(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0099126
(43) 공개일자 2022년07월12일

(51) 국제특허분류(Int. Cl.)
G16H 50/50 (2018.01) *A61B 5/00* (2021.01)
G06T 5/00 (2019.01) *G06T 7/11* (2017.01)
G16H 30/20 (2018.01) *G16H 30/40* (2018.01)
G16H 50/20 (2018.01) *G16H 50/70* (2018.01)
(52) CPC특허분류
G16H 50/50 (2018.01)
A61B 5/4504 (2013.01)
(21) 출원번호 10-2020-0182821
(22) 출원일자 2020년12월24일
심사청구일자 2020년12월24일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이유미
서울특별시 서대문구 연세로 50-1 연세의료원 내과학교실
홍남기
서울특별시 서대문구 연세로 50-1 연세의료원 내과학교실
박혜정
서울특별시 관악구 남부순환로194길 68 조은터빌라 401호
(74) 대리인
설인보, 배지현

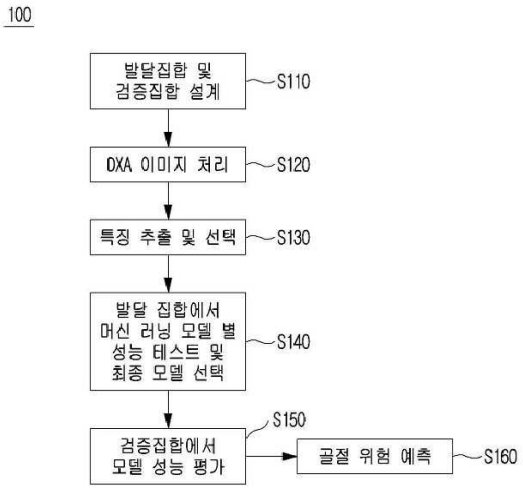
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 **골절 위험 예측 방법 및 장치**

(57) 요 약

본 발명의 실시예들은 골절 위험을 예측하기 위한 방법들에 관한 것으로서, 머신 러닝 기반의 본 라디오믹스 스코어(Bone Radiomic Score) 모델을 개발함으로써 골절 위험을 예측을 향상시키기 위한 것이다. 본 발명의 일 실시예로서 골절 위험을 예측하기 위한 방법은, 발달 집합을 설계하는 단계, 발달 집합에 포함된 복수의 대상들에 대한 뼈 이미지들을 처리하는 단계, 뼈 이미지들로부터 특성 변수들을 추출하는 단계, 추출된 특성 변수들 중 골절 위험을 예측하기 위해 필요한 최적의 특성 변수들을 선택하는 단계, 최적의 특성 변수들을 발달 집합의 훈련 집합을 이용하여 머신러닝을 수행하는 단계, 골절 위험을 예측하기 위한 본 라디오믹스 스코어 모델을 설계하는 단계를 수행하도록 구성된다.

대 표 도 - 도1



(52) CPC특허분류
A61B 5/4528 (2013.01)
G06T 5/002 (2013.01)
G06T 7/11 (2017.01)
G16H 30/20 (2018.01)
G16H 30/40 (2018.01)
G16H 50/20 (2018.01)
G16H 50/70 (2018.01)
G06T 2207/20032 (2013.01)
G06T 2207/30008 (2013.01)

명 세 서

청구범위

청구항 1

골절 위험을 예측하는 방법에 있어서,

발달 집합을 설계하는 단계;

상기 발달 집합에 포함된 복수의 대상들에 대한 뼈 이미지들을 처리하는 단계;

상기 뼈 이미지들로부터 특성 변수들을 추출하는 단계;

상기 추출된 특성 변수들 중 상기 골절 위험을 예측하기 위해 필요한 최적의 특성 변수들을 선택하는 단계;

상기 최적의 특성 변수들을 상기 발달 집합의 훈련 집합을 이용하여 머신러닝을 수행하는 단계;

상기 골절 위험을 예측하기 위한 본 라디오믹스 스코어 모델을 설계하는 단계를 포함하는, 골절 위험 예측 방법.

청구항 2

제1항에 있어서,

상기 머신러닝은 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델 중 하나를 이용해 수행되는, 골절 위험 예측 방법.

청구항 3

제2항에 있어서,

상기 머신러닝을 수행하는 단계에서,

상기 최적의 특성 변수들을 상기 랜덤 포레스트 모델, 상기 엘라스틱 넷 모델, 상기 그라디언트 부스팅 결정 트리 모델 및 상기 서포트 벡터 머신 모델에 각각 적용하여 머신 러닝을 수행하고,

그 결과를 검증 집합에 적용하여 하나의 머신 러닝 모델을 선택하는, 골절 위험 예측 방법.

청구항 4

제1항에 있어서,

상기 뼈 이미지는 DXA 고관절 이미지이고, 상기 DXA 고관절 이미지로부터 대퇴 경부(Femoral Nect), 전자(trochanter), 전자 간부(intertochanter) 및 전체 고관절(total hip)에 해당하는 4개의 관심영역들을 설정하고, 4개의 관심영역들에 대하여 특성 변수를 추출하기 위한 전처리 과정을 거치는 것을 특징으로 하는, 골절 위험 예측 방법.

청구항 5

제4항에 있어서, 상기 전처리 과정은 상기 관심영역들에 대해 생성된 가이드 라인을 따라 각 영역 별로 서로 다른 색상으로 마스킹을 실시하고, 상기 가이드 라인을 제거한 다음 이미지 히스토그램 평활화 및 메디안 필터링을 거쳐 기질 패턴을 향상시키는 과정으로 이루어지는, 골절 위험 예측 방법.

청구항 6

제1항에 있어서, 상기 발달 집합은 환자군 데이터 베이스와 대조군 데이터 베이스로부터 환자군 및 대조군이 1:A로 매칭된 발달 집합을 설계하고, 상기 매칭된 발달 집합으로부터 사용이 불가능한 뼈 이미지를 제거하여 최종 발달 집합을 설계하는, 골절 위험 예측 방법.

청구항 7

제1항에 있어서,

상기 최적의 특성 변수는 GLRLM(Gray Level Run Length Matrix) Run Entropy, GLRLM GLNN (Gray Level Non-uniformity Normalized), GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), GLCM IDMN (Inverse Difference Moment Normalized), GLCM Cluster Prominence, GLSZM SZN(Size-Zone Non-Uniformity), GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis), GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity), GLSZM GLV(Gray Level Variance), GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized)인, 골절 위험 예측 방법.

청구항 8

제1항에 있어서,

상기 뼈 이미지는 이중 에너지 X선 흡수 계측법(DXA) 이미지이며,

검사 대상의 DXA 이미지를 상기 본 라디오믹스 스코어 모델을 이용하여 골절 위험이 있는지 여부 또는 골절 위험의 정도를 예측하는 단계로, 본 라디오 믹스 스코어가 미리 설정된 임계값을 초과하는 경우 고위험군으로 예측하는 단계를 더 수행하는, 골절 위험 예측 방법.

청구항 9

골절 위험을 예측하기 위한 컴퓨터 프로그램 코드를 포함하는 컴퓨터로 읽을 수 있는 매체에 있어서,

상기 컴퓨터 프로그램 코드가 컴퓨터에서 수행되면 상기 컴퓨터가:

발달 집합을 설계하는 단계;

상기 발달 집합에 포함된 복수의 대상들에 대한 뼈 이미지들을 처리하는 단계;

상기 뼈 이미지들로부터 특성 변수들을 추출하는 단계;

상기 추출된 특성 변수들 중 상기 골절 위험을 예측하기 위해 필요한 최적의 특성 변수들을 선택하는 단계;

상기 최적의 특성 변수들을 상기 발달 집합의 훈련 집합을 이용하여 머신러닝을 수행하는 단계; 및

상기 골절 위험을 예측하기 위한 본 라디오믹스 스코어 모델을 설계하는 단계를 수행하도록 구성된, 컴퓨터 기록 매체.

청구항 10

제9항에 있어서,

상기 머신러닝은 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델 중 하나를 이용해 수행되는, 컴퓨터 기록 매체.

청구항 11

제10항에 있어서,

상기 머신러닝을 수행하는 단계에서,

상기 최적의 특성 변수들을 상기 랜덤 포레스트 모델, 상기 엘라스틱 넷 모델, 상기 그라디언트 부스팅 결정 트리 모델 및 상기 서포트 벡터 머신 모델에 각각 적용하여 머신 러닝을 수행하고,

그 결과를 검증 집합에 적용하여 하나의 머신 러닝 모델을 선택하는, 컴퓨터 기록 매체.

청구항 12

제9항에 있어서,

상기 뼈 이미지는 DXA 고관절 이미지이고, 상기 DXA 고관절 이미지로부터 대퇴 경부(Femoral Nect), 전자

(trochanter), 전자 간부(intertochanter) 및 전체 고관절(total hip)에 해당하는 4개의 관심영역들을 설정하고, 4개의 관심영역들에 대하여 특성 변수를 추출하기 위한 전처리 과정을 거치는 것을 특징으로 하는, 컴퓨터 기록 매체.

청구항 13

제12항에 있어서, 상기 전처리 과정은 상기 관심영역들에 대해 생성된 가이드 라인을 따라 각 영역 별로 서로 다른 색상으로 마스킹을 실시하고, 상기 가이드 라인을 제거한 다음 이미지 히스토그램 평활화 및 메디안 필터링을 거쳐 기질 패턴을 향상시키는 과정으로 이루어지는, 컴퓨터 기록 매체.

청구항 14

제9항에 있어서, 상기 발달 집합은 환자군 데이터 베이스와 대조군 데이터 베이스로부터 환자군 및 대조군이 1:A로 매칭된 발달 집합을 설계하고, 상기 매칭된 발달 집합으로부터 사용이 불가능한 뼈 이미지를 제거하여 최종 발달 집합을 설계하는, 컴퓨터 기록 매체.

청구항 15

제9항에 있어서,

상기 최적의 특성 변수는 GLRLM(Gray Level Run Length Matrix) Run Entropy, GLRLM GLNN (Gray Level Non-uniformity Normalized), GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), GLCM IDMN (Inverse Difference Moment Normalized), GLCM Cluster Prominence, GLSZM SZN(Size-Zone Non-Uniformity), GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis), GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity), GLSZM GLV(Gray Level Variance), GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized)인, 컴퓨터 기록 매체.

청구항 16

제9항에 있어서,

상기 뼈 이미지는 이중 에너지 X선 흡수 계측법(DXA) 이미지이며,

검사 대상의 DXA 이미지를 상기 본 라디오믹스 스코어 모델을 이용하여 골절 위험이 있는지 여부 또는 골절 위험의 정도를 예측하는 단계로, 본 라디오 믹스 스코어가 미리 설정된 임계값을 초과하는 경우 고위험군으로 예측하는 단계를 더 수행하는, 컴퓨터 기록 매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 골절 위험을 예측하기 위한 방법에 관한 것으로서, 머신 러닝 기반의 본 라디오믹스 스코어(Bone Radiomic Score) 모델을 개발함으로써 향상된 골절 위험 예측 성능을 제공하기 위한 것이다.

배 경 기 술

[0002] 전 세계적으로 고령화가 진행되는 시기에 고관절 골절은 매우 심각한 건강 문제가 되고 있다. 고관절 골절은 사망율, 질병률 및 경제적 부담을 증가시킬 수 있다. 이중 에너지 X선 흡수법(DXA: Dual energy X-ray Absorpthionmetry)를 이용한 면적 골밀도(aBMD: areal Bone Mineral Density) 측정 방법은 현재 골다공증(osteoporosis)을 진단하는 표준 방법이다. 그러나 DXA만을 이용한 골밀도 측정 방법으로 골다공증이 없는 것으로 판명된 개인들 중 절반 정도에서 취약한 골절이 발생하고 있다. 따라서, 골절 위험 예측 방법을 향상시키는 방안이 필요하다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) 1. 대한민국 공개번호 제2020-0085470호: AI 골밀도 판독 장치 및 방법

(특허문헌 0002) 2. 대한민국 공개번호 제10-2017-0034381: 뼈 장애, 연골 장애, 또는 이 둘다의 관리 또는 개선을 위한 조성물 및 방법

발명의 내용

해결하려는 과제

[0004] 본 발명의 목적은 향후 골절이 발생할 가능성과 관련된 골절 위험 예측 방법을 제공하는 것이다.

[0005] 본 발명의 다른 목적은 골절 위험 예측을 위해 머신 러닝에 기반한 본 라디오믹스 스코어 모델을 제공하는 것이다.

[0006] 본 발명의 또 다른 목적은 복수의 머신 러닝 모델들이 사용되는 경우에 골절 위험 예측에 최적의 머신 러닝 모델을 찾는 방법을 제공하는 것이다.

[0007] 본 발명의 또 다른 목적은 이중 에너지 X선 흡수법(DXA)으로 촬영한 뼈 이미지로부터 본 라디오믹스 스코어를 추정하는 방법을 제공하는 것이다.

[0008] 본 발명의 또 다른 목적은 영상 이미지로부터 관심 영역에 대한 기질의 특성 변수(texture feature)을 도출하고, 이에 대해 러닝 머신을 적용하여 고관절 골절 위험을 예측 방법을 개선하는 것이다.

[0009] 본 발명에서 이루고자 하는 기술적 목적들은 이상에서 언급한 사항들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 이하 설명할 본 발명의 실시예들로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 고려될 수 있다.

과제의 해결 수단

[0010] 이하에서는 본 발명의 실시예들로서 골절 위험을 예측하기 위한 방법에 관한 것으로서, 머신 러닝 기반의 본 라디오믹스 스코어(Bone Radiomic Score) 모델을 개발함으로써 골절 위험을 예측을 향상시키기 위한 것이다.

[0011] 본 라디오믹스 스코어(Bone Radiomic Score)란 뼈 이미지로부터 얻어진 기질의 특성 변수 중 선택된 최적의 기질 변수를 사용하여 골절 위험도를 예측할 수 있도록 제시된 점수 체계이며, 테스트 집합에서의 최적 성능을 보여주는 머신 러닝 모델로부터 골절에 대한 가능성을 계산한 0 에서부터 100까지의 점수이다.

[0012] 본 발명의 일 실시예로서 골절 위험을 예측하기 위한 방법은, 발달 집합을 설계하는 단계, 발달 집합에 포함된 복수의 대상들에 대한 뼈 이미지들을 처리하는 단계, 뼈 이미지들로부터 기질의 특성 변수들을 추출하는 단계, 추출된 특성 변수들 중 골절 위험을 예측하기 위해 필요한 최적의 특성 변수들을 선택하는 단계, 최적의 특성 변수들을 발달 집합의 훈련 집합을 이용하여 머신러닝을 수행하는 단계, 골절 위험을 예측하기 위한 본 라디오믹스 스코어 모델을 설계하는 단계를 수행하도록 구성된다.

[0013] 이때, 머신러닝은 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델 중 하나를 이용해 수행될 수 있다.

[0014] 상기 머신러닝을 수행하는 단계에서, 최적의 특성 변수들을 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델에 각각 적용하여 머신 러닝을 수행하고, 그 결과를 검증 집합에 적용하여 하나의 머신 러닝 모델을 선택할 수 있다. 이때, 본 라디오믹스 스코어 모델의 성능을 평가하기 위해 훈련 집합의 일부분을 성능 평가 데이터로 사용할 수 있다.

[0015] 특히, 본 발명의 또 다른 실시예에 따르면, 고관절 부위에서 선택된 관심 영역에서의 기질의 특성 변수를 바탕으로 고관절 골절의 위험성을 본 라디오믹스 스코어를 통해 예측하는 것에 또 다른 목적이 있다.

[0016] 이를 위해, 상기 뼈 이미지는 DXA 고관절 이미지이고, 상기 DXA 고관절 이미지로부터 대퇴 경부(Femoral Nect), 전자(trochanter), 전자 간부(intertochanter) 및 전체 고관절(total hip)에 해당하는 4개의 관심영역들을 설정하고, 4개의 관심영역들에 대하여 특성 변수를 추출하기 위한 전처리 과정을 거치도록 구성할 수 있다.

[0017] 이러한 전처리 과정은 상기 관심영역들에 대해 생성된 가이드 라인을 따라 각 영역 별로 서로 다른 색상으로 마스킹을 실시하고, 상기 가이드 라인을 제거한 다음 이미지 히스토그램 평활화 및 메디안 필터링을 거쳐 기질 패턴을 향상시키는 과정으로 이루어질 수 있다.

[0018] 이때, 발달 집합은 환자군 데이터 베이스와 대조군 데이터 베이스로부터 환자군 및 대조군이 1:A로 매칭된 발달 집합을 설계할 수 있다. 이때, A는 0이 아닌 자연수 중 선택될 수 있다.

[0019] 매칭된 발달 집합으로부터 사용이 불가능한 뼈 이미지를 제거하여 최종 발달 집합을 설계할 수 있다.

[0020] 최적의 특성 변수는 GLRLM(Gray Level Run Length Matrix) Run Entropy, GLRLM GLNN (Gray Level Non-uniformity Normalized), GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), GLCM IDMN (Inverse Difference Moment Normalized), GLCM Cluster Prominence, GLSZM SZN(Size-Zone Non-Uniformity), GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis), GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity), GLSZM GLV(Gray Level Variance), GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized)일 수 있다.

[0021] 뼈 이미지는 이중 에너지 X선 흡수 계측법(DXA) 이미지일 수 있다. 본 발명의 실시예들 및 컴퓨터 프로그램 코드는 검사 대상의 DXA 이미지를 본 라디오믹스 스코어 모델을 이용하여 골절 위험이 있는지 여부 또는 골절 위험의 정도를 예측하는 단계로, 본 라디오 믹스 스코어가 미리 설정된 임계값을 초과하는 경우 고위험군으로 예측하는 단계를 더 수행하도록 구성될 수 있다.

[0022] 상술한 본 발명의 양태들은 본 발명의 바람직한 실시예들 중 일부에 불과하며, 본원 발명의 기술적 특징들이 반영된 다양한 실시예들이 당해 기술분야의 통상적인 지식을 가진 자에 의해 이하 상술할 본 발명의 상세한 설명을 기반으로 도출되고 이해될 수 있다.

발명의 효과

[0023] 본 발명의 실시예들에 따르면 다음과 같은 효과를 얻을 수 있다.

[0024] 본 발명의 실시예들을 통해 향상된 골절 위험 예측 방법이 제공된다.

[0025] 머신 러닝에 기반한 골절 위험 예측 방법을 제공함으로써 영상 이미지만을 판독하더라도 보다 정확하게 골절 여부를 예측할 수 있다.

[0026] 복수의 머신 러닝 모델들이 사용되는 경우에 골절 위험 예측에 최적의 머신 러닝 모델을 찾는 방법을 제공할 수 있다. 예를 들어, 대상 뼈마다 적용될 머신러닝 방법이 다를 수 있으며, 각 대상에 따라 적합한 머신러닝 방법을 제안할 수 있다.

[0027] 본 라디오믹스 스코어 모델을 이용하여 이중 에너지 X선 흡수 계측법(DXA)으로 촬영한 뼈 이미지로부터를 해당 뼈가 골절의 위험이 있는지 여부를 추정할 수 있다.

[0028] 특히, 본 발명의 실시예들을 통해, 고관절 부위에서 선택된 관심 영역에서의 특성 변수를 바탕으로 고관절 골절의 위험성을 본 라디오믹스 스코어를 통해 정확하게 예측할 수 있는 바, 임상 경험이 부족한 의료진의 고관절 골절 위험도 판단에 효과적인 보조 지표를 제공할 수 있는 효과가 있다.

[0029] 본 발명의 실시예들에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 이하의 본 발명의 실시예들에 대한 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 도출되고 이해될 수 있다. 즉, 본 발명을 실시함에 따른 의도하지 않은 효과들 역시 본 발명의 실시예들로부터 당해 기술분야의 통상의 지식을 가진 자에 의해 도출될 수 있다.

도면의 간단한 설명

[0030] 본 발명에 관한 이해를 돕기 위해 상세한 설명의 일부로 포함되고, 첨부된 도면들은 본 발명에 대한 다양한 실시예들을 제공한다. 또한, 첨부된 도면들은 상세한 설명과 함께 본 발명의 실시 형태들을 설명하기 위해 사용된다.

도 1은 본 발명의 실시예로서 골절 위험을 예측하는 방법을 설명하기 위한 도면이다.

도 2는 발달 집합 및 검증 집합을 설계하는 방법을 설명하기 위한 도면이다.

도 3은 DXA 이미지 처리 과정을 설명하기 위한 도면이다.

도 4는 DXA 이미지 처리 과정을 설명하기 위한 또 다른 도면이다.

도 5는 선택된 소정 개수의 특성 변수들을 설명하기 위한 도면이다.

도 6은 후속된 고관절 골절을 겪은 사람과 그렇지 않은 사람들 간의 최종 모델에서 선택된 특성 변수의 정규화된 평균 차이를 도시한 것이다.

도 7는 도 1 내지 도 6에서 설명한 골절 위험 예측 방법이 수행될 수 있는 장치 및 컴퓨터 기록 매체를 설명하기 위한 도면이다.

도 8은 본 라디오 스코어 모델과 골절 위험도 간의 관계를 설명하는 도면이다.

도 9은 고관절 골절 사고에 대한 카플란 마이어 생존 도표를 나타낸다.

도 10은 본 발명에 따른 본 라디오믹스 스코어가 높은 그룹과 낮은 그룹 들 간에 있어서 주요 골다공증성 골절 위험도 차이를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0031] 이하의 실시예들은 본 발명의 구성요소들과 특징들을 소정 형태로 결합한 것들이다. 각 구성요소 또는 특징은 별도의 명시적 언급이 없는 한 선택적인 것으로 고려될 수 있다. 각 구성요소 또는 특징은 다른 구성요소나 특징과 결합되지 않은 형태로 실시될 수 있다. 또한, 일부 구성요소들 및/또는 특징들을 결합하여 본 발명의 실시예를 구성할 수도 있다. 본 발명의 실시예들에서 설명되는 동작들의 순서는 변경될 수 있다. 어느 실시예의 일부 구성이나 특징은 다른 실시예에 포함될 수 있고, 또는 다른 실시예의 대응하는 구성 또는 특징과 교체될 수 있다.

[0032] 도면에 대한 설명에서, 본 발명의 요지를 흐릴 수 있는 부분, 장치 및/또는 구성 등은 기술하지 않았으며, 당업자의 수준에서 이해할 수 있을 정도의 부분, 장치 및/또는 구성 또한 기술하지 아니하였다. 또한, 도면에서 동일한 도면 부호를 사용하여 지칭하는 부분은 장치 구성 또는 방법에서 동일한 구성 요소 또는 단계를 의미한다.

[0033] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함(comprising 또는 including)"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "쪋부" 또는 "쪋기" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미한다. 또한, "일(a 또는 an)", "하나(one)", "그(the)" 및 유사 관련어는 본 발명을 기술하는 문맥에 있어서(특히, 이하의 청구항의 문맥에서) 본 명세서에 달리 지시되거나 문맥에 의해 분명하게 반박되지 않는 한, 단수 및 복수 모두를 포함하는 의미로 사용될 수 있다.

[0034] 또한, 본 발명의 실시예들에서 사용되는 특정(特定) 용어들 및/또는 기호들은 본 발명의 이해를 돕기 위해서 제공된 것이며, 이러한 특정 용어의 사용은 본 발명의 기술적 사상을 벗어나지 않는 범위에서 다른 형태로 변경될 수 있다.

[0035] 예를 들어, 기질의 특성 변수라는 용어는 간단히 특성 변수 또는 기질 특성이나 텍스처 특성이라는 용어로 쓰일 수 있고, 사례 그룹은 환자 집단과 같은 의미로 쓰일 수 있다.

[0036] 이하, 본 발명에 따른 바람직한 실시 형태를 첨부된 도면을 참조하여 상세하게 설명한다. 첨부된 도면과 함께 이하에 개시될 상세한 설명은 본 발명의 예시적인 실시형태를 설명하고자 하는 것이며, 본 발명이 실시될 수 있는 유일한 실시형태를 나타내고자 하는 것이 아니다.

[0038] **1. DXA 및 라디오믹스**

[0039] 이중 에너지 X선 흡수 계측법(DXA: Dual Energy X-ray Absorptionmetry)은 서로 다른 두 종류의 X선을 조사하여 골밀도를 측정하는 검사이다.

[0040] 라디오믹스(Radiomics)는 예측 가능한 성과 또는 개선된 진단을 통해 임상적 결정을 지원하기 위한 표준 치료(Standard-of-care) 의료 영상으로부터 육안으로 식별하기 어려운 질병의 특징들을 정량적 이미지 특성으로써 포괄적이고 자동화된 높은 처리 마이닝(mining)을 위한 방식이다.

[0041] 라디오믹스는 뼈의 공간 및 질감(texture) 이질성을 포함하여 방사선 표현형의 큰 패널(large panel)을 정량화할 수 있다. 골 조직 형태 계측법(bone histomorphometry), 정량적골염량측정법(QCT: Quantitative Computed Tomography), 또는 일반 X선 이미지에 기반한 여러 연구에서는 골 질량 단독뿐 아니라 고관절 골절의 취약성에 부분적으로 기여할 수 있는 골 분포 및 근위 대퇴골의 미세구조(micro-architecture)에서 노화 및 질병과 관련된 실질적인 공간 이질성을 제안하고 있다. 최근 연구에서는 처리되지 않은 DXA 고관절 이미지를 사용하여 노화

에 따른 픽셀 단위의 골밀도(BMD: Bone Mineral Density) 변화에 대한 시공간 이질성을 보여주고 있다.

[0042] 만약 적절하게 활용된다면 라디오믹스는 DXA 고관절 영상으로부터 고관절 골절과 관련된 픽셀 단위 수준의 텍스처 인덱스를 수집(mining)하는데 유용하게 사용될 수 있으며, 이는 표준 치료 기준에서 DXA 골밀도 검사에 추가되는(add-on) 정보로서 고관절 골절 위험을 예측하는 방법을 보다 개선할 수 있다.

[0043] 이하에서는 본 발명의 실시예들로서 DXA 및 머신 러닝에 기반하여 골절의 위험을 예측하는 방법에 대해서 설명한다.

[0045] **2. 머신러닝 기반의 골절 위험 예측 방법**

[0046] 본 발명의 실시예들에서 골절 위험을 예측하는 방법은 인체를 구성하는 모든 뼈들을 대상으로 수행될 수 있다. 다만, 이하에서는 설명의 편의를 위해 고관절 부위의 골절 위험도를 예측하는 방법을 개시한다. 또한, 다른 부위의 뼈에 대한 골절 위험을 예측하는 방법은 고관절 골절 위험도를 예측하는 방법과 동일하게 수행될 수 있다.

[0047] 도 1은 본 발명의 실시예로서 골절 위험을 예측하는 방법을 설명하기 위한 도면이다.

[0048] 본 발명의 실시예들에서, DXA 기반으로 고관절에 대한 골절 위험을 예측하는 방법을 위해 발달 집합(Development set) 및 검증 집합(Validation set)을 설계할 수 있다 (S110).

[0049] 먼저 발달 집합을 설계하는 방법에 대해서 설명한다.

[0050] 고관절 골절을 당한 개인들을 가장 구별하기 좋게 선택된 특성 변수를 기반으로 본 라디오믹스 스코어 모델를 개발하기 위해 발달 집합 즉, 사례-대조군(case-control study) 연구가 설계된다.

[0051] 발달 집합을 설계하는 방법을 설명하기 위해 도 2를 참조할 수 있다. 도 2의 왼편을 참조하면, 사례-대조군 연구를 위한 발달 집합은 환자군 데이터 베이스(case group data base; 201)로부터 표본을 추출하고, 대조군 데이터 베이스(control group date base; 203)로부터 표본을 추출한다 (S201, S203). 이때, '사례' 또는 '케이스 집단'은 고관절 골절을 겪는 환자군을 의미한다. 예를 들어, S201 단계에서 환자군은 관찰 기간 동안 취약한 고관절 골절을 겪은 65세 이상의 여성으로 설정될 수 있다.

[0052] 환자군 데이터 베이스(201)는 전자건강기록(EHR: Electronic Health Recode) 데이터 베이스일 수 있다. EHR 데이터 베이스란 환자의 건강 정보를 환자의 동의 하에 소정 기간 동안 저장하고 있는 정보의 보관소이다. 환자군 데이터 베이스(201)는 사례 데이터 베이스 또는 케이스 데이터 베이스라고 불릴 수 있다.

[0053] 대조군 데이터 베이스(203)는 한국 도시 농촌 노인(KURE: Korean Urban Rural Elderly) 코호트 데이터 베이스일 수 있다. 이때, 대조군은 관찰 기간 동안 취약한 고관절 골절을 겪지 않은 65세 이상의 여성으로 설정될 수 있다. 각 표본에 대한 DXA 영상 이미지는 전자건강기록(EHR) 데이터 베이스 및/또는 코호트 데이터 베이스에 이미 저장되어 있는 것으로 가정할 수 있다.

[0054] 각 데이터 베이스로부터 추출한 표본들을 1:A (환자군:대조군)로 매칭한다. 이때, A는 표본의 개수로서 2 이상의 자연수일 수 있다 (S205).

[0055] S205 단계에서 데이터 표본들의 비율은 보유한 데이터 셋에 따라 적절히 결정될 수 있으며, 예를 들어 골절위험 예측 방식의 목적에 따라 M:N 비율로도 설정할 수 있다. 이때, M, N은 1 이상의 자연수이다.

[0056] 매칭된 발달 집합 구성은 골절 환자군(또는, 사례군) 및 골절이 없는 대조군으로 구성된다 (S207).

[0057] 매칭된 발달 집합에서 사용이 불가능한 DXA 뼈 이미지는 제거될 수 있다. 예를 들어, 매칭된 발달 집합에서 DXA 고관절 이미지가 없는 개체 및 DXA 머신 타입이 달라 사용할 수 없는 DXA 고관절 이미지를 가진 개체는 매칭된 발달 집합에서 제거한다 (S209).

[0058] 이를 통해 환자군 및 대조군을 포함하는 최종 발달 집합이 구성될 수 있다. 최종 발달 집합에서 무작위로 선택된 75%는 훈련 집합으로 구성되고, 나머지 25%는 테스트 집합으로 구성된다. 훈련 집합은 본 라디오믹스 스코어 모델을 설계하기 위한 머신 러닝의 학습을 위해 사용되고, 테스트 집합은 훈련을 통해 학습한 본 라디오믹스 스코어 모델에 대한 최종 성능 평가를 위해 사용된다 (S211).

[0059] S211 단계를 통해 구성된 최종 발달 집합을 이용하여 골절 위험 예측을 위한 머신 러닝 기반의 DXA 본 라디오믹스 스코어 모델을 설계할 수 있다 (S213).

[0060] S211 단계 및 S213 단계에서 발달 집합 내 훈련 집합을 이용하여 훈련하여 본 라디오믹스 스코어 모델을 설계하는 보다 구체적인 방법에 대해서는 도 1의 나머지 설명과 관련하여 후술한다.

[0061] 검증 집합은 설계한 본 라디오믹스 스코어 모델의 성능을 평가하기 위해 훈련/학습시킨 모델이 잘 작동하는지 여부를 검증하기 위한 것이다. 검증 집합은 발달 집합과 독립적으로 구성되거나 최종 발달 집합의 훈련 집합의 일부를 이용하여 구성될 수 있다. 이하에서 설명할 본 발명의 양태들에서 검증 집합은 발달 집합과 독립적으로 구성되는 것을 가정하여 설명한다. 다만, 머신러닝 기법에 따라 검증 집합은 발달 집합을 기반으로 구성되는 훈련 집합의 일부를 희생하여 검증 집합으로 이용할 수 있다. 검증 집합의 모표본을 어떻게 설정하는지 만이 다를 뿐, 검증 집합으로서의 기능은 동일하게 적용될 수 있다.

[0062] 다시 도 2의 오른편을 참조하여 검증 집합을 설계하는 방법에 대해서 설명한다. 검증 집합은 사례-코호트(case-cohort) 연구를 위해 사용된다. 시간-사건 데이터 구조를 사용하는 지역 사회 기반 코호트에서 고관절 골절 위험 예측에 대한 본 라디오믹스 스코어 모델의 성능을 테스트하기 위해, 검증 집합은 KURE 코호트(Korean Urban Rural Elderly cohort)의 케이스 코호트 설계를 이용하여 무작위로 서브샘플링된다.

[0063] 현재 단계에서 본 라디오믹스 스코어 모델을 위한 마스크 생성의 노동 집약적이고 시간 소모적인 특성을 감안할 때, 사례-코호트 설계는 전체 공변량 데이터가 개별적 사례 및 하위 코호트에만 필요하다는 이점을 제공하는 동시에 단일 결과만을 분석하는 네스티드(nested) 사례-대조군 설계와 달리 여러 결과를 연구할 수 있다.

[0064] 이러한 검증 집합을 설계하기 위해, 대조군 데이터 베이스(203)에서 모표본 집단을 추출한다 (S215).

[0065] 추출한 모표본 집단에서 불필요한 대조군을 제거한다. 이때, 불필요한 대조군이란 연구 목적과 맞지 않는 표본 집단을 의미한다. 예를 들어, 65세 이상의 여성을 대상으로 설계된 본 라디오믹스 스코어 모델을 검증하기 위한 검증 집합은 남성으로 구성된 대조군 및 65세 미만의 여성 대조군은 불필요하다. 따라서 이러한 불필요한 대조군을 제거되는 것이 바람직하다 (S217).

[0066] 이후, 남은 표본 집단에서 후속 조치 기간 동안 연구 목적에 따라 대조군을 선정하고 (S219), 사례 코호트 집단(또는, 케이스 코호트 집단)을 구성한다 (S221).

[0067] S221 단계에서, 대조군의 10%는 서브 샘플링되어 사례 코호트 집단으로 구성될 수 있다. 서브 샘플링된 대조군과 서브 샘플링되지 않은 대조군 중 고관절 골절 환자를 더하여 사례 코호트 집단으로 구성할 수 있다. 이때, 서브 샘플링 비율은 골절 위험 예측 목적, 검증 집합의 크기 및/또는 연구 목적 등에 따라 달라질 수 있다.

[0068] 다음 표 1은 발달 집합과 검증 집합의 특성의 일례를 나타낸다.

표 1

[0069]

	발달 집합 (사례-대조군 연구: 병원 EHR 및 커뮤니티 코호트; 골절까지 시간의 중위값: 2.1 [0.9-3.9] 년)			검증 집합 (10% 서브 샘플링한 사례-코호트 연구 및 커뮤니티 코호트)		
	사례: 후속 조치 기간 동안 고관절 골절을 겪은 집단 (n=143)	대조군: 후속 조치 기간 동안 고관절 골절을 겪지 않은 집단(n=290)	P 값	후속 조치 기간 동안 고관절 골절을 겪은 집단 (n=34)	후속 조치 기간 동안 고관절 골절을 겪지 않은 집단(n=186)	P 값
나이	73 ± 6	73 ± 6	0.807	75 ± 5	71 ± 4	<0.001
BMI	23.2 ± 4.1	24.2 ± 3.1	0.004	23.6 ± 2.6	24.6 ± 3.0	0.091
Previous fracture	88 (62)	97 (33)	<0.001	25 (73)	58 (31)	<0.001
DXA T-score						
Lumbar spine	-2.1 ± 1.2	-2.0 ± 1.2	0.477	-1.9 ± 1.2	-1.8 ± 1.1	0.713
Femoral neck	-2.8 ± 1.0	-2.2 ± 1.0	<0.001	-2.8 ± 0.9	-2.0 ± 0.8	<0.001
Total hip	-2.0 ± 1.0	-1.3 ± 1.0	<0.001	-2.0 ± 0.8	-1.1 ± 0.8	<0.001
DXA 골다공증	98 (69)	155 (53)	0.003	26 (77)	82 (44)	0.001
Previous anti-osteoporosis medication use	31 (22)	67 (23)	0.739	9 (27)	36 (19)	0.344
Diabetes	46 (32)	54 (19)	0.002	8 (24)	46 (25)	0.881
Hypertension	93 (65)	187 (64)	0.910	26 (76)	116 (62)	0.114
Any cancer	16 (11)	21 (7)	0.167	4 (11.8)	17 (9)	0.632

[0070] 표 1에서 데이터는 중간값 ± 기준 편차 또는 확률(%)로 표기되었다. BMI는 Body mass index의 약자이고, EHR은 전자 건강 기록의 약자이다. n은 각 집단의 표본 개수(명)를 의미한다.

[0071] 골절 위험 예측 방법을 설명하기 위해 다시 도 1을 참조한다.

[0072] 도 1의 S110 단계 및 도 2에서 설명한 바와 같이 발달 집합 및 검증 집합을 설계할 수 있다. 이후, 발달 집합에서 훈련 집합에 속한 표본들에 대한 DXA 이미지를 처리하는 과정을 수행한다 (S120).

[0073] S120 단계에서 DXA 이미지는 소정 기관의 표준화된 프로토콜에 따라 획득되는 것이 바람직하다. S120 단계에서 DXA 이미지 처리 과정은 기질특성 추출을 위한 전처리 과정으로, 이하에서는 도 3 및 도 4를 이용하여 설명한다.

[0074] 영상 저장 전송 시스템(PACS: Picture Archival and Communication system) 또는 영상 데이터 베이스에서 훈련 집합에 속한 표본들에 대한 DXA 고관절 리포트를 추출할 수 있다. 이때, DXA 고관절 리포트는 의료용 디지털 영상 및 통신(DICOM: Digital Imaging and Communications in Medicine) 파일 포맷으로 출력될 수 있다 (S310).

[0075] DXA 고관절 리포트는 고관절에 대한 DXA 영상 이미지 자료를 의미한다. 출력된 DXA 고관절 이미지에서 머신러닝을 통해 분석하고자 하는 관심 영역(ROI: Region Of Interest)을 분석하고 영상 이미지를 다듬는 크로핑(Cropping)을 수행한다. 예를 들어, 고관절 부위 중 관심 부위인 대퇴 근위부(proximal femur)에서 소정 개수의 관심 영역을 크로핑할 수 있다 (S320). 예를 들어, 크로핑은 도 4에서와 같이 네모 형태로 관심 영역을 선택하고, 이를 크로핑함으로써 달성될 수 있다.

[0076] S320 단계에서 관심 영역에 대해 크로핑을 한 이후, 소정 개수의 관심 영역에 색상으로 표시를 할 수 있으며 이를 마스크(mask)한다고 한다. 예를 들어, 관심 영역이 대퇴 경부(Femoral Nect), 전자(trochanter), 전자 간부(intertochanter) 및 전체 고관절(total hip)인 경우에, DXA 이미지에서 각 관심 영역들이 구분되도록 DXA 머신에 의해 생성된 가드 라인을 따라 각 영역별로 서로 다른 색상의 마스크를 생성할 수 있다 (S330). 도 4에 도시되어 있듯이, 이러한 마스킹 작업은 DXA머신에 의해 생성되어 있는 가이드 라인을 따라 이루어지게 되며, 각 관심영역 별로 서로 다른 색상으로 마스킹될 수 있다. 관심 영역은 골절 위험을 측정하고자 하는 뼈의 종류에 따라 설정될 수 있으며, 사용자의 사용 목적에 따라 설정될 수 있다.

[0077] 이후 가이드 라인을 제거하고 (S340), 이미지 히스토그램 평활화(CLAHE: Contrast-Limited Adaptive Histogram Equalization) 단계를 수행하고 (S350), DXA 고관절 이미지에서 배경 노이즈를 줄이고 텍스처(기질) 패턴을 향상시키기 위한 메디안 필터링(Median Filter)을 수행한다 (S360). 특히, CLAHE 단계와 조합된 메디안 필터링 과정을 거침으로써 텍스처(기질) 패턴을 향상시키는 데 크게 기여할 수 있다.

[0078] 골절 위험 예측 방법을 설명하기 위해 도 1을 다시 참조한다.

[0079] 도 1의 S120 단계에서 처리된 DXA 이미지에서 머신러닝을 수행하기 위한 특징들을 추출하고 선택한다 (S130).

[0080] S130 단계에서 DXA 이미지에서 관심 영역들 각각에서 소정 개수의 그레이 레벨(gray-level) 기질의 특성 변수(texture feature)들을 추출한다. 예를 들어, 피라다이오믹스(pyradiomics) 프로그램을 이용하는 경우 DXA 이미지에서 관심 영역에 대해 최대 75개의 특성 변수들이 자동으로 추출될 수 있다. 만약, 관심 영역이 4개(예를 들어, 대퇴 경부, 전자, 전자 간부 및 전체 고관절 전체)라면, 하나의 DXA 이미지 당 300개 (75*4)의 특성 변수가 추출될 수 있다

[0081] 75개의 특성 변수들은 본 라디오믹스 스코어 모델을 이용하여 고관절 골절 여부를 분석하기 위한 특성 변수들로서 상세한 설명의 말미에 첨부된 표 2를 참조할 수 있다.

[0082] 다음으로 S130 단계에서 추출한 특성 변수들에서 골절 위험 예측 방법을 위한 머신 러닝 기반의 본 라디오믹스 스코어 모델에서 사용될 특성 변수들을 선택한다.

[0083] 이와 관련, 표 3은 본 발명의 바람직한 구현예로써, pyradiomics 3.0 을 사용하여 자동화된 프로세스로 75개의 특성 변수를 4개의 관심영역(대퇴 경부(Femoral Nect; FN), 전자(Trochanter; TR), 전자 간부(Intertochanter; IT) 및 전체 고관절(Total Hip; TH))에 대해 얻어진 300개의 특성 변수에 대한 급내 상관 계수(ICC: Interclass Correlation Coefficient)를 나타낸 것이다.

표 3

[0085] ('*'표시: 추출된 특성 변수)

	클래스 내 상관관계 (95% CI)			
특성 변수	Total hip	Femoral neck	Intertrochanter	Trochanter
Grey Level Co-occurrence Matrix (GLCM) features				
Autocorrelation	0.95 (0.90-0.97)	0.97 (0.94-0.98)	0.90 (0.79-0.95)	0.91 (0.81-0.95)
Joint Average	0.94 (0.88-0.97)	0.97 (0.93-0.98)	0.87 (0.73-0.93)	0.90 (0.79-0.95)
Cluster Prominence	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Cluster Shade	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Cluster Tendency	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Contrast	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Correlation	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Difference Average	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Difference Entropy	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Difference Variance	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Joint Energy	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Joint Entropy	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Informational Measure of Correlation (IMC) 1	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Informational Measure of Correlation (IMC) 2	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Inverse Difference Moment (IDM)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Maximal Correlation Coefficient (MCC)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Inverse Difference Moment Normalized (IDMN)	0.98 (0.95-0.99)	0.99 (0.98-0.99) (*)	0.95 (0.90-0.98) (*)	0.97 (0.94-0.99)
Inverse Difference (ID)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Inverse Difference Normalized (IDN)	0.98 (0.96-0.99)	0.98 (0.97-0.99)	0.96 (0.92-0.98)	0.98 (0.96-0.99)
Inverse Variance	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Maximum Probability	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Sum Average	0.94 (0.88-0.97)	0.97 (0.93-0.98)	0.87 (0.73-0.93)	0.90 (0.79-0.95)
Sum Entropy	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Sum of Squares	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Gray Level Size Zone Matrix (GLSZM) features				
Small Area Emphasis (SAE)	0.88 (0.76-0.94)	0.71 (0.40-0.86)	0.82 (0.63-0.91)	0.83 (0.65-0.92)
Large Area Emphasis (LAE)	0.99 (0.99-0.99)	0.98 (0.96-0.99)	0.99 (0.99-0.99)	0.98 (0.95-0.99)
Gray Level Non-Uniformity (GLN)	0.98 (0.95-0.99)	0.89 (0.78-0.95)	0.98 (0.96-0.99)	0.96 (0.93-0.98) (*)
Gray Level Non-Uniformity Normalized (GLNN)	0.99 (0.98-0.99)	0.97 (0.95-0.98)	0.99 (0.98-0.99)	0.96 (0.93-0.98)
Size-Zone Non-Uniformity (SZN)	0.95 (0.90-0.97) (*)	0.96 (0.92-0.98)	0.94 (0.87-0.97)	0.72 (0.40-0.86)
Size-Zone Non-Uniformity Normalized (SZNN)	0.92 (0.83-0.96)	0.91 (0.81-0.95)	0.88 (0.76-0.94)	0.62 (0.20-0.82)
Zone Percentage (ZP)	0.98 (0.97-0.99)	0.97 (0.93-0.98)	0.98 (0.97-0.99)	0.97 (0.93-0.98)
Gray Level Variance (GLV)	0.99 (0.99-0.99)	0.98 (0.96-0.99) (*)	0.99 (0.98-0.99)	0.98 (0.97-0.99)
Zone Variance (ZV)	0.99 (0.99-0.99)	0.98 (0.96-0.99)	0.99 (0.99-0.99)	0.98 (0.96-0.99)
Zone Entropy (ZE)	0.97 (0.95-0.98)	0.96 (0.91-0.98)	0.98 (0.97-0.99)	0.95 (0.97-0.88)
Low Gray Level Zone Emphasis (LGLZE)	0.88 (0.76-0.94)	0.98 (0.95-0.99)	0.81 (0.61-0.91)	0.71 (0.40-0.86)
High Gray Level Zone Emphasis (HGLZE)	0.96 (0.92-0.98)	0.97 (0.95-0.98)	0.94 (0.88-0.97)	0.91 (0.81-0.95)

Small Area Low Gray Level Emphasis (SALGLE)	0.91 (0.82-0.96)	0.75 (0.48-0.88)	0.88 (0.74-0.94)	0.81 (0.61-0.91)
Small Area High Gray Level Emphasis (SAHGLE)	0.90 (0.76-0.94) (*)	0.92 (0.84-0.96)	0.84 (0.67-0.92)	0.80 (0.58-0.90)
Large Area Low Gray Level Emphasis (LALGLE)	0.98 (0.97-0.99)	0.98 (0.97-0.99)	0.96 (0.93-0.98)	0.95 (0.91-0.97)
Large Area High Gray Level Emphasis (LAHGLE)	0.97 (0.94-0.98)	0.97 (0.95-0.98)	0.88 (0.75-0.94)	0.94 (0.87-0.97)
Gray Level Run Length Matrix (GLRLM) features				
Short Run Emphasis (SRE)	0.99 (0.98-0.99)	0.98 (0.96-0.99)	0.99 (0.98-0.99)	0.99 (0.98-0.99)
Long Run Emphasis (LRE)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Gray Level Non-Uniformity (GLN)	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.98-0.99)
Gray Level Non-Uniformity Normalized (GLNN)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Run Length Non-Uniformity (RLN)	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Run Length Non-Uniformity Normalized (RLNN)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Run Percentage (RP)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Gray Level Variance (GLV)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Run Variance (RV)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Run Entropy (RE)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Low Gray Level Run Emphasis (LGLRE)	0.92 (0.83-0.96)	0.93 (0.86-0.97)	0.74 (0.46-0.87)	0.81 (0.60-0.91)
High Gray Level Run Emphasis (HGLRE)	0.95 (0.90-0.97)	0.97 (0.94-0.98)	0.90 (0.80-0.95)	0.90 (0.80-0.95)
Short Run Low Gray Level Emphasis (SRLGLE)	0.90 (0.79-0.95)	0.94 (0.89-0.97)	0.75 (0.47-0.88)	0.72 (0.42-0.86)
Short Run High Gray Level Emphasis (SRHGLE)	0.97 (0.95-0.98)	0.98 (0.96-0.99)	0.95 (0.91-0.98)	0.94 (0.88-0.97)
Long Run Low Gray Level Emphasis (LRLGLE)	0.97 (0.95-0.99)	0.98 (0.97-0.99)	0.91 (0.82-0.95)	0.96 (0.92-0.98)
Long Run High Gray Level Emphasis (LRHGLE)	0.95 (0.90-0.97)	0.96 (0.92-0.98)	0.85 (0.70-0.93)	0.89 (0.78-0.95)
Gray Level Dependence Matrix (GLDM) features				
Small Dependence Emphasis (SDE)	0.99 (0.98-0.99)	0.97 (0.94-0.98)	0.99 (0.98-0.99)	0.97 (0.95-0.99)
Large Dependence Emphasis (LDE)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Gray Level Non-Uniformity (GLN)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Dependence Non-Uniformity (DN)	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Dependence Non-Uniformity Normalized (DNN)	0.99 (0.99-0.99) (*)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Gray Level Variance (GLV)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Dependence Variance (DV)	0.99 (0.99-0.99)	0.97 (0.95-0.98)	0.99 (0.98-0.99)	0.98 (0.97-0.99)
Dependence Entropy (DE)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)	0.99 (0.99-0.99)
Low Gray Level Emphasis (LGLE)	0.93 (0.86-0.96)	0.92 (0.84-0.96)	0.72 (0.41-0.86)	0.90 (0.79-0.95)
High Gray Level Emphasis (HGLE)	0.95 (0.90-0.97)	0.97 (0.94-0.98)	0.90 (0.79-0.95)	0.91 (0.81-0.95)
Small Dependence Low Gray Level Emphasis (SDLGLE)	0.91 (0.82-0.95)	0.90 (0.79-0.95)	0.70 (0.37-0.85)	0.83 (0.65-0.92)
Small Dependence High Gray Level Emphasis (SDHGLE)	0.96 (0.92-0.98)	0.98 (0.96-0.99)	0.94 (0.89-0.97)	0.91 (0.82-0.96)

Large Dependence Low Gray Level Emphasis (LDLGLE)	0.94 (0.88-0.97)	0.94 (0.89-0.97)	0.74 (0.46-0.87)	0.92 (0.84-0.96)
Large Dependence High Gray Level Emphasis (LDHGLE)	0.95 (0.89-0.97)	0.97 (0.94-0.98)	0.88 (0.74-0.94)	0.90 (0.80-0.95)
Neighbouring Gray Tone Difference Matrix (NGTDM) features				
Coarseness	0.99 (0.99-0.99)	0.98 (0.97-0.99)	0.99 (0.99-0.99)	0.99 (0.98-0.99)
Contrast	0.98 (0.97-0.99)	0.97 (0.93-0.98)	0.97 (0.95-0.98)	0.98 (0.97-0.99)
Busyness	0.95 (0.90-0.97)	0.97 (0.95-0.98)	0.90 (0.80-0.95)	0.90 (0.79-0.95)
Complexity	0.99 (0.98-0.99)	0.97 (0.95-0.98)	0.97 (0.95-0.99)	0.98 (0.96-0.99)
Strength	0.99 (0.99-0.99)	0.99 (0.98-0.99)	0.99 (0.98-0.99)	0.99 (0.98-0.99)

[0087] 골절 위험 예측을 위해 가장 성능이 좋은 특성 변수를 선택하기 위한 과정은, 1) 추출한 특성 변수들에 대해 스케일링하고, 2) 이를 정규화(normalization)한 후, 3) 정규화된 특성 변수들 중 과도하게 상관된 특성 변수들(예를 들어, correlation coefficient>0.80)은 제거하고, 4) 단변량분석에서 FDR (False Discovery Rate)기반 다중분석보정 이후 유의하게 연관성이 있는 (보정 p value < 0.20) 특성변수를 선택하고, 5) 선택된 변수들 중에서 교차검증 기반 재귀변수제거법 (recursive feature elimination with cross validation) 을 통해 F1 score 를 최대화하는 최소 중요 변수 조합을 데이터에 기반하여 선택할 수 있다.

[0088] 상술한 특성 변수를 선택하기 위한 1) 내지 5) 단계는 변수를 선택하는 과정으로서, 특정 사람(또는 사용자)의 주관적인 의견으로 변수를 선택하지 않고, 수학 알고리즘 모델들을 통하여 중요한 변수들을 선택할 수 있도록 하는 선택 과정이다. 위 1) 내지 5) 단계를 통해 상관계수가 높은(즉 correlation coefficient 가 높은) 변수를 제거하고 교차검증 기반 재귀변수제거법 (recursive feature elimination with cross validation) 을 통해 변수들 중에서 F1 score 를 최대화하는 최소 중요 변수 조합을 데이터에 기반하여 선택할 수 있다.

[0089] 이러한 선택 과정을 통해서 소정 개수의 최적의 특성 변수들이 선택된다. 예를 들어, S130 단계에서 75개의 특성 변수들이 추출되는 경우를 가정하는 경우, 앞서의 과정을 거쳐 표 3의 특성변수들 중 표시(*)된 14개의 특성 변수들이 추출되었다.

[0090] 65세 이상의 고관절 골절을 겪은 여성들에 대한 골절 위험 예측을 위해서는 14개의 특성 변수들이 선택될 수 있다. 선택된 14개의 특성 변수들에 대해서는 도 5를 참조할 수 있다.

[0091] 도 5를 참조하면, 세로축은 선택된 특성 변수들을 나타내고, 가로축은 특성 변수간 상대적 특성 중요도를 나타낸다. 도 5에서 GLRLM Run Entropy (TH)의 중요도가 0.40으로 가장 높고, GLSZM GLV (FN)은 0.02~0.03 정도의 중요도를 갖는다.

[0092] 여기서 선택된 기질특성은, 4개의 관심영역, 대퇴 경부(Femoral Nect; FN), 전자(Trochanter; TR), 전자 간부(Intertochanter; IT) 및 전체 고관절(Total Hip; TH)을 기준으로, 바람직하게는 전체 고관절(TH)의 GLRLM(Gray Level Run Length Matrix) Run Entropy, 대퇴 경부(FN)의 GLRLM(Gray Level Run Length Matrix) Run Entropy, 대퇴 경부(FN)의 GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), 전체 고관절(TH)의 GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), 대퇴 경부(FN)의 GLCM IDMN (Inverse Difference Moment Normalized), 전체 고관절(TH)의 GLCM Cluster Prominence, 전체 고관절(TH)의 GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis), 전자 간부(IT)의 GLCM IDMN(inverse difference moment normalized), 대퇴 경부(FN)의 GLCM Cluster Prominence, 전체 고관절(TH)의 GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized), 전체 고관절(TH)의 GLRLM GLNN (Gray Level Non-uniformity Normalized), 전자(TR)의 GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity), 전체 고관절(TH)의 GLSZM SZN(Size-Zone Non-Uniformity), 대퇴 경부(FN)의 GLSZM GLV(Gray Level Variance)의 14개 특성 변수일 수 있다.

[0093] 위 14개의 특성 변수들은 육안으로 검정하기 어려운 텍스쳐 패턴을 수량화하기 위해 채용된 것이다. 나타난 패턴을 요약하면, 골절 고위험군은, 동일한 골밀도에서도 픽셀 간 신호 강도의 차이가 감소(즉, 흑백의 차이가 줄

어듬)하는 반면, 이러한 흑백 픽셀이 배열되어 있는 패턴은 더욱 불규칙하다.

[0094] 한편, 골절 위험도가 낮은 사람들은, 동일한 골밀도에서도 픽셀 간 신호 강도 차이가 큰 반면 각 픽셀들 간 패턴 배열을 규칙적이다.

[0096] 이때, GLRLM(Gray Level Run Length Matrix)은 픽셀의 개수 단위의 길이로 정의되는 그레이 레벨 런(gray level run)을 정량화한다. 또한, GLRLM Run entropy는 그레이 레벨 및 런 길이의 분포에 대한 불확실성 및 임의성을 측정한다. GLRLM GLNN (Gray Level Non-uniformity Normalized)는 이미지 내에서 그레이 레벨 강도 값의 유사성을 측정하기 위해 사용된다.

[0097] 또한, GLSZM(Gray Level Size Zone Matrix)은 DXA 이미지에서 그레이 존 영역을 정량화할 수 있다. GLSZM GLV(Gray level Variance) 특성은 해당 존들에 대한 그레이 레벨 조명도를 측정하고, GLSZM GLN(Gray Level Non-Uniformity)는 그레이 레벨 강도값의 변동성을 측정하며, GLSZM SZN(Size-Zone Non-Uniformity)는 이미지 내에서 크기 영역의 변동성을 측정한다.

[0098] GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis)는 더 높은 그레이 레벨 값을 가진 더 작은 크기 영역의 조인트 분포 이미지에서의 비율을 측정한 값을 나타낸다.

[0099] GLCM (Grey Level Co-occurrence Matrix) IMC1 (informal measure of correlation 1)은 이미지의 텍스처의 특성을 나타내고, GLCM IDMN (inverse difference moment normalized)은 이미지의 동질성을 측정한다. GLCM Cluster Prominence는 GLCM의 왜곡 및 비대칭성에 대한 측정 값을 나타낸다.

[0100] GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized)은 이미지 전반에 종속성의 유사성을 측정하는 값으로, 낮을수록 종속성 간에 더 많은 동질성을 나타낸다.

[0101] 상술한 특성 변수들 및 나머지 특성 변수들에 대해서는 하기 표 2에 기술된 설명을 참조할 수 있다.

[0103] 다시 도 1을 참조하면, S110 단계에서 설계한 발달 집합을 이용하여 머신 러닝 모델 별 성능을 테스트함으로써 고관절 골절 위험을 예측하기 위한 라디오믹스 스코어 모델에 사용될 머신 러닝 모델을 선택할 수 있다 (S140).

[0104] 발달 집합의 트레이닝 세트에서 사용될 수 있는 머신 러닝 모델들은 대략 4 종류가 있다. 예를 들어, 랜덤 포레스트(Random Forest) 모델, 엘라스틱 넷(Elastic net) 모델, 그라디언트 부스팅 결정 트리(Gradient Boosting Decision Tree) 모델 및 서포트 벡터 머신(Support Vector Machine) 모델이 있다. 각 모델들에 대한 하이퍼파라미터(hyperparameter)는 3 폴드 교차 확인 (3-fold cross validation)을 5회 반복한 격자 탐색법(grid search method)을 이용하여 조율된다.

[0105] 발달 집합의 테스트 세트에서는 상술한 머신 러닝 모델들이 AUROC (Area Under the Receiver-Operating characteristics Curve) 및 AUPRC (Area Under the Precision-Recall Curve) 등의 지표를 이용하여 평가된다. 머신 러닝 모델 칼리브레이션(calibration)은 신뢰성(예측된 확률과 실제 확률의 차이)과 결단성(예측된 평균과의 조건적 확률의 차이)을 분리하는 Brier Score를 이용하여 평가된다.

[0106] S140 단계에서, 발달 집합에서 DXA 이미지에 대한 각 머신 러닝 모델에 대한 성능 평가 및 계측 결과는 다음 표 4와 같다. 다음 표 4는 고관절 DXA 이미지로부터 특성 변수를 이용한 고관절 골절을 예측하기 위한 머신 러닝 모델들의 성능 및 계측 매트릭스를 나타낸다.

표 4

[0108]

	Accuracy	Precision	Recall	F1-score	AUROC	AUPRC	Brier score
Test set							
Random forest	0.73	0.58	0.72	0.64	0.784	0.664	0.175
GBDT	0.72	0.61	0.47	0.53	0.768	0.638	0.180
SVC	0.74	0.59	0.72	0.65	0.759	0.607	0.182
Elastic net	0.73	0.56	0.86	0.62	0.758	0.607	0.182

[0109] 표 4를 참조하면 네 개의 DXA 고관절 라디오믹스 모델들의 Brier score는 0.175 내지 0.182 범위에 있고, 위 모델들은 유사한 정확도를 나타내고 있다. 다만, AUROC 및 AUPRC의 값들은 랜덤 포레스트 모델이 가장 성능이 좋은 것을 알 수 있다.

[0110] 따라서, 성능이 가장 좋은 것으로 판단된 랜덤 포레스트 모델이 본 라디오믹스 스코어를 측정하기 위해 선택된다. 즉, S130 단계에서 선택된 14개의 특성 변수들은 랜덤 포레스트 모델을 훈련시키기 위해 사용된다.

[0111] 14개의 특성 변수들 중에서, 전체 고관절(TH)의 GLRLM(Gray Level Run Length Matrix) Run Entropy, 대퇴 경부(FN)의 GLRLM(Gray Level Run Length Matrix) Run Entropy, 대퇴 경부(FN)의 GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), 전체 고관절(TH)의 GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), 대퇴 경부(FN)의 GLCM IDMN (Inverse Difference Moment Normalized), 전체 고관절(TH)의 GLCM Cluster Prominence의 6개 특성 변수의 경우, 도 5에 나타나듯이, 랜덤 포레스트 모델에 기여하는 높은 중요도를 가진 6개의 특성 변수로 취급될 수 있다. 도 6을 참고하면, 고관절 골절을 겪은 사람들은, 고관절 골절을 겪지 않은 사람들에 비해, 발달 집합과 검증 집합 모두에서 더 높은 GLRLM(Gray Level Run Length Matrix) Run Entropy, 더 낮은 GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1)을 보이는데, 둘 모두 기질 패턴들에서 더 큰 이질성(heterogeneity)을 나타낸다. 또한, 고관절 골절을 겪은 사람들은, 그렇지 않은 사람들에 비해, 발달 집합과 검증 집합 모두에서 더 높은 GLCM IDMN (Inverse Difference Moment Normalized)과 더 낮은 GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity)을 보이며, 둘 모두 그레이-레벨 픽셀 강도 값에서 더 적은 편차를 나타낸다.

[0112] S110 단계에서 설계한 검증 집합에서 결정된 랜덤 포레스트 모델을 평가할 수 있다 (S150).

[0113] 검증 집합에서 S140 단계에서 선택한 모델의 성능이 검증되면, 골절 위험 예측을 위해 트레이닝을 받고 실제 예측에 사용된다 (S160).

[0114] 상술한 S140 단계 및 S150 단계는 수행되지 않을 수 있다. 예를 들어, 여러 개의 머신 러닝 모델들을 사용할 수 있는 경우에는 추출된 특성 변수에 기반하여 머신 러닝 모델들을 선택할 수 있지만, 하나의 머신 러닝 모델만이 고려되는 경우에는 해당 단계들은 수행되지 않을 수 있다. 하지만 이러한 경우에도 해당 머신 러닝 모델이 검증 집합에서 적합한지 여부에 대한 성능 평가가 이뤄지는 것이 바람직하다.

[0115] 도 7은 도 1 내지 도 6에서 설명한 골절 위험 예측 방법이 수행될 수 있는 장치 및 컴퓨터 기록 매체를 설명하기 위한 도면이다.

[0116] 도 7(A)는 골절 위험을 예측하는 방법이 수행되는 장치 구조의 일례를 나타내고, 도 7(B)는 골절 위험을 예측하는 프로그램이 기록된 기록 매체의 일례를 나타낸다.

[0117] 도 7(A)에 도시된 장치는 유무선 통신을 위한 송수신기(510), 장치를 제어하기 위한 프로세서(520) 및 메모리(530)를 포함할 수 있다. 송수신기(510)는 유선 및/또는 무선을 통해 외부 또는 내부 데이터 베이스들과 통신을 수행할 수 있다. 해당 통신을 통해 송수신기(510)는 DXA 이미지 데이터를 데이터 베이스들로부터 수신 및 획득할 수 있다. 프로세서는 송수신기(520) 및 메모리(530)를 제어하며, 메모리(530)에 저장된 명령어들을 이용하여 송수신기(520)이 외부 또는 내부 데이터 베이스들과 통신을 수행하도록 제어할 수 있다. 또한, 프로세서(520)는 송수신기이 수신한 DXA 이미지 파일을 디코딩할 수 있으며, 메모리에 저장된 명령어들을 제어하여 도 1 내지 도 6에서 설명한 골절 위험 예측 방법이 수행할 수 있다.

[0118] 도 7(A)에서 송수신기(510)는 환자군 데이터 베이스(201) 및 대조군 데이터 베이스(203)과 유선 또는 무선 통신을 수행할 수 있다. 환자군 데이터 베이스(201) 및 대조군 데이터 베이스(203)는 장치의 외부에 구성될 수 있으나, 장치의 내부에 구성되어 있을 수 있다.

[0119] 도 7(B)에서 골절 위험을 예측하기 위한 컴퓨터 프로그램 코드(555)를 포함하는 컴퓨터로 읽을 수 있는 매체(550)는, 컴퓨터 프로그램 코드가 컴퓨터에서 수행되면 컴퓨터가 발달 집합을 설계하는 단계, 발달 집합에 포함된 복수의 대상들에 대한 뼈 이미지들을 처리하는 단계, 뼈 이미지들로부터 특성 변수들을 추출하는 단계, 추출된 특성 변수들 중 골절 위험을 예측하기 위해 필요한 최적의 특성 변수들을 선택하는 단계, 최적의 특성 변수들을 발달 집합의 훈련 집합을 이용하여 머신러닝을 수행하는 단계, 골절 위험을 예측하기 위한 본 라디오믹스 스코어 모델을 설계하는 단계를 수행하도록 구성된다.

[0120] 도 7(B)에서 컴퓨터로 읽을 수 있는 매체(550)는 컴팩트 디스트(CD: Compact Disk) 모양으로 도시를 하였지만,

이에 국한되지 않고 USB, DVD, 플레쉬 메모리, 램, 롬 등의 형태일 수 있다. 또한, 컴퓨터에서 온라인으로 프로그램 코드(555)를 다운받아 저장할 수 있는 모든 저장매체를 의미할 수 있다. 또한, 상기 컴퓨터 프로그램 코드(555)는 스마트폰 등 휴대 단말 기기 또는 고정 단말 기기에서 어플리케이션 형태로 다운받는 형태를 포함할 수 있다.

[0121] 머신러닝은 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델 중 하나를 이용해 수행될 수 있다.

[0122] 상기 머신러닝을 수행하는 단계에서, 최적의 특성 변수들을 랜덤 포레스트 모델, 엘라스틱 넷 모델, 그라디언트 부스팅 결정 트리 모델 및 서포트 벡터 머신 모델에 각각 적용하여 머신 러닝을 수행하고, 그 결과를 검증 집합에 적용하여 하나의 머신 러닝 모델을 선택할 수 있다. 이때, 본 라디오믹스 스코어 모델의 성능을 평가하기 위해 훈련 집합의 일부분을 성능 평가 데이터로 사용할 수 있다.

[0123] 이때, 발달 집합은 환자군 데이터 베이스와 대조군 데이터 베이스로부터 환자군 및 대조군이 1:A로 매칭된 발달 집합을 설계할 수 있다. 이때, A는 0이 아닌 자연수 중 선택될 수 있다.

[0124] 매칭된 발달 집합으로부터 사용이 불가능한 뼈 이미지를 제거하여 최종 발달 집합을 설계할 수 있다.

[0125] 예를 들어, 최적의 특성 변수는 GLRLM(Gray Level Run Length Matrix) Run Entropy, GLRLM GLNN (Gray Level Non-uniformity Normalized), GLCM (Grey Level Co-occurrence Matrix) IMC1 (Information Measure of Correlation 1), GLCM IDMN (Inverse Difference Moment Normalized), GLCM Cluster Prominence, GLSZM SZN(Size-Zone Non-Uniformity) , GLSZM (Gray Level Size Zone Matrix) SAHGLE (Small Area High Gray Level Emphasis), GLSZM (Gray Level Size Zone Matrix) GLN (Gray Level Non-Uniformity), GLSZM GLV(Gray Level Variance), GLDM (Gray Level Dependence Matrix) DNN(Dependence Non-Uniformity Normalized) 중 하나 이상을 포함할 수 있으며, 각각의 기질특성은 복수의 관심영역(ROI) 별로 추출될 수 있다. 특성 변수에 대한 자세한 설명은 표 2를 참조할 수 있다. 또한, 바람직하게는 앞서 살핀 것처럼, 랜덤 포레스트 모델에 따라 선정된 14개의 특성 변수가 적용될 수 있다.

[0126] 뼈 이미지는 이중 에너지 X선 흡수 계측법(DXA) 이미지일 수 있다. 본 발명의 실시예들 및 컴퓨터 프로그램 코드는 검사 대상의 DXA 이미지를 본 라디오믹스 스코어 모델을 이용하여 골절 위험이 있는지 여부 또는 골절 위험의 정도를 예측하는 단계를 더 수행하도록 구성될 수 있다.

[0128] **3. 실시예**

[0129] 이하에서는 도 1 내지 도 7에서 설명한 본 발명의 양태들에 대한 실시예들에 대해서 설명한다.

[0130] 도 2에서 발달 집합의 환자군 데이터 베이스는 세브란스 병원의 EHR 데이터 베이스이다. 이때, 환자군은 65세 이상의 여성이며 관찰기간인 2010년 1월 부터 2019년 12월까지 취약한 고관절 골절을 겪은 사람들이다. 환자군 표본 개수(N)은 172이다.

[0131] 발달 집합의 대조군 데이터 베이스는 KURE 코호트 데이터 베이스이며, 대조군은 관찰기간 2012년 1월부터 2018년 12월까지의 기간에서 65세 이상의 여성들 중 취약한 고관절 골절을 겪지 않은 사람들로 구성된다.

[0132] 환자군과 대조군은 1:2의 비율로 매칭되며, 추적 조사 기간은 0.9 년에서 3.9년으로 중간값은 2.1년이다. 매칭된 발달집합은 441명의 표본을 포함한다. 이때, 고관절 골절을 겪은 표본의 수는 147이며 대조군의 표본은 294이다.

[0133] 그 중에서 DXA 고관절 이미지가 없거나 DXA 머신의 종류가 달라 DXA 영상 이미지 판독이 불가한 표본은 환자군에서 4명, 대조군에서 4명씩 제외되었다.

[0134] 따라서 최종 발달 집합은 총 표본 433으로 환자군 143명 대조군 290명이다.

[0135] 총 표본에서 DXA 이미지에 대한 머신 러닝을 수행하기 위한 트레이닝 집합은 무작위로 75%를 선정하고, 나머지 25%는 성능을 평가하기 위한 테스트 집합으로 구성하였다.

[0136] 도 2에서 검증 집합은 머신 러닝을 통해 학습한 DXA 기반의 본 라디오믹스 스코어 모델이 성능을 평가하기 위한 것이다. KURE 코호트 데이터 베이스에서 2012년부터 2015년 동안 등록된 자들 중 2018년 12월 31일까지 추적조사를 한 인원 중에서 표본 집단을 추출했다. 즉, 총 표본은 3517명이다.

[0137] 이때, 65세 이상의 여성이 조사 대상이므로 남성 표본 1163명은 표본에서 제외하고, 매칭된 대조군에 포함된 여성 294명 또한 제외하였다. 마지막으로 DXA 고관절 이미지를 사용할 수 없는 인원 7명을 제외하였다.

[0138] 따라서, 추적기간 4.5년 내지 5.7년 동안 (중간값 5.4년) 대조군으로 구성될 수 있는 65세 이상 여성은 2053명이다.

[0139] 사례 코호트(case cohort)는 대조군 중 무작위 10%를 서브 샘플링하여 설정할 수 있다. 이때, 참가자 중 고관절 골절을 갖지 않는 186명과 고관절 골절을 갖는 3명으로 10%의 서브코호트를 구성하고, 구성한 서브코호트 외에서 고관절 골절을 겪는 31명을 사례 코호트로 구성하였다. 따라서 총 220명의 표본을 갖는 사례 코호트가 설계되었다.

[0140] 상기 표 1은 상술한 실시예에서 구성한 발달 집합 및 검증 집합을 기반으로 나온 결과값이다.

[0141] 상기 표 4는 상술한 실시예에서 구성한 발달 집합을 토대로 본 발명의 골절 위험 예측 방법을 위한 본 라디오믹스 스코어 모델을 구성하기 적합한 머신 러닝 모델들을 훈련시켜 테스트한 결과값이다.

[0142] 해당 결과를 도출하기 위해 표 2에 적시한 75개의 특성 변수들을 이용하였고, 그중 최적의 14개의 특성 변수들을 기반으로 머신 러닝 모델들을 각각 훈련 시켰다.

[0144] **3.1 결과**

[0145] 이하 결과는 표 1을 참조하여 설명한다.

[0146] 발달 집합에서 측정 대상의 나이 중간값은 73세(환자군 73±6 vs 대조군 73±6, p=0.807)이고, 고관절 골절 또는 DXA 테스팅 시간부터 관찰 종료 시간은 2.1년이다. 환자 집단 즉 사례 그룹(표본수 143)은 대조군(n=290)과 비교시 심각히 낮은 대퇴부 경부 T 스코어(-2.8 ± 1.0 vs. -2.2 ± 1.0, p<0.001) 및 더 취약한 골절 이력(62% vs 33%, p<0.001)을 갖는다.

[0147] 검증 집합 즉 검증 코호트(사례 코호트)에서, 후속 조치 기간 동안 고관절 골절을 겪은 대상 34명은 어떤 고관절 골절을 겪지 않은 집단(n=186)과 비교시 더 많은 나이를 가지며(75 ± 5 vs. 71 ± 4 year, p<0.001), 이전 골절의 더 높은 유병률(73 % vs 31 %, p<0.001), DXA 정의된 골다공증(77% vs. 44%, p=0.001), 및 낮은 대퇴부 경부 T 스코어(-2.8 ± 0.9 vs. -2.0 ± 0.8, p<0.001)를 갖는다.

[0149] **임상적 파라미터와 본 라디오믹스 스코어의 상관관계**

[0150] 검증 집합 및 검증 코호트에서, 본 라디오믹스 스코어는 12 내지 74의 범위를 가지며, 중간 값은 29이고 표준편차는 13이다. 본 라디오믹스 스코어는 나이(r=0.13, p=0.049), 키(r=0.18, p=0.009), 대퇴부 경부 BMD(r=-0.31, p<0.001), 총 고관절 BMD(r=-0.30, p<0.001), 및 FRAX 고관절 골절 확률들(r=0.21, p=0.002)과 약한 보통의 상관관계를 갖는다. 이때, 체중(r=-0.07, p=0.330) 및 요추골 BMD(r=0.029, p=0.671)는 본 라디오믹스 스코어 모델과 상관관계가 없다. 이전에 취약한 골절 이력(32 ± 15 vs. 27 ± 12, p=0.012)을 갖거나 DXA 기반의 골다공증(32 ± 15 vs. 26 ± 10, p<0.001)을 갖는 개개인에게는 이러한 이력이 없는 사람과 비교시 높은 본 라디오믹스 스코어가 관찰된다.

[0152] **본 라디오믹스 스코어 모델을 이용한 고관절 골절 예측 결과**

[0153] 추적 조사 기간 동안 고관절 골절 사고의 위험은 본 라디오믹스 스코어의 십분위수로 계단식으로 증가한다. 도 8은 본 라디오믹스 스코어 모델을 이용하여 고관절 골절 예측 결과를 설명하기 위한 도면이다.

[0154] 본 라디오믹스 스코어 모델은 10 단위로 표시될 수 있다. 도 8을 참조하면 세로축은 고관절 골절의 위험도를 나타내고, 가로축은 본 라디오믹스 스코어를 나타낸다. 도 8을 참조하면, 본 라디오믹스 스코어가 높아질수록 고관절의 부상률(골절 위험)이 높아짐을 알 수 있다.

[0155] 도 9는 고관절 골절 사고에 대한 카플란 마이어 생존 도표를 나타낸다.

[0156] 도 9(A)는 DXA 기반의 골다공증 도표를 나타내고, 도 9(B)는 본 라디오믹스 스코어를 나타내며, 도 9(C)는 검증 코호트 내에서 DXA 기반의 골다공증 및 본 라디오믹스 스코어 조합을 나타낸다. 도 9에서 세로축은 고관절 골절

없는 생존 확률을 나타내며, 가로축은 일 기준의 시간을 나타낸다.

[0157] 도 9에서 DXA 골다공증 또는 높은 본 라디오믹스 스코어(32 이상)가 있는 개인은 그렇지 않은 경우에 비해 누적 고관절 골절 없는 대상에 비해 생존 확률이 낮다 (도 9: DXA 골다공증 vs 없는 경우: log-rank p = 0.001, 위험비율 [HR] 4.09, 95% CI 1.76-9.47, p = 0.001; 높은 본 라디오믹스 스코어 vs 없는 경우: log-rank p <0.001, HR 4.51, 95 % CI 2.10-9.70, p <0.001). DXA 골다공증이 없고 본 라디오믹스 스코어가 낮은 사람에 비해 DXA 골다공증이 없지만 본 라디오믹스 점수가 높은 사람은 고관절 골절 위험이 약 5 배 높다(HR 5.07 [95% CI 1.14-22.55], p=0.033). 이는 DXA 골다공증 및 낮은 본 라디오믹스 스코어(HR 4.31, 95% CI 1.12-16.62, p=0.034)를 가진 피험자의 위험과 비슷하다. DXA 골다공증과 높은 본 라디오믹스 스코어를 가진 개체는 FRAX 고관절 골절 확률(조정된 HR 8.25, 95% CI 2.41-28.23, p=0.001, 표 5 참조)에 독립적인 모든 그룹(log-rank P <0.001)에서 고관절 골절 위험이 가장 높다(HR 15.73, 95% CI 4.33-57.15, p<0.001).

표 5

[0159]

	단변량 모델		다변량 모델	
고관절 골절 사고 예측	조정되지 않은 HR (95% CI)	P 값	조정된 HR (95% CI)	P 값
DXA 골다공증 / 본 라디오믹스 스코어				
No / Low	1.00 (reference)		1.00 (reference)	
No / High	5.07 (1.14-22.55)	0.033	4.49 (1.03-19.50)	0.045
Yes / Low	4.31 (1.12-16.62)	0.034	2.52 (0.71-9.02)	0.155
Yes / High	15.73 (4.33-57.15)	<0.001	8.25 (2.41-28.23)	0.001
FRAX 고관절 골절 확률 (>3%)	7.25 (2.14-24.50)	0.001	3.68 (1.06-12.74)	0.040

표 6

[0161]

	단변량 모델		다변량 모델 1		다변량 모델 2	
고관절 골절 사고 예측	조정되지 않은 HR (95% CI)	P 값	조정된 HR (95% CI)	P 값	조정된 HR (95% CI)	P 값
나이 (5세 증가시마다)	2.07 (1.29-3.32)	0.003	1.36 (0.80-2.28)	0.252	1.14 (0.64-2.04)	0.655
골절 과거력 (예 또는 아니오)	5.81 (2.57-13.12)	<0.001	3.79 (1.58-9.04)	0.003	3.15 (1.21-8.17)	0.018
대퇴경부 T점수 (0.1 T점수 감소 당)	1.13 (1.07-1.20)	<0.001	1.09 (1.03-1.15)	0.002	1.08 (1.01-1.15)	0.020
본 라디오믹스 스코어 (1포인트 증가 당)	1.07 (1.04-1.11)	<0.001	-	-	1.04 (1.00-1.09)	0.043
해럴의 C지수 (95% CI)	본 라디오믹스 스코어 단독: 0.73 (0.63-0.83)		모델 1: 0.80 (0.72-0.88)		모델 2: 0.85 (0.78-0.92)*	

[0163] 표 5는 DXA 골다공증 및 FRAX 스코어와 독립적인 고관절 골절 예측에 대한 본 라디오믹스 스코어 값을 분석한 것이다.

[0164] 본 라디오믹스 스코어의 1점 증가는 고관절 골절 위험이 7% 증가한 것과 관련이 있었으며(HR 1.07, 95 % CI 1.04-1.11, p <0.001) 공변량에 대한 조정 후에도 견고하게 유지된다 (조정 된 HR 1.04, 95 % CI 1.00). -1.09, p = 0.043; 표 6 참조). 본 라디오믹스 스코어만으로는 고관절 골절 (Harrel 's C index 0.73, 95 % CI 0.63 ~ 0.83)에 대해 약간의 차별성을 보였으며, 이는 연령, 이전 취약성 골절 이력 및 대퇴 경부 T 점수 요소가 더해졌을 때 상당히 개선될 수 있다. 표 6에서 다변량 모델1은 골절 예측 인자로서 나이, 골절 과거력, 대퇴경부 T 점수가 고려된 결과값을 나타낸다. 즉, 다변량 모델1은 본 라디오믹스 스코어가 적용되지 않은 경우의

골절 위험 예측 결과 값을 나타내며, 다변량 모델2는 본 라디오믹스 스코어가 다변량 모델1에 적용된 경우의 결과값을 나타낸다. 따라서, 본 라디오믹스 스코어가 다변량 모델1에 추가될 때 고관절 골절 위험 예측에 대한 C 지수 (판별력, 1.00에 가까울수록 좋은 모델) 가 0.80에서 0.85(p=0.040)로 개선되는 것을 나타낸다 (다변량 모델 2 참조).

[0165] 도 10은 본 발명에 따른 본 라디오믹스 스코어가 높은 그룹과 낮은 그룹 들 간에 있어서 대퇴골절을 포함한 주요 골다공증성 골절 위험도(척추, 손목, 위팔, 대퇴골절)를 대비한 것이고, 아래 표 7은 골다공증성 골절 발생과 관련하여 단일 변수 모델과 다중 변수 모델들 간의 차이를 비교한 표이다.

[0166] 도 10에 도시된 것처럼, 본 라디오믹스 스코어가 높은 군에서는 척추, 손목, 위팔, 대퇴골절과 같은 주요 골다공증성 골절 위험도가 34%로, 본 라디오믹스 스코어가 낮은 군의 17%에 비해 유의하게 높았다.

[0167] 또한, 표 7에서와 같이, 본 라디오믹스 스코어가 1점 높아질수록, 골다공증성 골절 (척추, 손목, 상완, 대퇴골절) 위험도가 4% 증가함을 확인할 수 있다. 본라디오믹스 스코어 증가는 나이, 골다공증 여부, 기존 골절여부와 독립적인 예측인자였다(adjusted odds ratio 1.03, 95% CI 1.01-1.06, p=0.019).

표 7

[0169]

	Univariate (단변량)			Multivariable (다변량)		
	Odds ratio	95% CI	P value	Odds ratio	95% CI	P value
본 라디오믹스 스코어 (1점 증가시)	1.04	1.02-1.07	<0.001	1.03	1.01-1.06	0.019
나이	1.10	1.06-1.13	<0.001	1.11	1.02-1.20	
골다공증	2.21	1.66-2.94	<0.001	1.66	0.78-3.51	0.182
골절 이력	2.59	1.96-3.43	<0.001	3.85	1.89-7.85	<0.001

표 2

[0171]

Based methods	Parameter	Formula	Description
Grey Level Co-occurrence Matrix (GLCM) : A Gray Level Co-occurrence Matrix (GLCM) of size Ng×Ng describes the second-order joint probability function of an image region constrained by the mask and is defined as P(i,j|δ,Θ). The (i,j)th element of this matrix represents the number of times the combination of levels i and j occur in two pixels in the image, that are separated by a distance of δ pixels along angle Θ.			
GLCM Features	Autocorrelation	$autocorrelation = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)ij$	Autocorrelation is a measure of the magnitude of the fineness and coarseness of texture.
	Joint Average	$joint\ average = \mu_x = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)i$	The mean gray level intensity of the I distribution.
	Cluster Prominence	$cluster\ prominence = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^4 p(i,j)$	Cluster Prominence is a measure of the skewness and asymmetry of the GLCM. A higher values implies more asymmetry about the mean while a lower value indicates a peak near the mean value and less variation about the mean.
	Cluster Shade	$cluster\ shade = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^3 p(i,j)$	Cluster Shade is a measure of the skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.

	Cluster Tendency	$cluster\ tendency = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j-\mu_x-\mu_y)^2 p(i,j)$	Cluster Tendency is a measure of groupings of voxels with similar gray-level values.
	Contrast	$cluster\ tendency = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i+j)^2 p(i,j)$	Contrast is a measure of the local intensity variation, favoring values away from the diagonal. A larger value correlates with a greater disparity in intensity values among neighboring voxels.
	Correlation	$correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)ij - \mu_x\mu_y}{\sigma_x(i)\sigma_y(j)}$	Correlation is a value between 0 (uncorrelated) and 1 (perfectly correlated) showing the linear dependency of gray level values to their respective voxels in the GLCM.
	Difference Average	$difference\ average = \sum_{k=0}^{N_g-1} k p_{x-y}(k)$	Difference Average measures the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values.
	Difference Entropy	$difference\ entropy = \sum_{k=0}^{N_g-1} p_{x-y}(k)\log_2(p_{x-y}(k)+\epsilon)$	Difference Entropy is a measure of the randomness/variability in neighborhood intensity value differences.
	Difference Variance	$difference\ variance = \sum_{k=0}^{N_g-1}(k-DA)^2 p_{x-y}(k)$	Difference Variance is a measure of heterogeneity that places higher weights on differing intensity level pairs that deviate more from the mean.
	Joint Energy	$joint\ entropy = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(p(i,j))^2$	Energy is a measure of homogeneous patterns in the image. A greater Energy implies that there are more instances of intensity value pairs in the image that neighbor each other at higher frequencies.
	Joint Entropy	$joint\ entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log_2(p(i,j)+\epsilon)$	Joint entropy is a measure of the randomness/variability in neighborhood intensity values.

	Informational Measure of Correlation (IMC) 1	$IMC1 = \frac{HXY - HXY1}{\max\{HX, HY\}}$	This class of features characterizes the textures of an image / object by creating a new matrix GLCM based on counting how often pairs of pixels with specific gray-level values and in a specified spatial relationship (distance and direction) occur in the image / object and then computing statistics from GLCM.IMC1 assesses the correlation between each and every probability distribution (quantifying the complexity of the texture). It represents the information of a single distribution.
	Informational Measure of Correlation (IMC) 2	$IMC2 = \sqrt{1 - e^{-2(HXY2 - HXY)}}$	IMC2 also assesses the correlation between each and every probability distribution (quantifying the complexity of the texture). It represents the information of two of the distributions.
	Inverse Difference Moment (IDM)	$IDM = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + k^2}$	IDM is a measure of the local homogeneity of an image. IDM weights are the inverse of the Contrast weights.
	Maximal Correlation Coefficient (MCC)	$MCC = \sqrt{second\ largest\ eigenvalue\ of\ Q}$ $Q(i,j) = \sum_{k=0}^{N_g} \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$	The Maximal Correlation Coefficient is a measure of complexity of the texture and 0≤MCC≤1.
	Inverse Difference Moment Normalized (IDMN)	$IDMN = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + \left(\frac{k^2}{N_g^2}\right)}$	IDMN (inverse difference moment normalized) is a measure of the local homogeneity of an image. IDMN weights are the inverse of the Contrast weights.
	Inverse Difference (ID)	$ID = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + k}$	ID is another measure of the local homogeneity of an image. With more uniform gray levels, the denominator will remain low, resulting in a higher overall value.
	Inverse Difference Normalized (IDN)	$IDN = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{1 + \left(\frac{k}{N_g}\right)}$	IDN is another measure of the local homogeneity of an image
	Inverse Variance	$\in verse\ \ variancc = \sum_{k=0}^{N_g - 1} \frac{p_{x-y}(k)}{k^2}$	Inverse of the variance calculated is taken

	Maximum Probability	$\max imum\ probability = \max\ (p(i,j))$	Maximum Probability is occurrences of the most predominant pair of neighboring intensity values.
	Sum Average	$\sum average = \sum_{k=2}^{2N_g} p_{x-y}(k)k$	Sum Average measures the relationship between occurrences of pairs with lower intensity values and occurrences of pairs with higher intensity values.
	Sum Entropy	$\sum entropy = \sum_{k=2}^{2N_g} p_{x-y}(k)\log_2(p_{x+y}(k)+\epsilon)$	Sum Entropy is a sum of neighborhood intensity value differences.
	Sum of Squares	$\sum squares = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu_x)^2 p(i,j)$	Sum of Squares or Variance is a measure in the distribution of neigboring intensity level pairs about the mean intensity level in the GLCM.
Gray Level Size Zone Matrix (GLSZM): A Gray Level Size Zone (GLSZM) quantifies gray level zones in an image. A gray level zone is defined as a the number of connected voxels that share the same gray level intensity. A voxel is considered connected if the distance is 1 according to the infinity norm (26-connected region in a 3D, 8-connected region in 2D). In a gray level size zone matrix P(i,j) the (i,j)th element equals the number of zones with gray level i and size j appear in image.			
GLSZM features	Small Area Emphasis (SAE)	$SAE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z}\frac{P(i,j)}{j^2}}{N_z}$	SAE is a measure of the distribution of small size zones, with a greater value indicative of more smaller size zones and more fine textures.
	Large Area Emphasis (LAE)	$LAE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z}P(i,j)j^2}{N_z}$	LAE is a measure of the distribution of large area size zones, with a greater value indicative of more larger size zones and more coarse textures.
	Gray Level Non-Uniformity (GLN)	$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_z}P(i,j)\right)^2}{N_z}$	GLN measures the variability of gray-level intensity values in the image, with a lower value indicating more homogeneity in intensity values.
	Gray Level Non-Uniformity Normalized (GLNN)	$GLNN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_z}P(i,j)\right)^2}{N_z^2}$	GLNN measures the variability of gray-level intensity values in the image, with a lower value indicating a greater similarity in intensity values. This is the normalized version of the GLN formula.
	Size-Zone Non-Uniformity (SZN)	$SZN = \frac{\sum_{j=1}^{N_z}\left(\sum_{i=1}^{N_g}P(i,j)\right)^2}{N_z}$	SZN measures the variability of size zone volumes in the image, with a lower value indicating more homogeneity in size zone volumes.

	Size-Zone Non-Uniformity Normalized (SZNN)	$SZNN = \frac{\sum_{j=1}^{N_z}\left(\sum_{i=1}^{N_g} P(i,j)\right)^2}{N_z^2}$	SZNN measures the variability of size zone volumes throughout the image, with a lower value indicating more homogeneity among zone size volumes in the image. This is the normalized version of the SZN formula.
	Zone Percentage (ZP)	$ZP = \frac{N_z}{N_p}$	ZP measures the coarseness of the texture by taking the ratio of number of zones and number of voxels in the ROI.
	Gray Level Variance (GLV)	$GLN = \sum_{i=1}^{N_g}\sum_{j=1}^{N_z} p(i,j)(i-\mu)^2$	GLV measures the variance in gray level intensities for the zones.
	Zone Variance (ZV)	$ZV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_z} p(i,j)(j-\mu)^2$	ZV measures the variance in zone size volumes for the zones.
	Zone Entropy (ZE)	$ZE = \sum_{i=1}^{N_g}\sum_{j=1}^{N_s} p(i,j)\log_2(p(i,j)+\epsilon)$	ZE measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneneity in the texture patterns.
	Low Gray Level Zone Emphasis (LGLZE)	$LGLZE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z} \frac{P(i,j)}{i^2}}{N_z}$	LGLZE measures the distribution of lower gray-level size zones, with a higher value indicating a greater proportion of lower gray-level values and size zones in the image.
	High Gray Level Zone Emphasis (HGLZE)	$HGLZE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z} P(i,j)i^2}{N_z}$	HGLZE measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image.
	Small Area Low Gray Level Emphasis (SALGLE)	$SALGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z} \frac{P(i,j)}{i^2 j^2}}{N_z}$	SALGLE measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values.
	Small Area High Gray Level Emphasis(SAHGLE)	$SAHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z} \frac{P(i,j)i^2}{j^2}}{N_z}$	SAHGLE measures the proportion in the image of the joint distribution of smaller size zones with higher gray-level values.

	Large Area Low Gray Level Emphasis (LALGLE)	$LALGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z}\frac{P(i,j)j^2}{i^2}}{N_z}$	LALGLE measures the proportion in the image of the joint distribution of larger size zones with lower gray-level values.
	Large Area High Gray Level Emphasis (LAHGLE)	$LAHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_z}P(i,j)i^2j^2}{N_z}$	LAHGLE measures the proportion in the image of the joint distribution of larger size zones with higher gray-level values.
Gray Level Run Length Matrix (GLRLM): A Gray Level Run Length Matrix (GLRLM) quantifies gray level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same gray level value. In a gray level run length matrix P(i,j|Θ), the (i,j)th element describes the number of runs with gray level i and length j occur in the image (ROI) along angle Θ.			
GLRLM features	Short Run Emphasis (SRE)	$SRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{P(i,j\|\theta)}{j^2}}{N_r(\theta)}$	SRE is a measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural textures.
	Long Run Emphasis (LRE)	$LRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}P(i,j\|\theta)j^2}{N_r(\theta)}$	LRE is a measure of the distribution of long run lengths, with a greater value indicative of longer run lengths and more coarse structural textures.
	Gray Level Non-Uniformity (GLN)	$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}P(i,j\|\theta)\right)^2}{N_r(\theta)}$	GLN measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
	Gray Level Non-Uniformity Normalized (GLNN)	$GLNN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}P(i,j\|\theta)\right)^2}{N_r(\theta)^2}$	GLNN measures the similarity of gray-level intensity values in the image, where a lower GLNN value correlates with a greater similarity in intensity values. This is the normalized version of the GLN formula.
	Run Length Non-Uniformity (RLN)	$RLN = \frac{\sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_g}P(i,j\|\theta)\right)^2}{N_r(\theta)}$	RLN measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image.

	Run Length Non-Uniformity Normalized (RLNN)	$RLNN = \frac{\sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_g} P(i,j\|\theta)\right)^2}{N_r(\theta)^2}$	RLNN measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. This is the normalized version of the RLN formula.
	Run Percentage (RP)	$RLNN = \frac{N_r(\theta)}{N_p}$	RP measures the coarseness of the texture by taking the ratio of number of runs and number of voxels in the ROI.
	Gray Level Variance (GLV)	$GLV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j\|\theta)(i-\mu)^2$	GLV measures the variance in gray level intensity for the runs.
	Run Variance (RV)	$RV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j\|\theta)(j-\mu)^2$	RV is a measure of the variance in runs for the run lengths.
	Run Entropy (RE)	$RE = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j\|\theta)\log_2(p(i,j\|\theta)+\epsilon)$	RE measures the uncertainty/randomness in the distribution of run lengths and gray levels. A higher value indicates more heterogeneity in the texture patterns.
	Low Gray Level Run Emphasis (LGLRE)	$LGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j\|\theta)}{i^2}}{N_r(\theta)}$	LGLRE measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
	High Gray Level Run Emphasis (HGLRE)	$HGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} P(i,j\|\theta)i^2}{N_r(\theta)}$	HGLRE measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
	Short Run Low Gray Level Emphasis (SRLGLE)	$SRLGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j\|\theta)}{i^2 j^2}}{N_r(\theta)}$	SRLGLE measures the joint distribution of shorter run lengths with lower gray-level values.
	Short Run High Gray Level Emphasis	$SBHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j\|\theta)i^2}{j^2}}{N_r(\theta)}$	SRHGLE measures the joint distribution of shorter run lengths with higher gray-level values.

	Long Run Low Gray Level Emphasis (LRLGLE)	$LRLGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j\|\theta)j^2}{i^2}}{N_r(\theta)}$	LRLGLRE measures the joint distribution of long run lengths with lower gray-level values.
	Long Run High Gray Level Emphasis (LRHGLE)	$LRHGLRE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} P(i,j\|\theta)i^2j^2}{N_r(\theta)}$	LRHGLRE measures the joint distribution of long run lengths with higher gray-level values.
Gray Level Dependence Matrix (GLDM) : A Gray Level Dependence Matrix (GLDM) quantifies gray level dependencies in an image. A gray level dependency is defined as a the number of connected voxels within distance δ that are dependent on the center voxel. A neighbouring voxel with gray level j is considered dependent on center voxel with gray level i if $\|i-j\| \leq \alpha$. In a gray level dependence matrix P(i,j) the (i,j)th element describes the number of times a voxel with gray level i with j dependent voxels in its neighbourhood appears in image.			
GLDM features	Small Dependence Emphasis (SDE)	$SDE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j)}{i^2}}{N_z}$	A measure of the distribution of small dependencies, with a greater value indicative of smaller dependence and less homogeneous textures.
	Large Dependence Emphasis (LDE)	$LDE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} P(i,j)j^2}{N_z}$	A measure of the distribution of large dependencies, with a greater value indicative of larger dependence and more homogeneous textures.
	Gray Level Non-Uniformity (GLN)	$GLN = \frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_d} P(i,j)\right)^2}{N_z}$	Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
	Dependence Non-Uniformity (DN)	$DN = \frac{\sum_{j=1}^{N_d}\left(\sum_{i=1}^{N_g} P(i,j)\right)^2}{N_z}$	Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image.
	Dependence Non-Uniformity Normalized (DNN)	$DNN = \frac{\sum_{j=1}^{N_d}\left(\sum_{i=1}^{N_g} P(i,j)\right)^2}{N_z^2}$	Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image. This is the normalized version of the DLN formula.
	Gray Level Variance (GLV)	$GLV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d} p(i,j)(i-\mu)^2,$ $where\mu = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d} ip(i,j)$	Measures the variance in grey level in the image.

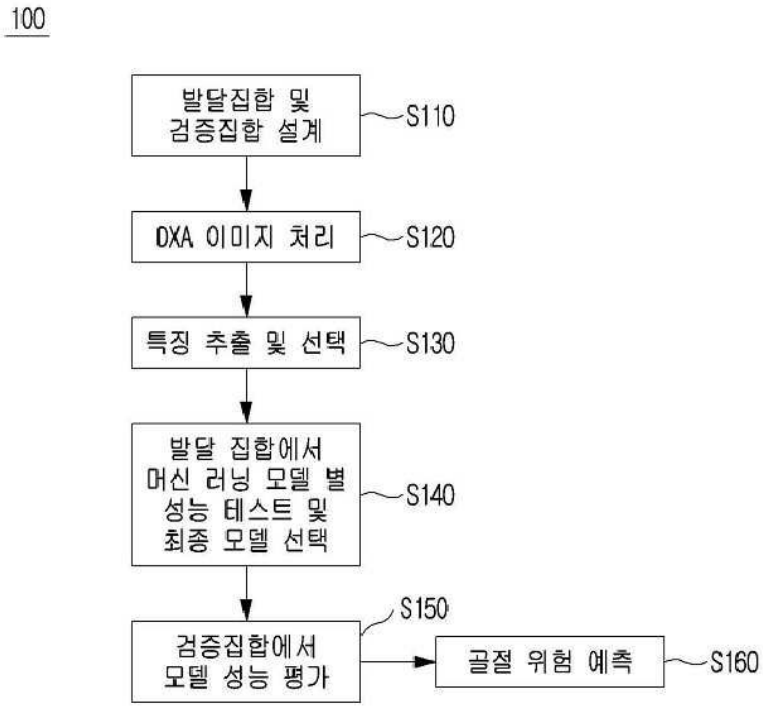
	Dependence Variance (DV)	$DV = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d} p(i,j)(j-\mu)^2,$ $where \mu = \sum_{i=1}^{N_g}\sum_{j=1}^{N_d} jp(i,j)$	Measures the variance in dependence size in the image.
	Dependence Entropy (DE)	$Dependence\ Entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)\log_2(p(i,j)+\epsilon)$	The randomness of GLDM. Higher Dependence Entropy implies more complex texture
	Low Gray Level Emphasis (LGLE)	$LGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} \frac{P(i,j)}{i^2}}{N_z}$	Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
	High Gray Level Emphasis (HGLE)	$HGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} P(i,j)i^2}{N_z}$	Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
	Small Dependence Low Gray Level Emphasis (SDLGLE)	$SDLGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} \frac{P(i,j)}{i^2j^2}}{N_z}$	Measures the joint distribution of small dependence with lower gray-level values.
	Small Dependence High Gray Level Emphasis (SDHGLE)	$SDHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} \frac{P(i,j)i^2}{j^2}}{N_z}$	Measures the joint distribution of small dependence with higher gray-level values.
	Large Dependence Low Gray Level Emphasis (LDLGLE)	$LDLGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} \frac{P(i,j)j^2}{i^2}}{N_z}$	Measures the joint distribution of large dependence with lower gray-level values.
	Large Dependence High Gray Level Emphasis (LDHGLE)	$LDHGLE = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_d} P(i,j)i^2j^2}{N_z}$	Measures the joint distribution of large dependence with higher gray-level values.
Neighbouring Gray Tone Difference Matrix (NGTDM): A Neighbouring Gray Tone Difference Matrix quantifies the difference between a gray value and the average gray value of its neighbours within distance δ.			
NGTDM features	Coarseness	$Coarseness = \frac{1}{\sum_{i=1}^{N_g} p_i s_i}$	Coarseness is a measure of average difference between the center voxel and its neighbourhood and is an indication of the spatial rate of change. A higher value indicates a lower spatial change rate and a locally more uniform texture.

	Contrast	$Contrast = \left(\frac{1}{N_{g,p}(N_{g,p}-1)}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_i p_j (i-j)^2\right)\left(\frac{1}{N_{v,p}}\sum_{i=1}^{N_g} s_i\right)$,where $p_i \neq 0,\ p_j \neq 0$	Contrast is a measure of the spatial intensity change, but is also dependent on the overall gray level dynamic range. Contrast is high when both the dynamic range and the spatial change rate are high, i.e. an image with a large range of gray levels, with large changes between voxels and their neighbourhood.
	Busyness	$Busyness = \frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i p_i - j p_j \rvert}$,where $p_i \neq 0,\ p_j \neq 0$	A measure of the change from a pixel to its neighbour. A high value for busyness indicates a 'busy' image, with rapid changes of intensity between pixels and its neighbourhood.
	Complexity	$Complexity = \frac{1}{N_{v,p}}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i-j \rvert \frac{p_i s_i + p_j s_j}{p_i + p_j}$, where $p_i \neq 0,\ p_j \neq 0$	An image is considered complex when there are many primitive components in the image, i.e. the image is non-uniform and there are many rapid changes in gray level intensity.
	Strength	$Strength = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} s_i}$, where $p_i \neq 0,\ p_j \neq 0$	Strength is a measure of the primitives in an image. Its value is high when the primitives are easily defined and visible, i.e. an image with slow change in intensity but more large coarse differences in gray level intensities.

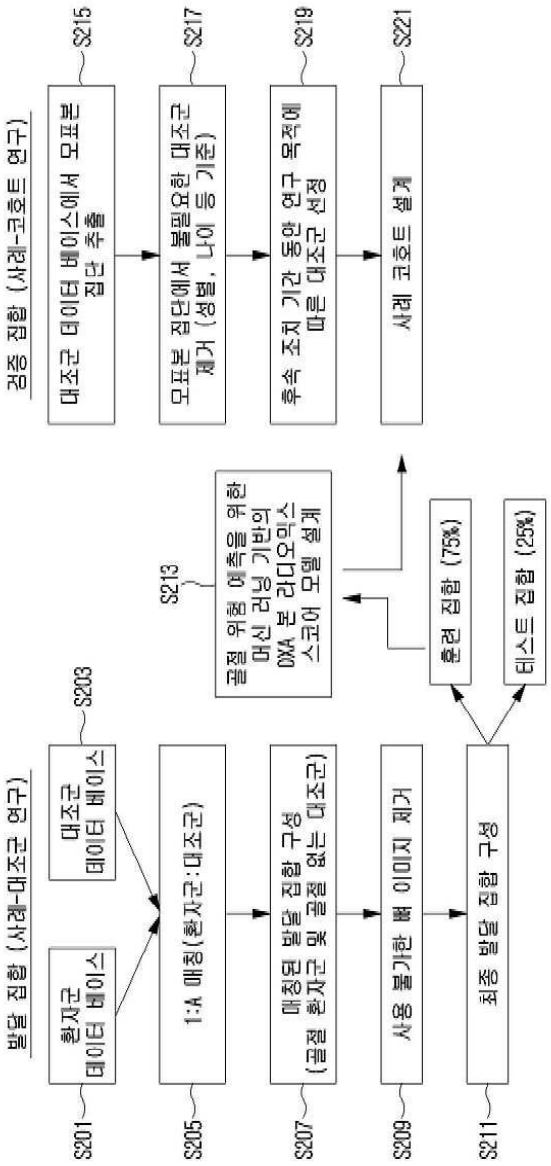
[0172] 상술한 본 발명의 실시예들은 본 발명의 필수적 특징을 벗어나지 않는 범위에서 다른 특정한 형태로 구체화될 수 있다. 따라서, 상술한 상세한 설명은 모든 면에서 제한적으로 해석되어서는 아니되고 예시적인 것으로 고려되어야 한다. 본 발명의 범위는 첨부된 청구항의 합리적 해석에 의해 결정되어야 하고, 본 발명의 등가적 범위 내에서의 모든 변경은 본 발명의 범위에 포함된다. 또한, 특허청구범위에서 명시적인 인용 관계가 있지 않은 청구항들을 결합하여 실시예를 구성하거나 출원 후의 보정에 의해 새로운 청구항으로 포함할 수 있다.

도면

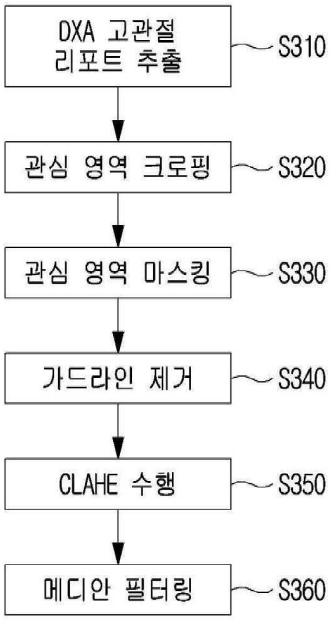
도면1


100
발달집합 및 검증집합 설계
S110
OXA 이미지 처리
S120
특징 추출 및 선택
S130
발달 집합에서 머신 러닝 모델 별 성능 테스트 및 최종 모델 선택
S140
S150
검증집합에서 모델 성능 평가
골절 위험 예측
S160

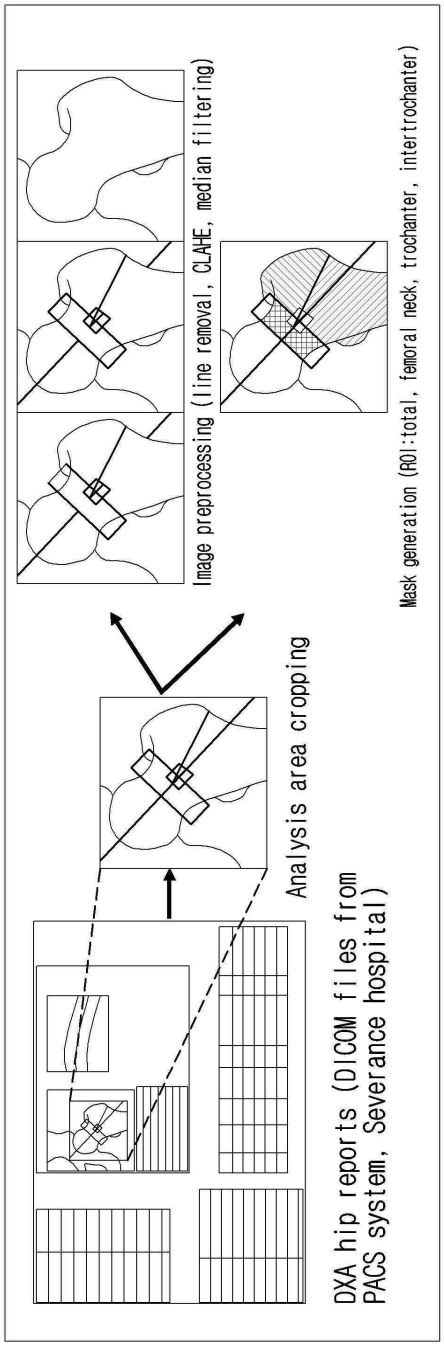
도면2


발달 집합 (사례-대조군 연구)
환자군 데이터 베이스
S201
대조군 데이터 베이스
S203
1:A 매칭(환자군:대조군)
S205
매칭된 발달 집합 구성 (골절 환자군 및 골절 없는 대조군)
S207
사용 불가한 뼈 이미지 제거
S209
최종 발달 집합 구성
S211
S213
골절 위험 예측을 위한 머신 러닝 기반의 DXA 본 라디오믹스 스코어 모델 설계
훈련 집합 (75%)
테스트 집합 (25%)
검증 집합 (사례-코호트 연구)
대조군 데이터 베이스에서 모표본 집단 추출
S215
모표본 집단에서 불필요한 대조군 제거 (성별, 나이 등 기준)
S217
추속 조치 기간 동안 연구 목적에 따른 대조군 선정
S219
사례 코호트 설계
S221

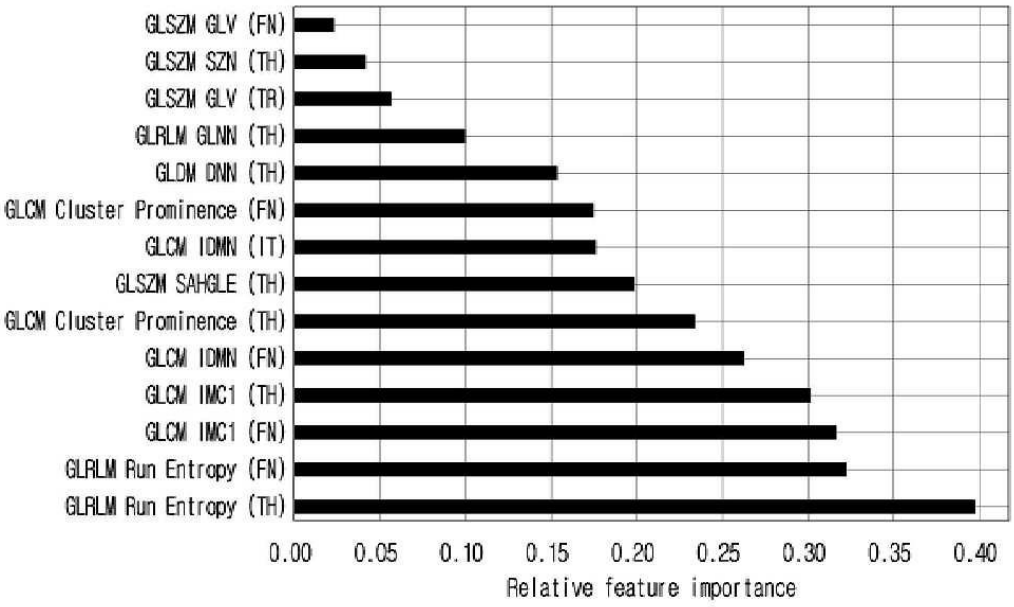
도면3


DXA 고관절
리포트 추출
S310
관심 영역 크로핑
S320
관심 영역 마스킹
S330
가드라인 제거
S340
CLAHE 수행
S350
메디안 필터링
S360

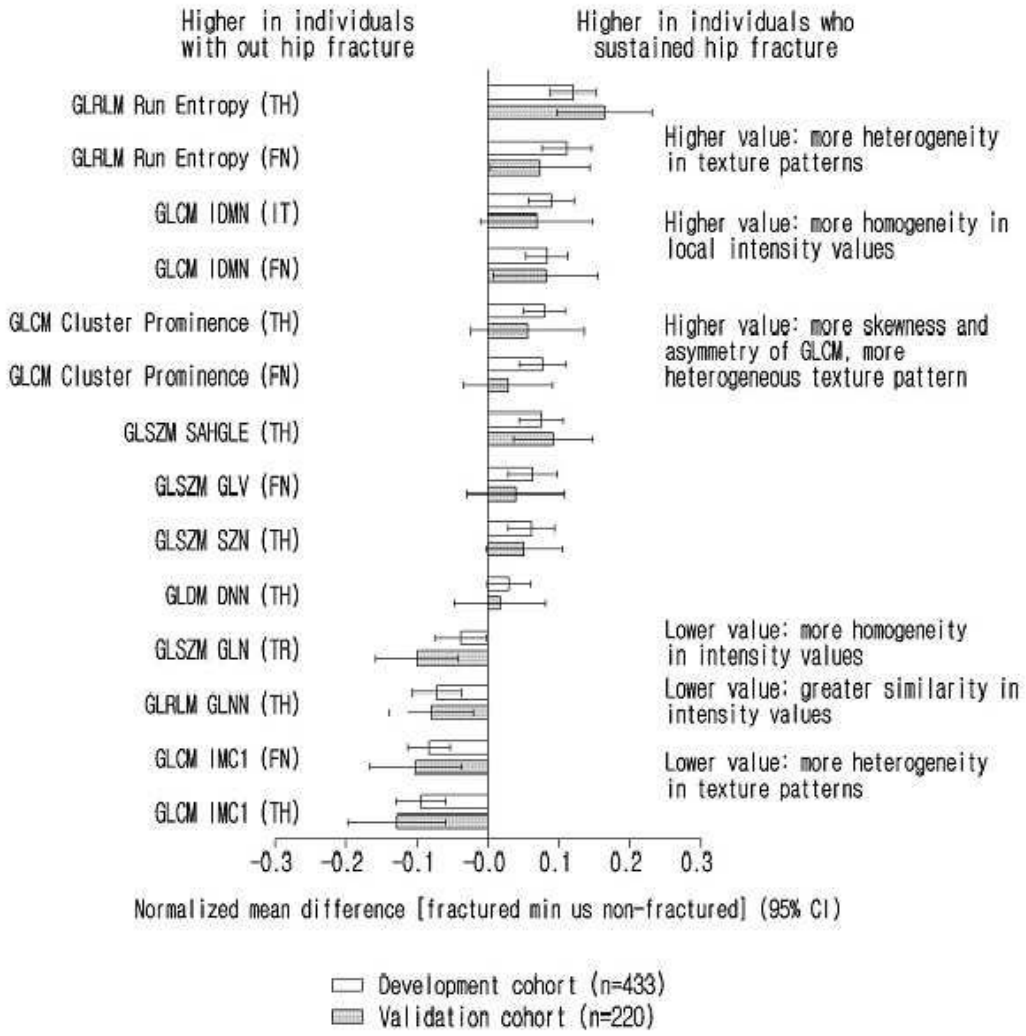
도면4


DXA hip reports (DICOM files from
PACS system, Severance hospital)
Analysis area cropping
Image preprocessing (line removal, CLAHE, median filtering)
Mask generation (ROI:total, femoral neck, trochanter, intertrochanter)

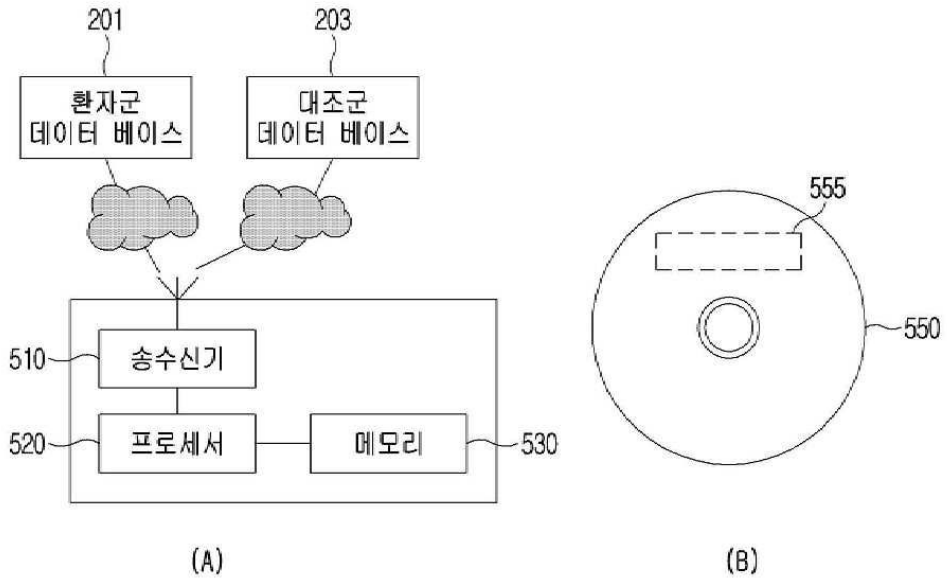
도면5


GLSZM GLV (FN)
GLSZM SZN (TH)
GLSZM GLV (TR)
GLRLM GLNN (TH)
GLDM DNN (TH)
GLCM Cluster Prominence (FN)
GLCM IDMN (IT)
GLSZM SAHGLE (TH)
GLCM Cluster Prominence (TH)
GLCM IDMN (FN)
GLCM IMC1 (TH)
GLCM IMC1 (FN)
GLRLM Run Entropy (FN)
GLRLM Run Entropy (TH)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
Relative feature importance

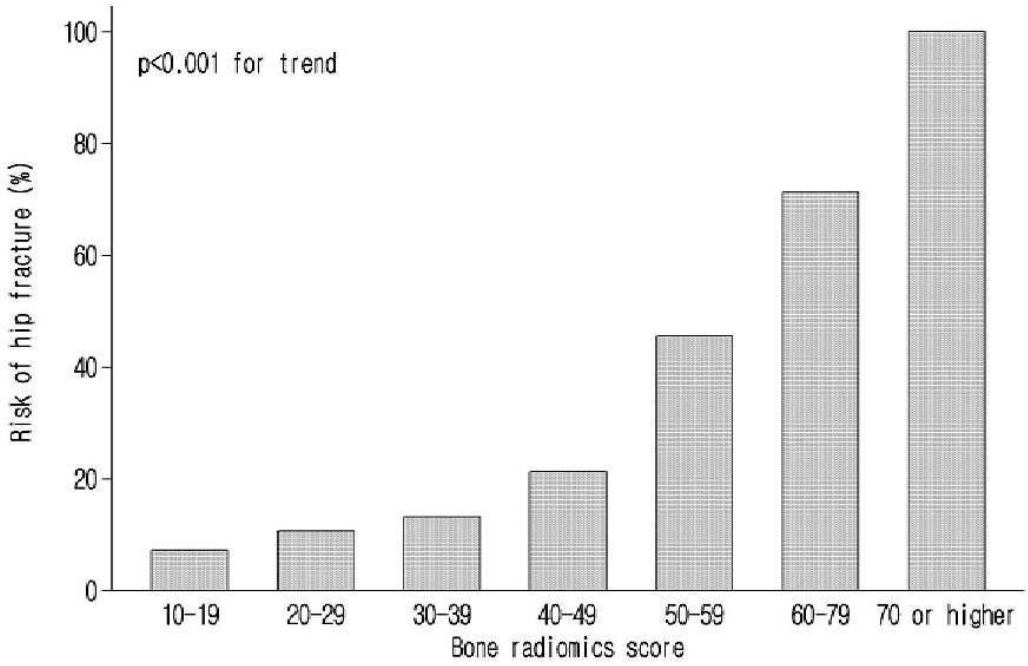
도면6


Higher in individuals with out hip fracture
Higher in individuals who sustained hip fracture
GLRLM Run Entropy (TH)
GLRLM Run Entropy (FN)
GLCM IDMN (IT)
GLCM IDMN (FN)
GLCM Cluster Prominence (TH)
GLCM Cluster Prominence (FN)
GLSZM SAHGLE (TH)
GLSZM GLV (FN)
GLSZM SZN (TH)
GLDM DNN (TH)
GLSZM GLN (TR)
GLRLM GLNN (TH)
GLCM IMC1 (FN)
GLCM IMC1 (TH)
Higher value: more heterogeneity in texture patterns
Higher value: more homogeneity in local intensity values
Higher value: more skewness and asymmetry of GLCM, more heterogeneous texture pattern
Lower value: more homogeneity in intensity values
Lower value: greater similarity in intensity values
Lower value: more heterogeneity in texture patterns
-0.3 -0.2 -0.1 -0.0 0.1 0.2 0.3
Normalized mean difference [fractured min us non-fractured] (95% CI)
Development cohort (n=433)
Validation cohort (n=220)

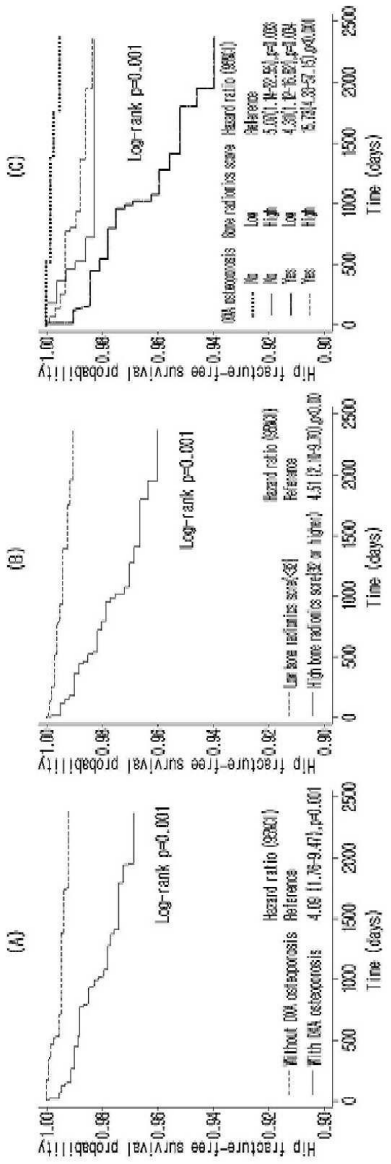
도면7


201
203
환자군
데이터 베이스
대조군
데이터 베이스
510
송수신기
520
프로세서
메모리
530
555
550
(A)
(B)

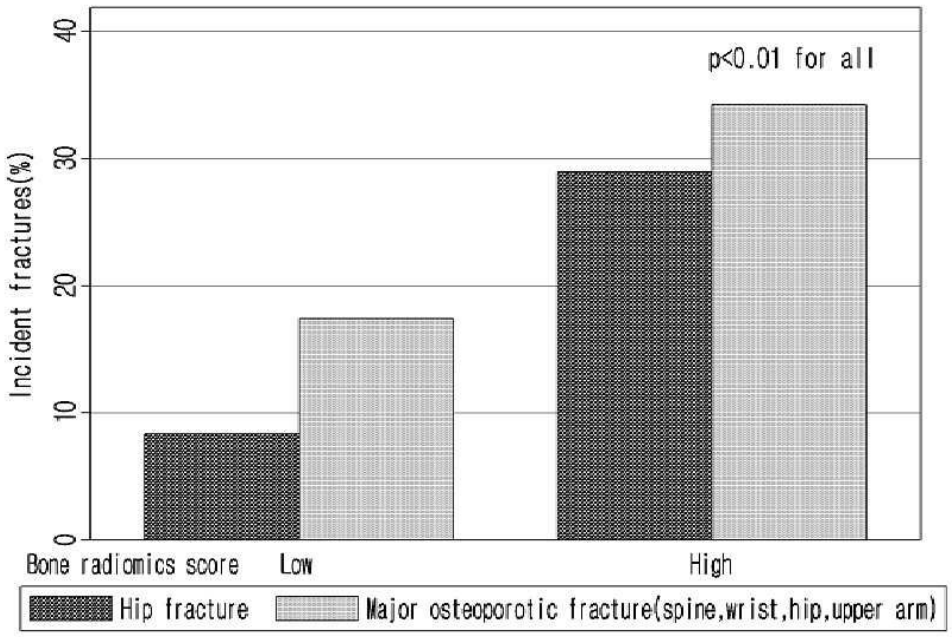
도면8


p<0.001 for trend
Risk of hip fracture (%)
100
80
60
40
20
0
10-19
20-29
30-39
40-49
50-59
60-79
70 or higher
Bone radiomics score

도면9


(A)
Hip fracture-free survival probability
1.00
0.98
0.96
0.94
0.92
0.90
0
500
1000
1500
2000
2500
Time (days)
Log-rank p<0.001
Without DXA osteoporosis
With DXA osteoporosis
Hazard ratio (95%CI)
Reference
4.09 (1.76-9.47), p=0.001
(B)
Hip fracture-free survival probability
Time (days)
Log-rank p<0.001
Low bone radiomics score(<32)
High bone radiomics score(32 or higher)
Hazard ratio (95%CI)
Reference
4.51 (2.19-9.30), p<0.001
(C)
Hip fracture-free survival probability
Time (days)
Log-rank p<0.001
DXA osteoporosis
Bone radiomics score
Hazard ratio (95%CI)
No
Low
Reference
No
High
5.07 (1.14-22.55), p=0.033
Yes
Low
4.38 (1.28-15.00), p=0.018
Yes
High
9.74 (4.33-21.92), p<0.001

도면10


p<0.01 for all
Incident fractures(%)
0
10
20
30
40
Bone radiomics score
Low
High
Hip fracture
Major osteoporotic fracture(spine,wrist,hip,upper arm)