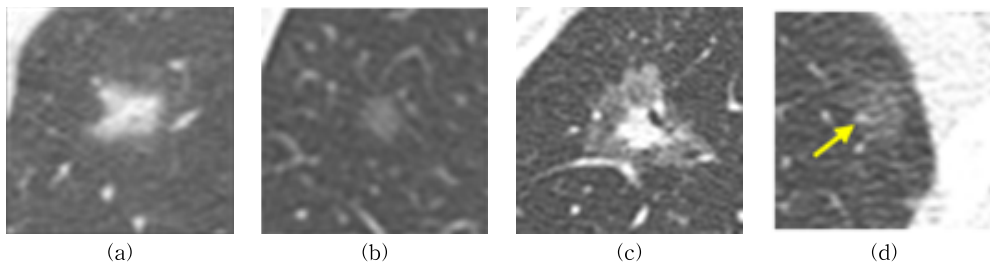


Fig. 1. Examples of GGNs in chest CT images. (a) solid nodule, (b) pure GGN, (c) part-solid GGN, and (d) part-solid GGN with small solid component.

별 기법을 이용하여 결절 타입을 분류하는데 의미 있는 특징값을 선별한다.

본 논문의 구성은 다음과 같다. 2장에서는 흉부 CT 영상에서 폐 실질 영역으로부터 간유리음영 결절을 분할하고 분할된 결절 영역의 다중 뷰 영상을 생성하는 방법, 생성된 다중 뷰 영상으로부터 특징벡터를 추출 및 선별하는 방법, 머신 러닝 기반의 분류기를 이용하여 간유리음영 결절을 분류하는 방법에 대해 설명한다. 3장에서는 제안방법을 실험한 결과 및 오분류 데이터에 대해 분석하고, 4장에서 제안방법의 효과와 결과를 요약하며 결론을 맺는다.

## 2. 제안방법

본 논문에서는 간유리음영 결절 영역의 텍스처 분석 및 분류를 위해 간유리음영 결절 자동분할을 수행하고, 분할된 볼륨 데이터로부터 다양한 정보를 이용하기 위하여 다중 뷰 영상을 생성하고, 생성된 다중 뷰 영상으로부터 결절영역의 특징벡터를 추출 및 선별하여 머신 러닝 기반의 분류기를 이용한 결절 분류를 수행한다. Fig. 2는 제안하는 폐 간유리음영 결절 분류 알고리즘의 순서도이다.

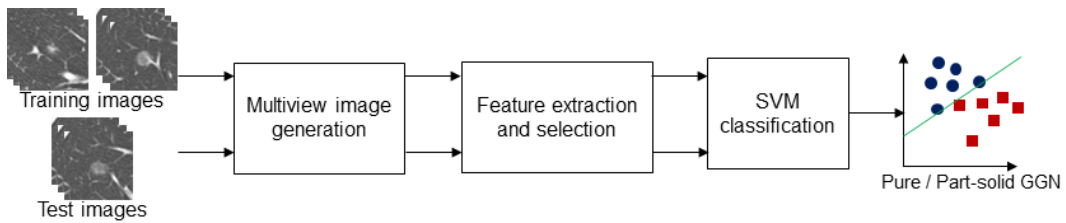


Fig. 2. A pipeline of the proposed method.

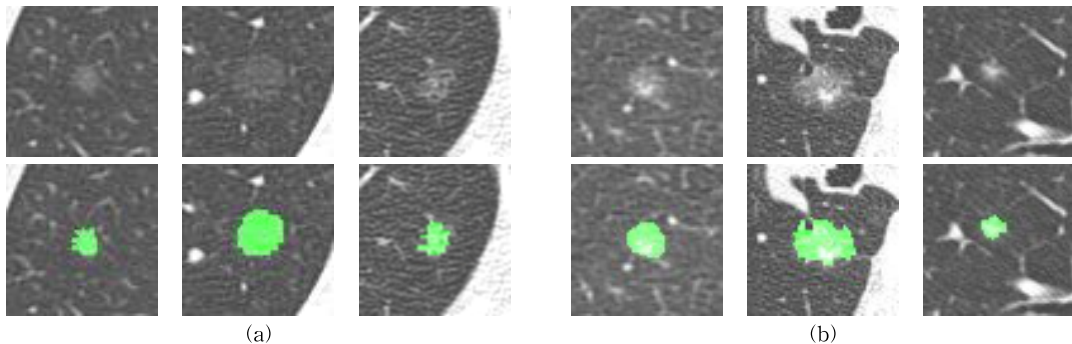


Fig. 3. Examples of segmented GGNs. (a) pure GGNs, and (b) part-solid GGNs.

## 2.1 간유리음영 결절 분할 및 다중 뷰 영상 생성

간유리음영 결절의 정보를 이용하여 순수 간유리음영 결절과 혼합 간유리음영 결절의 특성을 분석하기 위해 흉부 CT 영상으로부터 간유리음영 결절 영역을 분할하고 분할된 결절 영역의 다양한 평면으로부터 특징을 분석하기 위해 다중 뷰 영상을 생성한다.

간유리음영 결절을 분할하기 위하여 밝기값 히스토그램 모델링과 밝기값 제약 기반의 변형모델을 제안한 Jung의 방법을 이용하여 결절의 자동분할을 수행한다[13]. 첫째, 간유리음영 결절은 간유리음영 성분의 흐린 정도에 따라 밝기값이 다양하기 때문에 각 데이터마다 적합한 임계값을 산정하기 위해 밝기값 히스토그램을 이용한 적응적 임계값을 산정하여 초기 간유리음영 결절 영역을 추출한다. 둘째, 흐릿한 간유리음영 성분과 폐 실질을 잘 분리할 수 있도록 밝기값 제약을 사용해 제안한 변형모델을 이용하여 간유리음영 결절을 최종 분할한다. Fig. 3은 흉부 CT 영상에서 순수 간유리음영 결절과 혼합 간유리음영 결절을 분할한 결과이다.

분할된 결절 영역의 정보를 보다 많이 이용하기 위하여 축상면 뷰 영상(axial view image) 외에 다중 뷰 영상을 생성한다. 다중 뷰 영상은 Fig. 4와 같이

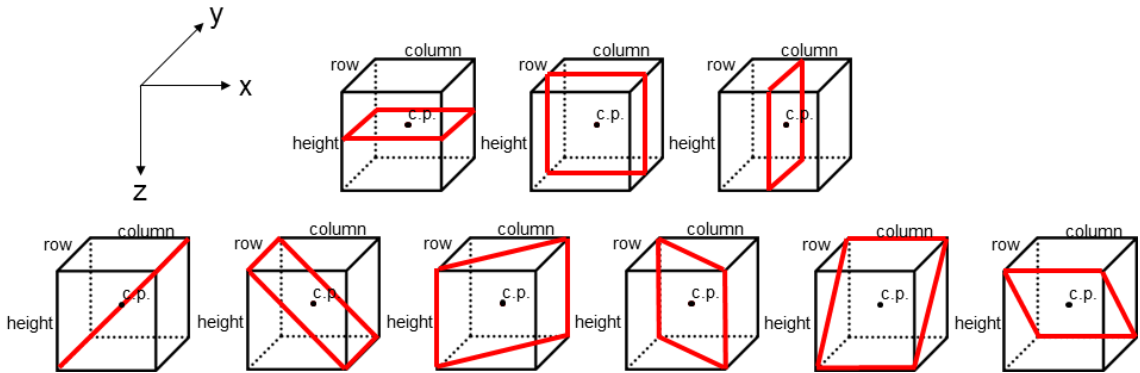


Fig. 4. Examples of generated multiview images.

왼쪽 상단에서부터 오른쪽 하단으로  $xy$ ,  $zx$ ,  $yz$ 축을 고려하는 3개의 2D 축상면, 관상면, 시상면 평면과  $xyz$ 축을 모두 고려하는 6개의 평면들로 총 9개의 영상으로 이루어져 있다. 간유리음영 결절의 무게중심을 중앙점(center point; c. p.)으로 하는 정육면체를 생성한다. 각 평면을  $xy-z(0^\circ)$ 와 같이 두 축으로 이루어진 평면과 다른 한 축과의 각으로 나타내면 Fig. 4에서 위 3개의 2D 평면들은 축상면 평면을  $xy-z(0^\circ)$ , 관상면 평면을  $zx-y(0^\circ)$ , 시상면 평면을  $yz-x(0^\circ)$ 로 나타낼 수 있고, 아래 6개의 2.5D 평면들은 왼쪽부터 오른쪽으로 순서대로  $yz-x(45^\circ)$ ,  $yz-x(135^\circ)$ ,  $zx-y(45^\circ)$ ,  $zx-y(135^\circ)$ ,  $xy-z(45^\circ)$ ,  $xy-z(135^\circ)$ 으로 나타낼 수 있다.

이처럼 간유리음영 결절 영역을 폐 실질 영역으로부터 분할하고, 분할된 간유리음영 결절로부터 다중

뷰 영상을 생성함으로써 간유리음영 결절의 모양이나 위치의 영향을 고려하여 결절의 특성을 분석할 수 있다.

### 2.2 특징벡터 추출 및 선별

생성된 다중 뷰 영상으로부터 결절 타입을 분류하는데 의미 있는 특징을 찾아내기 위하여 각 영상별 특징벡터를 추출 및 선별한다. 각 영상별 추출되는 특징벡터는 Table 1과 같이 픽셀의 밝기값을 고려하는 8개 히스토그램 특징벡터, 12개 초과임계값(above thresholding value), 5개 백분위 CT 밝기값(percentile CT attenuation value), 픽셀의 밝기값과 위치정보를 고려하는 명암도 동시발생 행렬(gray level co-occurrence matrix; GLCM) 기반의 56개 Haralick 특징벡터 그리고 픽셀의 밝기값과 연속성

Table 1. Features extracted from GGN

No.	Categories	Features
1-8	Histogram	Mean, Variance, Skewness, Kurtosis, Energy, Entropy, Min value(HU), Max value(HU)
9-20	Threshold	Above thresholding value(HU): -1000, -900, -800, -700, -600, -500, -400, -300, -200, -100, 0, 100
21-25	Percentile	Percentile CT attenuation value: 2.5th, 25th, 50th, 75th, 97.5th
26-169	Texture	GLCM(16/32level, degree: $0^\circ/45^\circ/90^\circ/135^\circ$ , distance: 1): Energy, Correlation, Contrast, Absolute value, Inverse difference, Entropy, Maximum probability GLRLM(16/32level, degree: $0^\circ/45^\circ/90^\circ/135^\circ$ ): Short run emphasis(SRE), Long run emphasis(LRE), Gray level non-uniformity(GLN), Run length non-uniformity(RLN), Run percentage(RP), Low gray-level run emphasis(LGRE), High gray-level run emphasis(HGRE), Short run low gray-level emphasis(SRLGE), Short run high gray-level emphasis(SRHGE), Long run low gray-level emphasis(LRLGE), Long run high gray-level emphasis(LRHGE)

을 고려하는 명암도 작용길이 행렬(gray level run-length matrix; GLRLM) 기반의 88개 특징벡터로 총 169개이다[14,15]. 텍스처 특징을 추출하기 전에 -3071~1024(HU) 밝기값 범위의 영상을 0~255 범위의 8비트 값으로 정규화하고, GLCM 특징과 GLRLM 특징 계산 시에는 영상의 밝기값 레벨을 16 레벨과 32레벨로 정규화하여 특징을 추출한다.

9개 다중 뷰 영상으로부터 각각 169개씩 추출된 총 1521개의 특징 중 결절의 타입을 분류하는 데 중요한 특징을 찾아내기 위하여 ReliefF 특징 선별 기법을 이용하여 특징을 선별한다[16]. 특징 선별 기법에는 ReliefF, sequential forward selection(SFS) 등이 있으나 ReliefF 기법이 이진 분류에서 많이 쓰이는 방법 중 하나이고, 다항식 차수가 낮아 시간 효율성이 높으며, 잡음에 예민하게 반응하지 않는 특성이 있기 때문에 본 논문에서는 ReliefF 기법을 적용하여 상위 특징을 선별한다. ReliefF 특징 선별 기법에서는 분류에 우수한 특징일수록 높은 가중치를 가지기 때문에 높은 가중치를 가지는 특징 순으로 상위 특징을 선별 할 수 있다. Fig. 5는 특징벡터 히트맵(heatmap)으로 추출된 특징의 수치를 시각화하여 보여주며, 히트맵 오른쪽의 컬러바는 특징의 수치를 0부터 1까지 정규화한 후 대치된 색을 보여준다. 분류하려는 데이터 그룹 간 특징이 다르게 나타날수록 그 특징을 이용했을 때 분류 성능이 우수하므로 분류하려는 데이터 그룹 간 색이 다르게 나타날수록 분류 성능에 우수한 특징이라는 것을 시각적으로 확인 할 수 있다. Fig. 5(a)는 시상면 영상에서 추출된 전체 169개 특징의 히트맵이고, Fig. 5(b)는 ReliefF 특징 선별 기법을 이용하여 전체 특징으로부터 선별된 의

미 있는 상위 10개 특징의 히트맵이다. X축은 특징벡터를 나타내므로 Fig. 5(a)에서 시상면 영상에서 추출된 총 169개의 특징과 Fig. 5(b)에서 전체 특징으로부터 ReliefF 특징 선별 기법을 이용해 선별된 총 10개의 특징을 나타내고, Y축은 데이터를 나타내는데 상위 10개 데이터는 순수 간유리음영 결절이고 하위 10개 데이터는 혼합 간유리음영 결절이다. Fig. 5(b)가 상대적으로 Fig. 5(a)보다 순수/혼합 간유리음영 결절 간 특징 대조가 상대적으로 극명하게 나타나는 것을 알 수 있다. 이는 특징 선별 기법을 이용하면 결절을 분류하는데 의미 있는 특징들을 찾아낼 수 있음을 보여준다.

### 2.3 SVM 기반의 간유리음영 결절 분류

간유리음영 결절을 분류하기 위하여 서포트 벡터 머신(support vector machine; SVM) 분류기를 이용한다. SVM은 주어진 특징벡터를 이용해 학습하며 데이터를 서로 다른 클래스로 나누는 초평면을 만드는 이진 분류기이다[17]. 학습 데이터로 분류기를 학습시킨 후, 테스트 데이터를 이용해 분류기의 분류 성능을 확인 할 수 있다[18]. 커널 함수로는 원형기반 함수(radial basis function; RBF)를 사용해 분류 모델의 학습을 수행한다.

### 3. 실험 및 결과

제안방법을 평가하기 위하여 10개 순수 간유리음영 결절과 10개 혼합 간유리음영 결절 총 20개의 간유리음영 결절이 나타나는 흉부 CT 영상을 이용해 실험하였고, 혼합 간유리음영 결절의 경우 고형성분

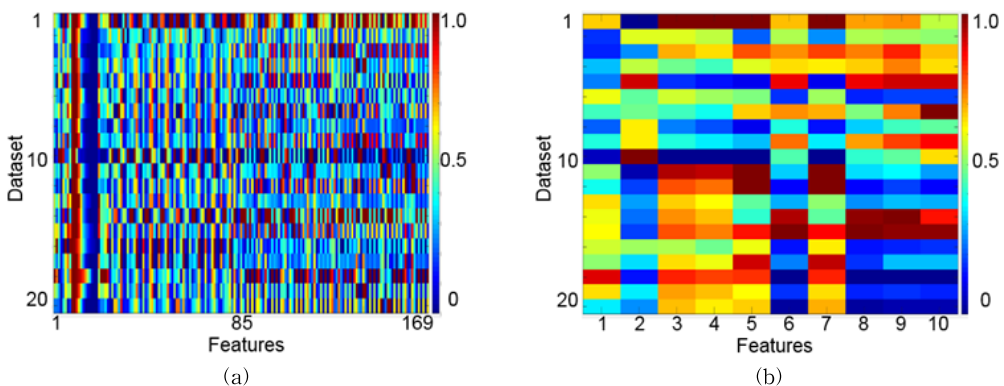


Fig. 5. Examples of feature selection. (a) heatmap of all features, and (b) heatmap of selected features.

의 직경이 5mm 이하인 결절만 전문의에 의해 선별되었다. 실험 영상으로는 서울대학교병원 영상의학과에서 SIEMENS Sensation 16 CT Scanner와 Philips Brilliance 64 CT Scanner로 촬영한 데이터를 사용하였으며, SIEMENS CT Scanner에서 사용된 커널은 B30f, B50f, B60f이고 Philips CT Scanner에서 사용된 커널은 YC이다. 영상 해상도는 512×512, 화소 크기는 0.50×0.50mm~0.76×0.76mm, 슬라이스 간격은 1mm이다.

분류기의 학습 및 검증을 위해 학습 데이터집합과 테스트 데이터집합으로 단일잔류(leave-one-out) 교차검증(cross validation) 방법을 통해 SVM 분류 모델의 학습 및 테스트를 수행하고, 분류 정확도 검증을 위하여 식 (1)을 이용해 혼합 간유리음영 결절 기준으로 정확도(accuracy), 민감도(sensitivity), 특이도(specificity), 양성예측도(positive predictive value, PPV) 그리고 음성예측도(negative predictive value, NPV)를 측정한다.

$$Accuracy = \frac{TP + TN}{TN + FP + FN + TP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$PPV = \frac{TP}{TP + FP}$$

$$NPV = \frac{TN}{TN + FN}$$

이 때, TP는 실제 양성인 실험결과 양성으로 옳게 예측된 샘플의 수, TN은 실제 음성이 실험결과 음성으로 옳게 예측된 샘플의 수, FP는 실제 음성이 실험결과 양성으로 잘못 예측된 샘플의 수 그리고 FN은 실제 양성인 실험결과 음성으로 잘못 예측된 샘플의 수이다.

Fig. 6은 특징벡터 히트맵으로 Fig. 6(a)-(d)는 순서대로 축상면 뷰 영상에서 추출된 전체 169개 특징의 히트맵, 축상면 뷰 영상에서 추출된 전체 특징들

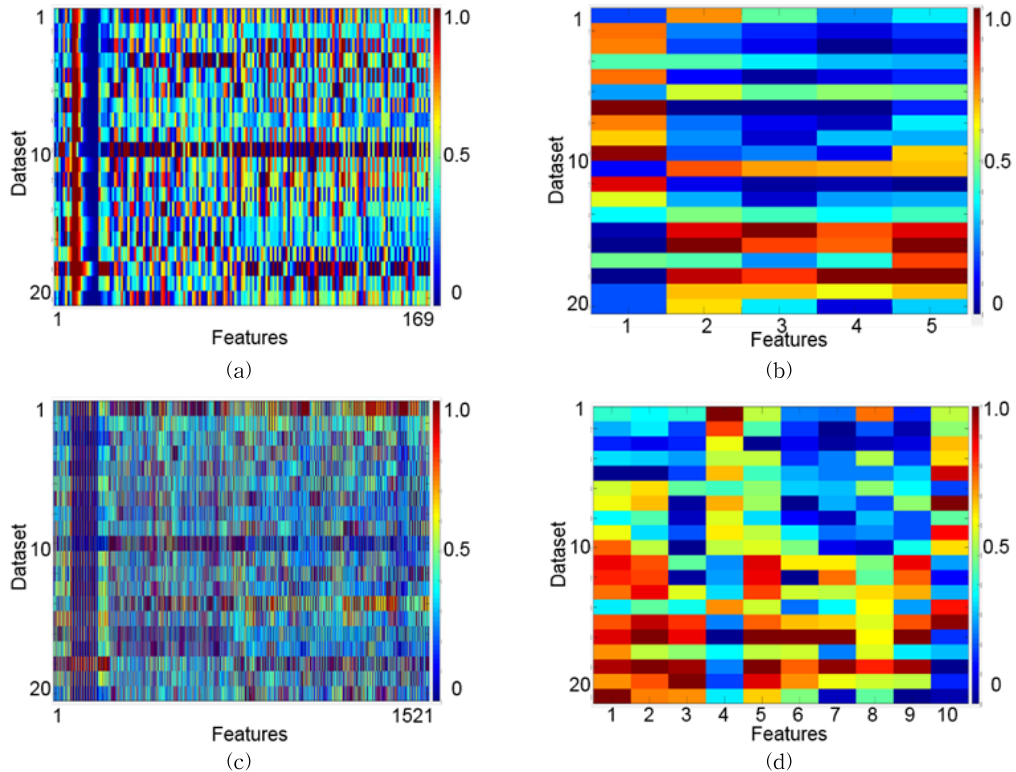


Fig. 6. Heatmaps of features. (a) all features of axial view, (b) selected top5 features of axial view, (c) all features of multiview, and (d) selected top10 features of multiview.

중 특징 선별 기법을 이용하여 선별한 상위 5개 특징의 히트맵, 다중 뷰 영상에서 추출된 전체 1521개 특징의 히트맵 그리고 다중 뷰 영상에서 추출된 전체 특징들 중 특징 선별 기법을 이용하여 선별한 상위 10개 특징의 히트맵이다. 측상면 뷰 영상에서 추출된 169개 특징을 특징 선별 기법을 이용하여 가중치가 급격히 떨어지지 않는 상위 특징들을 분석한 결과 상위 5개 특징이 선별되었고, 다중 뷰 영상에서 추출된 1521개 특징에 대해서도 동일한 방법으로 상위 10개 특징이 선별되었다. X축은 특징벡터를 나타내므로 Fig. 6(a)에서 시상면 영상에서 추출된 총 169개 특징, Fig. 6(b)에서 시상면 영상에서 추출된 169개 특징 중 선별된 상위 5개 특징, Fig. 6(c)에서 다중 뷰 영상에서 추출된 총 1521개 특징, Fig. 6(d)에서 다중 뷰 영상에서 추출된 총 1521개 특징 중 선별된 상위 10개 특징을 나타내고, Y축은 데이터를 나타내며 상위 10개 데이터는 순수 간유리음영 결절, 하위 10개 데이터는 혼합 간유리음영 결절이다. Fig. 6(a)보다 Fig. 6(b)가 순수/혼합 간유리음영 결절 간 특징 값 대조를 잘 보여주고 있고, Fig. 6(c)보다 Fig. 6(d)가 순수/혼합 간유리음영 결절 간 특징 값 대조를 잘 보여주고 있으며 Fig. 6(b)보다 Fig. 6(d)에서 상대적으로 특징 값이 더 대조되어 보인다. 이러한 특징 히트맵을 통해 Fig. 6(a)의 경우보다 Fig. 6(b)가, Fig. 6(c)의 경우보다 Fig. 6(d)가, 그리고 Fig. 6(b)의 경우보다 Fig. 6(d)가 특징 값 대조가 크므로 순수/혼합 간유리음영 결절 분류에 더 우수한 성능을 보이는 특징 집합이라고 볼 수 있다.

Table 2는 선별된 특징을 나타내며, 측상면 뷰 영상에서의 상위 5개 특징은 에너지, 엔트로피, 분산, 97.5 CT 밝기값, 최대 밝기값이고, 다중 뷰 영상에서의 상위 10개 특징은  $xy-z(135^\circ)$  영상에서 최대 밝기값,  $yz-x$

( $45^\circ$ ) 영상에서 97.5th CT 밝기값,  $yz-x(45^\circ)$  영상에서 에너지,  $zx-y(45^\circ)$  영상에서 최대 밝기값,  $yz-x(45^\circ)$  영상에서 최대 밝기값,  $yz-x(135^\circ)$  영상에서 97.5th CT 밝기값,  $xy-z(135^\circ)$  영상에서 초과 임계값 - 600(HU),  $yz-x(135^\circ)$  영상에서 최대 밝기값 그리고  $yz-x(0^\circ)$  영상에서 16레벨  $135^\circ$  방향의 long run low gray-level emphasis(LRLGE)이다. 측상면 뷰 영상에서 선별된 상위 5개 특징과 다중 뷰 영상에서 선별된 상위 10개 특징에는 히스토그램 특징인 분산, 에너지, 엔트로피, 최대 밝기값, 초과임계값인 - 600(HU), 백분위 CT 밝기값인 97.5th, 텍스처 특징인 LRLGE가 있다. 분산은 영상 밝기값의 분산된 정도를 나타내므로 고형 성분을 가지는 혼합 간유리음영 결절이 밝기값 범위가 넓게 나타나 상대적으로 순수 간유리음영 결절보다 높은 수치를 보였다. 에너지는 영상에서 동일한 밝기값 빈도가 높을수록 큰 값을 가지므로 영상의 패턴이 복잡한 혼합 간유리음영 결절이 상대적으로 영상 패턴이 단순한 순수 간유리음영 결절보다 낮은 수치를 보였다. 엔트로피는 에너지와 반대로 밝기값 빈도가 다양할수록 큰 값을 가지므로 영상의 패턴이 복잡한 혼합 간유리음영 결절이 상대적으로 순수 간유리음영 결절보다 높은 수치를 보였으며 최대 밝기값은 혼합 간유리음영 결절이 간유리음영 성분보다 밝기값이 높은 고형 성분을 포함하고 있으므로 높은 최대 밝기값 수치를 보였다. 초과임계값 수치인 - 600(HU)는 밝기값 수치가 - 600(HU)이상인 픽셀의 수를 의미하는데 혼합 간유리음영 결절의 간유리음영 영역이 순수 간유리음영 결절의 간유리음영 영역보다 상대적으로 밝은 값을 가지는 경향이 있어 높은 수치를 보였다. 백분위 CT 밝기값 수치인 97.5th는 히스토그램에서 밝기값 범위 상위 97.5%에 해당하는 밝기값을 의미하는데 혼합 간유리음영 결절이 밝기값이 높은 고형 성분을

Table 2. Selected features in axial view image and multiview images

Categories	Top 5 features in axial view image	Top 10 features in multiview images
Histogram	Variance, Energy, Entropy, Max value	$xy-z(135^\circ)$ : Max value, $yz-x(45^\circ)$ : Energy, $zx-y(45^\circ)$ : Max value, $yz-x(45^\circ)$ : Max value, $yz-x(135^\circ)$ : Max value
Threshold		$xy-z(135^\circ)$ : -600
Percentile	97.5th	$xy-z(135^\circ)$ : 97.5th, $yz-x(45^\circ)$ : 97.5th, $yz-x(135^\circ)$ : 97.5th
Texture		$yz-x(0^\circ)$ : LRLGE(16level, direction: $135^\circ$ )

포함하고 있으므로 상대적으로 순수 간유리음영 결절보다 높은 수치를 보였다. 텍스처 특징인 LRLGE는 gray level run-length matrix(GLRLM) 특징 중 하나로 낮은 밝기값이 연속적으로 나오는 정도를 나타내므로 혼합 간유리음영 결절이 상대적으로 패턴이 단순하고 고형 성분을 포함하지 않아 어두운 밝기값을 가지는 순수 간유리음영 결절보다 낮은 수치를 보였다.

Table 3은 각 축상면 뷰 영상에서 추출된 전체 특징벡터, 선별된 특징벡터를 이용해 분류한 정확도 수치, 다중 뷰 영상에서 추출된 전체 특징벡터, 선별된 특징벡터를 이용해 분류한 정확도 수치이다. 특징 선별 기법을 사용하지 않았을 때 축상면 뷰 영상의 특징만을 사용했을 때보다 다중 뷰 영상에서 추출된 특징을 사용했을 때 10% 더 높은 분류 성능을 보이는 것을 확인 할 수 있었고 축상면 뷰 영상에서 특징 선별 기법을 사용하지 않았을 때보다 사용했을 때 25% 더 높은 분류 성능을 보이고 다중 뷰 영상에서 특징 선별 기법을 사용하지 않았을 때보다 사용했을 때 40% 더 높은 분류 성능을 보이는 것을 확인 할 수 있었다. 이는 축상면 뷰 영상만을 사용할 때 보다

다중 뷰 영상을 사용할 때 간유리음영 결절의 더 많은 정보를 이용할 수 있고, 특징을 선별하지 않고 모두 사용할 때 보다 특징 선별 기법을 통해 선별된 특징을 사용할 때 간유리음영 결절로부터 추출된 특징 중 순수 간유리음영 결절과 혼합 간유리음영 결절의 특징 간 차이가 큰 특징들이 선별되기 때문에 분류 성능을 높일 수 있다는 것을 보여준다.

Fig. 7(a)은 축상면 뷰 영상에서 선별된 상위 5개 특징을 이용해 분류를 수행했을 때 오분류된 데이터, Fig. 7(b)는 다중 뷰 영상에서 선별된 상위 10개 특징을 이용해 분류를 수행했을 때 오분류된 데이터의 영상이다. Fig. 7(a)의 8개 오분류된 데이터 중 대다수가 Fig. 7(b)에서는 옳게 분류되었으나 3개 데이터는 여전히 오분류 되었다. Fig. 7(b)의 첫 번째 데이터와 두 번째 데이터는 모두 순수 간유리음영 결절이며 세 번째 데이터는 혼합 간유리음영 결절이다. Fig. 7(b)에서 오분류된 첫 번째 데이터는 타 순수 간유리음영 결절에 비해 상대적으로 밝기값과 엔트로피가 높고, 두 번째 데이터는 타 순수 간유리음영 결절에 비해 결절 중심부의 밝기값과 밀도가 상대적으로 높아 특징벡터가 혼합 간유리음영 결절과 유사하게 나

Table 3. Performance evaluation of GGNs classification results

	Axial view without feature selection	Axial view with feature selection	Multiview without feature selection	Multiview with feature selection
Accuracy[%]	35.00	60.00	45.00	85.00
Sensitivity[%]	30.00	50.00	50.00	90.00
Specificity[%]	40.00	70.00	40.00	80.00
PPV[%]	33.33	62.50	45.45	81.82
NPV[%]	36.36	58.33	44.44	88.89
AUC	37.00	58.00	41.00	88.00

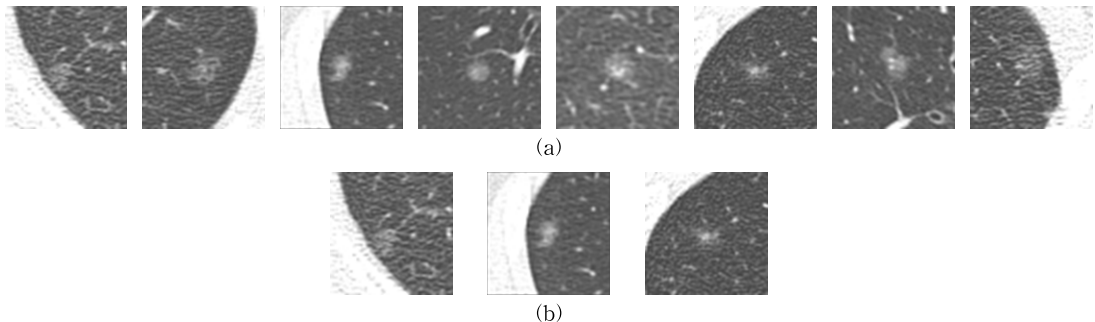


Fig. 7. Misclassified GGNs from SVM classifier. (a) misclassified GGNs from top 5 features in axial view, (b) misclassified GGNs from top 10 features in multiview.



타나는 것을 알 수 있었고 세 번째 데이터는 타 혼합 간유리음영 결절에 비해 고형 성분의 크기가 상대적으로 작아 특징벡터가 순수 간유리음영 결절과 유사하게 나타나는 것을 알 수 있었다.

#### 4. 결 론

본 논문에서는 순수 간유리음영 결절과 고형 성분이 5mm 미만으로 작은 혼합 간유리음영 결절을 분류하기 위하여 다중 뷰 영상에서 텍스처 특징 추출 및 특징 선별 기법을 통해 SVM 분류기를 학습시키고 분류 성능을 평가하였다. 분류 정확도는 다중 뷰 영상에서 60%의 정확도를 보이고 다중 뷰 영상에서 85%의 정확도를 보여 단일 뷰를 사용했을 때보다 다중 뷰를 사용했을 때 정확도 수치가 42% 향상되었다. 결과적으로 텍스처를 분석하여 간유리음영 결절을 분류할 때, 하나의 2D 대표단면만 사용했을 때보다 다중 뷰 영상을 사용했을 때 상대적으로 높은 분류 성능을 보이고, 모든 특징을 이용했을 때보다 특징 선별 기법을 통해 상위 특징을 이용했을 때 상대적으로 높은 분류 성능을 보이는 것을 확인 할 수 있었다.

#### REFERENCE

- [ 1 ] C.I. Henschke, D.F. Yankelevitz, R. Mirtcheva, G. McGuinness, D. McCauley, and O.S. Miettinen, "CT Screening for Lung Cancer: Frequency and Significance of Part-solid and Nonsolid Nodules," *American Journal of Roentgenology*, Vol. 178, No. 5, pp. 1053-1057, 2002.
- [ 2 ] K.H. Lee, J.M. Goo, S.J. Park, J.Y. Wi, D.H. Chung, H. Go, et al., "Correlation between the Size of the Solid Component on Thin-section CT and the Invasive Component on Pathology in Small Lung Adenocarcinomas Manifesting as Ground-glass Nodules," *Journal of Thoracic Oncology*, Vol. 9, No. 1, pp. 74-82, 2014.
- [ 3 ] J.H. Lee, C.M. Park, S.M. Lee, H. Kim, H.P. McAdams, and J.M. Goo, "Persistent Pulmonary Subsolid Nodules with Solid Portions of 5mm or Smaller: Their Natural Course and Predictors of Interval Growth," *European Radiology*, Vol. 26, No. 6, pp. 1529-1537, 2015.
- [ 4 ] T.N. Shewaye and A.A. Mekonnen, "Benign-Malignant Lung Nodule Classification with Geometric and Appearance Histogram Features," *arXiv Preprint arXiv:1605.08350*, 2016.
- [ 5 ] D. Kumar, A. Wong, and D.A. Clausi, "Lung Nodule Classification Using Deep Features in CT images," *Proceeding of Computer and Robot Vision*, pp. 133-138, 2015.
- [ 6 ] S.K. Dilger, J. Uthoff, A. Judisch, E. Hammond, S.L. Mott, B.J. Smith, et al., "Improved Pulmonary Nodule Classification Utilizing Quantitative Lung Parenchyma Features," *Journal of Medical Imaging*, Vol. 2, No. 4, pp. 041004-041004, 2015.
- [ 7 ] A. Kamiya, S. Murayama, H. Kamiya, T. Yamashiro, Y. Oshiro, and N. Tanaka, "Kurtosis and Skewness Assessments of Solid Lung Nodule Density Histograms: Differentiating Malignant from Benign Nodules on CT," *Japanese Journal of Radiology*, Vol. 32, No. 1, pp. 14-21, 2014.
- [ 8 ] H.D. Chae, C.M. Park, S.J. Park, S.M. Lee, K.G. Kim, and J.M. Goo, "Computerized Texture Analysis of Persistent Part-solid ground-glass Nodules: Differentiation of Preinvasive Lesions from Invasive Pulmonary Adenocarcinomas," *Radiology*, Vol. 273, No. 1, pp. 285-293, 2014.
- [ 9 ] S.A. Mattonen, D.A. Palma, C.J. Haasbeek, S. Senan, and A.D. Ward, "Early Prediction of Tumor Recurrence Based on CT Texture Changes after Stereotactic Ablative Radiotherapy (SABR) for Lung Cancer," *Medical Physics*, Vol. 41, No. 3, 033502, 2014.
- [ 10 ] J.Y. Son, H.Y. Lee, K.S. Lee, J.H. Kim, J. Han, J.Y. Jeong, et al., "Quantitative CT Analysis of Pulmonary Ground-glass Opacity Nodules for the Distinction of Invasive Adenocarcinoma from Pre-invasive or Minimally Invasive Adenocarcinoma," *PLoS One*, Vol. 9, No. 8, pp.

e104066, 2014.

- [11] J.Y. Son, H.Y. Lee, J.H. Kim, J. Han, J.Y. Jeong, K.S. Lee, et al, "Quantitative CT Analysis of Pulmonary Ground-glass Opacity Nodules for Distinguishing Invasive Adenocarcinoma from Non-invasive or Minimally Invasive Adenocarcinoma: the Added Value of Using Iodine Mapping," *European Radiology*, Vol. 26, No. 1, pp. 43-54, 2016.
- [12] S.Y. Lee, J. Jung, H.S. Lee, and H.Hong, "Pulmonary Ground-glass Opacity Nodules with Small Solid Portions Classification Using Texture Analysis in Chest CT Images," *Proceeding of The 2016 Fall Conference of Korea Multimedia Society*, pp. 234-236, 2016.
- [13] J. Jung and H. Hong, "Automatic Segmentation of Ground-glass Opacity Nodule using Multi-phase Deformable Model with Intensity Constraint in Chest CT Images," *Journal of KIISE : Software and Applications*, Vol. 40, No. 10, pp. 582-591, 2013.
- [14] R.M. Haralick and K. Shanmugam, "Texture Features for Image Classification," *IEEE Transactions on System, Man, and Cybernetics*, Vol. 3, No. 6, pp. 610-621, 1973.
- [15] X. Tang, "Texture Information in Run-length Matrices," *IEEE Transactions on Image Processing*, Vol. 7, No. 11, pp. 1602-1609, 1998.
- [16] M. Robnik-Šikonja and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF," *Machine Learning*, Vol. 53, No. 1-2, pp. 23-69, 2003.
- [17] C. Cortes and V. Vapnik, "Support-vector Networks," *Machine Learning*, Vol. 20, No. 3, pp. 273-297, 1995.
- [18] Y. Kim, J. Jung, H. Hong and S.I. Hwang, "Automated Prostate Cancer Detection on Multi-parametric MR imaging via Texture Analysis," *Journal of Korea Multimedia Society*, Vol. 19, No. 4, pp. 736-746, 2016.



**이 선 영**

2016년 서울여자대학교 컴퓨터학과 학사  
 2016년~현재 서울여자대학교 컴퓨터학과 석사과정 재학중  
 관심분야: 의료영상처리, 머신러닝, 인공지능



**정 주 립**

2007년 서울여자대학교 멀티미디어학과 학사  
 2009년 서울여자대학교 컴퓨터학과 석사  
 2009년~현재 서울여자대학교 컴퓨터학과 박사수료

관심분야: 의료영상처리, 영상정합, 영상분할



**이 한 상**

2011년 한국과학기술원 전기 및 전자공학과 학사  
 2013년 한국과학기술원 전기 및 전자공학과 석사  
 2013년~현재 한국과학기술원 전기 및 전자공학과 박사과정 재학 중

관심분야: 의료영상처리, 기계학습, 컴퓨터비전



**홍 헬 렌**

1994년 이화여자대학교 전자계산학과 학사  
 1996년 이화여자대학교 전자계산학과 석사  
 2001년 이화여자대학교 컴퓨터학과 박사

2001년~2003년 서울대학교 컴퓨터공학부 BK 박사후연구원

2003년~2006년 서울대학교 컴퓨터공학부 BK 조교수  
 2006년~현재 서울여자대학교 소프트웨어융합학과 교수  
 관심분야: 의료영상처리, 영상분할, 영상정합, 컴퓨터비전