



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0117391
(43) 공개일자 2023년08월08일

(51) 국제특허분류(Int. Cl.)
G16H 50/50 (2018.01) G06N 3/09 (2023.01)
G06T 7/00 (2017.01) G06V 10/25 (2022.01)
G06V 10/40 (2022.01) G16H 20/00 (2018.01)
G16H 30/00 (2018.01) G16H 50/20 (2018.01)
(52) CPC특허분류
G16H 50/50 (2018.01)
G06N 3/09 (2023.01)
(21) 출원번호 10-2023-7022202
(22) 출원일자(국제) 2021년11월30일
심사청구일자 없음
(85) 번역문제출일자 2023년06월30일
(86) 국제출원번호 PCT/IB2021/061125
(87) 국제공개번호 WO 2022/118190
국제공개일자 2022년06월09일
(30) 우선권주장
63/120,102 2020년12월01일 미국(US)

(71) 출원인
존슨 앤드 존슨 엔터프라이즈 이노베이션 인코포레이티드
미국 뉴저지주 08933 뉴 브런스윅 원 존슨 앤드 존슨 플라자
(72) 발명자
와시코 제이알. 조지
미국 뉴저지주 08901 뉴 브런스윅 원 존슨 앤드 존슨 플라자
에스테파르 라울 샌 호세
미국 뉴저지주 08901 뉴 브런스윅 원 존슨 앤드 존슨 플라자
(뒷면에 계속)
(74) 대리인
장훈

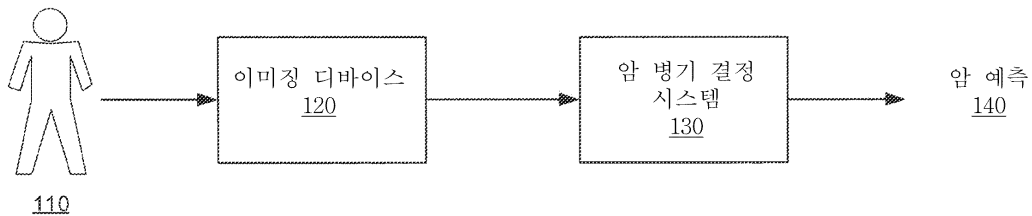
전체 청구항 수 : 총 118 항

(54) 발명의 명칭 림프절의 암 침범을 예측하기 위한 기계 학습

(57) 요약

1) 개별 림프절 침범의 림프절 레벨 위험 및/또는 2) 림프절 침범의 대상체 레벨 위험을 결정하기 위해 모델들의 훈련 및/또는 전개를 수반하는 전이성 암의 대상체 레벨 위험을 결정하기 위한 방법들이 본 명세서에 개시된다. 따라서, 본 방법들은 결절 질환을 가진 위험이 높거나 낮은 환자들을 식별하고, 선택적으로 예를 들어 치료를 통해 암 환자들의 안내식 개입을 가능하게 할 수 있다.

대표도 - 도1a



(52) CPC특허분류

G06T 7/0012 (2013.01)

G16H 20/00 (2021.08)

G16H 30/00 (2021.08)

G16H 50/20 (2018.01)

G06T 2207/10081 (2013.01)

G06T 2207/10104 (2013.01)

G06T 2207/20081 (2013.01)

(72) 발명자

킨제이 찰스 매튜

미국 뉴저지주 08901 뉴 브런스윅 원 존슨 앤드 존
슨 플라자

스티븐슨 크리스토퍼 스캇

미국 뉴저지주 08901 뉴 브런스윅 원 존슨 앤드 존
슨 플라자

명세서

청구범위

청구항 1

대상체(subject)의 전이성 암(metastatic cancer)의 대상체 레벨 위험을 결정하는 방법으로서,

상기 대상체의 복수의 림프절(lymph node)들을 포함하는, 상기 대상체로부터 캡처된 하나 이상의 이미지를 획득하는 단계; 및

상기 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델(risk model)을 적용함으로써 상기 전이성 암의 대상체 레벨 위험을 예측하는 단계를 포함하고,

상기 위험 모델은 기준 개체(reference individual)의 림프절들의 세트에 대한 전이성 암의 대체(imputed) LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 상기 림프절들의 세트의 상기 대체 LN-레벨 위험들은 훈련 이미지들의 암성(cancerous) 림프절과 비암성(non-cancerous) 림프절을 구별하는 대체 모델(imputation model)을 적어도 사용하여 결정되는, 방법.

청구항 2

제1항에 있어서, 상기 전이성 암의 대상체 레벨 위험을 예측하는 단계는,

상기 위험 모델에 의해 예측된 전이성 암의 LN-레벨 위험에 기초하여 상기 복수의 림프절들 중 하나 이상을 선택하는 단계; 및

상기 하나 이상의 림프절에 대해 예측된 상기 전이성 암의 LN-레벨 위험을 사용하여 상기 전이성 암의 대상체 레벨 위험을 결정하는 단계를 더 포함하는, 방법.

청구항 3

제2항에 있어서, 상기 복수의 림프절들 중 하나 이상을 선택하는 단계는 LN-레벨 위험의 최고 확률을 갖는 림프절을 식별하는 단계를 포함하고, 상기 전이성 암의 대상체 레벨 위험을 결정하는 단계는 상기 LN-레벨 위험의 최고 확률을 상기 전이성 암의 대상체 레벨 위험으로서 할당하는 단계를 포함하는, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 전이성 암의 LN-레벨 위험들을 예측하는, 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 기준 개체의 상기 림프절들의 세트는,

중앙(median) 위험 값을 결정하는 단계; 및

상기 기준 개체의 상기 림프절들의 대체 LN-레벨 위험들을 상기 중앙 위험 값과 비교하는 단계에 의해 선택되는, 방법.

청구항 6

제5항에 있어서, 기준 개체의 상기 림프절들의 세트는,

상기 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는 하나 이상의 림프절을 상기 림프절들의 세트에 포함시키는 단계에 의해 또한 선택되는, 방법.

청구항 7

제5항 또는 제6항에 있어서, 상기 림프절들의 세트에서 상기 중앙 위험 값보다 더 작은 LN-레벨 위험들을 갖는

하나 이상의 림프절이 상기 림프절들의 세트로부터 배제되는, 방법.

청구항 8

제5항 내지 제7항 중 어느 한 항에 있어서, 상기 중앙 위험 값은 상기 기준 개체의 중앙 대상체내 (intrasubject) LN-레벨 위험인, 방법.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 하나 이상의 이미지는 컴퓨터 단층촬영(computed tomography, CT) 이미지를 포함하는, 방법.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 상기 하나 이상의 이미지는 흉부 CT 스캔으로부터 획득되는, 방법.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 상기 암의 병기 (stage)를 결정하는 단계를 더 포함하는, 방법.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서, 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 상기 대상체에 대한 진단 또는 치료를 선택하는 단계를 더 포함하는, 방법.

청구항 13

제12항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우, 외과적 종양 절제 또는 조합된 기관지경 또는 내시경 진단 및 치료를 선택하는 단계를 포함하는, 방법.

청구항 14

제12항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 초과인 경우, PET 또는 PET-CT 스캔의 수행 또는 림프절 생검(biopsy)의 수행 중 하나 또는 둘 모두로부터 추가 진단 검사를 선택하는 단계를 포함하는, 방법.

청구항 15

제12항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우, PET 또는 PET-CT 스캔의 추가 진단 검사를 선택하는 단계를 포함하는, 방법.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서, 상기 대체 모델은 흉부 CT 스캔들로부터 획득된 훈련 이미지들을 사용하여 훈련되는, 방법.

청구항 17

제16항에 있어서, 상기 흉부 CT 스캔들로부터 획득된 훈련 이미지들은 하나 이상의 림프절을 포함하는, 방법.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 상기 대체 모델은 국가폐암검진연구(National Lung Screening Trial, NLST)의 훈련 이미지들을 사용하여 훈련되는, 방법.

청구항 19

제1항 내지 제17항 중 어느 한 항에 있어서, 상기 대체 모델은 맞춤형 데이터세트의 훈련 이미지들을 사용하여

훈련되는, 방법.

청구항 20

제16항 내지 제19항 중 어느 한 항에 있어서, 상기 대체 모델은 상기 훈련 이미지들 중 적어도 하나로부터 2개 이상의 라디오믹 패널(radiomic panel)들을 생성하는 단계에 의해 훈련되는, 방법.

청구항 21

제20항에 있어서, 상기 2개 이상의 라디오믹 패널들은 상기 훈련 이미지의 가우스 라플라시안(Laplacian of Gaussians) 변환 및 상기 훈련 이미지의 웨이블릿(wavelet) 변환을 포함하는, 방법.

청구항 22

제20항 또는 제21항에 있어서, 상기 대체 모델은 상기 2개 이상의 라디오믹 패널들의 각각으로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 방법.

청구항 23

제20항 또는 제21항에 있어서, 상기 대체 모델은,

상기 라디오믹 패널들에서, 림프절을 포함하는 관심 영역(region of interest, ROI)을 한정하는 단계;

상기 라디오믹 패널들 내의 상기 ROI로부터 특징들을 추출하는 단계; 및

적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 방법.

청구항 24

제22항 또는 제23항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(gray level co-occurrence matrix 1, GLCM) 특징들, 그레이 레벨 작용 길이 행렬(gray level run length matrix, GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(gray level size zone matrix, GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(neighborhood gray tone difference matrix, NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 방법.

청구항 25

제16항 내지 제24항 중 어느 한 항에 있어서, 상기 대체 모델은 훈련 이미지들에서의 전이의 존재 또는 부재를 나타내는 기준 실측 자료(reference ground truth) 값들을 사용하여 또한 훈련되는, 방법.

청구항 26

제25항에 있어서, 상기 기준 실측 자료 값들은 상기 암의 병기를 나타내는 N기 공변량(covariable) 값들로부터 도출되고, "0" 또는 "1"의 암 병기는 전이의 부재를 나타내고, "2" 또는 "3"의 암 병기는 전이의 존재를 나타내는, 방법.

청구항 27

제26항에 있어서, 상기 기준 실측 자료 값들은 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는, 방법.

청구항 28

제27항에 있어서, 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는 상기 기준 실측 자료 값들은 종양 또는 림프절들의 특징들에 기초하여 상기 N기 공변량 값들로부터 역산되는(back-calculated), 방법.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 상기 대체 모델은 랜덤 포레스트 분류기(random forest classifier) 또는 LASSO 분류기인, 방법.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 상기 위험 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기인, 방법.

청구항 31

제30항에 있어서, 상기 위험 모델은 지도형 학습 기법(supervised learning technique)들을 사용하여 훈련되는, 방법.

청구항 32

제1항 내지 제31항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 21개의 림프절들을 포함하는, 방법.

청구항 33

제1항 내지 제31항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 8 내지 20개의 림프절들, 10 내지 19개의 림프절들, 12 내지 18개의 림프절들, 또는 14 내지 16개의 림프절들을 포함하는, 방법.

청구항 34

제1항 내지 제33항 중 어느 한 항에 있어서, 상기 전이성 암은 전이성 폐암인, 방법.

청구항 35

제1항 내지 제34항 중 어느 한 항에 있어서, 상기 위험 모델은 적어도 AUC = 78%, 감도 = 82%, 또는 특이도 (Specificity) = 60%의 성능을 나타내는, 방법.

청구항 36

제1항 내지 제35항 중 어느 한 항의 위험 모델을 훈련시키는 방법으로서,

훈련 이미지들 내의 복수의 림프절들에 대한 대체 LN-레벨 위험들을 생성하기 위해 대체 모델을 적용하는 단계;
림프절들의 세트를 선택하는 단계로서, 상기 세트 내의 상기 림프절들은 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는, 상기 선택하는 단계;

상기 위험 모델을 훈련시키기 위한 기준 실측 자료들로서 상기 림프절들의 세트 내의 상기 림프절들의 상기 LN-레벨 위험들을 사용하는 단계로서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 LN-레벨 위험들을 예측할 수 있는, 상기 사용하는 단계를 포함하는, 방법.

청구항 37

제36항에 있어서, 상기 위험 모델은 하나 이상의 라디오믹 패널로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 위험 모델을 훈련시키는 단계에 의해 또한 훈련되는, 방법.

청구항 38

제37항에 있어서, 2개 이상의 라디오믹 패널들로부터 특징들을 추출하는 단계는,

상기 하나 이상의 라디오믹 패널에서, 림프절을 포함하는 관심 영역을 한정하는 단계;

상기 하나 이상의 라디오믹 패널 내의 상기 ROI로부터 특징들을 추출하는 단계를 포함하는, 방법.

청구항 39

제37항 또는 제38항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 방법.

청구항 40

대상체의 전이성 압의 대상체 레벨 위험을 결정하기 위한 시스템으로서,

상기 대상체의 하나 이상의 이미지를 캡처하도록 구성된 이미징 디바이스; 및

컴퓨팅 디바이스를 포함하고,

상기 컴퓨팅 디바이스는,

상기 대상체의 복수의 림프절들을 포함하는, 상기 대상체로부터 캡처된 하나 이상의 이미지를 획득하는 단계; 및

상기 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 압의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델을 적용함으로써 상기 전이성 압의 대상체 레벨 위험을 예측하는 단계를 수행하도록 구성되며,

상기 위험 모델은 기준 개체의 림프절들의 세트에 대한 전이성 압의 대체 LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 상기 림프절들의 세트의 상기 대체 LN-레벨 위험들은 훈련 이미지들의 양성 림프절과 비양성 림프절을 구별하는 대체 모델을 적어도 사용하여 결정되는, 시스템.

청구항 41

제40항에 있어서, 상기 전이성 압의 대상체 레벨 위험을 예측하는 단계는,

상기 위험 모델에 의해 예측된 전이성 압의 LN-레벨 위험에 기초하여 상기 복수의 림프절들 중 하나 이상을 선택하는 단계; 및

상기 하나 이상의 림프절에 대해 예측된 상기 전이성 압의 LN-레벨 위험을 사용하여 상기 전이성 압의 대상체 레벨 위험을 결정하는 단계를 더 포함하는, 시스템.

청구항 42

제41항에 있어서, 상기 복수의 림프절들 중 하나 이상을 선택하는 단계는 LN-레벨 위험의 최고 확률을 갖는 림프절을 식별하는 단계를 포함하고, 상기 전이성 압의 대상체 레벨 위험을 결정하는 단계는 상기 LN-레벨 위험의 최고 확률을 상기 전이성 압의 대상체 레벨 위험으로서 할당하는 단계를 포함하는, 시스템.

청구항 43

제40항 내지 제42항 중 어느 한 항에 있어서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 전이성 압의 LN-레벨 위험들을 예측하는, 시스템.

청구항 44

제40항 내지 제43항 중 어느 한 항에 있어서, 기준 개체의 상기 림프절들의 세트는,

중양 위험 값을 결정하는 단계; 및

상기 기준 개체의 상기 림프절들의 대체 LN-레벨 위험들을 상기 중양 위험 값과 비교하는 단계에 의해 선택되는, 시스템.

청구항 45

제44항에 있어서, 기준 개체의 상기 림프절들의 세트는,

상기 중양 위험 값보다 더 큰 LN-레벨 위험들을 갖는 하나 이상의 림프절을 상기 림프절들의 세트에 포함시키는 단계에 의해 또한 선택되는, 시스템.

청구항 46

제44항 또는 제45항에 있어서, 상기 림프절들의 세트에서 상기 중양 위험 값보다 더 작은 LN-레벨 위험들을 갖는 하나 이상의 림프절이 상기 림프절들의 세트로부터 배제되는, 시스템.

청구항 47

제44항 내지 제46항 중 어느 한 항에 있어서, 상기 중앙 위험 값은 상기 기준 개체의 중앙 대상체내 LN-레벨 위험인, 시스템.

청구항 48

제40항 내지 제47항 중 어느 한 항에 있어서, 상기 이미징 디바이스는 컴퓨터 단층촬영(computed tomography, CT) 스캐너, 자기공명 이미징(magnetic resonance imaging, MRI) 스캐너, 양전자방출단층촬영(positron emission tomography, PET) 스캐너, X-선 스캐너, 또는 초음파 이미징 디바이스 중 하나인, 시스템.

청구항 49

제40항 내지 제48항 중 어느 한 항에 있어서, 상기 이미징 디바이스는 CT 스캐너이고, 상기 하나 이상의 이미지는 컴퓨터 단층촬영(CT) 이미지를 포함하는, 시스템.

청구항 50

제40항 내지 제49항 중 어느 한 항에 있어서, 상기 하나 이상의 이미지는 흉부 CT 이미지를 포함하는, 시스템.

청구항 51

제40항 내지 제50항 중 어느 한 항에 있어서, 상기 컴퓨팅 디바이스는 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 상기 암의 병기를 결정하는 단계를 수행하도록 또한 구성되는, 시스템.

청구항 52

제40항 내지 제51항 중 어느 한 항에 있어서, 상기 컴퓨팅 디바이스는 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 상기 대상체에 대한 진단 또는 치료를 선택하는 단계를 수행하도록 또한 구성되는, 시스템.

청구항 53

제52항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우, 외과적 종양 절제 또는 조합된 기관지경 또는 내시경 진단 및 치료를 선택하는 단계를 포함하는, 시스템.

청구항 54

제53항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 초과인 경우, PET 또는 PET-CT 스캔의 수행 또는 림프절 생검(biopsy)의 수행 중 하나 또는 둘 모두의 추가 진단 검사를 선택하는 단계를 포함하는, 시스템.

청구항 55

제53항에 있어서, 상기 대상체에 대한 진단 또는 치료를 선택하는 단계는, 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우, PET 또는 PET-CT 스캔의 추가 진단 검사를 선택하는 단계를 포함하는, 시스템.

청구항 56

제40항 내지 제54항 중 어느 한 항에 있어서, 상기 대체 모델은 흉부 CT 스캔들로부터 획득된 훈련 이미지들을 사용하여 훈련되는, 시스템.

청구항 57

제56항에 있어서, 상기 흉부 CT 스캔들로부터 획득된 훈련 이미지들은 하나 이상의 림프절을 포함하는, 시스템.

청구항 58

제40항 내지 제57항 중 어느 한 항에 있어서, 상기 대체 모델은 국가폐암검진연구(NLST)의 훈련 이미지들을 사용하여 훈련되는, 시스템.

청구항 59

제40항 내지 제57항 중 어느 한 항에 있어서, 상기 대체 모델은 맞춤형 데이터세트의 훈련 이미지들을 사용하여 훈련되는, 시스템.

청구항 60

제56항 내지 제59항 중 어느 한 항에 있어서, 상기 대체 모델은 상기 훈련 이미지들 중 적어도 하나로부터 2개 이상의 라디오믹 패널들을 생성하는 단계에 의해 훈련되는, 시스템.

청구항 61

제60항에 있어서, 상기 2개 이상의 라디오믹 패널들은 상기 훈련 이미지의 가우스 라플라시안 변환 및 상기 훈련 이미지의 웨이블릿 변환을 포함하는, 시스템.

청구항 62

제60항 또는 제61항에 있어서, 상기 대체 모델은 상기 2개 이상의 라디오믹 패널들의 각각으로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 시스템.

청구항 63

제60항 또는 제61항에 있어서, 상기 대체 모델은,

상기 라디오믹 패널들에서, 림프절을 포함하는 관심 영역(ROI)을 한정하는 단계;

상기 라디오믹 패널들 내의 상기 ROI로부터 특징들을 추출하는 단계; 및

적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 시스템.

청구항 64

제62항 또는 제63항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 시스템.

청구항 65

제56항 내지 제64항 중 어느 한 항에 있어서, 상기 대체 모델은 훈련 이미지들에서의 전이의 존재 또는 부재를 나타내는 기준 실측 자료 값들을 사용하여 또한 훈련되는, 시스템.

청구항 66

제65항에 있어서, 상기 기준 실측 자료 값들은 상기 암의 병기를 나타내는 N기 공변량 값들로부터 도출되고, "0" 또는 "1"의 암 병기는 전이의 부재를 나타내고, "2" 또는 "3"의 암 병기는 전이의 존재를 나타내는, 시스템.

청구항 67

제66항에 있어서, 상기 기준 실측 자료 값들은 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는, 시스템.

청구항 68

제67항에 있어서, 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는 상기 기준 실측 자료 값들은 종양 또는 림프절들의 특징들에 기초하여 상기 N기 공변량 값들로부터 역산되는, 시스템.

청구항 69

제40항 내지 제68항 중 어느 한 항에 있어서, 상기 대체 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기인, 시

스텝.

청구항 70

제40항 내지 제69항 중 어느 한 항에 있어서, 상기 위험 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기인, 시스템.

청구항 71

제70항에 있어서, 상기 위험 모델은 지도형 학습 기법들을 사용하여 훈련되는, 시스템.

청구항 72

제40항 내지 제71항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 21개의 림프절들을 포함하는, 시스템.

청구항 73

제40항 내지 제72항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 8 내지 20개의 림프절들, 10 내지 19개의 림프절들, 12 내지 18개의 림프절들, 또는 14 내지 16개의 림프절들을 포함하는, 시스템.

청구항 74

제40항 내지 제73항 중 어느 한 항에 있어서, 상기 전이성 암은 전이성 폐암인, 시스템.

청구항 75

제40항 내지 제74항 중 어느 한 항에 있어서, 상기 위험 모델은 적어도 AUC = 78%, 감도 = 82%, 또는 특이도 = 60%의 성능을 나타내는, 시스템.

청구항 76

제40항 내지 제75항 중 어느 한 항에 있어서, 상기 위험 모델은,

훈련 이미지들 내의 복수의 림프절들에 대한 대체 LN-레벨 위험들을 생성하기 위해 대체 모델을 적용하는 단계;
림프절들의 세트를 선택하는 단계로서, 상기 세트 내의 상기 림프절들은 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는, 상기 선택하는 단계;

상기 위험 모델을 훈련시키기 위한 기준 실측 자료들로서 상기 림프절들의 세트 내의 상기 림프절들의 상기 LN-레벨 위험들을 사용하는 단계로서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 LN-레벨 위험들을 예측할 수 있는, 상기 사용하는 단계에 의해 훈련되는, 시스템.

청구항 77

제76항에 있어서, 상기 위험 모델은 하나 이상의 라디오믹 패널로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 위험 모델을 훈련시키는 단계에 의해 또한 훈련되는, 시스템.

청구항 78

제77항에 있어서, 2개 이상의 라디오믹 패널들로부터 특징들을 추출하는 단계는,

상기 하나 이상의 라디오믹 패널에서, 림프절을 포함하는 관심 영역(ROI)을 한정하는 단계;

상기 하나 이상의 라디오믹 패널 내의 상기 ROI로부터 특징들을 추출하는 단계를 포함하는, 시스템.

청구항 79

제77항 또는 제78항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 시스템.

청구항 80

명령어들을 포함하는 비밀시적 컴퓨터 판독가능 매체로서,

상기 명령어들은 프로세서에 의해 실행될 때 상기 프로세서로 하여금,

대상체의 복수의 림프절들을 포함하는, 상기 대상체로부터 캡처된 하나 이상의 이미지를 획득하게 하고,

상기 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델을 적용함으로써 전이성 암의 대상체 레벨 위험을 예측하게 하며,

상기 위험 모델은 기준 개체의 림프절들의 세트에 대한 전이성 암의 대체 LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 상기 림프절들의 세트의 상기 대체 LN-레벨 위험들은 훈련 이미지들의 양성 림프절과 비양성 림프절을 구별하는 대체 모델을 적어도 사용하여 결정되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 81

제80항에 있어서, 상기 프로세서로 하여금 상기 전이성 암의 대상체 레벨 위험을 예측하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때, 상기 프로세서로 하여금,

상기 위험 모델에 의해 예측된 전이성 암의 LN-레벨 위험에 기초하여 상기 복수의 림프절들 중 하나 이상을 선택하게 하고,

상기 하나 이상의 림프절에 대해 예측된 상기 전이성 암의 LN-레벨 위험을 사용하여 상기 전이성 암의 대상체 레벨 위험을 결정하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 82

제81항에 있어서, 상기 프로세서로 하여금 상기 복수의 림프절들 중 하나 이상을 선택하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때 상기 프로세서로 하여금 LN-레벨 위험의 최고 확률을 갖는 림프절을 식별하게 하는 명령어들을 더 포함하고; 상기 프로세서로 하여금 상기 전이성 암의 대상체 레벨 위험을 결정하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때 상기 프로세서로 하여금 상기 LN-레벨 위험의 최고 확률을 상기 전이성 암의 대상체 레벨 위험으로서 할당하게 하는 명령어들을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 83

제80항 내지 제82항 중 어느 한 항에 있어서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 전이성 암의 LN-레벨 위험들을 예측하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 84

제80항 내지 제83항 중 어느 한 항에 있어서, 기준 개체의 상기 림프절들의 세트는,

중양 위험 값을 결정하는 단계; 및

상기 기준 개체의 상기 림프절들의 대체 LN-레벨 위험들을 상기 중양 위험 값과 비교하는 단계에 의해 선택되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 85

제84항에 있어서, 기준 개체의 상기 림프절들의 세트는 상기 중양 위험 값보다 더 큰 LN-레벨 위험들을 갖는 하나 이상의 림프절을 상기 림프절들의 세트에 포함시키는 단계에 의해 또한 선택되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 86

제84항 또는 제85항에 있어서, 상기 림프절들의 세트에서 상기 중양 위험 값보다 더 작은 LN-레벨 위험들을 갖는 하나 이상의 림프절이 상기 림프절들의 세트로부터 배제되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 87

제84항 내지 제86항 중 어느 한 항에 있어서, 상기 중앙 위험 값은 상기 기준 개체의 중앙 대상체내 LN-레벨 위험인, 비밀시적 컴퓨터 판독가능 매체.

청구항 88

제80항 내지 제87항 중 어느 한 항에 있어서, 상기 하나 이상의 이미지는 컴퓨터 단층촬영(CT) 이미지를 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 89

제80항 내지 제88항 중 어느 한 항에 있어서, 상기 하나 이상의 이미지는 흉부 CT 스캔으로부터 획득되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 90

제80항 내지 제89항 중 어느 한 항에 있어서, 상기 프로세서에 의해 실행될 때 상기 프로세서로 하여금 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 암의 병기를 결정하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 91

제80항 내지 제90항 중 어느 한 항에 있어서, 상기 프로세서에 의해 실행될 때 상기 프로세서로 하여금 상기 전이성 암의 예측된 대상체 레벨 위험에 따라 상기 대상체에 대한 진단 또는 치료를 선택하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 92

제91항에 있어서, 상기 프로세서로 하여금 상기 대상체에 대한 진단 또는 치료를 선택하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우에 상기 프로세서로 하여금 외과적 종양 절제 또는 조합된 기관지경 또는 내시경 진단 및 치료를 선택하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 93

제91항에 있어서, 상기 프로세서로 하여금 상기 대상체에 대한 진단 또는 치료를 선택하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 초과인 경우에 상기 프로세서로 하여금 PET 또는 PET-CT 스캔의 수행 또는 림프절 생검의 수행 중 하나 또는 둘 모두의 추가 진단 검사를 선택하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 94

제91항에 있어서, 상기 프로세서로 하여금 상기 대상체에 대한 진단 또는 치료를 선택하게 하는 상기 명령어들은, 상기 프로세서에 의해 실행될 때 상기 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우에 상기 프로세서로 하여금 PET 또는 PET-CT 스캔의 추가 진단 검사를 선택하게 하는 명령어들을 더 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 95

제80항 내지 제94항 중 어느 한 항에 있어서, 상기 대체 모델은 흉부 CT 스캔들로부터 획득된 훈련 이미지들을 사용하여 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 96

제95항에 있어서, 상기 흉부 CT 스캔들로부터 획득된 훈련 이미지들은 하나 이상의 림프절을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 97

제80항 내지 제96항 중 어느 한 항에 있어서, 상기 대체 모델은 국가폐암검진연구(NLST)의 훈련 이미지들을 사

용하여 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 98

제80항 내지 제96항 중 어느 한 항에 있어서, 상기 대체 모델은 맞춤형 데이터세트의 훈련 이미지들을 사용하여 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 99

제95항 내지 제98항 중 어느 한 항에 있어서, 상기 대체 모델은 상기 훈련 이미지들 중 적어도 하나로부터 2개 이상의 라디오믹 패널들을 생성하는 단계에 의해 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 100

제99항에 있어서, 상기 2개 이상의 라디오믹 패널들은 상기 훈련 이미지의 가우스 라플라시안 변환 및 상기 훈련 이미지의 웨이블릿 변환을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 101

제99항 또는 제100항에 있어서, 상기 대체 모델은 상기 2개 이상의 라디오믹 패널들의 각각으로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 102

제99항 또는 제100항에 있어서, 상기 대체 모델은,

상기 라디오믹 패널들에서, 림프절을 포함하는 관심 영역(ROI)을 한정하는 단계;

상기 라디오믹 패널들 내의 상기 ROI로부터 특징들을 추출하는 단계; 및

적어도 상기 추출된 특징들을 사용하여 상기 대체 모델을 훈련시키는 단계에 의해 또한 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 103

제101항 또는 제102항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 104

제95항 내지 제103항 중 어느 한 항에 있어서, 상기 대체 모델은 훈련 이미지들에서의 전이의 존재 또는 부재를 나타내는 기준 실측 자료 값들을 사용하여 또한 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 105

제104항에 있어서, 상기 기준 실측 자료 값들은 상기 암의 병기를 나타내는 N기 공변량 값들로부터 도출되고, "0" 또는 "1"의 암 병기는 전이의 부재를 나타내고, "2" 또는 "3"의 암 병기는 전이의 존재를 나타내는, 비밀시적 컴퓨터 판독가능 매체.

청구항 106

제105항에 있어서, 상기 기준 실측 자료 값들은 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는, 비밀시적 컴퓨터 판독가능 매체.

청구항 107

제106항에 있어서, 개별 림프절들이 양성 또는 비양성인지 여부를 나타내는 상기 기준 실측 자료 값들은 중앙 또는 림프절들의 특징들에 기초하여 상기 N기 공변량 값들로부터 역산되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 108

제80항 내지 제107항 중 어느 한 항에 있어서, 상기 대체 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기인, 비밀시적 컴퓨터 판독가능 매체.

청구항 109

제80항 내지 제108항 중 어느 한 항에 있어서, 상기 위험 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기인, 비밀시적 컴퓨터 판독가능 매체.

청구항 110

제109항에 있어서, 상기 위험 모델은 지도형 학습 기법들을 사용하여 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 111

제80항 내지 제110항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 또는 21개의 림프절들을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 112

제80항 내지 제111항 중 어느 한 항에 있어서, 상기 복수의 림프절들은 8 내지 20개의 림프절들, 10 내지 19개의 림프절들, 12 내지 18개의 림프절들, 또는 14 내지 16개의 림프절들을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 113

제80항 내지 제112항 중 어느 한 항에 있어서, 상기 전이성 암은 전이성 폐암인, 비밀시적 컴퓨터 판독가능 매체.

청구항 114

제80항 내지 제113항 중 어느 한 항에 있어서, 상기 위험 모델은 적어도 AUC = 78%, 감도 = 82%, 또는 특이도 = 60%의 성능을 나타내는, 비밀시적 컴퓨터 판독가능 매체.

청구항 115

제80항 내지 제114항 중 어느 한 항에 있어서, 상기 위험 모델은,

훈련 이미지들 내의 복수의 림프절들에 대한 대체 LN-레벨 위험들을 생성하기 위해 대체 모델을 적용하는 단계;

림프절들의 세트를 선택하는 단계로서, 상기 세트 내의 상기 림프절들은 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는, 상기 선택하는 단계;

상기 위험 모델을 훈련시키기 위한 기준 실측 자료들로서 상기 림프절들의 세트 내의 상기 림프절들의 상기 LN-레벨 위험들을 사용하는 단계로서, 상기 위험 모델은 상기 대체 모델에 의해 예측된 상기 대체 LN-레벨 위험들보다 더 큰 확률 범위로 LN-레벨 위험들을 예측할 수 있는, 상기 사용하는 단계에 의해 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 116

제115항에 있어서, 상기 위험 모델은 하나 이상의 라디오믹 패널로부터 특징들을 추출하는 단계 및 적어도 상기 추출된 특징들을 사용하여 상기 위험 모델을 훈련시키는 단계에 의해 또한 훈련되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 117

제116항에 있어서, 2개 이상의 라디오믹 패널들로부터 특징들을 추출하는 단계는,

상기 하나 이상의 라디오믹 패널에서, 림프절을 포함하는 관심 영역(ROI)을 한정하는 단계;

상기 하나 이상의 라디오믹 패널 내의 상기 ROI로부터 특징들을 추출하는 단계를 포함하는, 비일시적 컴퓨터 판독가능 매체.

청구항 118

제115항 또는 제116항에 있어서, 상기 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함하는, 비일시적 컴퓨터 판독가능 매체.

발명의 설명

기술 분야

[0001] 관련 출원과의 상호 참조

[0002] 본 출원은 2020년 12월 1일에 출원된 미국 가특허 출원 제63/120,102호에 대한 우선권의 이익을 주장하며, 이의 각각의 전체 개시내용은 모든 목적을 위해 전체적으로 참고로 본 명세서에 포함된다.

배경 기술

[0003] 흉부의 컴퓨터 단층촬영(computed tomographic, CT) 이미징은 폐암 위험이 있는 사람들의 평가를 위한 임상 표준이 되었다. 이들 노력은 그의 가장 빠르고 가장 치유가능한 병기(stage)들에서 질환을 나타낼 수 있는 폐 결절(lung nodule)의 검출에 초점을 맞추고 있다. 그러한 폐 결절의 검출은, 그것이 암성(cancerous)인지 여부를 결정하고, 그렇다면 그것이 신체의 흉부 또는 심지어 원격 영역에서의 다른 구조물로 퍼졌는지 여부를 결정하기 위하여 더 광범위한 임상 평가를 유발시킨다. 이는 폐암의 병기를 결정하기 위해 수행되며, 이는 결국 치료적 개입을 위한 계획을 지시한다.

[0004] 폐암 병기 결정(staging)의 일 태양은 흉부 림프절(lymph node)에서의 그의 존재 또는 부재(결절 상태)의 결정이다. 현재의 치료 표준은, 결절이 먼저 양성 또는 악성인 것으로 결정되고, 이어서 림프절의 양전자방출단층촬영(positron emission tomography, PET) 스캐닝 및 외과적, 내시경 또는 기관지경 기반 샘플링과 같은 추가 검사가 완료되는 단계적 접근법을 수반한다. 조직 샘플링 절차는 정밀 검사(work up)에 위험, 비용 및 시간을 더한다. 이들 위험을 완화하고 진단과 치료 사이의 시간을 단축하기 위해 새로운 비침습 방법이 필요하다. 본 명세서에서 더 상세히 기술되는 바와 같이, 폐암 검진(screening)에 사용되는 초기 흉부 CT 스캔에 적용되는 기계 학습(machine learning) 및 딥 러닝(deep learning)-기반 기술들이 이들 필요성을 해결할 수 있다.

[0005] 폐암 검출, 병기 결정 및 치료 계획 입안은 수주 또는 수개월이 걸릴 수 있는 다단계 과정이다. 그러한 임상 정밀 검사의 위험 및 비용에 더하여, 이러한 과정이 필요로 하는 추가 시간은 상당한 환자 불안 및 높아진 간격 전이(interval metastases) 가능성으로 이어진다. 실제로, 더 길어진 진단부터 치료까지의 시간은 더 나쁜 총체적 생존과 연관된다. 이들 우려를 완화시키기 위한 예시적인 접근법은 거의 동시적(near simultaneous) 암 진단, 병기 결정, 및 요법에 대한 기관지경 접근법의 경우에서의 치료를 허용하는 것이다.

발명의 내용

[0006] 본 명세서에 개시된 본 발명의 실시 형태들은 고위험 집단에서 폐암을 검진하는 데 사용되는 흉부 CT 스캔의 분석에 기초한다. 일부 태양에서, 이는 흉부 내의 적어도 하나의 림프절(LN)로 전이되었고 따라서 PET 스캐닝 및 림프절 생검(biopsy)과 같은 추가 검사를 필요로 하는 암에 대한 높은 위험에 있는 환자들을 식별한다. 이는 임상 제공자가 가장 필요한 사람들에 대한 검사를 우선순위화하게 한다. 역으로, 림프절 전이를 가질 위험이 낮은 의심되는 폐암이 있는 환자들은 외과적 절제 또는 조합된 기관지경 진단 및 치료와 같은 즉각적인 요법으로 직접 진행할 수 있다. 이는 질환 진단과 요법 사이의 시간을 제거할 것이다. 구체적으로, 본 명세서에 개시된 본 발명의 실시 형태들은 결절 질환을 가질 위험이 높거나 낮은 환자들의 식별을 가능하게 함으로써, 임상 의가 고위험 개체들에서 추가 검사를 우선순위화하게 하거나, 낮은 전이 위험을 갖는 폐암 환자들이 기관지경 전달 요법이 바로 뒤따르는 확증 기관지경 생검으로 직접 진행하게 한다.

[0007] 대상체(subject)의 전이성 암(metastatic cancer)의 대상체 레벨 위험을 결정하는 방법이 본 명세서에 개시되는 데, 본 방법은 대상체의 복수의 림프절들을 포함하는, 대상체로부터 캡처된 하나 이상의 이미지를 획득하는 단

계; 및 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델(risk model)을 적용함으로써 전이성 암의 대상체 레벨 위험을 예측하는 단계를 포함하고, 위험 모델은 기준 개체(reference individual)의 림프절들의 세트에 대한 전이성 암의 대체(imputed) LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 림프절들의 세트의 대체 LN-레벨 위험들은 훈련 이미지들의 양성(cancerous) 림프절과 비양성(non-cancerous) 림프절을 구별하는 대체 모델(imputation model)을 적어도 사용하여 결정된다. 다양한 실시 형태들에서, 전이성 암의 대상체 레벨 위험을 예측하는 단계는 위험 모델에 의해 예측된 전이성 암의 LN-레벨 위험에 기초하여 복수의 림프절들 중 하나 이상을 선택하는 단계; 및 하나 이상의 림프절에 대해 예측된 전이성 암의 LN-레벨 위험을 사용하여 전이성 암의 대상체 레벨 위험을 결정하는 단계를 더 포함한다.

[0008] 다양한 실시 형태들에서, 복수의 림프절들 중 하나 이상을 선택하는 단계는 LN-레벨 위험의 최고 확률을 갖는 림프절을 식별하는 단계를 포함하고, 전이성 암의 대상체 레벨 위험을 결정하는 단계는 LN-레벨 위험의 최고 확률을 전이성 암의 대상체 레벨 위험으로서 할당하는 단계를 포함한다. 다양한 실시 형태들에서, 위험 모델은 대체 모델에 의해 예측된 대체 LN-레벨 위험들보다 더 큰 확률 범위를 갖는 전이성 암의 LN-레벨 위험들을 예측한다. 다양한 실시 형태들에서, 기준 개체의 림프절들의 세트는 중앙(median) 위험 값을 결정하는 단계; 및 기준 개체의 림프절들의 대체 LN-레벨 위험들을 중앙 위험 값과 비교하는 단계에 의해 선택된다. 다양한 실시 형태들에서, 기준 개체의 림프절들의 세트는 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는 하나 이상의 림프절을 림프절들의 세트에 포함시키는 단계에 의해 또한 선택된다. 다양한 실시 형태들에서, 림프절들의 세트에서 중앙 위험 값보다 더 작은 LN-레벨 위험들을 갖는 하나 이상의 림프절이 림프절들의 세트로부터 배제된다.

[0009] 다양한 실시 형태들에서, 중앙 위험 값은 기준 개체의 중앙 대상체내(intrasubject) LN-레벨 위험이다. 다양한 실시 형태들에서, 하나 이상의 이미지는 컴퓨터 단층촬영(CT) 이미지를 포함한다. 다양한 실시 형태들에서, 하나 이상의 이미지는 흉부 CT 스캔으로부터 획득된다. 다양한 실시 형태들에서, 본 명세서에 개시된 방법들은 전이성 암의 예측된 대상체 레벨 위험에 따라 암의 병기를 결정하는 단계를 더 포함한다. 다양한 실시 형태들에서, 본 명세서에 개시된 방법들은 전이성 암의 예측된 대상체 레벨 위험에 따라 대상체에 대한 진단 또는 치료를 선택하는 단계를 더 포함한다. 다양한 실시 형태들에서, 대상체에 대한 진단 또는 치료를 선택하는 단계는, 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 미만인 경우, 외과적 종양 절제 또는 조합된 기관지경 또는 내시경 진단 및 치료를 선택하는 단계를 포함한다. 다양한 실시 형태들에서, 대상체에 대한 진단 또는 치료를 선택하는 단계는, 전이성 암의 예측된 대상체 레벨 위험이 임계 위험 값 초과인 경우, PET 또는 PET-CT 스캔의 수행 또는 림프절 생검의 수행 중 하나 또는 둘 모두의 추가 진단 검사를 선택하는 단계를 포함한다.

[0010] 다양한 실시 형태들에서, 대체 모델은 흉부 CT 스캔들로부터 획득된 훈련 이미지들을 사용하여 훈련된다. 다양한 실시 형태들에서, 흉부 CT 스캔들로부터 획득된 훈련 이미지들은 하나 이상의 림프절을 포함한다. 다양한 실시 형태들에서, 대체 예측 모델은 국가폐암검진연구(National Lung Screening Trial, NLST)의 훈련 이미지들을 사용하여 훈련된다. 다양한 실시 형태들에서, 대체 모델은 맞춤형 데이터세트의 훈련 이미지들을 사용하여 훈련된다.

[0011] 다양한 실시 형태들에서, 대체 모델은 훈련 이미지들 중 적어도 하나로부터 2개 이상의 라디오믹 패널(radiomic panel)들을 생성하는 단계에 의해 훈련된다. 다양한 실시 형태들에서, 2개 이상의 라디오믹 패널들은 훈련 이미지의 가우스 라플라시안(Laplacian of Gaussians) 변환 및 훈련 이미지의 웨이블릿(wavelet) 변환을 포함한다. 다양한 실시 형태들에서, 대체 모델은 2개 이상의 라디오믹 패널들의 각각으로부터 특징들을 추출하는 단계 및 적어도 추출된 특징들을 사용하여 대체 모델을 훈련시키는 단계에 의해 또한 훈련된다. 다양한 실시 형태들에서, 대체 모델은, 라디오믹 패널들에서, 림프절을 포함하는 관심 영역(region of interest, ROI)을 한정하는 단계; 라디오믹 패널들 내의 ROI로부터 특징들을 추출하는 단계; 및 적어도 추출된 특징들을 사용하여 대체 모델을 훈련시키는 단계에 의해 또한 훈련된다. 다양한 실시 형태들에서, 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(gray level co-occurrence matrix 1, GLCM) 특징들, 그레이 레벨 작용 길이 행렬(gray level run length matrix, GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(gray level size zone matrix, GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(neighborhood gray tone difference matrix, NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함한다.

[0012] 다양한 실시 형태들에서, 대체 모델은 훈련 이미지들에서의 전이의 존재 또는 부재를 나타내는 기준 실측 자료 값들을 사용하여 또한 훈련된다. 다양한 실시 형태들에서, 기준 실측 자료 값들은 암의 병기를 나타내는 N기 공변량(covariable) 값들로부터 도출되고, "0" 또는 "1"의 암 병기는 전이의 부재를 나타내고, "2" 또는 "3"의 암 병기는 전이의 존재를 나타낸다. 다양한 실시 형태들에서, 기준 실측 자료 값들은 개별 림프절들이 양성 또

는 비암성인지 여부를 나타낸다. 다양한 실시 형태들에서, 개별 림프절들이 양성 또는 비암성인지 여부를 나타내는 기준 실측 자료 값들은 중앙 또는 림프절들의 특징들에 기초하여 N기 공변량 값들로부터 역산된다(back-calculated).

[0013] 다양한 실시 형태들에서, 대체 모델은 랜덤 포레스트 분류기(random forest classifier) 또는 LASSO 분류기이다. 다양한 실시 형태들에서, 위험 모델은 랜덤 포레스트 분류기 또는 LASSO 분류기이다. 다양한 실시 형태들에서, 위험 모델은 지도형 학습 기법(supervised learning technique)들을 사용하여 훈련된다. 다양한 실시 형태들에서, 복수의 림프절들은 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21개의 림프절들을 포함한다. 다양한 실시 형태들에서, 복수의 림프절들은 8 내지 20개의 림프절들, 10 내지 19개의 림프절들, 12 내지 18개의 림프절들, 또는 14 내지 16개의 림프절들을 포함한다. 다양한 실시 형태들에서, 전이성 암은 전이성 폐암이다. 다양한 실시 형태들에서, 위험 모델은 적어도 AUC = 78%, 감도 = 82%, 또는 특이도(Specificity) = 60%의 성능을 나타낸다.

[0014] 청구항 1 내지 청구항 35 중 어느 한 청구항의 위험 모델을 훈련시키는 방법이 본 명세서에 부가적으로 개시되는데, 본 방법은 훈련 이미지들 내의 복수의 림프절들에 대한 대체 LN-레벨 위험들을 생성하기 위해 대체 모델을 적용하는 단계; 림프절들의 세트를 선택하는 단계로서, 세트 내의 림프절들은 중앙 위험 값보다 더 큰 LN-레벨 위험들을 갖는, 상기 선택하는 단계; 위험 모델을 훈련시키기 위한 기준 실측 자료들로서 림프절들의 세트 내의 림프절들의 LN-레벨 위험들을 사용하는 단계로서, 위험 모델은 대체 모델에 의해 예측된 대체 LN-레벨 위험들보다 더 큰 확률 범위로 LN-레벨 위험들을 예측할 수 있는, 상기 사용하는 단계를 포함한다. 다양한 실시 형태들에서, 위험 모델은 하나 이상의 라디오믹 패널로부터 특징들을 추출하는 단계 및 적어도 추출된 특징들을 사용하여 대체 모델을 훈련시키는 단계에 의해 또한 훈련된다. 다양한 실시 형태들에서, 2개 이상의 라디오믹 패널들로부터 특징들을 추출하는 단계는, 하나 이상의 라디오믹 패널에서, 림프절을 포함하는 관심 영역(ROI)을 한정하는 단계; 하나 이상의 라디오믹 패널 내의 ROI로부터 특징들을 추출하는 단계를 포함한다. 다양한 실시 형태들에서, 추출된 특징들은 1차 특징들, 형상 특징들, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징들, 그레이 레벨 작용 길이 행렬(GLRLM) 특징들, 그레이 레벨 크기 구역 행렬(GLSZM) 특징들, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징들 중 하나 이상의 특징 카테고리를 포함한다.

[0015] 부가적으로, 전술된 방법들을 수행하기 위한 시스템이 본 명세서에 개시된다. 다양한 실시 형태들에서, 대상체의 전이성 암의 대상체 레벨 위험을 결정하기 위한 시스템은 대상체의 하나 이상의 이미지를 캡처하도록 구성된 이미징 디바이스; 및 컴퓨팅 디바이스를 포함하고, 컴퓨팅 디바이스는, 대상체의 복수의 림프절들을 포함하는, 대상체로부터 캡처된 하나 이상의 이미지를 획득하는 단계; 및 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델을 적용함으로써 전이성 암의 대상체 레벨 위험을 예측하는 단계를 수행하도록 구성되며, 위험 모델은 기준 개체의 림프절들의 세트에 대한 전이성 암의 대체 LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 림프절들의 세트의 대체 LN-레벨 위험들은 훈련 이미지들의 양성 림프절과 비암성 림프절을 구별하는 대체 모델을 적어도 사용하여 결정된다.

[0016] 전술된 방법을 수행하기 위한 비일시적 컴퓨터 판독가능 매체가 본 명세서에 부가적으로 개시된다. 다양한 실시 형태들에서, 비일시적 컴퓨터 판독가능 매체는 명령어들을 포함하는데, 명령어들은 프로세서에 의해 실행될 때 프로세서로 하여금, 대상체의 복수의 림프절들을 포함하는, 대상체로부터 캡처된 하나 이상의 이미지를 획득하게 하고,

[0017] 획득된 하나 이상의 이미지의 추출된 특징들에, 이미지들 내의 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하도록 훈련되는 위험 모델을 적용함으로써 전이성 암의 대상체 레벨 위험을 예측하게 하며, 위험 모델은 기준 개체의 림프절들의 세트에 대한 전이성 암의 대체 LN-레벨 위험들로부터 도출된 라벨들을 적어도 사용하여 훈련되고, 림프절들의 세트의 대체 LN-레벨 위험들은 훈련 이미지들의 양성 림프절과 비암성 림프절을 구별하는 대체 모델을 적어도 사용하여 결정된다.

도면의 간단한 설명

[0018] 본 발명의 이들 및 다른 특징부들, 태양들, 및 장점들은 하기 설명 및 첨부 도면들을 참고하여 더 잘 이해될 것이다.

도 1a는 일 실시 형태에 따른, 대상체의 전이성 암의 대상체 레벨 위험의 결정에 의한 암 병기 결정을 위한 시스템 환경 개요를 도시한다.

도 1b는 일 실시 형태에 따른 암 병기 결정 시스템의 블록도를 도시한다.

도 2a는 일 실시 형태에 따른, 림프절군(lymph node station)들을 식별하는 예시적인 림프절 맵을 도시한다.

도 2b는 일 실시 형태에 따른 예시적인 림프절 주석(annotation)이다.

도 3a는 일 실시 형태에 따른, 대체(imputation) 모델 및 위험 모델의 예시적인 구현예를 도시한다.

도 3b는 일 실시 형태에 따른 위험 모델의 구현예를 더 상세히 도시한다.

도 4는 도 1a, 도 1b, 도 2a, 도 2b, 도 3a 및 도 3b에 도시된 엔터티(entity)들을 구현하기 위한 예시적인 컴퓨터를 예시한다.

도 5a 내지 도 5r은 상이한 림프절 라디오믹 특징들의 예시적인 상관 열 맵(correlation heat map)을 도시한다.

도 6은 대체 모델의 림프절 레벨 성능을 도시한다.

도 7a 및 도 7b는 LN-레벨 분류기에 의해 예측된 최고 확률 점수를 사용하는 대상체 레벨 성능을 도시한다.

도 8은 2단계 부트스트랩 접근법(bootstrap approach)에서의 대체 모델의 림프절 레벨 성능을 도시한다.

도 9a 및 도 9b는 부트스트랩 접근법을 사용하는 대상체 레벨 성능을 도시한다.

실행가능한 경우는 언제나, 유사하거나 동일한 도면 부호가 도면에 사용될 수 있으며, 유사하거나 동일한 기능을 나타낼 수 있다는 것에 유의한다. 예를 들어, "개체 315A"와 같은 도면 부호 뒤의 문자는 텍스트가 그 특정 도면 부호를 갖는 요소를 구체적으로 지칭함을 나타낸다. "개체 315"와 같은 뒤따르는 문자가 없는 텍스트 내의 도면 부호는 그 도면 부호를 지니는 도면들 내의 요소들 중 임의의 것 또는 전부를 지칭한다(예컨대, 텍스트 내의 "개체 315"는 도면들에서 도면 부호 "개체 315A", "개체 315B", "개체 315C" 및 "개체 315D"를 지칭한다).

발명을 실시하기 위한 구체적인 내용

I. 정의

청구범위 및 명세서에 사용된 용어는 달리 명시되지 않는 한 하기에 제시된 바와 같이 정의된다.

용어 "대상체"는 생체내, 생체외, 또는 시험관내, 남성 또는 여성에 상관없이 세포, 조직, 또는 장기, 인간 또는 비-인간을 포함한다.

용어 "포유류"는 인류 및 비-인류 둘 모두를 포함하고, 인류, 비-인류 영장류, 개, 고양이, 쥐, 소, 말, 및 돼지를 포함하지만 이에 한정되지 않는다.

용어 "샘플" 또는 "검사 샘플"은 정맥천자, 배설, 사정, 마사지, 생검, 바늘 흡인, 세척 샘플, 찰과, 외과적 절개, 또는 중재시술을 포함하는 수단들 또는 기타 당업계에 공지된 수단들에 의해 대상체로부터 채취된, 단일 세포 또는 다중 세포들 또는 세포 조각 또는 혈액 샘플과 같은 체액의 일정부분을 포함할 수 있다. 체액의 일정부분의 예들은 양수, 방수, 담즙, 림프, 모유, 간질액, 혈액, 혈장, 귀지(cerumen(earwax)), 쿠퍼액(사정전 액체), 암죽, 미즙, 여성 사정액, 월경, 점액, 타액, 오줌, 구토물, 눈물, 질 윤활액, 땀, 혈청, 정액, 피지, 고름, 흉수, 뇌척수액, 윤활액, 세포내액, 및 유리체액을 포함한다.

용어 "기준 개체"는 알려진 암 병기 결정을 갖는 개체를 지칭한다. 예를 들어, 기준 개체는 건강한 개체(예컨대, 대조 개체)를 포함할 수 있다. 다른 예로서, 기준 개체는 암으로 이전에 진단된 개체를 포함할 수 있다. 다른 예로서, 기준 개체는 림프절로 전이된 암, 예컨대 4기 암으로 이전에 진단된 개체를 포함한다.

용어 "하나 이상의 이미지를 획득하는 것"은 대상체로부터 캡처된 하나 이상의 이미지를 획득하는 것 또는 대상체로부터 캡처된 하나 이상의 이미지로부터 획득된 샘플을 획득하는 것을 포함한다. 하나 이상의 이미지를 획득하는 것은 대상체 또는 대상체로부터 획득된 샘플로부터 하나 이상의 이미지를 캡처하는 단계들을 수행하는 것을 포함할 수 있다. 문구는 또한, 예컨대, 대상체 또는 대상체로부터 획득된 샘플로부터 하나 이상의 이미지를 캡처하는 단계들을 수행한 제3자로부터 하나 이상의 이미지를 수신하는 것을 포함할 수 있다. 하나 이상의 이미지는 통상의 기술자에 의해 저장 메모리에 저장되는 것을 포함하는 다양한 공지된 방법들을 통해 획득될 수 있다.

용어 "림프절 침범(lymph node involvement)", "림프절 전이(lymph node metastasis)" 또는 "림프절 암"은 대상체의 하나 이상의 림프절에서 전이된 암의 존재 또는 부재를 지칭한다.

- [0027] 용어 "림프절 전이 위험", "림프절 레벨 위험" 및 "LN-레벨 위험"은 상호교환적으로 사용되며, 일반적으로 특정 림프절에서의 전이성 암의 위험을 지칭한다.
- [0028] 어구 "전이성 암의 대상체 레벨 위험" 또는 "전이성 암의 환자 레벨 위험"은 상호교환적으로 사용되며, 대상체에서 암이 전이된 가능성을 지칭한다. 다양한 실시 형태들에서, 대상체에서의 암은 대상체의 하나 이상의 림프절에서 전이된 폐암이다.
- [0029] 용어 "암 병기" 또는 "암의 병기"는 대상체에서의 암의 병기를 지칭한다. 암 병기는 0기 내지 4기를 포함하는 TNM 병기 결정 시스템을 지칭할 수 있다. 추가적으로, 암 병기는 결절암 0기 내지 3기를 지칭할 수 있다. 결절암 0기는 결절 침범이 없음을 나타내며, 결절암 1기는 폐문(hilar) 결절 침범을 나타내고, 결절 2기 또는 3기는 종격(mediastinal) 침범을 나타낸다.
- [0030] 본 명세서에 사용되는 바와 같이, 단수 형태("a", "an" 및 "the")는 그 내용이 명확하게 달리 지시하지 않는 한 복수의 지시 대상을 포함하는 것에 유의하여야 한다.
- [0031] **II. 시스템 환경 개요**
- [0032] 도 1a는 일 실시 형태에 따른, 대상체의 전이성 암의 대상체 레벨 위험의 결정에 의한 암 병기 결정을 위한 시스템 환경 개요(100)를 도시한다. 시스템 환경(100)은 대상체(110), 이미징 디바이스(120), 및 암 예측(140)을 결정하기 위한 암 병기 결정 시스템(130)을 소개하기 위한 맥락을 제공한다. 다양한 실시 형태들에서, 암 병기 결정 시스템(130)에 의해 결정된 암 예측(140)은 대상체(110)에 대한 암 병기 결정으로서 역할한다. 다양한 실시 형태들에서, 암 병기 결정 시스템(130)에 의해 결정된 암 예측(140)은 대상체(110)에 대한 예측된 위험 레벨이고, 대상체(110)에 대한 암 병기 결정을 확정하기 위한 기초로서 역할한다.
- [0033] 다양한 실시 형태들에서, 대상체는 건강하다. 따라서, 본 명세서에 기술된 암 병기 결정을 위한 방법은 건강한 대상체에서 암의 조기 검출에 유리할 수 있다. 다양한 실시 형태들에서, 대상체는 이전에 암으로 진단되었다. 다양한 실시 형태들에서, 대상체는 건강하지만 암 형태를 갖는 것으로 의심된다(예컨대, 암과 연관된 증상을 나타낸 대상체). 특정 실시 형태들에서, 대상체의 암의 유형은 폐암이다. 따라서, 본 명세서에 기술된 암 병기 결정을 위한 방법들은 림프절 침범을 포함한 전이성 암의 높은 위험이 있는 대상체들, 및 림프절 전이를 가질 위험이 낮은, 폐암이 의심되는 대상체들을 식별하고 구별하는 데 유리할 수 있다. 전이성 암의 위험이 높은 대상체들은 PET 스캐닝 및 림프절 생검과 같은 추가 검사를 겪을 수 있는 반면, 림프절 전이를 가질 위험이 낮은 대상체들은 폐암의 외과적 절제 또는 조합된 기관지경 진단 및 치료와 같은 즉각적인 요법을 겪을 수 있다.
- [0034] 다양한 실시 형태들에서, 이미징 디바이스(120)는 추가 분석을 위해 대상체(110)로부터 이미지를 캡처한다. 다양한 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)로부터 획득된 검사 샘플로부터 이미지를 캡처한다. 이미지 및/또는 샘플은 개체에 의해 또는 제3자, 예컨대 의료인 또는 의료 전문가에 의해 획득될 수 있다. 의료 전문가들의 예에는 의사, 응급 의료 기술자, 간호사, 응급 의료요원, 심리학자, 채혈사, 의학 물리학 인원, 전담간호사, 외과의사, 치과의사, 및 통상의 기술자에 공지될 임의의 기타 명백한 의료 전문가를 포함한다.
- [0035] 일부 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 해부학적 위치의 이미지를 캡처한다. 대상체의 예시적인 해부학적 위치들은 폐, 흉강, 신장, 간, 췌장, 뇌, 위, 장, 엉덩이, 무릎, 다리, 팔, 및 얼굴을 포함할 수 있다. 다양한 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 흉강의 이미지를 캡처한다. 다양한 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 하나 이상의 림프절의 이미지를 캡처한다. 다양한 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 하나 이상의 림프절을 포함하는 흉강의 이미지를 캡처한다.
- [0036] 다양한 실시 형태들에서, 이미징 디바이스(120)는 컴퓨터 단층촬영(CT) 스캐너, 자기공명 이미징(MRI) 스캐너, 양전자방출단층촬영(PET) 스캐너, X-선 스캐너, 또는 초음파 이미징 디바이스 중 하나이다. 특정 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 하나 이상의 이미지를 캡처하는 CT 스캐너이다. 특정 실시 형태들에서, 이미징 디바이스(120)는 대상체(110)의 하나 이상의 림프절을 포함하는 흉강의 하나 이상의 CT 이미지를 캡처하는 CT 스캐너이다.
- [0037] 일반적으로, 암 병기 결정 시스템(130)은 하나 이상의 이미지(예컨대, 이미징 디바이스(120)에 의해 캡처된 이미지)를 분석하고 암 예측(140)을 생성한다. 다양한 실시 형태들에서, 암 예측(140)은 대상체(110)에서의 암의 예측된 병기일 수 있다. 다양한 실시 형태들에서, 암 예측(140)은 대상체(110)에서의 암 림프절 침범의 예측된 가능성일 수 있다. 다양한 실시 형태들에서, 암 예측(140)은 권장된 임상 접근법(예컨대, PET 스캐닝 및 림프

절 생검과 같은 추가 검사) 또는 요법(예컨대, 외과적 절제 및/또는 약물 요법)을 포함할 수 있다. 암 병기 결정 시스템(130)은 도 4에 관련하여 아래 논의되는 바와 같이 컴퓨터 시스템(400)으로서 구현된, 하나 이상의 컴퓨터를 포함할 수 있다. 따라서, 다양한 실시 형태들에서, 암 병기 결정 시스템(130)에 관련하여 기재된 단계들은 인 실리코(in silico)로 수행된다.

[0038] 다양한 실시 형태들에서, 이미징 디바이스(120) 및 암 병기 결정 시스템(130)은 상이한 당사자들에 의해 채용된다. 예를 들어, 제1 당사자는 이미징 디바이스(120)를 동작시켜 대상체(110)로부터 하나 이상의 이미지를 캡처하고 이어서 암 병기 결정 시스템(130)을 구현하여 암 예측(140)을 결정하는 제2 당사자에게 캡처된 하나 이상의 이미지를 제공한다.

[0039] 이제 일 실시 형태에 따른 암 병기 결정 시스템(130)의 컴퓨터 로직 컴포넌트들을 예시하는 블록도를 도시하는 도 1b를 참조한다. 구체적으로, 암 병기 결정 시스템(130)은 림프절 특징 모듈(145), 대체 모듈(150), 위험 훈련 모듈(155), 위험 전개 모듈(160), 암 예측 모듈(165), 및 훈련 데이터 저장소(170)를 포함한다.

[0040] 일반적으로, 림프절 특징 모듈(145)은 대상체들 또는 훈련 대상체들로부터 유래된 이미지들로부터 특징들을 추출한다. 대체 모듈(150)은 대체 모델을 훈련시키고 전개한다. 대체 모델은 추출된 특징들을 분석하고, 이미지들 내의 각각의 림프절에 대한 암의 림프절 레벨 위험을 결정한다. 위험 훈련 모듈(155)은 대체 모델에 의해 결정된 암의 림프절 레벨 위험을 적어도 포함하는 훈련 데이터를 사용하여 위험 모델을 훈련시킨다. 위험 전개 모듈(160)은 대상체(110)에 대한 대상체 레벨 전이 위험 또는 암 병기의 예측과 같은 암 예측을 결정하기 위하여 대상체(예컨대, 도 1a의 대상체(110))로부터 획득된 이미지들로부터 추출된 특징들을 분석하기 위하여 위험 예측 모델들을 구현한다. 달리 말하면, 암 병기 결정 시스템(130)은 적어도 2개의 모델(예컨대, 대체 모델 및 위험 모델)을 훈련시키고 구현한다. 따라서, 적어도 2개의 모델을 수반하는 이러한 다단계 구현을 통해, 암 병기 결정 시스템(130)은 대상체에 대해 대상체 레벨 전이 위험 또는 암 병기의 예측과 같은 암 예측을 생성한다.

[0041] 대체 모델 및 위험 모델 둘 모두의 구현은 개선된 암 예측이 얻어지게 한다. 구체적으로, 대체 모델은 훈련 데이터의 대상체-레벨 N기부터 학습하고, 림프절들의 추출된 특징들에 적용되어 어느 림프절들이 양성 N기의 원인이 될 가능성이 있는지를 한정한다. 이어서, 위험 모델은 대체 데이터(imputed data)로부터 더 정제된 림프절 전이 위험 확률을 학습한다. 여기서, 위험 모델은 대체 모델에 의해 예측된 대체 LN-레벨 위험들보다 더 큰 확률 범위를 갖는 전이성 암의 LN-레벨 위험들을 예측한다. 이는 전이성 암의 대상체 레벨 위험의 더 정확한 결정을 가능하게 한다.

[0042] 암 병기 결정 시스템(130)의 컴포넌트들은 이하에서 2개의 단계를 참조하여 기술된다: 1) 훈련 단계 및 2) 전개 단계. 더 구체적으로는, 훈련 단계는 기준 개체들(예컨대, 알려진 암 병기 결정을 갖는 개체들)로부터 캡처된 훈련 이미지들과 같은 훈련 데이터에 기초한 하나 이상의 모델의 구축 및 훈련을 지칭한다. 따라서, 모델들은 전개 단계 동안, 모델들의 구현이 대상체(예컨대, 도 1a의 대상체(110))에 대한 암 병기 결정의 예측을 가능하게 하도록 훈련 데이터를 사용하여 훈련된다. 다양한 실시 형태들에서, 대체 모듈(150)은 또한 이하에서 대체 모델로 지칭되는 모델을 훈련시키고 전개하도록 구현될 수 있다. 따라서, 대체 모듈(150)은 훈련 단계 동안 대체 모델을 훈련시키고 전개 단계 동안 대체 모델을 전개한다.

[0043] 일부 실시 형태들에서, 암 병기 결정 시스템(130)의 컴포넌트들은 훈련 단계 및 전개 단계 중 하나 동안 적용된다. 예를 들어, 위험 훈련 모듈(155) 및 훈련 데이터 저장소(170)는 훈련 단계 동안 적용되어 위험 모델을 훈련시킨다. 추가적으로, 위험 전개 모듈(160)은 전개 상태 동안 적용된다. 다양한 실시 형태들에서, 암 병기 결정 시스템(130)의 컴포넌트들은 컴포넌트들이 훈련 단계 동안 적용되는지 아니면 전개 단계 동안 적용되는지에 따라 상이한 당사자들에 의해 수행될 수 있다. 그러한 시나리오들에서, 예측 모델의 훈련 및 전개는 상이한 당사자들에 의해 수행된다. 예를 들어, 훈련 단계 동안 적용되는 위험 훈련 모듈(155) 및 훈련 데이터 저장소(170)는 (예컨대, 위험 모델을 훈련시키기 위해) 제1 당사자에 의해 채용될 수 있고 전개 상태 동안 적용되는 위험 전개 모듈(160)은 (예컨대, 위험 모델을 전개하기 위해) 제2 당사자에 의해 수행될 수 있다. 유사하게, 대체 모듈(150)은 제1 당사자에 의해 대체 모델을 훈련시키기 위해 적용될 수 있고, 대체 모듈(150)은 제2 당사자에 의해 대체 모델을 전개하기 위해 적용될 수 있다.

[0044] III. 암 병기 결정을 위한 방법

[0045] 본 명세서에 기술된 실시 형태들은 대상체로부터 캡처된 하나 이상의 이미지를 분석하고 하나 이상의 모델을 적용하여 림프절을 침범하는 전이성 암 위험의 대상체 레벨을 결정하는 것에 의한 암 병기 결정을 위한 방법을 포함한다. 그러한 방법들은 도 1b에 묘사된 암 병기 결정 시스템(130)에 의해 수행될 수 있다.

- [0046] 도 1b의 암 병기 결정 시스템(130)의 모듈들을 참조하면, 림프절 특징 모듈(145)은 하나 이상의 이미지를 분석하고 하나 이상의 이미지로부터 특징들을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 기준 개체들로부터 획득된 훈련 이미지들로부터 특징들을 추출한다. 훈련 이미지들로부터 추출된 이들 특징은 하기에 더 상세히 기술되는 바와 같이 대체 모델과 같은 하나 이상의 모델을 훈련시키는 데 사용될 수 있다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 대상체(예컨대, 도 1a의 대상체(110))로부터 획득된 이미지들로부터 특징들을 추출한다. 대상체로부터 획득된 이미지들로부터 추출된 이들 특징은 하기에 더 상세히 기술되는 바와 같이 위험 모델과 같은 모델을 전개함으로써 분석될 수 있다.
- [0047] 다양한 실시 형태들에서, 하나 이상의 이미지 또는 훈련 이미지는 이미지에서 하나 이상의 림프절을 포함한다. 예를 들어, 이미지 또는 훈련 이미지는 하나 이상의 림프절을 포함하는 흉강의 이미지일 수 있다. 다양한 실시 형태들에서, 이미지는 적어도 2개의 림프절을 포함한다. 다양한 실시 형태들에서, 이미지는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21개의 림프절을 포함한다. 다양한 실시 형태들에서, 이미지는 1 내지 10개의 림프절, 1 내지 15개의 림프절, 2 내지 10개의 림프절, 2 내지 15개의 림프절, 2 내지 20개의 림프절, 4 내지 10개의 림프절, 4 내지 15개의 림프절, 4 내지 20개의 림프절, 5 내지 10개의 림프절, 5 내지 15개의 림프절, 5 내지 20개의 림프절, 8 내지 15개의 림프절, 8 내지 20개의 림프절, 5 내지 20개의 림프절, 8 내지 15개의 림프절, 8 내지 20개의 림프절, 10 내지 15개의 림프절, 10 내지 20개의 림프절, 12 내지 15개의 림프절, 12 내지 20개의 림프절, 또는 15 내지 20개의 림프절을 포함한다. 다양한 실시 형태들에서, 다수의 이미지가 단일 대상체 또는 기준 개체로부터 캡처된다. 다수의 이미지는 대상체 또는 기준 개체의 상이한 림프절들을 캡처할 수 있다. 다양한 실시 형태들에서, 다수의 이미지는 적어도 2개의 림프절을 포함한다. 다양한 실시 형태들에서, 다수의 이미지는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 또는 21개의 림프절을 포함한다. 다양한 실시 형태들에서, 다수의 이미지는 8 내지 20개의 림프절을 포함한다. 다양한 실시 형태들에서, 다수의 이미지는 10 내지 19개의 림프절, 12 내지 18개의 림프절, 또는 14 내지 16개의 림프절을 포함한다.
- [0048] 도 2a는 일 실시 형태에 따른, 림프절군들을 식별하는 예시적인 림프절 맵을 도시한다. 보다 구체적으로, 도 2a는 미국 흉부 학회(American Thoracic Society) 매핑 방식(mapping scheme)으로부터 조정된 국소 림프절 분류들의 도시이다. 여기서, 림프절군들은 각각의 림프절의 위치 및 분류를 나타내기 위해 주석이 달린다. 구체적으로, 주석들은 다음과 같다: (1) 쇄골상 림프절(Supraclavicular node), (2R/2L) 상부 기관주위 림프절(Upper paratracheal node), (3A) 혈관전 림프절(Pre-vascular node), (3P) 척추전 림프절(Pre-vertebral node), (4R/4L) 하부 기관주위 림프절, (5) 대동맥하 림프절(Subaortic node), (6) 대동맥주위 림프절(para-aortic node), (7) 용골하 림프절(subcarinal node), (8) 식도주위 림프절(paraesophageal node), (9) 폐간막 림프절(pulmonary ligament node), (10) 폐문 림프절, (11) 엽간 림프절(interlobar node), (12) 엽성 림프절(lobar node), (13) 분절 림프절(segmental node), (14) 세분절 림프절(sub-segmental node).
- [0049] 림프절 특징 모듈(145)은 이미지 내의 하나 이상의 림프절에 주석이 달린다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 림프절 주위의 영역을 한정함으로써 하나 이상의 림프절에 주석이 달린다. 다양한 실시 형태들에서, 림프절 주위의 한정된 영역은 구형 영역이다. 다양한 실시 형태들에서, 구형 영역은 X mm의 반경을 갖는다. 다양한 실시 형태들에서, X는 5 mm이다. 다양한 실시 형태들에서, X는 7.5 mm이다. 일부 실시 형태들에서, X는 0.5 mm, 1 mm, 1.5 mm, 2 mm, 2.5 mm, 3 mm, 3.5 mm, 4 mm, 4.5 mm, 5.5 mm, 6 mm, 6.5 mm, 7 mm, 8 mm, 8.5 mm, 9 mm, 9.5 mm, 10 mm, 10.5 mm, 11 mm, 11.5 mm, 12 mm, 12.5 mm, 13 mm, 13.5 mm, 14 mm, 14.5 mm, 또는 15 mm이다. 다양한 실시 형태들에서, X는 1 내지 15 mm이다. 다양한 실시 형태들에서, X는 2 내지 12 mm이다. 다양한 실시 형태들에서, X는 3 내지 10 mm이다. 다양한 실시 형태들에서, X는 4 내지 8 mm이다. 다양한 실시 형태들에서, X는 5 내지 7.5 mm이다. 다양한 실시 형태들에서, X는 4 내지 6 mm이다. 다양한 실시 형태들에서, X는 6 내지 8 mm이다. 다양한 실시 형태들에서, X는 7 내지 8 mm이다. 다양한 실시 형태들에서, X는 10 내지 15 mm이다.
- [0050] 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 림프절들 주위에 2개 이상의 구형 영역을 한정함으로써 하나 이상의 림프절에 주석이 달린다. 예를 들어, 림프절 특징 모듈(145)은 림프절 주위의 5 mm의 제1 구형 영역 및 림프절 주위의 7.5 mm의 제2 구형 영역을 한정함으로써 하나 이상의 림프절에 주석이 달린다. 도 2b는 일 실시 형태에 따른 예시적인 림프절 주석이다. 구체적으로, 도 2b는 2개의 2R 상부 기관주위 림프절 주위의 2개의 한정된 구형 영역을 도시한다.
- [0051] 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 림프절 주위의 각각의 한정된 영역으로부터 특징들을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 원본 입력의 변환을 수행하여 변환된 이미지를 생성하

고, 이로부터 추가 특징들이 이어서 추출된다. 일 실시 형태에서, 림프절 특징 모듈(145)은 원본 이미지의 웨이블릿 변환을 수행한다. 일 실시 형태에서, 림프절 특징 모듈(145)은 원본 이미지의 가우스 라플라시안(LoG)을 수행한다. 따라서, 림프절 특징 모듈(145)은 변환된 이미지 내의 림프절들 주위의 한정된 영역들로부터 특징들을 추출할 수 있다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 원본 이미지 및 변환된 이미지 내의 림프절들 주위의 한정된 영역들로부터 특징들을 추출한다.

[0052] 특정 실시 형태들에서, 림프절 특징 모듈(145)은 원본 이미지의 2개의 변환(예컨대, 원본 이미지의 웨이블릿 변환 및 원본 이미지의 가우스 라플라시안(LoG) 둘 모두)을 수행한다. 따라서, 림프절 특징 모듈(145)은 제1 변환된 이미지 및 제2 변환된 이미지 둘 모두에서 림프절들 주위의 한정된 영역들로부터 특징들을 추출할 수 있다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 원본 이미지, 제1 변환된 이미지 및 제2 변환된 이미지 내의 림프절들 주위의 한정된 영역들로부터 특징들을 추출한다.

[0053] 원본 이미지, 제1 변환된 이미지, 및/또는 제2 변환된 이미지로부터의 예시적인 특징 카테고리들은 1차 특징, 형상 특징, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징, 그레이 레벨 작용 길이 행렬(GLRLM) 특징, 그레이 레벨 크기 구역 행렬(GLSZM) 특징, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징을 포함할 수 있다. 1차 특징들은 강도(intensity) 값들의 1차 통계들의 특징들을 지칭한다. GLCM 특징들은 이미지에 걸쳐 주어진 오프셋에서의 동시 발생 그레이스케일 값들의 히스토그램으로부터 도출된다. GLRLM 특징들은 그레이 레벨 작용 길이 행렬의 통계들을 수반한다. GLSZM 특징들은 그레이 레벨 크기 구역 행렬의 통계들을 수반한다. NGTDM 특징들은 이웃 그레이 톤 차이 행렬의 통계들을 수반한다. 상이한 특징 카테고리들에서의 특징들의 예들은 상관 특징, LMC2, MCC, 엔트로피 차이, 결합 엔트로피(Joint Entropy), 엔트로피 합, 클러스터 프로미넌스(Cluster prominence), 평균 차이, 결합 평균, 평균 합, 자기상관(autocorrelation), 분산 차이, 콘트라스트, 클러스터 경향(cluster tendency), 제곱 합(sum of squares), LDMN, LDN, 클러스터 셰이드(Cluster shade), LMC1, 역분산(Inverse variance), LD, LDM, 결합 에너지, 및 최대 확률을 포함한다.

[0054] 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 이미지(210)로부터 적어도 50개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 이미지로부터 적어도 100개의 특징, 적어도 150개의 특징, 적어도 200개의 특징, 적어도 250개의 특징, 적어도 300개의 특징, 적어도 350개의 특징, 적어도 400개의 특징, 적어도 450개의 특징, 적어도 500개의 특징, 적어도 550개의 특징, 적어도 600개의 특징, 적어도 650개의 특징, 적어도 700개의 특징, 적어도 750개의 특징, 적어도 800개의 특징, 적어도 850개의 특징, 적어도 900개의 특징, 적어도 950개의 특징, 적어도 1000개의 특징, 적어도 1100개의 특징, 적어도 1200개의 특징, 적어도 1300개의 특징, 적어도 1400개의 특징, 적어도 1500개의 특징, 적어도 1600개의 특징, 적어도 1700개의 특징, 적어도 1800개의 특징, 적어도 1900개의 특징, 적어도 2000개의 특징, 적어도 3000개의 특징, 적어도 4000개의 특징, 적어도 5000개의 특징, 적어도 6000개의 특징, 적어도 7000개의 특징, 적어도 8000개의 특징, 적어도 9000개의 특징, 또는 적어도 10,000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 10개의 특징 내지 10,000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 100개의 특징 내지 1000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 300개의 특징 내지 900개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 500개의 특징 내지 1000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 500개의 특징 내지 900개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 700개의 특징 내지 1000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 700개의 특징 내지 900개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 800개의 특징 내지 1000개의 특징을 추출한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 800개의 특징 내지 900개의 특징을 추출한다.

[0055] 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 대체 모델을 훈련시키거나 전개하기 위해, 추출된 특징들을 대체 모듈(150)에 제공한다. 예를 들어, 추출된 특징들이 기준 개체들로부터 캡처된 훈련 이미지들로부터 도출되는 경우, 림프절 특징 모듈(145)은 추출된 특징들을 대체 모듈(150)에 제공한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 위험 모델을 훈련시키기 위해, 추출된 특징들을 위험 훈련 모듈(155)에 제공한다. 예를 들어, 추출된 특징들이 기준 개체들로부터 캡처된 훈련 이미지들로부터 도출되는 경우, 림프절 특징 모듈(145)은 위험 모델을 훈련시키기 위해, 추출된 특징들을 위험 훈련 모듈(155)에 제공한다. 다양한 실시 형태들에서, 림프절 특징 모듈(145)은 위험 모델을 전개하기 위해, 추출된 특징들을 위험 전개 모듈(160)에 제공한다. 예를 들어, 추출된 특징들이 대상체(예컨대, 도 1a의 대상체(110))로부터 캡처된 이미지들로부터 도출되는 경우, 림프절 특징 모듈(145)은 추출된 특징들을 위험 전개 모듈(160)에 제공한다.

[0056] III.A. 대체 모델의 훈련

[0057] 대체 모듈(150)은 대체 모델을 훈련시키고 전개한다. 일반적으로, 대체 모델은 하나 이상의 림프절을 포함하는 이미지로부터 추출된 특징들을 분석하고, 림프절이 양성인 하나 이상의 림프절에 대한 확률을 예측하도록 구성된다. 따라서, 하나 이상의 림프절을 포함하는 이미지로부터의 추출된 특징들에 기초하여, 대체 모델은 림프절들의 각각에 대해 확률을 대체하며, 대체 확률은 림프절이 양성(예컨대, 전이성 암)일 가능성을 나타낸다. 달리 말하면, 림프절 레벨에서 전이성 암의 위험을 결정함으로써 양성 림프절과 비양성(non-cancerous) 림프절을 구별한다.

[0058] 다양한 실시 형태들에서, 대체 모델은 회귀 모델(예컨대, 선형 회귀, 로지스틱 회귀, 또는 다항식 회귀), 결정 트리, 랜덤 포레스트, 지원 벡터 머신, 나이브 베이즈 모델(Naïve Bayes model), k-평균 클러스터, 또는 신경망(예컨대, 순방향-공급 네트워크, 콘볼루션 신경망(convolutional neural network, CNN), 딥 신경망(deep neural network, DNN), 오토인코더 신경망, 생성적 적대 네트워크, 또는 순환 네트워크(예컨대, 장단기 메모리 네트워크(long short-term memory network, LSTM), 양방향 순환 네트워크, 딥 양방향 순환 네트워크), 또는 이들의 임의의 조합 중 어느 하나이다. 특정 실시 형태들에서, 대체 모델은 랜덤 포레스트 분류기 또는 최소 절대 수축 및 선택 연산자(least absolute shrinkage and selection operator, LASSO) 분류기이다. 다양한 실시 형태들에서, 대체 모델은 기계 학습된 모델이다. 다양한 실시 형태들에서, 대체 모델은 부스팅된 그래디언트 기계 학습 모델(boosted gradient machine learning model)(예컨대, xgboost)이다.

[0059] 대체 모델은 기계 학습 구현 방법, 예를 들어 선형 회귀 알고리즘, 로지스틱 회귀 알고리즘, 결정 트리 알고리즘, 지원 벡터 머신 분류, 나이브 베이즈 분류, K-최근접 이웃 분류, 랜덤 포레스트 알고리즘, 딥 러닝 알고리즘, 그래디언트 부스팅 알고리즘, 및 차원 축소 기법들, 예를 들어 다양체 학습(manifold learning), 주 성분 분석, 인자 분석, 오토인코더 정규화(autoencoder regularization), 및 독립 성분 분석, 또는 이들의 조합 중 어느 하나를 사용하여 훈련될 수 있다. 다양한 실시 형태들에서, 대체 모델은 지도형 학습 알고리즘, 비지도형 학습 알고리즘, 반-지도형 학습 알고리즘(예컨대, 부분 지도), 약한 지도, 전달, 멀티-태스크 학습, 또는 이들의 임의의 조합을 이용하여 훈련된다.

[0060] 다양한 실시 형태들에서, 대체 모델은 하나 이상의 파라미터, 예를 들어 하이퍼파라미터 또는 모델 파라미터를 갖는다. 하이퍼파라미터들은 일반적으로 훈련 이전에 확립된다. 하이퍼파라미터들의 예들은 학습 속도, 결정 트리의 깊이 또는 리프(leaf)들, 딥 신경망에서의 은닉 층들의 개수, k-평균 클러스터에서의 클러스터의 개수, 회귀 모델에서의 페널티, 및 비용 함수와 연관된 정규화 파라미터를 포함한다. 모델 파라미터들은 일반적으로 훈련 동안 조정된다. 모델 파라미터들의 예들은 신경망의 층들에서의 노드들, 지원 벡터 머신에서의 지원 벡터들, 및 회귀 모델의 계수들과 연관된 가중치들을 포함한다. 대체 모델의 모델 파라미터들은 대체 모델의 예측 능력을 개선을 개선하기 위해 훈련 데이터를 이용하여 훈련(예컨대, 조정)된다.

[0061] 대체 모듈(150)은 훈련 데이터를 사용하여 대체 모델을 훈련시킨다. 훈련 데이터는 훈련 데이터 저장소(170)에 저장되고/되거나 그로부터 검색될 수 있다. 다양한 실시 형태들에서, 훈련 데이터는 기준 개체들로부터 획득된 하나 이상의 림프절을 포함하는 훈련 이미지들로부터의 추출된 특징들을 포함한다. 다양한 실시 형태들에서, 기준 개체는 건강한 개체(예컨대, 대조군 개체)를 포함한다. 다양한 실시 형태들에서, 기준 개체는 암으로 이전에 진단된 개체를 포함한다. 다양한 실시 형태들에서, 기준 개체는 림프절로 전이된 암으로 이전에 진단된 개체를 포함한다. 따라서, 그러한 개체로부터 캡처된 이미지는 하나 이상의 양성 림프절을 포함한다.

[0062] 다양한 실시 형태들에서, 대체 모델을 훈련시키는 데 사용되는 훈련 데이터는 림프절이 양성(이하 "양성" 또는 "+")로도 지칭됨)인지 여부 또는 림프절이 비양성(이하 "음성" 또는 "-")로도 지칭됨)인지 여부를 나타내는 기준 실측 자료(reference ground truth)를 포함한다. 다양한 실시 형태들에서, 훈련 데이터의 기준 실측 자료는 "1" 또는 "0"과 같은 이진 값들이다. 예를 들어, 양성인 양성 림프절은 "1"의 값을 갖는 훈련 데이터에서 식별될 수 있는 반면, 비양성인 음성 림프절은 "0"의 값을 갖는 훈련 데이터에서 식별될 수 있다. 함께, 대체 모듈(150)은 훈련 데이터를 사용하여 대체 모델을 훈련시켜 손실 함수를 최소화하여, 대체 모델은 입력(예컨대, 림프절의 추출된 특징들)에 기초하여 결과(예컨대, 림프절에 대한 림프절 전이성 암의 확률)를 더 잘 예측할 수 있다. 다양한 실시 형태들에서, 손실 함수는 최소 절대 수축 및 선택 연산자(LASSO) 회귀, 리지(Ridge) 회귀, 또는 엘라스틱넷(ElasticNet) 회귀 중 임의의 것에 대해 구성된다.

[0063] 다양한 실시 형태들에서, 훈련 데이터는 공개적으로 이용가능한 데이터베이스로부터 획득 및/또는 도출될 수 있다. 예를 들어, 훈련 데이터는 기준 개체에 대한 결절 병기(N기) 암 정보를 포함하는, 국가폐암검진연구(NLS

T)로부터 획득 및/또는 도출될 수 있다. 일부 실시 형태들에서, 훈련 데이터는, 예컨대 복수의 기준 개체로부터 하나 이상의 림프절의 이미지들을 캡처함으로써 공개적으로 이용가능한 데이터베이스들에 독립적으로 획득 및 수집될 수 있다. 예를 들어, 그러한 훈련 데이터는 맞춤형 데이터세트일 수 있다. 맞춤형 데이터세트의 훈련 데이터는 환자 치료 동안 획득될 수 있고, 따라서 림프절의 대응하는 외과적 샘플링 정보를 가질 수 있다. 훈련 데이터는 기준 개체들에 대한 N기 암 정보, 예컨대 기준 개체들에 대한 이미징 및/또는 생검 결과들에 기초하여 임상이에 의해 이전에 진단된 N기 정보를 포함할 수 있다.

[0064] 다양한 실시 형태들에서, 림프절이 양성인지 여부를 나타내는 기준 실측 자료는 임상 병기 데이터로부터 도출된 결절 병기(N기) 정보이다. 예를 들어, 공개적으로 이용가능한 NLST 데이터세트는 기준 개체들에 대한 임상 병기 정보(예컨대, 병기 1a, 1b, 2a, 2b, 3a 등)를 포함하지만, 림프절 특정 정보를 포함하지 않는다. 다시 말하면, NLST 데이터세트는 개별 림프절이 양성 또는 비양성인지 여부를 구별하지 않는다. 그러한 실시 형태들에서, 림프절이 양성인지 여부를 나타내는 기준 실측 자료는 N기 정보 및 기준 개체의 암의 다른 특성들(예컨대, 종양 크기, 종양 위치, 종양 형상 및/또는 종양 밀도) 및/또는 림프절의 다른 특성들(예컨대, 림프절 크기, 림프절 밀도, 및/또는 림프절 형상)에 기초하여 다시 계산될 수 있다. 제1 예로서, 0 또는 1의 N기를 갖는 기준 개체는 암 전이에 아직 직면하지 않았고, 따라서 이 기준 개체의 림프절에는 "음성" 또는 비양성의 기준 실측 자료가 할당된다. 다른 예로서, 2 또는 3의 N기를 갖는 기준 개체는 암 전이에 직면하였다. 따라서, 종양 및/또는 림프절의 특성들에 기초하여, 림프절에는 "양성" 또는 "음성"의 기준 실측 자료가 할당된다.

[0065] III.B. 대체 모델의 전개

[0066] 대체 모듈(150)은 훈련된 대체 모델을 전개시켜 하나 이상의 림프절을 포함하는 이미지로부터 추출된 특징들을 분석하고, 림프절이 양성인 하나 이상의 림프절에 대한 확률을 예측한다. 다양한 실시 형태들에서, 훈련된 대체 모델은 시험 데이터세트에 대해 전개된다. 다양한 실시 형태들에서, 시험 데이터세트 및 대체 모델을 훈련시키는 데 사용되는 훈련 데이터세트는 공통 데이터세트로부터 도출될 수 있다. 예를 들어, 공통 데이터세트는 대체 모델을 훈련시키고 시험하기 위한 2개의 서브세트로 분할될 수 있다. 일부 실시 형태들에서, 공통 데이터세트는 50:50 훈련:시험 데이터세트 분할을 겪는다. 일부 실시 형태들에서, 공통 데이터세트는 60:40 훈련:시험 데이터세트 분할을 겪는다. 일부 실시 형태들에서, 공통 데이터세트는 80:20 훈련:시험 데이터세트 분할을 겪는다.

[0067] 훈련된 대체 모델의 전개는 림프절들에 대한 확률들(예컨대, 0과 1 사이의 연속 확률들)을 산출하는데, 각각의 확률은 대응하는 림프절이 양성일 가능성이 있는지 여부를 나타낸다. 다양한 실시 형태들에서, 대체 모듈(150)은 위험 모델을 훈련시키는 데 사용될 대체 모델에 의해 예측된 림프절들에 대한 림프절 전이 위험들 중 하나 이상을 제공한다. 다양한 실시 형태들에서, 대체 모듈(150)은 위험 모델을 훈련시키는 데 사용될 대체 모델에 의해 예측된 림프절들에 대한 림프절 전이 위험들 전부를 제공한다. 다양한 실시 형태들에서, 대체 모듈(150)은 대체 모델에 의해 예측된 림프절들에 대한 림프절 전이 위험들의 서브세트를 식별하기 위해 선택 프로세스를 수행한다. 따라서, 림프절 전이 위험들의 서브세트는 후속적으로 위험 모델을 훈련시키는 데 사용된다.

[0068] 일부 실시 형태들에서, 대체 모듈(150)은 림프절 전이 위험들을 임계 값과 비교함으로써 림프절 전이 위험들의 서브세트를 선택한다. 림프절에 대한 림프절 전이 위험이 임계 값 초과인 경우, 림프절에 대한 림프절 전이 위험은 위험 모델을 훈련시키기 위한 서브세트에 포함되도록 선택된다. 림프절에 대한 림프절 전이 위험이 임계 값 미만인 경우, 림프절에 대한 림프절 전이 위험은 서브세트로부터 배제되고, 따라서 위험 모델을 훈련시키기 위해 사용되지 않는다.

[0069] 다양한 실시 형태들에서, 임계 값은 고정 값이다. 다양한 실시 형태들에서, 임계 값은 전이성 암의 50% 위험이다. 다양한 실시 형태들에서, 임계 값은 전이성 암의 30%, 35%, 40%, 45%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% 또는 99% 위험이다. 일부 실시 형태들에서, 임계 값은 30 내지 70%이다. 일부 실시 형태들에서, 임계 값은 40 내지 60%이다. 일부 실시 형태들에서, 임계 값은 45 내지 55%이다. 일부 실시 형태들에서, 임계 값은 48 내지 52%이다. 일부 실시 형태들에서, 임계 값은 49 내지 51%이다. 다양한 실시 형태들에서, 임계 값은 가변적이다. 다양한 실시 형태들에서, 임계 값은 암에 걸리는 것에 대한 중앙 대상체내 확률이다.

[0070] 대체 모듈(150)은 위험 모델을 훈련시키기 위해 림프절 전이 위험들의 서브세트를 위험 훈련 모듈(155)에 제공한다.

[0071] III.C. 위험 모델의 훈련

- [0072] 위험 훈련 모듈(155)은 대체 모델에 의해 예측된 림프절들에 대한 림프절 전이 위험들 중 적어도 하나 이상을 포함하는 훈련 데이터를 사용하여 위험 모델을 훈련시킨다. 여기서, 하나 이상의 림프절 전이 위험은 전술된 바와 같이, 대체 모듈(150)에 의해 선택된 림프절 전이 위험들의 서브세트일 수 있다. 위험 모델을 훈련시키기 위한 림프절 전이 위험들의 서브세트는 대체 모델을 훈련하는 데 사용되었던 훈련 데이터세트와 비교하여 정제된 훈련 데이터세트를 나타낸다. 구체적으로, 정제된 훈련 데이터세트는 양성일 가능성이 더 높은 림프절들의 서브세트를 포함하는 반면, 양성일 가능성이 더 적은(그러나 달리 위양성을 나타낼 수 있거나 위험 모델의 훈련을 혼란시키고 있을 수 있는) 림프절들은 정제된 훈련 데이터세트로부터 제거된다. 따라서, 이러한 정제된 훈련 데이터세트로 위험 모델을 훈련시킴으로써, 위험 모델은 전이성 질환을 갖는 대상체들에서 모든 림프절들에 대해 더 큰 확률 범위를 예측한다. 보다 구체적으로, 위험 모델은 암에 걸릴 가능성이 있는 림프절들에 더 높은 림프절 전이 위험을 할당하고 암에 걸릴 가능성이 없는 림프절들에 더 낮은 림프절 전이 위험을 할당할 가능성이 더 크다.
- [0073] 다양한 실시 형태들에서, 위험 모델은 하나 이상의 림프절을 포함하는 이미지로부터 추출된 특징들을 분석하고, 림프절이 양성인 하나 이상의 림프절의 각각에 대해 연속 확률을 예측한다. 따라서, 하나 이상의 림프절을 포함하는 이미지로부터의 추출된 특징들에 기초하여, 위험 모델은 림프절들의 각각에 대해 확률을 결정하며, 결정된 확률은 림프절이 양성(예컨대, 전이성 암)일 가능성을 나타낸다. 여기서, 위험 모델은 림프절 레벨에서 전이성 암의 위험들을 결정한다. 다양한 실시 형태들에서, 위험 모델은, 둘 모두의 모델이 이미지로부터 추출된 특징들을 분석하고 림프절 레벨에서 전이성 암의 위험들을 예측한다는 점에서 대체 모델과 유사하다. 그러나, 위험 모델은 대체 모델의 예측들, 예컨대 대체 모델에 의해 예측된 하나 이상의 림프절 레벨 위험을 사용하여 훈련된다. 함께, 이는 위험 모델이 대체 모델에 의해 예측된 대체 LN-레벨 위험들보다 더 큰 확률 범위를 갖는 전이성 암의 LN-레벨 위험들을 예측할 수 있게 한다.
- [0074] 위험 모델은 대체 모델에 의해 예측된 적어도 하나 이상의 림프절 레벨 위험을 포함하는 훈련 데이터를 사용하여 훈련된다. 여기서, 하나 이상의 림프절 전이 위험은 위험 모델을 훈련시키기 위한 기준 실측 자료로서 역할할 수 있다. 다양한 실시 형태들에서, 훈련 데이터는 기준 개체들로부터 획득된 하나 이상의 림프절을 포함하는 훈련 이미지들로부터의 추출된 특징들을 추가적으로 포함할 수 있다. 일부 실시 형태들에서, 위험 모델을 훈련시키는 데 사용되는 훈련 이미지들은 대체 모델을 훈련시키는 데 이전에 사용된 동일한 훈련 이미지들일 수 있다. 일부 실시 형태들에서, 위험 모델을 훈련시키는 데 사용되는 훈련 이미지들은 대체 모델을 훈련시키는 데 이전에 사용된 훈련 이미지들과 상이하다.
- [0075] 다양한 실시 형태들에서, 훈련 데이터는 위험 모델이 각각의 훈련 예를 사용하여 훈련되도록 개별 훈련 예들을 포함한다. 예를 들어, 각각의 훈련 예는 특정 훈련 이미지의 하나 이상의 림프절의 추출된 특징들뿐만 아니라 기준 실측 자료(예컨대, 특정 훈련 이미지의 동일한 하나 이상의 림프절에 대한 림프절 전이 위험)를 포함할 수 있다. 따라서, 각각의 훈련 예에 대해, 위험 모델은 손실 함수(예컨대, 최소 절대 수축 및 선택 연산자(LASSO) 회귀, 리지 회귀, 또는 엘라스틱넷 회귀 중 임의의 것에 대한 손실 함수)를 최소화하도록 훈련된다.
- [0076] 다양한 실시 형태들에서, 위험 모델은 회귀 모델(예컨대, 선형 회귀, 로지스틱 회귀, 또는 다항식 회귀), 결정 트리, 랜덤 포레스트, 지원 벡터 머신, 나이브 베이즈 모델, k-평균 클러스터, 또는 신경망(예컨대, 순방향-공급 네트워크, 콘볼루션형 신경망(CNN), 딥 신경망(DNN), 오토인코더 신경망, 생성적 적대 네트워크, 또는 순환 네트워크(예컨대, 장단기 메모리 네트워크(LSTM), 양방향 순환 네트워크, 딥 양방향 순환 네트워크), 또는 이들의 임의의 조합 중 어느 하나이다. 특정 실시 형태들에서, 위험 모델은 랜덤 포레스트 분류기 또는 최소 절대 수축 및 선택 연산자(LASSO) 분류기이다. 다양한 실시 형태들에서, 위험 모델은 기계 학습된 모델이다.
- [0077] 위험 모델은 기계 학습 구현 방법, 예를 들어 선형 회귀 알고리즘, 로지스틱 회귀 알고리즘, 결정 트리 알고리즘, 지원 벡터 머신 분류, 나이브 베이즈 분류, K-최근접 이웃 분류, 랜덤 포레스트 알고리즘, 딥 러닝 알고리즘, 그래디언트 부스팅 알고리즘, 및 차원 축소 기법들, 예를 들어 다양체 학습, 주성분 분석, 인자 분석, 오토인코더 정규화, 및 독립 성분 분석, 또는 이들의 조합 중 어느 하나를 사용하여 훈련될 수 있다. 다양한 실시 형태들에서, 위험 모델은 지도형 학습 알고리즘, 비지도형 학습 알고리즘, 반-지도형 학습 알고리즘(예컨대, 부분 지도), 약한 지도, 전달, 멀티-태스크 학습, 또는 이들의 임의의 조합을 이용하여 훈련된다.
- [0078] 다양한 실시 형태들에서, 위험 모델은 하나 이상의 파라미터, 예를 들어 하이퍼파라미터 또는 모델 파라미터를 갖는다. 하이퍼파라미터들은 일반적으로 훈련 이전에 확립된다. 하이퍼파라미터들의 예들은 학습 속도, 결정 트리의 깊이 또는 리프들, 딥 신경망에서의 은닉 층들의 개수, k-평균 클러스터에서의 클러스터의 개수, 회귀 모델에서의 페널티, 및 비용 함수와 연관된 정규화 파라미터를 포함한다. 모델 파라미터들은 일반적으로 훈련

동안 조정된다. 모델 파라미터들의 예들은 신경망의 층들에서의 노드들, 지원 벡터 머신에서의 지원 벡터들, 및 회귀 모델의 계수들과 연관된 가중치들을 포함한다. 위험 모델의 모델 파라미터들은 위험 모델의 예측 능력을 개선을 개선하기 위해 훈련 데이터를 이용하여 훈련(예컨대, 조정)된다.

[0079] III.C. 위험 모델의 전개

[0080] 위험 전개 모듈(160)은 대상체(예컨대, 도 1a에 묘사된 대상체(110))에 대한 전이성 암의 대상체 레벨 위험을 결정하기 위해 위험 모델을 전개한다. 여기서, 위험 전개 모듈(160)은 훈련된 위험 모델에 대한 입력으로서 대상체로부터 캡처된 하나 이상의 림프절을 포함하는 이미지들의 추출된 특징들을 제공한다. 훈련된 위험 모델은 추출된 특징들을 분석하고, 개별 림프절이 양성일 가능성이 있는지 여부를 나타내는 림프절들에 대한 확률들을 출력한다.

[0081] 다양한 실시 형태들에서, 위험 모델의 전개 후에, 위험 전개 모듈(160)은 위험 모델에 의해 예측된 전이성 암의 개별 림프절 레벨 위험들로부터 대상체에 대한 전이성 암의 대상체 레벨 위험으로 이행한다. 다양한 실시 형태들에서, 위험 전개 모듈(160)은 전이성 암의 개별 림프절 레벨 위험들을 조합하여 전이성 암의 대상체 레벨 위험을 결정한다. 다양한 실시 형태들에서, 위험 전개 모듈(160)은 전이성 암의 림프절 레벨 위험들에 걸쳐 통계 연산을 수행한다. 통계 연산은 전이성 암의 개별 림프절 레벨 위험들에 걸쳐 최대 값, 최소 값, 평균 값, 중앙 값 또는 모드 값을 식별할 수 있다. 특정 실시 형태들에서, 통계 연산은 전이성 암의 개별 림프절 레벨 위험들에 걸친 최대 값이다. 따라서, 전이성 질환의 최고 림프절 레벨 위험을 갖는 림프절은 대상체에 대한 전이성 질환의 대상체 레벨 위험으로서 할당된다.

[0082] 암 병기 결정 모듈(165)은 대상체에 대한 전이성 질환의 대상체 레벨 위험에 기초하여 대상체에 대한 암 병기를 결정한다. 다양한 실시 형태들에서, 암 병기 결정 모듈(165)은 대상체에 대한 전이성 질환의 대상체 레벨 위험을 하나 이상의 임계 위험 값과 비교함으로써 대상체에 대한 암 병기를 결정한다. 이들 임계 위험 값은 대상체에서의 암이 0기, 1기, 2기, 또는 3기 중 임의의 것일 가능성이 있는지 여부를 결정하기 위한 컷오프들로서 역할할 수 있다.

[0083] 일부 실시 형태들에서, 대상체 레벨 위험이 30% 초과인 경우, 대상체는 적어도 3기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 40% 초과, 50% 초과, 60% 초과, 70% 초과, 80% 초과, 또는 90% 초과인 경우, 대상체는 적어도 3기 암에 걸린 것으로서 분류된다.

[0084] 다양한 실시 형태들에서, 임계 위험 값들은 하나 이상의 암 병기에 대응하는 위험 범위들을 한정한다. 따라서, 임계 위험 값들은 위험 범위들의 상한 및 하한을 설정한다. 일부 실시 형태들에서, 대상체 레벨 위험이 30% 내지 100%인 경우, 대상체는 결절 3기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 40% 내지 100%인 경우, 대상체는 결절 3기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 50% 내지 100%인 경우, 대상체는 결절 3기 암에 걸린 것으로서 분류된다.

[0085] 일부 실시 형태들에서, 대상체 레벨 위험이 10% 내지 40%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 15% 내지 40%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 20% 내지 40%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 30% 내지 40%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 15% 내지 30%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 20% 내지 30%인 경우, 대상체는 결절 2기 암에 걸린 것으로서 분류된다.

[0086] 일부 실시 형태들에서, 대상체 레벨 위험이 5% 내지 20%인 경우, 대상체는 결절 0기 또는 결절 1기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 10% 내지 20%인 경우, 대상체는 결절 0기 또는 결절 1기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 5% 내지 15%인 경우, 대상체는 결절 0기 또는 결절 1기 암에 걸린 것으로서 분류된다. 일부 실시 형태들에서, 대상체 레벨 위험이 8% 내지 12%인 경우, 대상체는 결절 0기 또는 결절 1기 암에 걸린 것으로서 분류된다.

[0087] IV. 암 병기 결정을 위한 예시적인 방법

[0088] 도 3a는 일 실시 형태에 따른, 대체 모델 및 위험 모델의 예시적인 구현예를 도시한다. 여기서, 기준 개체(305A, 305B)들은 특정 N기로 식별될 수 있다. 예를 들어, 기준 개체(305A)는 N기 음성으로서 식별되는데, 이는 기준 개체(305A)가 림프절 암 침범을 나타내지 않음을 나타낸다. 기준 개체(305B)는 N기 양성으로서 식별되는데, 이는 기준 개체(305B)가 림프절 암 침범을 나타냄을 나타낸다. 그러나, 대상체 레벨 N기는 특정 림프절

들을 구별하지 않고 개별 림프절들이 양성 또는 비양성인지 여부를 구별하지 않는다.

[0089] 대체 모델(310)은 각각의 림프절에 대한 전이 위험(예컨대, 림프절 전이 위험)을 대체함으로써, 전이성 암을 갖는 림프절들과 비양성인 림프절들 사이의 구별을 가능하게 한다. 다양한 실시 형태들에서, 대체 모델(310)은 림프절이 양성일 가능성을 나타내는 각각의 림프절에 대한 확률을 예측한다.

[0090] 도 3a에 도시된 바와 같이, 대체 모델(310)의 출력은 개체(305C)(기준 개체(305A)에 대응함)가 림프절 침범을 나타내지 않음을 식별할 수 있다. 다시 말하면, 개체(305C)의 림프절들 중 어느 것도 양성일 가능성이 없다. 이는 N기 음성 식별을 확인한다. 추가적으로, 대체 모델(310)의 출력은 개체(305D)(기준 개체(305B)에 대응함)가 하나 이상의 림프절에 대한 림프절 침범을 나타냄을 식별할 수 있다. 예를 들어, 대체 모델(310)은 개체(305D)에서의 정확히 하나의 림프절이 양성으로서 식별되는 반면 개체(305D)에서의 다른 림프절들이 비양성이도록 확률들을 대체할 수 있다.

[0091] 다양한 실시 형태들에서, 대체 모델(310)에 의해 예측된 림프절들에 대한 림프절 전이 위험들 중 하나 이상은 위험 모델(320)을 훈련시키는 데 사용된다. 다양한 실시 형태들에서, 대체 모델(310)에 의해 예측된 림프절들에 대한 림프절 전이 위험들 전부는 위험 모델(320)을 훈련시키는 데 사용된다. 다양한 실시 형태들에서, 림프절들의 서브세트에 대응하는 림프절 전이 위험의 서브세트는 위험 모델(320)을 훈련시키는 데 사용됨으로써, 위험 모델이 전이성 질환을 갖는 대상체에서 대상체들에서의 림프절들 전부에 대한 더 큰 확률 범위를 예측하게 할 수 있다.

[0092] 일단 훈련되면, 위험 모델(320)은 하나 이상의 대상체(315)에 대한 전이성 암의 대상체 레벨 위험을 예측하기 위해 전개된다. 따라서, 위험 모델(320)은 전이성 질환의 림프절 레벨 위험들로부터 전이성 질환의 대상체 레벨 위험으로 이행하도록 전개된다. 도 3a에 도시된 바와 같이, 위험 모델(320)은 개체(315)들로부터 캡처된 이미지들로부터 추출된 특징들을 분석하도록 전개된다. 위험 모델(320)은 개체(315)들의 범위에 대해 전이성 암의 대상체 레벨 위험들을 예측한다. 예를 들어, 위험 모델(320)은 개체(315A)에 대한 전이성 암의 20% 위험, 개체(315B)에 대한 전이성 암의 30% 위험, 개체(315C)에 대한 전이성 암의 50% 위험, 및 개체(315D)에 대한 전이성 암의 60% 위험을 예측한다. 따라서, 개체(315A, 315B)들은 초기 암(예컨대, 0 또는 1기)으로 분류될 수 있는 반면, 개체(315C, 315D)들은 말기 암(예컨대, 2, 3, 또는 4기)으로 분류될 수 있다.

[0093] 개체(315)들은 전이성 질환의 그들의 대상체 레벨 위험에 따라 상이한 치료들을 받을 수 있다. 예를 들어, 림프절 전이를 가질 위험이 낮은 개체(315A, 315B)들은 외과적 절제 또는 조합된 기관지경 진단 및 치료와 같은 즉각적인 요법으로 바로 진행할 수 있다. 다른 예로서, 림프절 전이를 갖는 위험이 더 높은 개체(315C, 315D)들은 PET 스캐닝 및 림프절 생검과 같은 추가 검사를 받을 수 있다.

[0094] 도 3b는 일 실시 형태에 따른 위험 모델(320)의 구현예를 더 상세히 도시한다. 여기서, 위험 모델(320)은 대상체(350)로부터 캡처된 이미지로부터의 추출된 특징들을 분석하고, 대상체에 대해 하나 이상의 림프절의 각각에 대한 확률들을 출력한다. 도 3b에 도시된 바와 같이, 위험 모델(320)은 3개의 림프절에 대한 확률들을 출력한다. 다른 실시 형태들에서, 위험 모델(320)은 대상체(350)에 대해 추가적인 또는 더 적은 림프절 확률들을 출력할 수 있다.

[0095] 도 3b에 도시된 바와 같이, 위험 모델(320)은 제1 림프절이 20%의 양성 확률 및 역으로 80%의 비양성 확률을 갖는다고 예측한다. 추가적으로, 위험 모델(320)은 제2 림프절이 30%의 양성 확률 및 역으로 70%의 비양성 확률을 갖는다고 예측한다. 추가적으로, 위험 모델(320)은 제3 림프절이 60%의 양성 확률 및 역으로 40%의 비양성 확률을 갖는다고 예측한다. 전이성 암의 대상체 레벨 위험을 식별하기 위해 개별 림프절 확률들이 분석된다. 이 예에서, 전이성 암의 최대 림프절 확률은 전이성 암의 대상체 레벨 확률로서 취해진다. 60%의 양성 확률을 갖는 제3 림프절은 림프절들 중에서 최고 확률이다. 따라서, 대상체(350)에 대한 전이성 암의 대상체 레벨 확률은 60%이다.

[0096] V. 암

[0097] 본 명세서에 기술된 방법들은 대상체의 암의 대상체 레벨 위험을 결정하는 것에 의한 암 병기 결정을 포함한다. 다양한 실시 형태들에서, 대상체의 암은 다음 중 하나 이상을 포함할 수 있다: 림프종, B 세포 림프종, T 세포 림프종, 균상식육종, 호지킨병, 골수성 백혈병, 방광암, 뇌암, 신경계 암, 두경부암, 두경부의 편평세포암종, 신장암, 폐암, 신경모세포종/교모세포종, 난소암, 췌장암, 전립선암, 피부암, 간암, 흑색종, 구강, 인후, 후두, 및 폐의 편평세포암종, 결장암, 자궁경부암, 자궁 암종, 유방암, 상피암, 위장암, 신장암, 비노생식암, 폐 암, 식도암종, 위암, 갑상선암, 두경부 암종, 큰 창자 암, 조혈 암, 고환암, 결장 및/또는 직장암, 자궁암, 또는 전

립선암. 다양한 실시 형태들에서, 암은 폐 기관지폐포 암종(bronchioloalveolar carcinoma, BAC), 방광암, 여성 생식관 악성종양(예컨대, 자궁 장액 암종, 자궁내막 암종, 외음부 편평세포 암종, 및 자궁 육종), 난소 표면 상피 암종(예컨대, 난소, 상피 난소암, 나팔관암, 및 원발성 복막암의 투명 세포 암종), 유방 암종, 비소세포 폐암(non-small cell lung cancer, NSCLC), 남성 생식관 악성종양(예컨대, 고환암), 후복막 또는 복막 암종, 위식도 선암, 식도위 접합부 암종, 간 간세포 암종, 식도 및 식도위 접합부 암종, 자궁경부암, 담관암종, 췌장 선암종, 간의 담관 선암종, 소장 악성종양, 위 선암종, 원발부위 불명암(cancer of unknown primary (CUP), 결 직장 선암종, 식도 암종, 전립선 선암종, 신장암, 두경부 편평상피 암종, 흉선암종, 비흑색종 피부암, 갑상선 암종(예컨대, 유두상 암종), 두경부암, 항문 암종, 비상피성 난소 암종(비-EOC), 포도막 흑색종, 악성 흉막 중 피종, 소세포 폐암(small cell lung cancer, SCLC), 중추 신경계 암, 신경내분비 종양, 및 연조직 종양 중 어느 하나일 수 있다. 소정 실시 형태들에서, 암은 유방암, 비소세포 폐암, 방광암, 신장암, 결장암, 및 흑색종이다. 일부 실시 형태들에서, 대상체의 암은 방광암, 유방암, 결장암, 신장암, 폐암, 흑색종, 난소암, 췌장암, 전립선암, 직장암, 위암, 갑상선암, 또는 자궁암 중 어느 하나를 포함하는 전이성 암일 수 있다. 특정 실시 형태들에서, 암은 폐암이다. 특정 실시 형태들에서, 암은 소세포 폐암, 비소세포 폐암, 선암종, 편평세포 암, 대 세포 암종, 소세포 암종, 복합 소세포 암종, 폐 육종, 폐 림프종, 기관지 유암종 중 어느 하나를 포함하는 폐암의 일 유형이다.

[0098] VI. 안내식 개입(Guided Intervention)

[0099] 본 명세서에 기술된 실시 형태들은 대상체에서 전이성 암의 대상체 레벨 위험을 결정하는 것에 의한 암 병기 결정의 확정을 수반한다. 다양한 실시 형태들에서, 대상체에 대해 예측된 전이성 암의 대상체 레벨 위험 및/또는 병기 결정에 기초하여 대상체에 개입이 제공된다. 이하에서 안내식 개입으로 지칭되는 그러한 개입은 진단의 적용, 치료제의 적용, 대상체의 후속 검사, 후속 생검(예컨대, 암 생검 또는 림프절 생검, 림프절의 기관지경 샘플링), 후속 이미지 스캐닝(예컨대, PET 스캐닝, MRI 스캐닝, 초음파 이미징, 또는 X-선 이미징), 종양 절제, 기관지경 진단, 치료(들)의 선택 및/또는 시행, 약제적 조성물의 선택/투여, 임상 시험에서의 등록, 임상 시험에서의 비등록, 또는 이들의 임의의 조합 중 어느 하나일 수 있다.

[0100] 다양한 실시 형태들에서, 대상체에 대해 예측된 전이성 암의 대상체 레벨 위험 또는 예측된 암 병기는 대상체가 전이성 암의 높은 위험에 있음을 나타낸다. 다양한 실시 형태들에서, 전이성 암의 높은 위험에 있는 대상체는 임계 점수 초과인 전이성 암의 예측된 대상체 레벨 위험을 갖는다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 20% 초과 위험이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 25% 초과 위험이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 30% 초과, 35% 초과, 40% 초과, 45% 초과, 50% 초과, 55% 초과, 60% 초과, 65% 초과, 70% 초과, 75% 초과, 80% 초과, 85% 초과, 9% 초과, 95% 초과, 또는 99% 초과 위험이다. 다양한 실시 형태들에서, 임계 점수는 20% 내지 80%이다. 일부 실시 형태들에서, 임계 값은 30 내지 70%이다. 일부 실시 형태들에서, 임계 값은 40 내지 60%이다. 일부 실시 형태들에서, 임계 값은 45 내지 55%이다. 일부 실시 형태들에서, 임계 값은 20 내지 60%이다. 일부 실시 형태들에서, 임계 값은 30 내지 55%이다. 일부 실시 형태들에서, 임계 값은 35 내지 50%이다. 일부 실시 형태들에서, 임계 값은 40 내지 45%이다.

[0101] 다양한 실시 형태들에서, 대상체에 대한 예측된 암 병기가 적어도 2기(예컨대, 2기 또는 3기)인 경우, 대상체는 전이성 암의 높은 위험에 있다. 그러한 실시 형태들에서, 안내식 개입은 대상체의 후속 검사, 대상체의 후속 생검, 또는 대상체의 후속 이미지 스캐닝 중 하나를 수반한다. 이는 임상 제공자들이 최고 위험에 있는 대상체들의 후속 검사, 생검, 또는 이미지 스캐닝을 우선순위화하게 한다.

[0102] 다양한 실시 형태들에서, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 말기 전이성 암에 부적절한 절차는 대상체에 대해 제외될 수 있다. 제외된 그러한 절차들의 예들은 종양 절제, 폐 절제, 방사선(예컨대, 외부 빔 방사선), 화학요법(예컨대, 국소 또는 전신 화학요법), 면역조절 요법, 또는 심지어 전신 마취를 포함한다.

[0103] 특정 실시 형태들에서, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 예측을 확인하기 위해 후속 이미지 스캐닝을 겪는다. 예를 들어, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 PET 스캐닝을 추가로 겪는다. 그러한 실시 형태들에서, PET 스캐닝이 음성 결과(예컨대, 림프절 침범 없음)를 제출하는 경우, 대상체는 추가 검사를 겪을 수 있다. 예를 들어, 대상체는 림프절을 명확하게 샘플링하기 위해 후속 생검을 추가로 겪을 수 있다. 일부 실시 형태들에서, 후속 생검은 림프절들의 기관지경 샘플링이다.

[0104] 특정 실시 형태들에서, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 예측을 확인하

기 위해 후속 생검을 직접 겪는다. 예를 들어, 대상체는 예측을 확인하기 위해 후속 이미지 스캐닝을 겪을 필요가 없고 대신에 후속 생검을 직접 겪는다. 예를 들어, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 림프절들의 기관지경 샘플링을 추가로 겪는다.

[0105] 본 명세서에 기술된 바와 같이, 전이성 암의 대상체 레벨 위험은 림프절에 걸쳐 개별 림프절 레벨 위험들로부터 결정될 수 있다. 따라서, 대상체 위험 레벨에 기초하여 전이성 암의 높은 위험을 갖는 것으로 예측된 대상체의 경우, 높은 대상체 위험 레벨로 이어지는 특정 림프절(들)이 식별될 수 있다. 따라서, 이들 특정 림프절은 후속 검사 또는 치료를 위해 표적화될 수 있다. 특정 실시 형태들에서, 이들 특정 림프절은 예측을 확인하기 위해 후속 이미지 스캐닝(예컨대, PET 스캐닝)을 겪을 수 있다. 특정 실시 형태들에서, 이들 특정 림프절은 특정 림프절들을 명확하게 샘플링하기 위해 후속 생검을 겪을 수 있다. 따라서, 후속 생검은 이들 특정 림프절에서 전이성 암의 존재 또는 부재를 확인할 수 있다.

[0106] 다양한 실시 형태들에서, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 임상 시험에 등록될 수 있다. 예를 들어, 그러한 임상 시험은 전이성 암에 대한 치료의 치료 효능을 시험하고 있을 수 있으며, 따라서 이는 등록자들이 전이성 암을 나타낼 것이라는 그의 적격성 기준에 있어서 명시할 수 있다. 다양한 실시 형태들에서, 대상체가 전이성 암의 높은 위험에 있는 것으로 예측되는 경우, 대상체는 임상 시험에서의 등록으로부터 보류될 수 있다. 예를 들어, 그러한 임상 시험은 암의 비전이성 형태에 대한 치료의 치료 효능을 시험하고 있을 수 있다.

[0107] 다양한 실시 형태들에서, 대상체에 대해 예측된 전이성 암의 대상체 레벨 위험은 대상체가 전이성 암의 낮은 위험에 있음을 나타낸다. 다양한 실시 형태들에서, 전이성 암의 낮은 위험에 있는 대상체는 임계 점수 미만인 예측된 대상체 레벨 위험을 갖는다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 20% 초과 미만이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 25% 초과 미만이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 30% 미만, 35% 미만, 40% 미만, 45% 미만, 50% 미만, 55% 미만, 60% 미만, 65% 미만, 70% 미만, 75% 미만, 80% 미만, 85% 미만, 90% 미만, 95% 미만, 또는 99% 미만 위험이다. 일부 실시 형태들에서, 임계 값은 5% 내지 40%이다. 일부 실시 형태들에서, 임계 값은 10 내지 30%이다. 일부 실시 형태들에서, 임계 값은 15 내지 25%이다. 일부 실시 형태들에서, 임계 값은 18 내지 22%이다. 다양한 실시 형태들에서, 대상체에 대한 예측된 암 병기가 2기 미만(예컨대, 0기 또는 1기)인 경우, 대상체는 전이성 암의 낮은 위험에 있다. 그러한 실시 형태들에서, 안내식 개입은 종양 절제, 기관지경 진단, 치료(들)의 선택 및/또는 시행, 및/또는 약제적 조성물의 선택/투여 중 하나를 수반한다. 이는 전이가 아직 발생하지 않은 대상체들의 신속한 식별 및 치료를 가능하게 한다.

[0108] 다양한 실시 형태들에서, 치료제는 전이암의 대상체 레벨 위험에 기초하여 선택되고/되거나 대상체에 투여될 수 있으며, 선택된 치료제는 암에 대해 효능을 보일 가능성이 높다. 예시적인 치료제는 화학요법, 방사선, 항원-특이적 단일클론 항체, 또는 면역요법을 포함한다. 다양한 실시 형태들에서, 전이성 암의 높은 위험에 있는 대상체에는 치료제가 제공된다. 다양한 실시 형태들에서, 전이성 암의 높은 위험에 있는 대상체는 임계 점수 초과인 전이성 암의 예측된 대상체 레벨 위험을 갖는다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 20% 초과 위험이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 25% 초과 위험이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 30% 초과, 35% 초과, 40% 초과, 45% 초과, 50% 초과, 55% 초과, 60% 초과, 65% 초과, 70% 초과, 75% 초과, 80% 초과, 85% 초과, 90% 초과, 95% 초과, 또는 99% 초과 위험이다. 다양한 실시 형태들에서, 대상체에 대한 예측된 암 병기가 적어도 2기(예컨대, 2기, 3기 또는 4기)인 경우, 대상체는 전이성 암의 높은 위험에 있다.

[0109] 다양한 실시 형태들에서, 전이성 암의 낮은 위험에 있는 대상체에는 치료제가 제공된다. 다양한 실시 형태들에서, 전이성 암의 낮은 위험에 있는 대상체는 임계 점수 미만인 예측된 대상체 레벨 위험을 갖는다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 20% 초과 미만이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 25% 초과 미만이다. 다양한 실시 형태들에서, 임계 점수는 전이성 암의 30% 미만, 35% 미만, 40% 미만, 45% 미만, 50% 미만, 55% 미만, 60% 미만, 65% 미만, 70% 미만, 75% 미만, 80% 미만, 85% 미만, 90% 미만, 95% 미만, 또는 99% 미만 위험이다. 다양한 실시 형태들에서, 대상체에 대한 예측된 암 병기가 2기 미만(예컨대, 0기 또는 1기)인 경우, 대상체는 전이성 암의 낮은 위험에 있다.

[0110] 다양한 실시 형태들에서, 대상체가 전이성 암의 낮은 위험에 있는 것으로 예측되는 경우, 대상체는 초기 암 치료로 안내될 수 있다. 초기 암 치료들의 예들은 종양 절제, 폐 절제, 화학요법(예컨대, 국소 또는 전신 화학요법), 국소 전달식 요법(예컨대, 경기관지 바늘 점적주입 또는 카테터 기반 제거), 또는 방사선을 포함한다.

- [0111] 특정 실시 형태들에서, 대상체가 전이성 암의 낮은 위험에 있는 것으로 예측되는 경우, 대상체는 예측을 확인하기 위해 후속 이미지 스캐닝을 겪는다. 예를 들어, 대상체가 전이성 암의 낮은 위험에 있는 것으로 예측되는 경우, 대상체는 PET 스캐닝을 추가로 겪는다. 그러한 실시 형태들에서, PET 스캐닝이 음성 결과(예컨대, 림프절 침범 없음)를 제출하는 경우, PET 스캔 결과는 예측과 일치한다. 그러나, PET 스캐닝이 양성 결과(예컨대, 림프절 침범)를 제출하는 경우, 대상체는 추가 검사를 겪을 수 있다. 예를 들어, 대상체는 림프절을 명확하게 샘플링하기 위해 후속 생검을 추가로 겪을 수 있다. 일부 실시 형태들에서, 후속 생검은 림프절들의 기관지경 샘플링이다.
- [0112] 다양한 실시 형태들에서, 치료제는 생물제제인데, 예컨대 사이토카인, 항체, 가용성 사이토카인 수용체, 안티센스 올리고뉴클레오타이드, siRNA 등이다. 그러한 생물학적 제제는 생물학적 제제의 뮤테인 및 유도체를 포함하며, 이 유도체는 예를 들어 당업계에 알려진 바와 같은 융합 단백질, 폐길화(PEGylated) 유도체, 콜레스테롤 복합 유도체 등을 포함할 수 있다. 또한 사이토카인의 길항제 및 사이토카인 수용체, 예컨대 트랩 및 단일클론 길항제가 포함된다. 또한 본 명세서에 제시된 활성제들과의 바이오시밀러 또는 생물학적동등 약제들이 포함된다.
- [0113] 폐암 치료제들은 도세탁셀(docetaxel), 시스플라틴(cisplatin), 카보플라틴(carboplatin), 젬시타빈(gemcitabine), Nab-파클리탁셀(Nab-paclitaxel), 파클리탁셀, 페메트렉시드(pemetrexed), 게피티닙(gefitinib), 엘로티닙(erlotinib), 브리가티닙(brigatinib)(알룬브릭®(Alunbrig®)), 카프마티닙(capmatinib)(타브렉타®(Tubrexta®)), 셀퍼카티닙(selpercatinib)(레테브모®(Retevmo®)), 엔트렉티닙(entrectinib)(로즐리트렉®(Rozlytrek®)), 롤라티닙(lorlatinib)(로브레나®(Lorbrena®)), 라로트렉티닙(larotrectinib)(비트락비®(Vitrakvi®)), 다코미티닙(dacomitinib)(비짐프로®(Vizimpro®)), 및 비노렐빈과 같은 화학요법을 포함할 수 있다. 폐암 치료제들은 더발루맵(durvalumab)(임핀지®(Imfinzi®)), 니볼루맵(nivolumab)(옵디보®(Opdivo®)), 펄브로리주맵(pembrolizumab)(키트루다®(Keytruda®)), 아테졸리주맵(atezolizumab)(티센트릭®(Tecentriq®)), 및 라무시루맵(ramucirumab)과 같은 항체 요법을 포함할 수 있다.
- [0114] 다양한 실시 형태들에서, 약제적 조성물은 전이암의 대상체 레벨 위험에 기초하여 대상체에 선택 및/또는 투여될 수 있고, 선택된 치료제는 암에 대해 효능을 보일 가능성이 높다. 개인에 투여된 약제적 조성물은 위에서 설명된 치료제와 같은 활성제를 포함한다. 활성 성분은 치료적 유효량, 즉, 질병 또는 그로부터 매개된 의학적 병태를 치료하기 위해 투여될 때 충분한 양으로 존재한다. 조성물들은 또한 전달 및 효능을 향상시키는, 예컨대 활성 성분들의 전달 및 안정성을 향상시키는 다양한 다른 제제들을 포함할 수 있다. 따라서, 예를 들어, 조성물들은 또한, 원하는 제형에 따라, 약제적으로 수용가능한, 무독성 운반체 또는 희석제들을 포함할 수 있고, 이는 동물 또는 인간 투여를 위한 약제적 조성물들을 제제화하는데 공통적으로 사용되는 매개체로서 정의된다. 희석제는 조합의 생물학적 활성에 영향을 미치지 않도록 선택된다. 이러한 희석제들의 예들은 증류수, 완충수, 생리식염수, PBS, 링거액, 포도당액, 및 헨크 용액(Hank's solution)이다. 또한, 약제적 조성물 또는 제제는 다른 담체, 항원보강제, 또는 무독성, 비치료, 비면역원성 안정화제, 부형제 등을 포함할 수 있다. 조성물들은 또한 pH 조정 및 완충제, 독성 조절제, 습윤제 및 세제와 같은 생리학적 조건에 근사화하기 위한 추가 물질들을 포함할 수 있다. 조성물은 항산화제와 같은 다양한 안정화제들 중 임의의 것을 또한 포함할 수 있다.
- [0115] 본 명세서에 기술된 약제적 조성물이 다양하고 상이한 방법들로 투여될 수 있다. 예들은 경구, 비강, 직장, 국소, 복강내, 정맥, 근육내, 피부밑, 피하, 경피, 수막공간내, 또는 두개내 방법을 통해 약제적으로 허용가능한 담체를 포함하는 조성물을 투여하는 것을 포함한다. 다양한 실시 형태들에서, 약제적 조성물은 카테터 기반 전개를 통해(예컨대, 가요성 또는 강성 로봇 기반 기관지경술을 사용하여) 투여될 수 있다.
- [0116] 그러한 약제적 조성물은 (예컨대, 폐암 환자의 진단 후) 치료를 위해 투여될 수 있다. 본 발명의 맥락에서 사용되는 바와 같은 질환 또는 장애의 예방하기, 예방법 또는 예방은 폐암의 발생, 발병, 진행 또는 재발, 폐암 증상의 일부 또는 전부를 예방하거나 폐암의 발병 가능성을 감소시키기 위한 조성물의 투여를 지칭한다. 폐암의 "치료하기", "치료", 또는 "치료법"은 본 발명에 따른 치료의 시행에 의한 암의 진행을 늦추거나, 중지시키거나 역전시키는 것을 의미할 것이다. 바람직한 실시 형태에서, 폐암의 치료는 암의 진행을, 이상적으로는 암 자체를 소멸시키는 지점까지 역전시키는 것을 의미한다.
- [0117] **VII. 컴퓨터 구현예**
- [0118] 대상체의 전이성 암의 대상체 레벨 위험을 결정하는 것에 의한 암 병기 결정을 위한 방법들을 포함하는 본 발명의 방법들이 일부 실시 형태들에서 하나 이상의 컴퓨터에서 수행된다.

- [0119] 예를 들어, 대체 모델 및/또는 위험 모델 및 데이터베이스 저장소의 구축 및 전개는 하드웨어 또는 소프트웨어, 또는 이 둘의 조합으로 구현될 수 있다. 본 발명의 일 실시 형태에서, 기계-판독가능 저장 매체가 제공되고, 매체는 기계 판독가능 데이터로 인코딩된 데이터 저장 재료를 포함하고, 이는, 상기 데이터를 사용하기 위한 명령어들로 프로그래밍된 기계를 사용할 때, 본 명세서에 기술된 모델들의 데이터세트들 및 실행 및 결과들 중 임의의 것을 표시할 수 있다. 본 발명은 프로세서, 데이터 저장 시스템(휘발성 및 비휘발성 메모리 및/또는 저장 소자들을 포함), 그래픽 어댑터, 포인팅 디바이스, 네트워크 어댑터, 적어도 하나의 입력 디바이스, 및 적어도 하나의 출력 디바이스를 포함하는 프로그램가능 컴퓨터들에서 실행되는 컴퓨터 프로그램들에서 구현될 수 있다. 그래픽 어댑터에 디스플레이가 결합된다. 프로그램 코드는 전술된 기능들을 수행하고 출력 정보를 생성하기 위해 입력 데이터에 적용된다. 출력 정보는 하나 이상의 출력 디바이스에 공지된 방식으로 적용된다. 컴퓨터는, 예를 들어, 종래 설계의 개인용 컴퓨터, 마이크로컴퓨터, 또는 워크스테이션일 수 있다.
- [0120] 각각의 프로그램은 컴퓨터 시스템과의 통신을 위하여 고급 절차적 또는 객체-지향적 프로그래밍 언어로 구현될 수 있다. 그러나, 프로그램들은, 원하는 경우, 어셈블리어 또는 기계어로 구현될 수 있다. 어느 경우든, 언어는 컴파일된 언어 또는 해석된 언어일 수 있다. 각각의 이러한 컴퓨터 프로그램은 바람직하게는, 전술된 절차들을 수행하기 위해 저장 매체 또는 디바이스가 컴퓨터에 의해 판독될 때 컴퓨터를 구성 및 동작시키기 위해, 범용 또는 특수목적 프로그램가능 컴퓨터에 의해 판독가능한 저장 매체 또는 디바이스(예컨대, ROM 또는 자기 디스크)에 저장된다. 시스템은 또한 컴퓨터 프로그램으로 구성된 컴퓨터-판독가능 저장 매체로서 구현되도록 고려될 수 있고, 그렇게 구성된 저장 매체는 본 명세서에 기재된 기능들을 수행하기 위해 컴퓨터로 하여금 특별하고 사전정의된 방식으로 동작하게 한다.
- [0121] 서명 패턴들 및 그것들의 데이터베이스들이 그것들의 사용을 용이하게 하기 위해 다양한 매체에서 제공될 수 있다. "매체"는 본 발명의 서명 패턴 정보를 포함하는 제품을 지칭한다. 본 발명의 데이터베이스들은 컴퓨터 판독가능 매체, 예컨대 컴퓨터에 의해 직접 판독 및 액세스될 수 있는 임의의 매체에 기록될 수 있다. 그러한 매체는 플로피 디스크, 하드 디스크 저장 매체 및 자기 테이프와 같은 자기 저장 매체; CD-ROM과 같은 광학 저장 매체; RAM 및 ROM과 같은 전기 저장 매체; 및 자기/광학 저장 매체와 같은 이들 카테고리의 하이브리드를 포함한다. 통상의 기술자는 현재 공지된 컴퓨터 판독가능 매체들 중 임의의 것이 어떻게 본 데이터베이스 정보의 기록을 포함하는 제품을 생성하는 데 사용될 수 있는지 쉽게 이해할 수 있다. "기록"은 해당 기술분야에서 공지된 바와 같이 임의의 이러한 방법들을 이용하여 컴퓨터 판독가능 매체에 정보를 저장하기 위한 프로세스를 지칭한다. 저장된 정보에 액세스하는 데 사용되는 수단에 기초하여, 임의의 간편한 데이터 저장 구조체가 선택될 수 있다. 다양한 데이터 프로세서 프로그램들 및 포맷들, 예컨대 워드 프로세싱 텍스트 파일, 데이터베이스 포맷 등이 저장에 사용될 수 있다.
- [0122] 일부 실시 형태들에서, 대상체의 전이성 압의 대상체 레벨 위험을 결정하는 것에 의한 압 병기 결정을 위한 방법들을 포함하는 본 발명의 방법들은 분산 컴퓨팅 시스템 환경(예컨대, 클라우드 컴퓨팅 환경)에서 하나 이상의 컴퓨터에서 수행된다. 이 설명에서, "클라우드 컴퓨팅"은 구성가능 컴퓨팅 리소스들의 공유 세트에 대한 주문형 네트워크 액세스를 가능하게 하는 모델로서 정의된다. 클라우드 컴퓨팅은 구성가능 컴퓨팅 리소스들의 공유 세트에 대한 주문형 액세스를 제공하도록 사용될 수 있다. 구성가능 컴퓨팅 리소스들의 공유 세트는 가상화를 통해 신속하게 발급되고 낮은 관리 노력 또는 서비스 사업자 상호작용으로 배포되고, 이어서 그에 따라 스케일링될 수 있다. 클라우드-컴퓨팅 모델은, 예를 들어, 주문형 셀프-서비스, 광범위 네트워크 액세스, 리소스 풀링, 신속한 탄력성, 도수제(measured service) 등과 같은 다양한 특징들로 구성될 수 있다. 클라우드-컴퓨팅 모델은 또한, 예를 들어, 서비스로서의 소프트웨어(Software as a Service, "SaaS"), 서비스로서의 플랫폼(Platform as a Service, "PaaS"), 서비스로서의 기반구조(Infrastructure as a Service, "IaaS")와 같은 다양한 서비스 모델들을 노출시킬 수 있다. 클라우드-컴퓨팅 모델은 또한 프라이빗 클라우드, 커뮤니티 클라우드, 퍼블릭 클라우드, 하이브리드 클라우드 등과 같은 상이한 전개 모델들을 이용하여 전개될 수 있다. 본 명세서 및 청구범위에서, "클라우드-컴퓨팅 환경"은 클라우드 컴퓨팅이 사용되는 환경이다.
- [0123] 도 4는 도 1a, 도 1b, 도 2a, 도 2b, 도 3a 및 도 3b에 도시된 엔터티들을 구현하기 위한 예시적인 컴퓨터를 예시한다. 컴퓨터(400)는 칩셋(404)에 결합된 적어도 하나의 프로세서(402)를 포함한다. 칩셋(404)은 메모리 제어기 허브(420) 및 입력/출력(I/O) 제어기 허브(422)를 포함한다. 메모리(406) 및 그래픽 어댑터(412)는 메모리 제어기 허브(420)에 결합되고, 디스플레이(418)는 그래픽 어댑터(412)에 결합된다. 저장 디바이스(408), 입력 디바이스(414), 및 네트워크 어댑터(416)는 I/O 제어기 허브(422)에 결합된다. 컴퓨터(400)의 다른 실시 형태들이 상이한 아키텍처들을 갖는다.
- [0124] 저장 디바이스(408)는 하드 드라이브, 콤팩트 디스크 판독전용 메모리(CD-ROM), DVD, 또는 솔리드-스테이트 메

모리 디바이스와 같은 비밀시적 컴퓨터-판독가능 저장 매체이다. 메모리(406)는 프로세서(402)에 의해 사용되는 명령어들 및 데이터를 유지한다. 입력 인터페이스(414)는 터치-스크린 인터페이스, 마우스, 트랙볼, 또는 다른 유형의 포인팅 디바이스, 키보드, 또는 이들의 일부 조합이고, 데이터를 컴퓨터(400)에 입력하는 데 사용된다. 일부 실시 형태들에서, 컴퓨터(400)는 사용자의 제스처들을 통해 입력 인터페이스(414)로부터 입력(예컨대, 명령들)을 수신하도록 구성될 수 있다. 그래픽 어댑터(412)는 디스플레이(418) 상에 이미지 및 기타 정보를 표시한다. 네트워크 어댑터(416)는 컴퓨터(400)를 하나 이상의 컴퓨터 네트워크들에 결합시킨다.

[0125] 컴퓨터(400)는 전술된 기능들을 제공하기 위해 컴퓨터 프로그램 모듈들을 실행하도록 구성된다. 본 명세서에 사용된 바와 같이, 용어 "모듈"은 특정 기능을 제공하는 데 사용되는 컴퓨터 프로그램 로직을 지칭한다. 따라서, 모듈은 하드웨어, 펌웨어, 및/또는 소프트웨어로 구현될 수 있다. 일 실시 형태에서, 프로그램 모듈들은 저장 디바이스(408)에 저장되고, 메모리(406)로 로딩되고, 프로세서(402)에 의해 실행된다.

[0126] 도 1a 또는 도 1b의 엔터티들에 의해 사용되는 컴퓨터(400)들의 유형들은 실시 형태 및 엔터티에 의해 요구되는 프로세싱 전력에 따라 달라질 수 있다. 예를 들어, 암 병기 결정 시스템(130)은 단일 컴퓨터(400) 또는 서버 팜에서와 같은 네트워크를 통해 서로 통신하는 다중 컴퓨터(400)들에서 실행될 수 있다. 컴퓨터(400)들은 그래픽 어댑터(412), 및 디스플레이(418)와 같은, 전술된 컴포넌트들 중 일부가 없을 수 있다.

[0127] VIII. 시스템

[0128] 대상체의 전이성 암의 대상체 레벨 위험을 결정하는 것에 의한 암 병기 결정을 위한 시스템이 본 명세서에 또한 기술된다. 다양한 실시 형태들에서, 그러한 시스템은 적어도 도 1a에 전술된 암 병기 결정 시스템(130)을 포함할 수 있다. 다양한 실시 형태들에서, 암 병기 결정 시스템(130)은, 도 4에 묘사된 예시적인 컴퓨터(400)를 갖는 컴퓨터 시스템과 같은, 컴퓨터 시스템으로서 구현된다.

[0129] 다양한 실시 형태들에서, 시스템은 도 1a에서 전술된 이미징 디바이스(120)와 같은 이미징 디바이스를 포함한다. 다양한 실시 형태들에서, 시스템은 암 병기 결정 시스템(130)(예컨대, 컴퓨터 시스템) 및 이미징 디바이스 둘 모두를 포함한다. 이러한 실시 형태들에서, 암 병기 결정 시스템(130)은 대상체로부터 캡처된 이미지들(예컨대, CT 스캔)을 수신하도록 이미징 디바이스(120)와 통신가능하게 결합될 수 있다. 컴퓨터 시스템은 이미지들을 분석하고 대상체에 대한 결절 질환의 대상체 레벨 위험을 결정하기 위해 대체 모델 및 위험 모델 중 하나 또는 둘 모두를 인 실리코로 구현한다.

[0130] 실시예

[0131] 하기는 본 발명을 수행하기 위한 특정 실시 형태들의 실시예들이다. 실시예들은 예시적인 목적만을 위해 제공되며 본 발명의 범주를 어떠한 방식으로도 제한하려고 의도되지 않는다. 사용된 숫자(예컨대, 양, 온도 등)에 관련하여 정확성을 보장하기 위한 노력이 실행되었지만, 일부 실험 오차 및 편차는 허용되어야 한다.

[0132] 실시예 1: 연구 집단 및 특징 추출

[0133] 이 연구는 국가폐암검진연구(NLST)의 CT 아암(arm)으로부터의 데이터를 사용하여 수행되었다. NLST의 이러한 하위 코호트(sub-cohort)는 658명의 연구 참가자, 즉 양성 결절을 갖는 301명 및 폐암을 갖는 357명으로부터의 데이터를 포함하였다. 658명의 연구 참가자의 각각의 CT 스캔을 살폈고 최대 21개의 림프절군(사람당 16 +/- 6)에 주석을 달았다(9,969 LN 주석 달림; 대조군에서 4,631 LN 및 폐암에 걸린 사람들에서 5,318 LN). 이 초기 샘플로부터, 폐암에 걸린 288명의 사람은 분석에 이용가능한 병기 결정을 가졌다. 균형잡힌 초기 코호트를 생성하기 위해, 287명의 대조군 대상체를 식별하였다. 림프절군들이 도 2a에 도시되어 있고 표 1에 상술되어 있다.

[0134] [표 1]

군별 림프절 주석

림프절군	#LN (대상체당 평균)	
	대조군 그룹	암 그룹
1 L	34	22
1R	29	13
2L	191	186
2R	732	879
3A	203	263
3P	133	139
4L	430	459
4R	837	987
5	386	457
6	189	132
7	477	541
8L	212	230
8R	140	215
9L	29	98
9R	8	14
10L	181	188
10R	92	60
11L	74	106
11R	57	59
12L	5	4
12R	5	8

[0135]

[0136] 각각의 LN군에 대해, 5 mm 및 7.5 mm 반경들을 갖는 구형 영역을 한정하였다. 구 중심을 주석 라벨에 의해 한정하였다. 2개의 LN 주석의 예가 도 2b에 도시되어 있다.

[0137]

이들 더 작은(5 mm) 및 더 큰(7.5 mm) 구형 ROI(region of interest, 관심 영역)를 사용하여 모델 구축에 사용될 이미지의 범위를 탐색하였다. 이들 구형 ROI로부터, 3개의 별개의 라디오믹 패널이 각각 6개의 특징 카테고리로 추출되었다. 각각의 패널은 이들이 이용한 메트릭의 개수가 달랐으며, 이는 CT의 상이한 변환에 기초하였다. 패널1은 특징 추출을 위한 입력으로서 원본 이미지를 사용하였다. 패널 2 및 패널 3은 입력 이미지의 웨이블릿 변환 및 가우스 라플라시안(LoG) 변환을 각각 채용하였다. 특징 카테고리들은 1차 특징, 형상 특징, 그레이 레벨 동시 발생 행렬 1(GLCM) 특징, 그레이 레벨 작용 길이 행렬(GLRLM) 특징, 그레이 레벨 크기 구역 행렬(GLSZM) 특징, 및 이웃 그레이 톤 차이 행렬(NGTDM) 특징을 포함하였다. 특징들의 내역이 표 3에 나타나 있다.

[0138]

[표 3]

CT 스캔 이미지로부터의 특징.

특징의 개수						
특징 카테고리						
이미지 카테고리	1 차	형상*	GLCM	GLRLM	GLSZM	NGTDM
원본	22	14	24	16	16	5
웨이블릿	144	0	192	128	128	40
LoG	18	0	24	16	16	5

*각각의 군이 고정된 구형 형상 영역들로 분석됨에 따라 LN 분석에 포함되지 않은 형상 특징들

[0139]

[0140]

도 5a 내지 도 5r은 GLCM 특징 카테고리에 대한 상이한 패널들에 걸친 상이한 림프절 라디오믹 특징들의 예시적인 상관 열 맵들을 도시한다. 추가적으로, 도 5a 내지 도 5r 각각은 특정 특징들을 도시한다. 도 5a 내지 도 5r에 도시된 그러한 특징들의 예들은 상관 특징, LMC2, MCC, 엔트로피 차이, 결합 엔트로피, 엔트로피 합, 클러스터 프로미넌스, 평균 차이, 결합 평균, 평균 합, 자기상관, 분산 차이, 콘트라스트, 클러스터 경향, 제곱 합, LDMN, LDN, 클러스터 세이드, LMC1, 역분산, LD, LDM, 결합 에너지, 및 최대 확률을 포함한다. 도 5c 내지 도

5r은 상이한 필터들에 대한 특징들을 도시한다. 구체적으로, 도 5c 내지 도 5r은 "XYZ"의 명칭을 제공하는데, 여기서 "X", "Y", 및 "Z" 각각은 이미지의 X, Y, 또는 Z 방향에 적용되는 필터를 지칭한다. "H"는 고역 통과 필터를 지칭하는 반면, "L"은 저역 통과 필터를 지칭한다. 따라서, "HHH"의 명칭은 X, Y, 및 Z 방향들의 각각에서 고역 통과 필터링을 겪은 3D 이미지로부터 도출된 특징들을 지칭한다. "HHL"의 명칭은 X 및 Y 방향들에서 고역 통과 필터링을 겪고 Z 방향에서 저역 통과 필터링을 겪은 3D 이미지로부터 도출된 특징들을 지칭한다.

[0141] 실시예 2: 암 병기 결정을 위한 대체 모델

[0142] 모델링은 각각의 림프절에 대한 암 침범의 예측으로 시작하였다. 이를 행하기 위해, 암 병기(NLST에서 보고됨)를 각각의 주석이 달린 LN군에 대체시켰다(개별 LN 데이터는 NLST에서 보고되지 않았으며, 따라서 LN 침범은 암 병기로부터 "빠졌다"). 이때, 데이터를 훈련을 위해 50:50으로 분할하였다: 표 2에 제공된 그들 데이터의 분포에 의한 시험.

[0143] [표 2]

데이터의 예시적인 50:50 분할				
N 기	훈련		시험	
	대조군	암	대조군	암
0	2337	2179	2263	1713
1	0	127	0	73
2	0	125	0	342
3	0	40	0	42

[0144] 모델의 결과는 이진(NLST 데이터세트에서 N기 공변량을 사용한 결절 전이 예/아니오. "아니오"는 N = 0 또는 N = 1이고, "예"는 N = 2 또는 N = 3)이었다. 메트릭들을 정규화하였고(z-점수), 랜덤 포레스트 및 LASSO 분류기들을 훈련시키고 시험하였다. 각각의 LN에 대해, 그것이 양성인 연속 확률을 생성하였다. 예를 들어, 20% 양성 확률을 갖는 LN이 있을 수 있다(뿔샘 수단에 의해, 그것에는 80% 암-없음 확률이 할당되었다). 연구 참가자들 내의 그리고 이들에 걸친 모든 LN을 독립적인 데이터인 것으로 간주하였다. 초기 모델의 성능이 도 6에 도시되어 있다. 구체적으로, 모델은 곡선 아래 면적(Area under the Curve, AUC) = 65.3%, 정확도(Accuracy, ACC) = 96.6%, 감도 = 61.2%, 및 특이도 = 60.8%를 달성하였다.

[0145] NLST 데이터에 대한 제한은 LN 특이적 정보가 이용가능하지 않았으며 각각의 사람에서 평균 16개의 LN이 주석이 달리고 라디오믹스 기반 평가를 거쳤다는 것이다. 따라서, 이러한 모델은 임상 의학에서 전형적으로 양성 세포들을 갖는 단지 1개의 LN(암이 모든 LN에 동시에 확산된 것이 아니라 대체로 단지 1개 또는 2개의 LN으로 확산됨)이 있을 때 전이성 암을 갖는 환자 내의 모든 LN이 "+"이었음을 가정하였다. 따라서, 대상체 레벨 암 병기로부터 개별 LN 침범을 대체하는 것은 LN 레벨에서 모델 성능을 혼란시키는 위양성들을 도입한다(역시, 그 사람의 CT 스캔 내의 모든 LN이 "+"인 것으로 가정되기 때문이다).

[0146] 이어서, 주어진 대상체의 모든 LN군들에 대해 LN-레벨 분류기에 의해 예측된 최고 확률 점수를 사용하여 N기 가능성을 결정하였다. 도 7a 및 도 7b는 N기 가능성들에 기초한 대상체 레벨 성능(AUC = 69.1%, 감도 = 72.7%, 및 특이도 = 61.6%)을 도시한다. 점수들(예컨대, 최대 확률들)의 분포가 또한 도 7b에 나타난 바와 같이 N기들에 걸쳐 플로팅되었다(plotted). 여기서, 3기 암을 갖는 경우들(좌측 패널)에 대해 더 높은 최대 확률이 관찰되었다.

[0147] 실시예 3: 스텝업(step-up) 림프절 분류 접근법: 환자내 LN들의 부트스트래핑(bootstrapping)

[0148] 2단계 부트스트랩 접근법을 수행하여 전이성 질환의 LN 레벨 위험으로부터 전이성 질환의 대상체 레벨 위험으로 이행하였다. 특히, 2단계 부트스트랩 접근법은 도 3a를 참조하여 전술된 바와 같이 대체 모델 및 위험 모델의 구현을 포함하였다. 여기서, 난제는 대상체의 N기가 분류기를 훈련시키기 위해 모든 LN군들에 대체된다는 것이었다. 이러한 대체는 대부분의 LN이 암이 없음을 고려하면 LN군에서 오류를 도입한다.

[0149] 2단계는 다음과 같았다:

[0150] 1. 대체: 알려진 전이성 질환을 갖는 각각의 대상체에 대해, 암이 침범할 가능성이 가장 큰 림프절들을 식별하였다. 이는 중앙 대상체내 값을 초과하는 (최종 모델로부터 출력된) 양성 확률을 갖는 LN들을 선택함으로써 각각의 대상체에 대해 수행되었다(도 3a에 도시된 대체 모델을 수반하는 대체 프로세스). 암에 걸리는 것에 대한 중앙 대상체내 확률 미만인 LN은 음성으로서 분류되었다.

- [0152] 2. 분류: 코호트는 50:50의 훈련:시험으로 분할되었다. 이어서, 각각의 대상체에 대한 존재 결절 질환을 예측하기 위하여 LN 결과들의 대체 풀(pool)을 사용하여 제2 분류기(예컨대, 도 3a에서의 위험 모델(320))을 훈련시켰다.
- [0153] 이 모델의 출력은 실시예 2에서 진술된 모델의 출력과 유사한, 각각의 LN이 양성일 연속 확률이었다. 그러나, 추가 대체 및 분류 단계들의 이득은, 이제 전이성 질환을 갖는 대상체들 내의 모든 림프절들에 대해 할당된 더 큰 확률 범위가 있었다는 것이었다. 보다 구체적으로, 암에 걸릴 가능성이 있는 그들 LN은 더 높은 "+" 확률을 가졌지만, 암에 걸릴 가능성이 없는 것들은 더 낮은 "+" 확률을 가졌다. 예측 단계를 위해, 최고 확률을 갖는 LN에 기초하여 LN 침범의 대상체-레벨 점수가 생성되었다(도 3b). 구체적으로, 이는 LN 특이적 확률을 취하는 것 및 그들의 폐암의 LN 침범을 갖는 대상체 특이적 확률을 제공하는 것을 수반하였다. 도 3b에 도시된 시험 대상체에서, 하부 LN에는 60% 암 보유 가능성(예컨대, 확률(+) = 0.6)이 할당되었다. 이는 3개의 LN 중에서 최고 "+" 확률을 갖는 LN이었다. 따라서, 대상체 레벨 확률에는 0.6의 값이 할당되었다.
- [0154] 랜덤 포레스트 분류기를 사용하여 스텝업 접근법을 훈련시킨 후에, 부트스트랩 모델의 대상체 레벨 성능은 도 9a에 예시된 바와 같이 AUC = 77.7%, 감도 = 81.8%, 및 특이도 = 59.9%를 달성하였다. 따라서, 2단계 부트스트랩 모델의 성능은 실시예 2에 기술된 접근법과 비교하여 개선되었다. LASSO 분류기를 사용하여 유사한 성능을 달성하였다. 마지막으로, 모델은 N기에 걸친 점수들(또는 최대 확률들)의 분포를 플로팅하는 데 사용될 수 있다(도 9b). 모델이 더 높은 N기를 갖는 그들 경우에 대해 더 높은 최대 확률을 산출함에 주목하는 것이 흥미롭다. 이러한 확률 점수는 상이한 임상 상황에서 사용되어 상이한 환자 분류 의사결정 도구들을 가능하게 할 수 있다. 훈련 및 시험을 위한 추가 데이터는 이진 공변량이라기보다는 범주 공변량으로서 N을 예측하는 모델을 제공할 수 있다.
- [0155] 본 출원에 인용된 모든 간행물, 특허, 특허 출원, 및 다른 문헌은 각각의 개별 간행물, 특허, 특허 출원, 또는 다른 문헌이 모든 목적을 위해 참고로 포함되는 것으로 개별적으로 나타낸 것과 동일한 정도로 모든 목적을 위해 전체적으로 본 명세서에 참고로 포함된다.
- [0156] 다양한 특정 실시 형태들이 도시되고 설명되었지만, 상기 명세서는 제한적이지 않다. 본 개시내용(들)의 사상 및 범주로부터 벗어남 없이 다양한 변형들이 이루어질 수 있음이 이해될 것이다. 본 명세서의 검토 시 많은 변형들이 통상의 기술자에게 명백할 것이다.
- [0157] **실시예 4: CT 이미지 기반 예측 디바이스를 사용한 예시적인 안내식 개입**
- [0158] 본 명세서에 기술된 방법을 구현하는(예컨대, 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험을 예측하기 위한 위험 모델을 전개시키는) 이미지 기반 디바이스는 PET 스캐닝을 수반하는 종래의 방법의 불량한 성능 때문에 임상 치료에 영향을 줄 것이다. 체계적인 림프절 절개/샘플링을 갖는 219명의 NSCLC 환자의 최근 조사는 수술전 PET 병기 결정의 성능을 조사하였다(예컨대, 그 전체가 전체적으로 참고로 포함된 문헌(Li et al., Implications of False Negative and False Positive Diagnosis in Lymph Node Staging of NSCLC by Means of 18F-FDG PET/CT. PLoS ONE 8(10):e78552). 폐문 및 종격 림프절 전이를 검출하기 위한, PET에 대한 감도, 특이도, PPV 및 NPV는 74.2%, 54.4%, 86.8% 및 73.5%였는데, 이때 위음성률은 13.2%이고 위양성률은 45.5%였다. 따라서, 본 명세서에 기술된 방법들은 림프절 침범을 초래한 전이성 질환을 검출하기 위한 비-침습 기법을 나타내며, 이는 의사결정 및 임상 치료를 개선할 것이다. 이러한 디바이스가 의사결정 및 임상 치료에 미칠 영향의 예들은 하기를 포함한다:
- [0159] 크기가 전이 위험이 낮은 종양을 갖는 대상체에서 추가 검사를 프롬프트함. 작은 폐 종양을 갖는 사람들에서 수술전 PET 스캐닝 또는 림프절 생검의 사용에 대한 어떠한 치료 표준도 없기 때문에, CT 디바이스에 의해 캡처된 CT 이미지들에 기초한 림프절 침범의 예측은 PET 스캔을 통해 또는 대상체의 림프절들의 기관지경 샘플링을 통해 후속 검사를 프롬프트한다. 비교하면, 통상적인 방법들은 대상체를 수술로 안내하였을 것이고, 여기서 더 진행된 질환 병기는 오직 수술 중에서만 인식될 것이다.
- [0160] 음성 PET 결과를 얻은 사람들에서 림프절의 조직 기반 샘플링을 프롬프트함. PET가 기능 부전의 위양성률 및 위음성률을 갖는다는 일반적인 인식이 있지만, 진단 검사는 PET 스캔의 결과에 의해 크게 영향을 받는다. 더 작은 종양 크기와 음성 PET를 갖는 환자는 종종 추가 LN 샘플링 없이 수술로 나아간다. 이러한 환자 서브세트는 수술 전에 추가 LN 샘플링으로부터 이득을 얻을 것이다.
- [0161] 위양성 PET 스캔을 갖는 환자를 병기 결정하고 수술 개입에 대한 그들의 적격성을 입증함. 현재, PET 스캔에 기초하여 진행된 병기의 질환을 갖는 것으로 강력하게 의심되는 환자는 추가 확증적 LN 샘플링 없이 비-치료적/

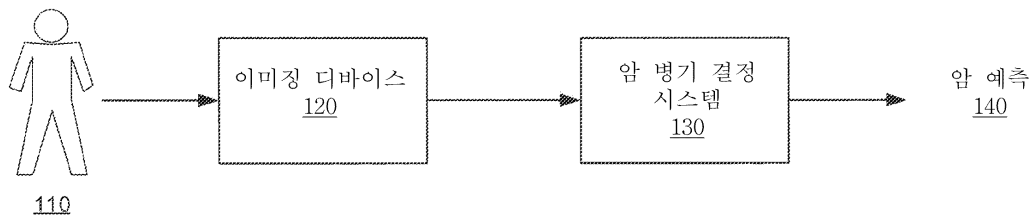
비-외과적 요법으로 나아간다. (예컨대, 림프절의 CT 이미지에 기초하여) 림프절 침범의 예측을 받은 환자 및 음성 PET 결과의 환자는 대신에, 수술에 대한 의사를 결정하기 전에 더 확정적인 림프절 샘플링으로 나아갈 것이다. 따라서, 이들 환자는 외과적 절제를 위한 후보일 수 있고, 그들의 치료 가능성을 증가시킬 것이다.

[0162] 어느 LN이 전이성 질환을 가질 가능성이 가장 큰지를 알린. 전이성 질환에 대해 높은 위험에 있는 환자는 림프절의 기관지경 또는 외과적 샘플링으로 나아간다. 전형적으로, 이 전략은 중앙에 근접한 그들 LN에 가장 중점을 두어, 더 먼 LN을 샘플링되지 않거나 과소 샘플링되는 상태로 둔다. 여기서, 본 명세서에 기술된 방법들은 림프절 침범의 예측 및 또한 침범된 림프절(들)의 식별을 가능하게 한다. 이는 조직을 획득하고 수술을 최대화하며 진단 누락 위험을 최소화하기 위하여 임상 팀들을 그들의 샘플링 전략에 적절히 중점을 두도록 안내한다.

[0163] LN 전이가 없는 환자에 중점을 둔 임상 시험에서의 등록을 정제한 국소 질환을 갖는 환자들은 종종 국소 전달 (화학요법 또는 제거의 정기관지 바늘 점적주입)에 중점을 둔 임상 시험에 대해 요청된다. 불행하게도, 그러한 환자들은 림프절 샘플링을 겪을 때 절차 동안 더 진행된 병기의 암에 걸린 것으로 종종 발견된다. 이는 적절한 치료를 지연시킨다. 따라서, 본 명세서에 기술된 방법들(예컨대, 림프절들에 대한 전이성 암의 림프절(LN) 레벨 위험들을 예측하기 위한 위험 모델을 전개함)의 구현은 림프절 침범을 나타내는 환자들의 식별을 가능하고, 따라서 이들 환자는 전이가 없는 환자에게 중점을 둔 임상 시험에 등록될 필요가 없다.

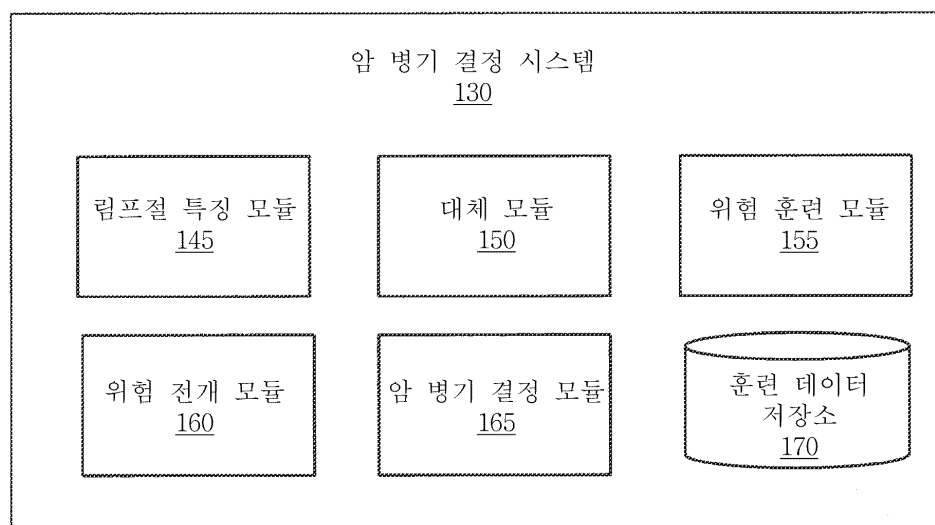
도면

도면1a

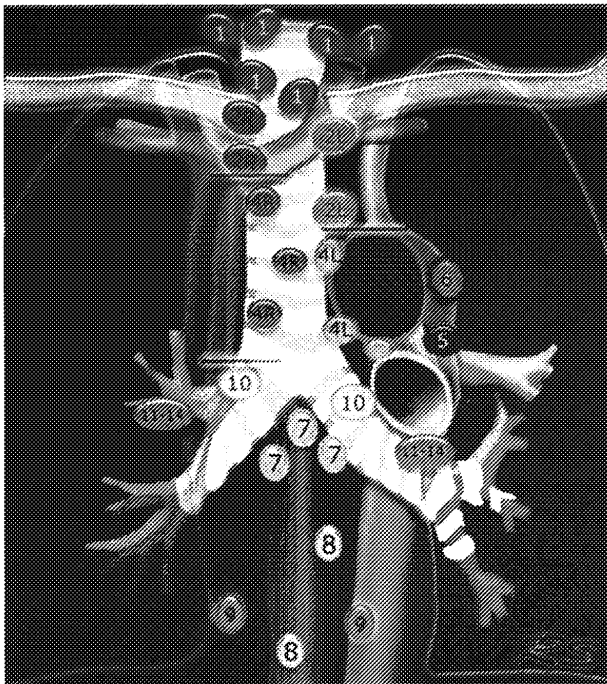


100

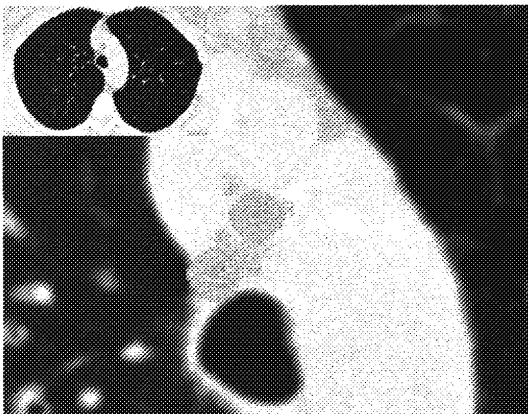
도면1b



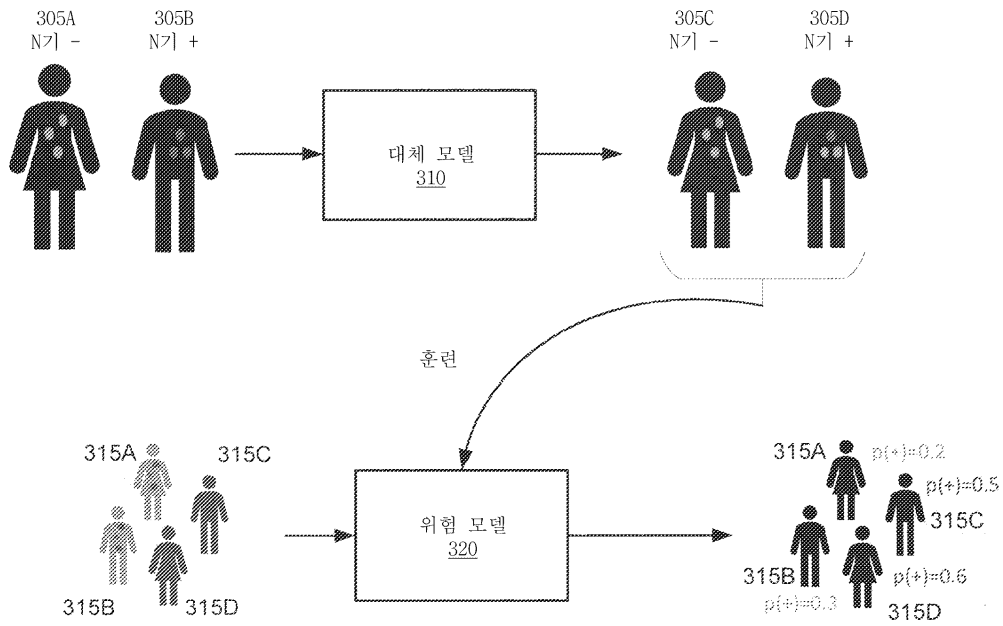
도면2a



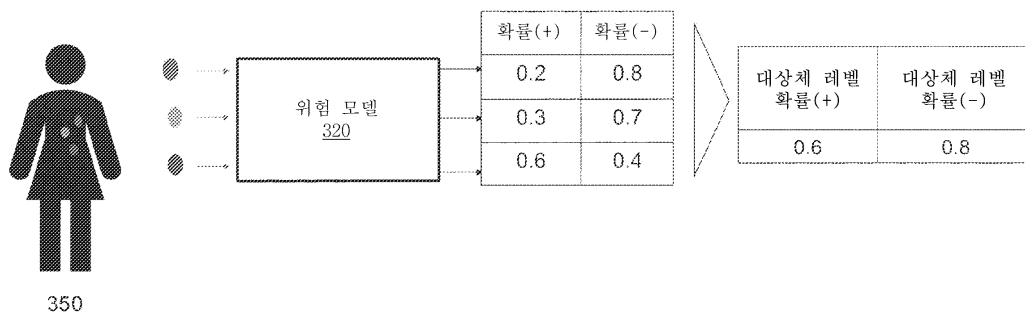
도면2b



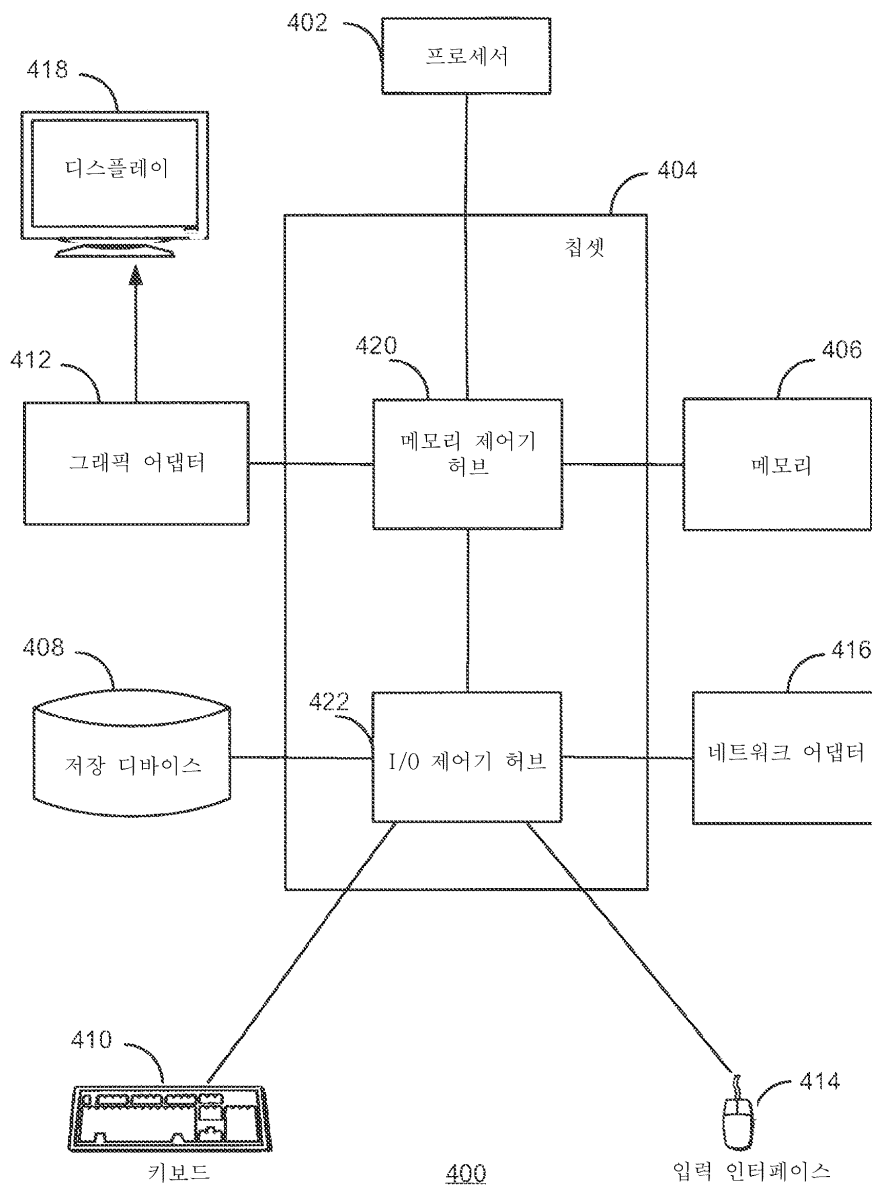
도면3a



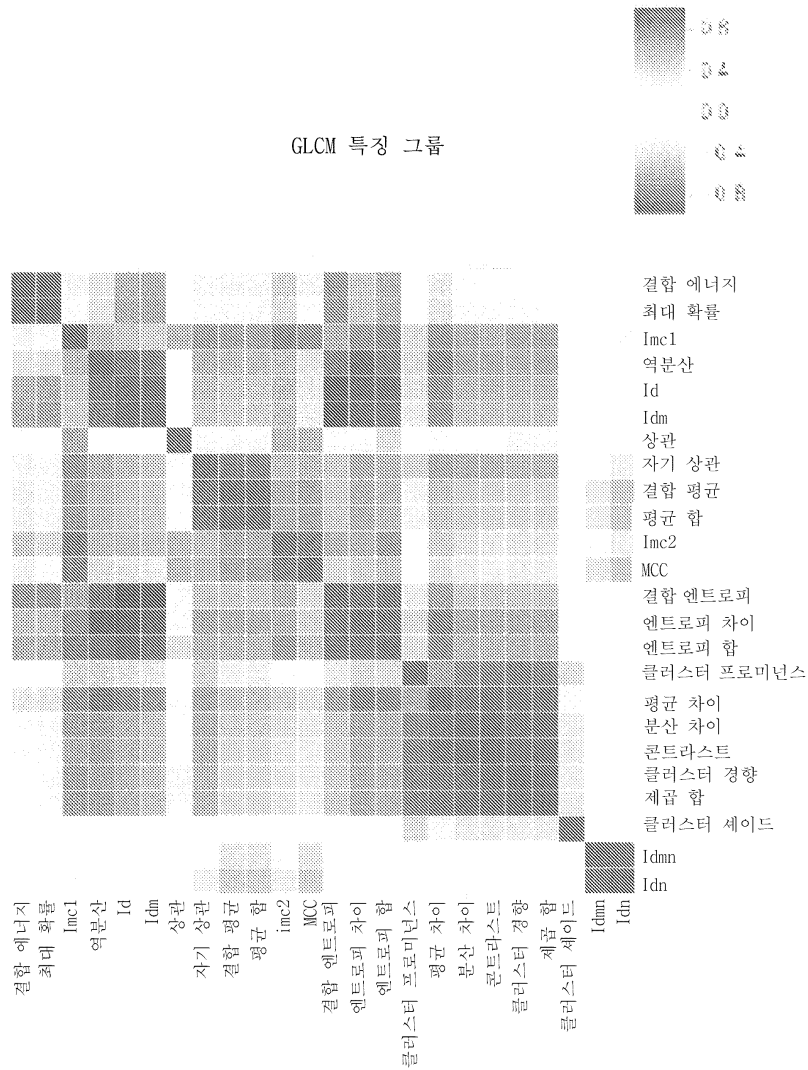
도면3b



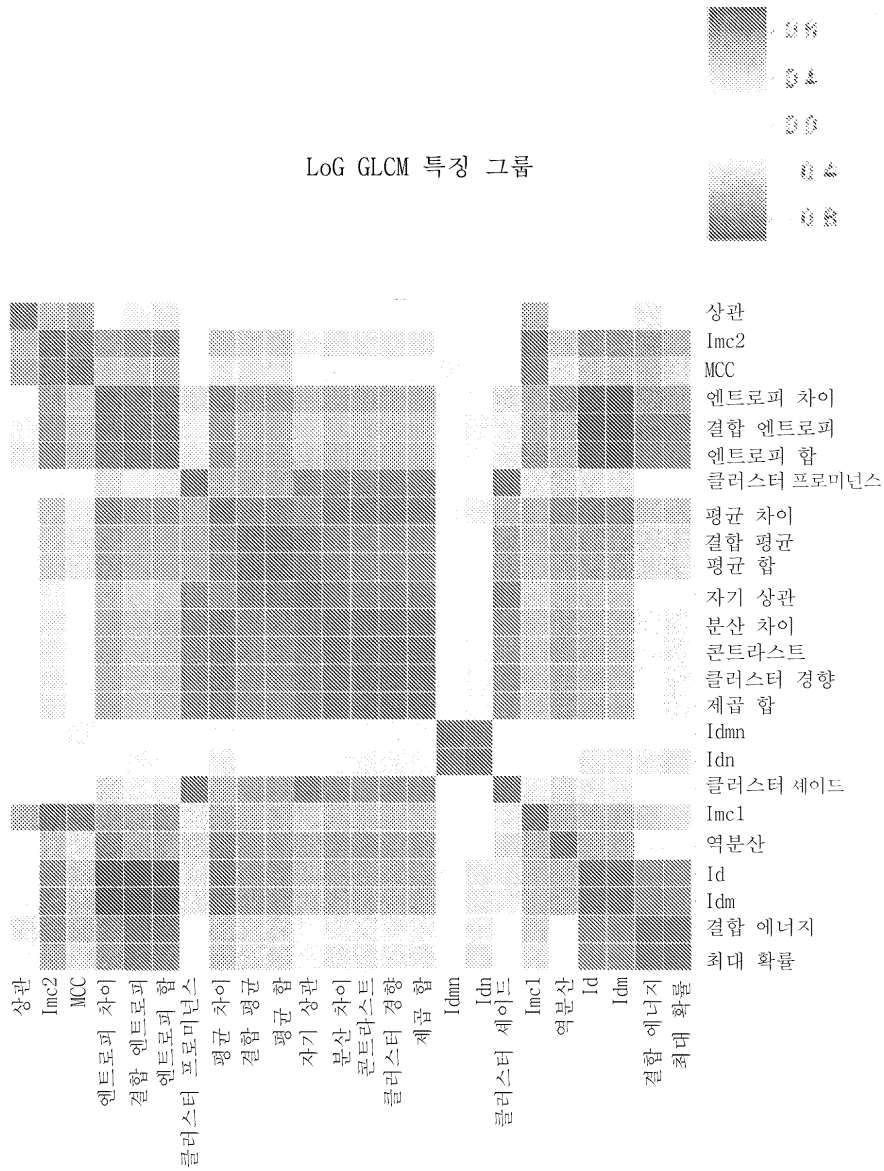
도면4



도면5a

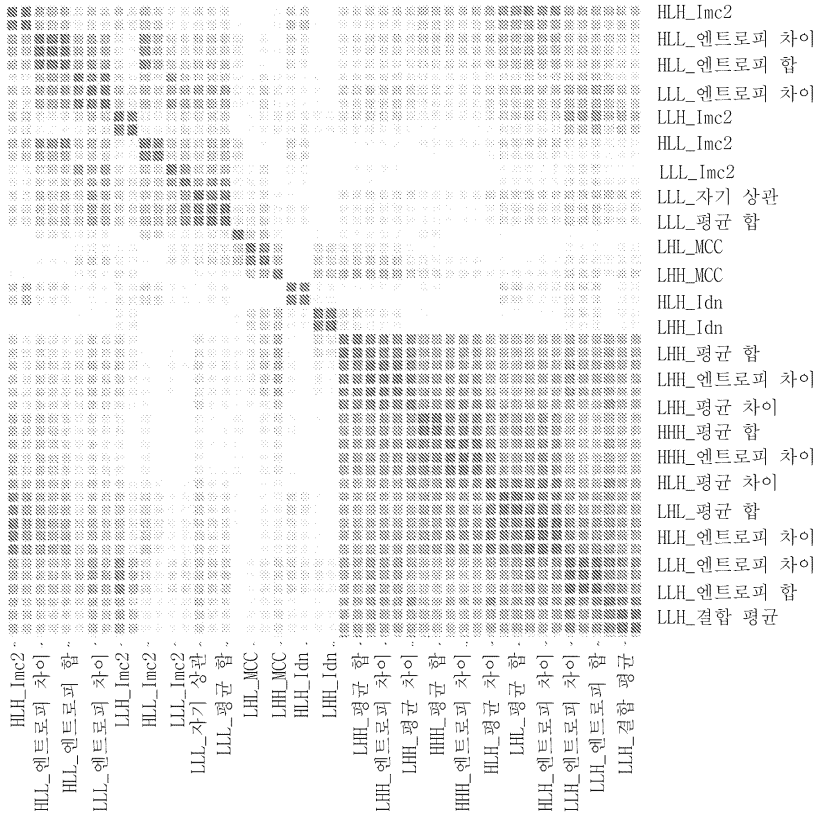


도면5b



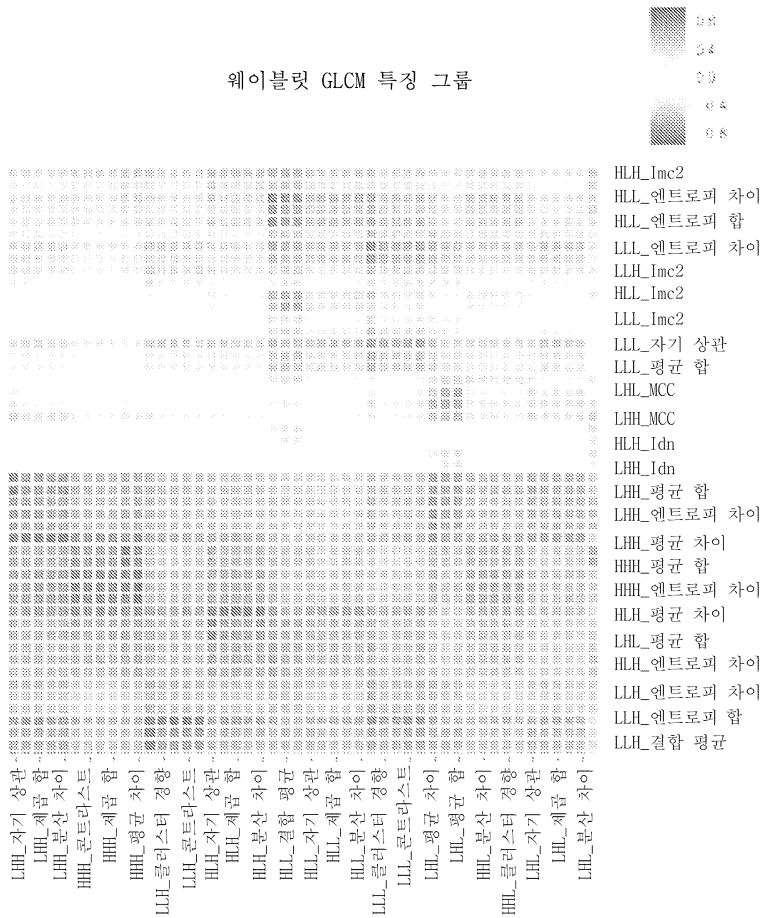
도면5c

웨이블릿 GLCM 특징 그룹



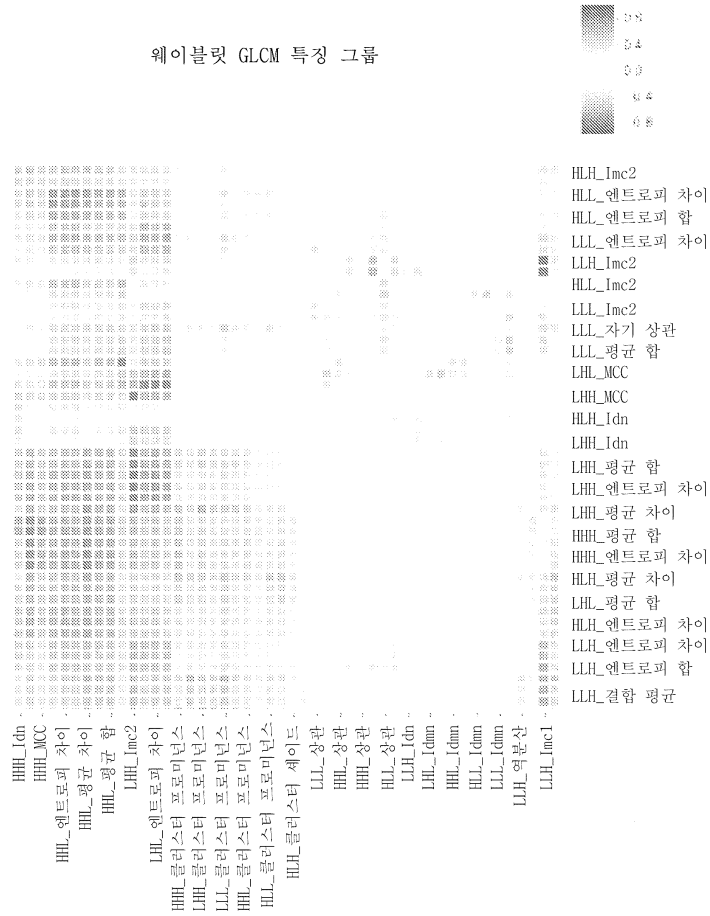
도면5d

웨이블릿 GLCM 특징 그룹



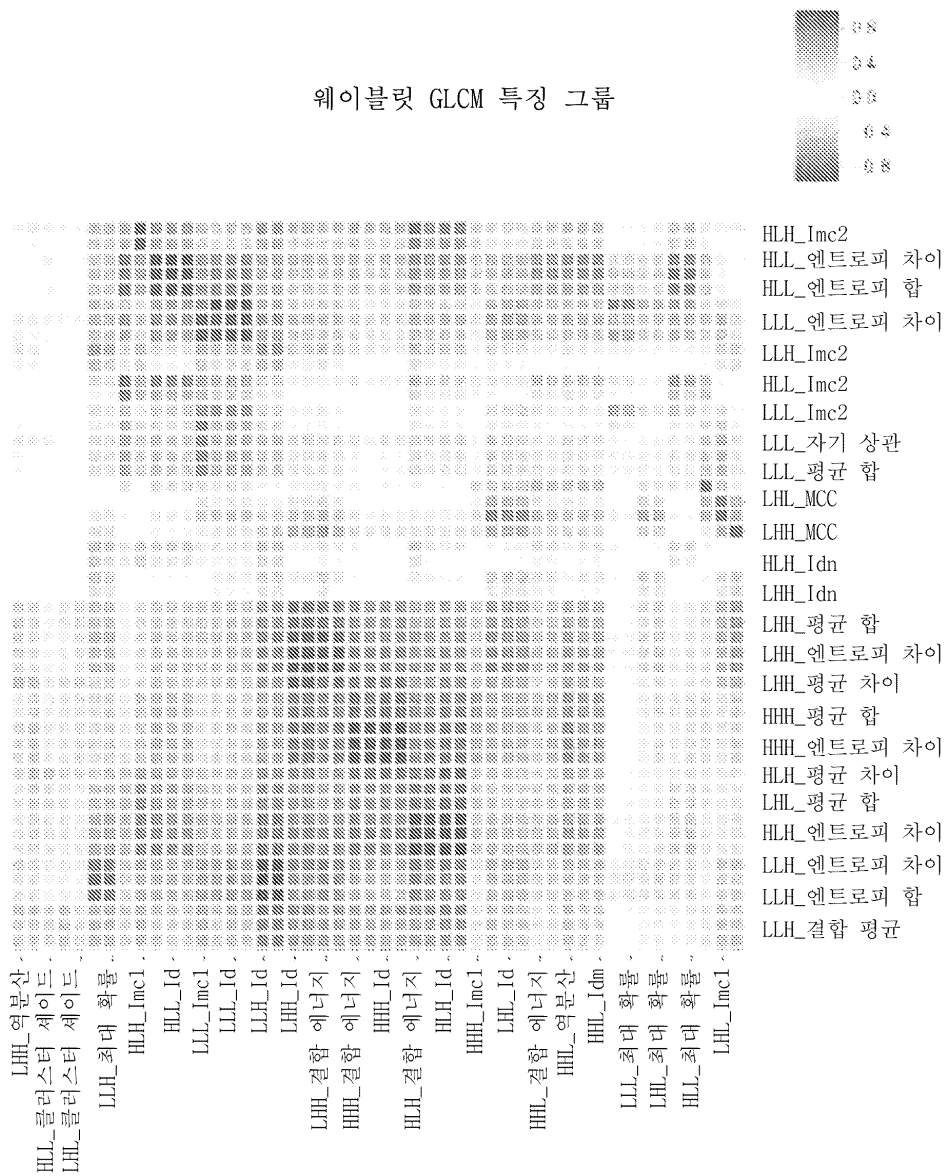
도면 5e

웨이블릿 GLCM 특징 그룹



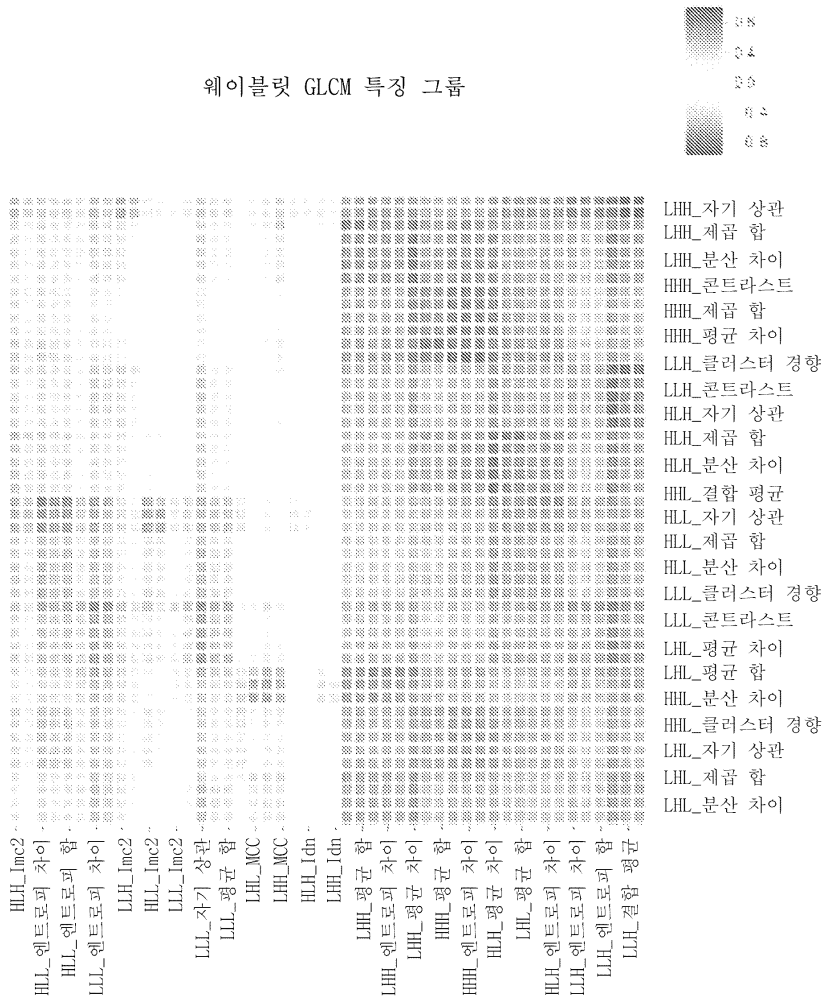
도면5f

웨이블릿 GLCM 특징 그룹



도면5g

웨이블릿 GLCM 특징 그룹



도면5h

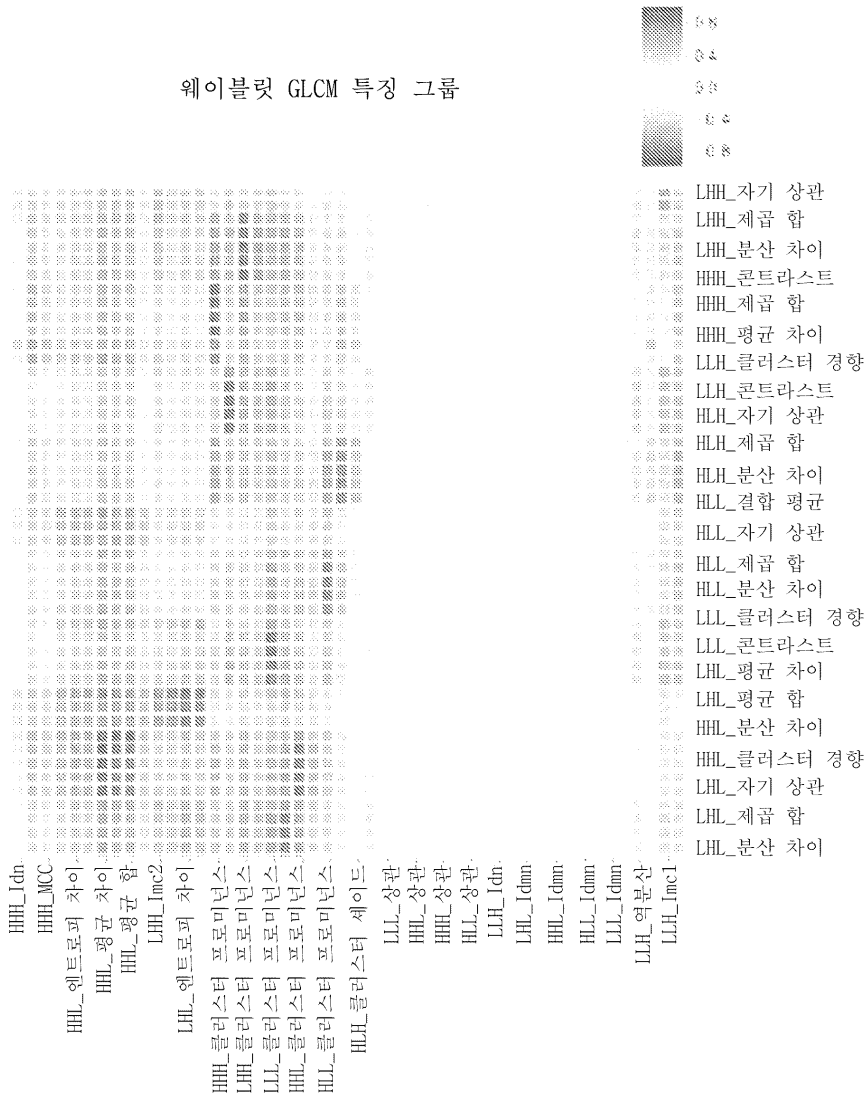
웨이블릿 GLCM 특징 그룹



LHH_자기 상관	LHH_자기 상관
LHH_제공 합	LHH_제공 합
LHH_분산 차이	LHH_분산 차이
HHH_콘트라스트	HHH_콘트라스트
HHH_제공 합	HHH_제공 합
HHH_평균 차이	HHH_평균 차이
LHL_클러스터 경향	LHL_클러스터 경향
LHL_콘트라스트	LHL_콘트라스트
HLH_자기 상관	HLH_자기 상관
HLH_제공 합	HLH_제공 합
HLH_분산 차이	HLH_분산 차이
HLL_결합 평균	HLL_결합 평균
HLL_자기 상관	HLL_자기 상관
HLL_제공 합	HLL_제공 합
HLL_분산 차이	HLL_분산 차이
LLL_클러스터 경향	LLL_클러스터 경향
LLL_콘트라스트	LLL_콘트라스트
LHL_평균 차이	LHL_평균 차이
LHL_평균 합	LHL_평균 합
HHH_분산 차이	HHH_분산 차이
HHH_클러스터 경향	HHH_클러스터 경향
LHL_자기 상관	LHL_자기 상관
LHL_제공 합	LHL_제공 합
LHL_분산 차이	LHL_분산 차이

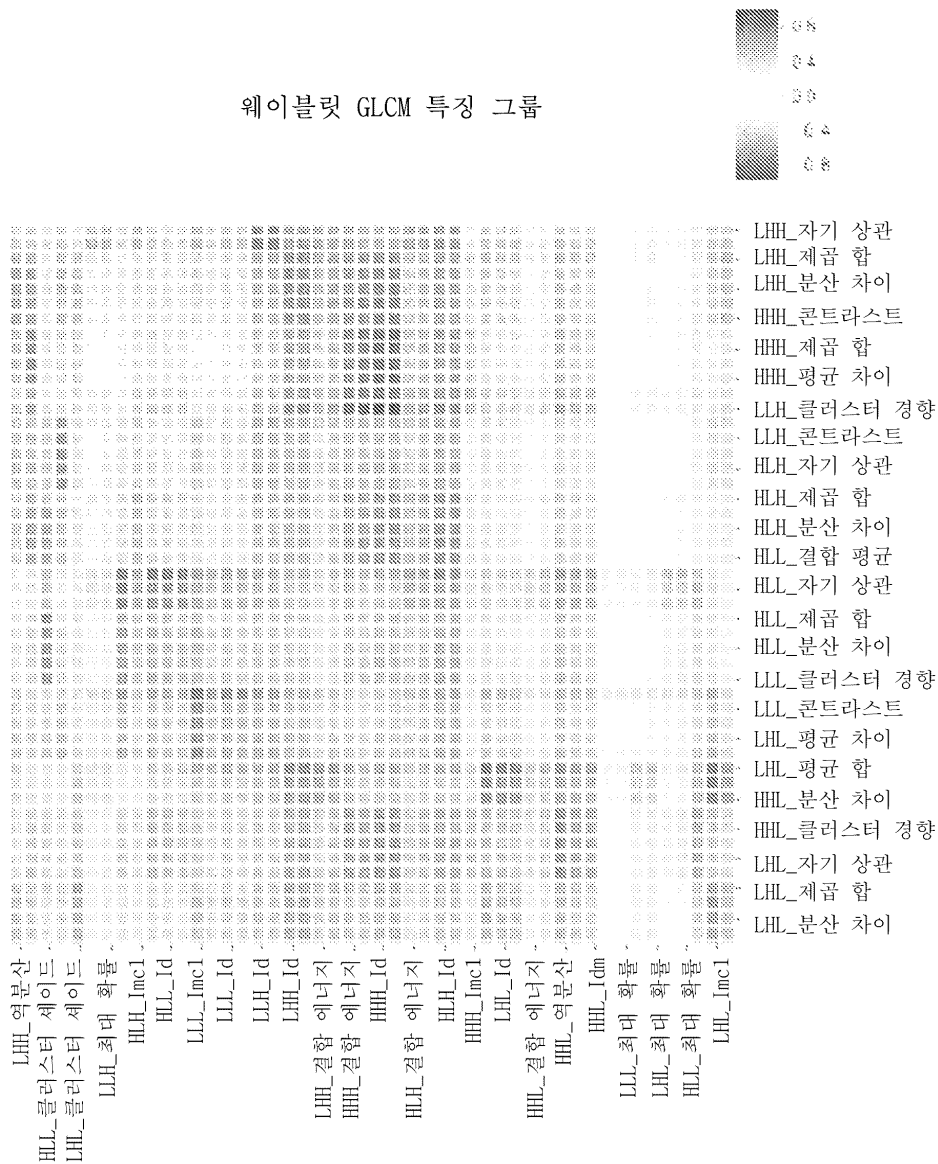
도면5i

웨이블릿 GLCM 특징 그룹



도면5j

웨이블릿 GLCM 특징 그룹



도면 5k

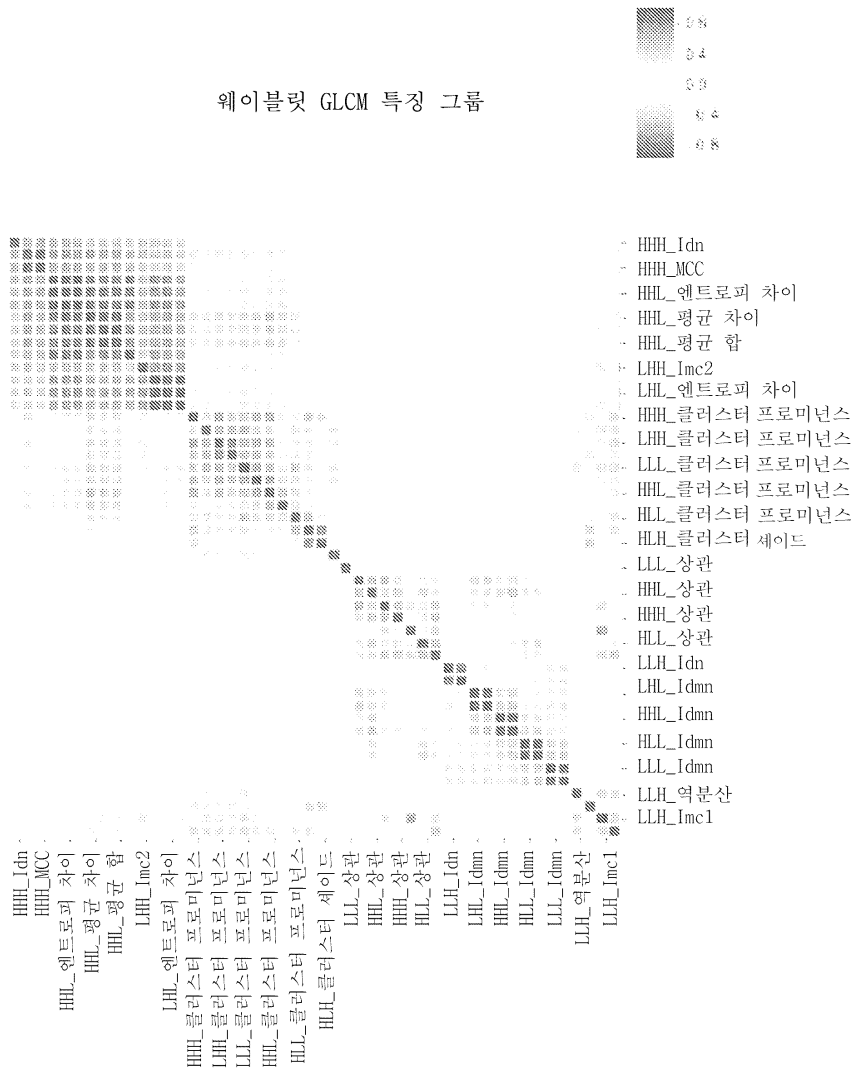


도면51



도면5m

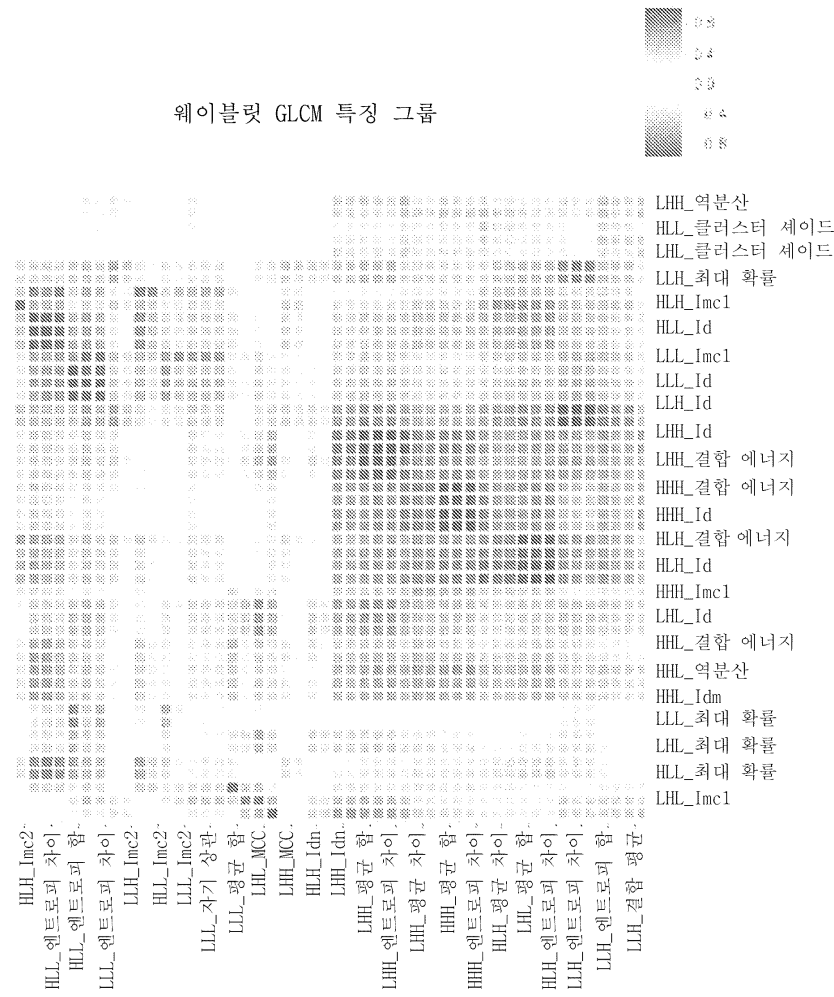
웨이블릿 GLCM 특징 그룹



도면 $5n$

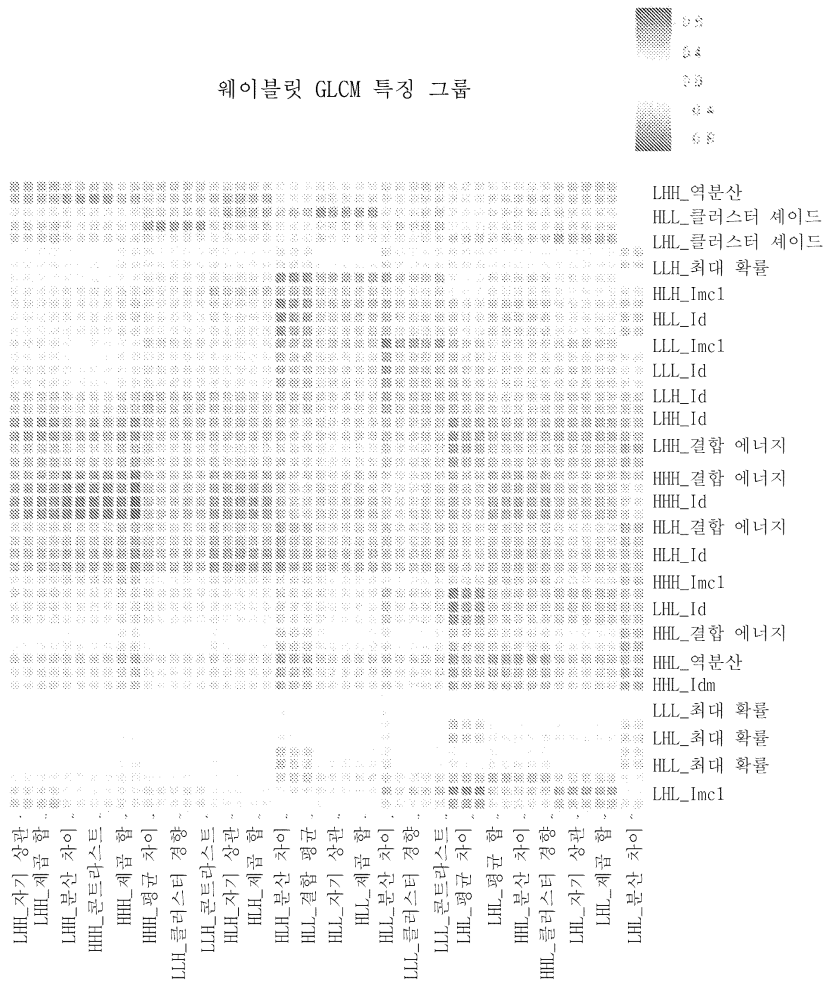


도면50



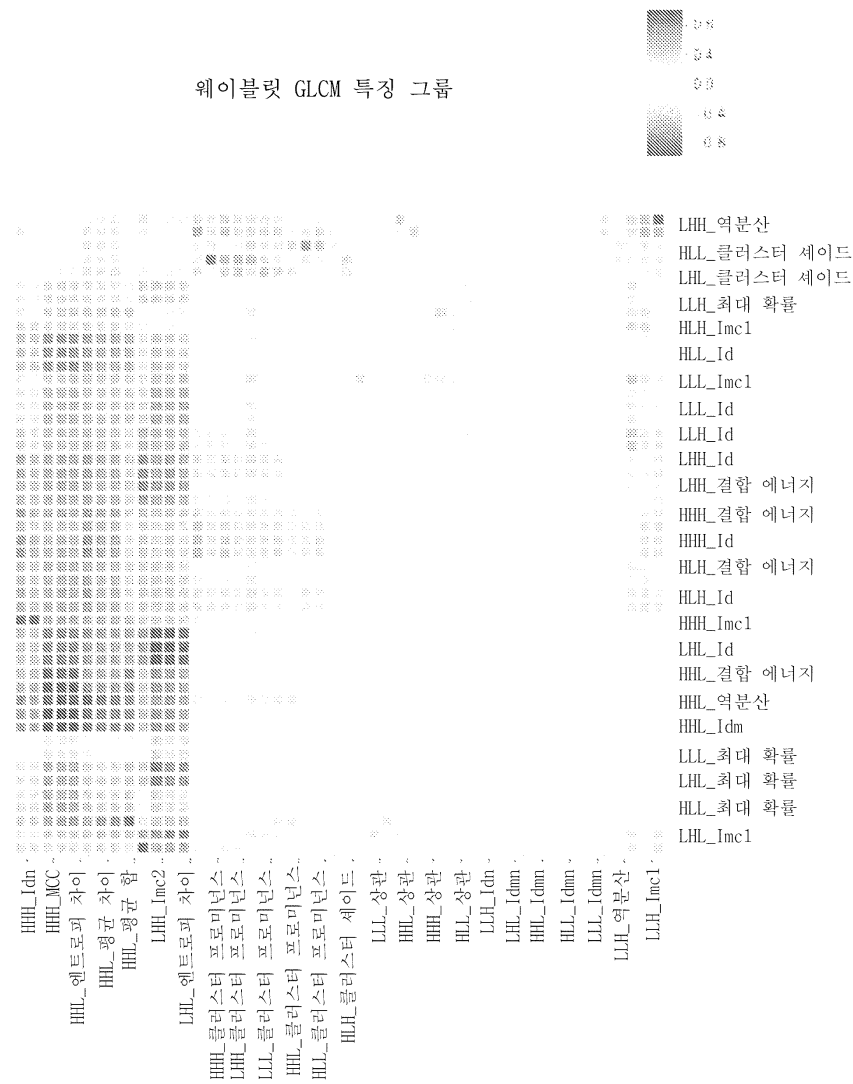
도면 5p

웨이블릿 GLCM 특징 그룹



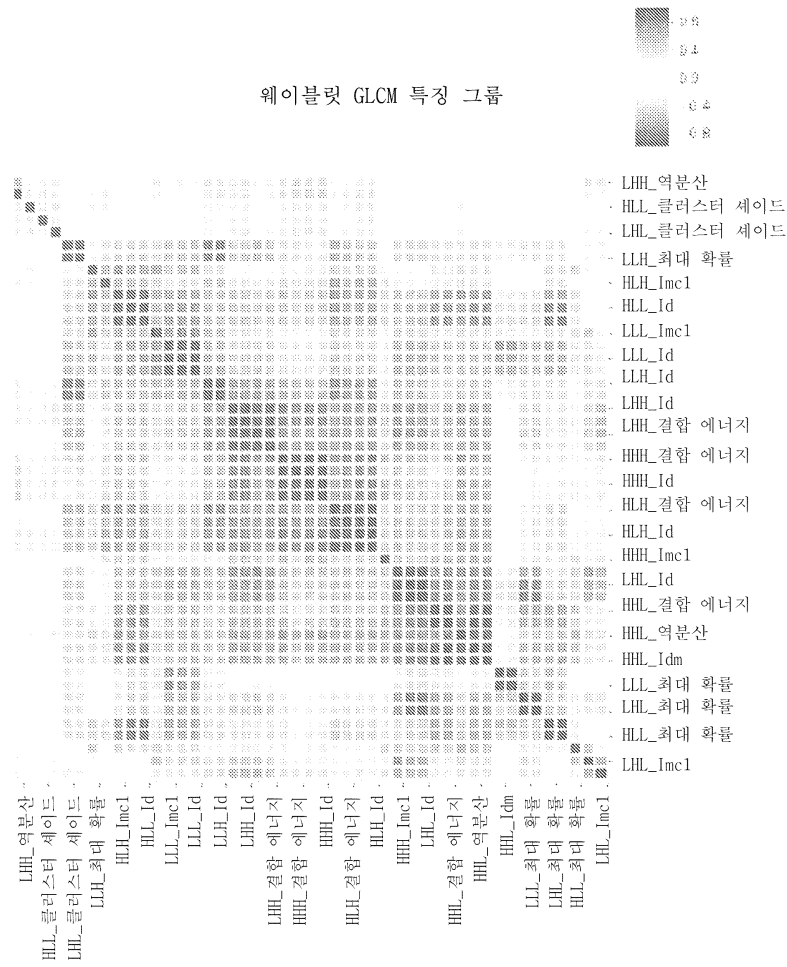
도면5q

웨이블릿 GLCM 특징 그룹



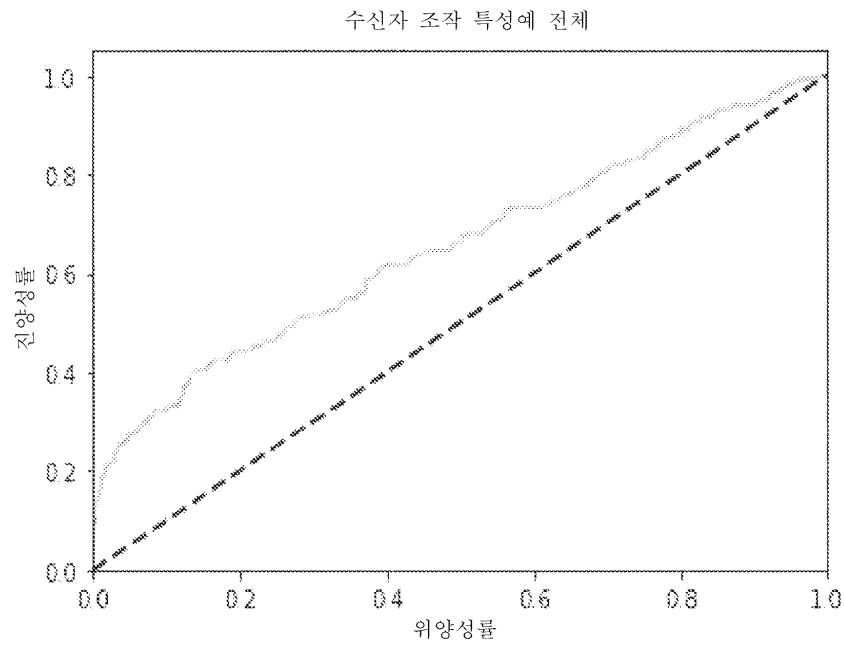
도면5r

웨이블릿 GLCM 특징 그룹

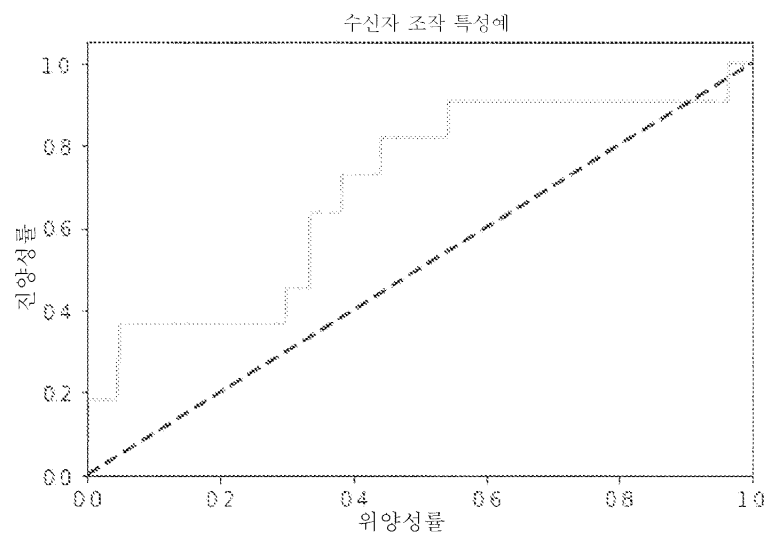


도면6

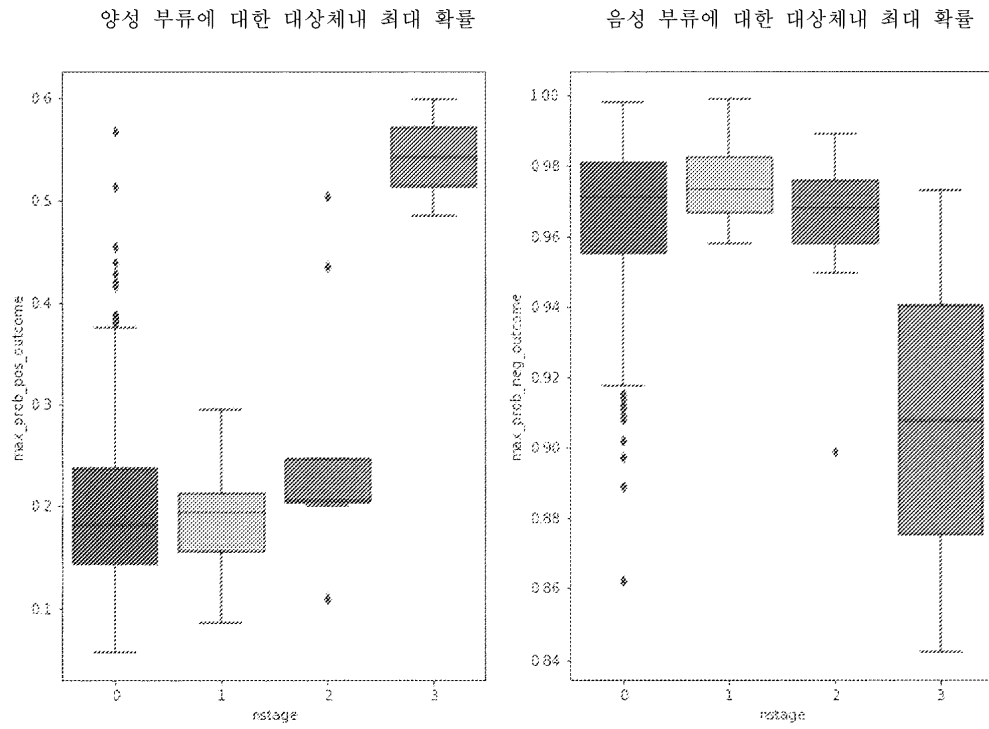
50-50 분할: LN 레벨 성능



도면7a

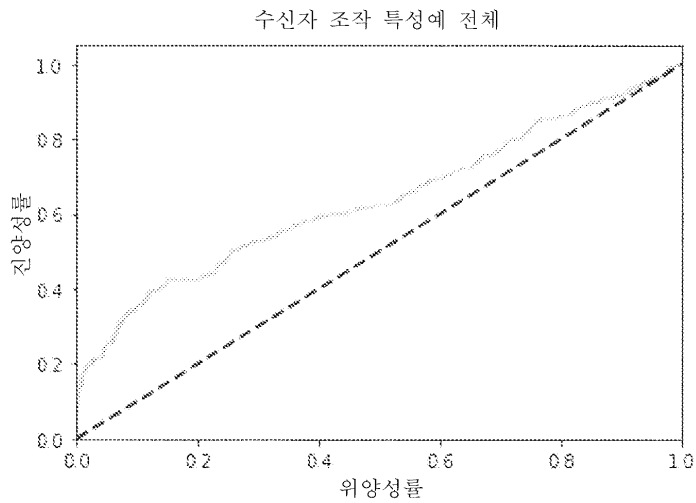


도면7b



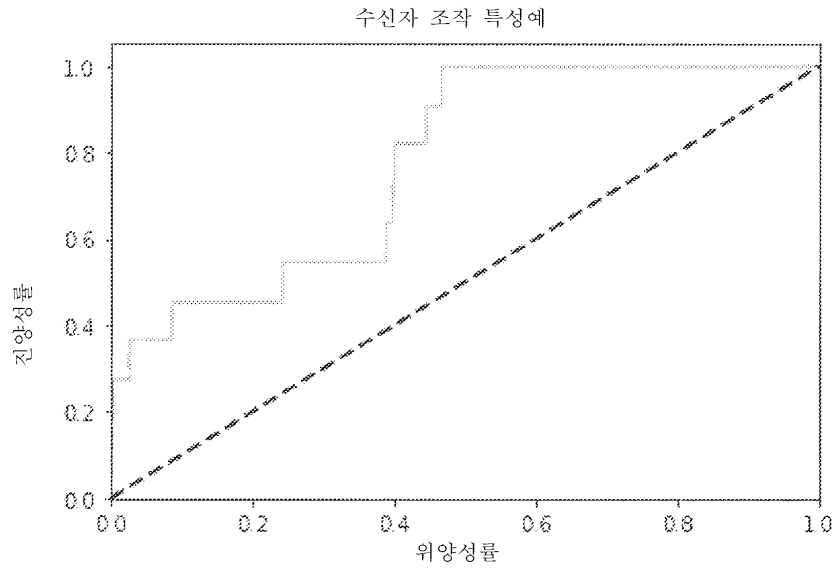
도면8

50-50 분할 LN 레벨 성능



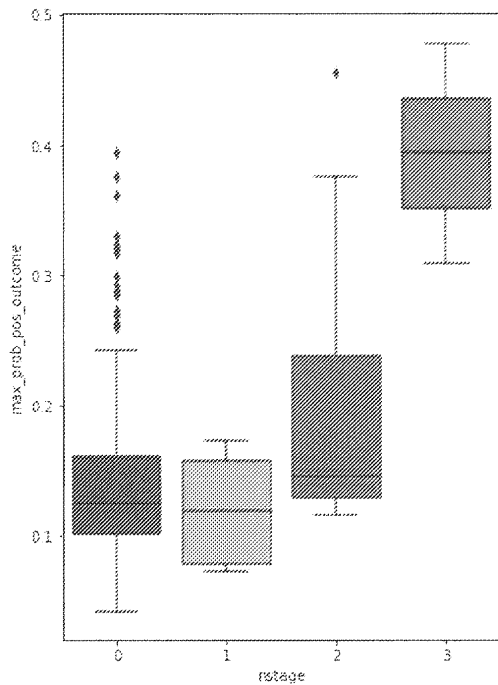
도면9a

부트스트랩 접근법: 대상체-레벨 성능 (50-50 분할)



도면9b

양성 부류에 대한 대상체내 최대 확률



음성 부류에 대한 대상체내 최대 확률

