

**(19) 대한민국특허청(KR)**
(12) 공개특허공보(A)**(11) 공개번호** 10-2024-0159562
(43) 공개일자 2024년11월05일

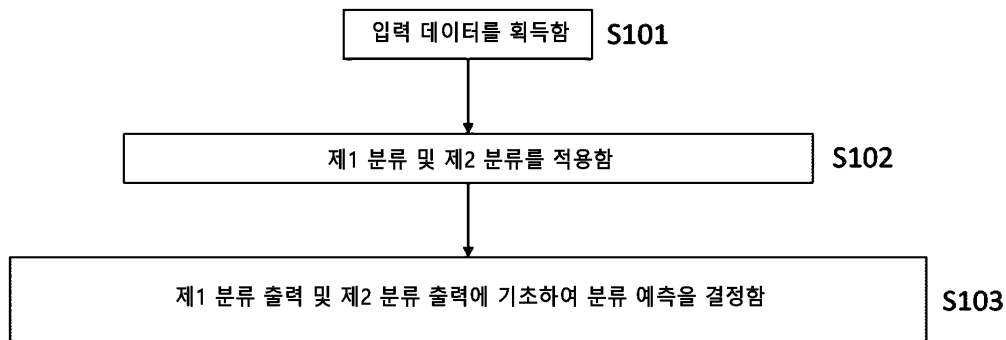
(51) 국제특허분류(Int. Cl.)
G06N 3/09 (2023.01) G06N 3/045 (2023.01)
G06N 3/0464 (2023.01)
(52) CPC특허분류
G06N 3/09 (2023.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2024-7017982
(22) 출원일자(국제) 2022년10월19일
심사청구일자 없음
(85) 번역문제출일자 2024년05월28일
(86) 국제출원번호 PCT/GB2022/052663
(87) 국제공개번호 WO 2023/073346
국제공개일자 2023년05월04일
(30) 우선권주장
21205556.0 2021년10월29일
유럽특허청(EPO)(EP)

(71) 출원인
파나케이아 테크놀로지스 리미티드
영국 캠브리지 씨비1 2엘에이 스테이션 로드 쉘리
스버리 하우스
(72) 발명자
라하르자, 무함마드 판두
영국 캠브리지 씨비1 2엘에이 스테이션 로드 쉘리
스버리 하우스 파나케이아 테크놀로지스 리미티드
내
리, 시우시
영국 캠브리지 씨비1 2엘에이 스테이션 로드 쉘리
스버리 하우스 파나케이아 테크놀로지스 리미티드
내
(74) 대리인
특허법인 무한

전체 청구항 수 : 총 15 항

(54) 발명의 명칭 분류의 방법**(57) 요약**

제1 클래스에 대응하는 제1 입력 데이터를 포함하는 제1 데이터셋을 획득하는 단계, 제1 클래스에 대응하는 제2 입력 데이터를 포함하는 제2 데이터셋을 획득하는 단계, 제1 데이터셋을 이용하여 적어도 하나의 분류기를 학습시키는 단계, 및 제2 데이터셋의 제2 입력 데이터를 적어도 하나의 분류기에 입력시키는 단계와, 제1 분류 및 제2 분류를 포함하는 분류 모델을 제공하는 단계를 포함하고, 제1 분류는 제2 분류보다 제1 클래스에 대응하는 제2 입력 데이터의 더 큰 비율이 제1 클래스 내에 속할 것으로 예측하는 것인 컴퓨터로 구현되는 방법이 제공된다.

대표도 - 도1

(52) CPC특허분류
G06N 3/0464 (2023.01)

명세서

청구범위

청구항 1

제1 클래스에 대응하는 제1 입력 데이터를 포함하는 제1 데이터셋을 획득하는 단계;

상기 제1 클래스에 대응하는 제2 입력 데이터를 포함하는 제2 데이터셋을 획득하는 단계;

상기 제1 데이터셋을 이용하여 적어도 하나의 분류기를 학습시키는 단계;

상기 제2 데이터셋의 상기 제2 입력 데이터를 상기 적어도 하나의 분류기에 입력시키는 단계, 및 제1 분류 및 제2 분류를 포함하는 분류 모델을 제공하는 단계를 포함하고, 상기 제1 분류는 상기 제2 분류보다 상기 제1 클래스에 대응하는 상기 제2 입력 데이터의 더 큰 비율이 상기 제1 클래스 내에 속할 것으로 예측하는 것인,

컴퓨터로 구현되는 방법.

청구항 2

제1항에 있어서,

상기 제1 분류는 제1 분류기를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고,

상기 제2 분류는 상기 제1 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함하고,

상기 제1 예측 임계치를 결정하는 단계는, 상기 제1 클래스에 대응하는 상기 제2 데이터셋의 입력에 대한 상기 제1 분류기의 최고 출력 값을 결정하는 단계 및 상기 제1 예측 임계치를 상기 최고 출력 값보다 작은 값 또는 상기 최고 출력 값과 같은 값으로 설정하는 단계를 포함하고,

상기 제2 예측 임계치를 결정하는 단계는, 제2 클래스에 대응하는 상기 제2 데이터셋의 입력에 대한 상기 최저 출력 값을 결정하는 단계 및 상기 제2 예측 임계치를 최저 출력 값보다 큰 값 또는 상기 최저 출력 값과 같은 값으로 설정하는 단계를 포함하는,

방법.

청구항 3

제1항에 있어서,

상기 적어도 하나의 분류기는 적어도 두 개의 분류기들을 포함하고,

상기 분류 모델을 제공하는 단계는 상기 적어도 두 개의 분류기들 중 제1 분류기를 결정하는 단계를 포함하고,

상기 제1 분류기는 상기 제1 클래스에 대응하는 상기 제2 입력 데이터의 더 큰 비율이 상기 적어도 두 개의 분류기들 중에서 상기 제1 클래스 내에 속할 것으로 예측하고,

상기 적어도 두 개의 분류기들 중 제2 분류기를 결정하는 단계를 포함하고,

상기 제2 분류기는 제2 클래스에 대응하는 상기 제2 입력 데이터의 더 큰 비율이 상기 적어도 두 개의 분류기들 중에서 상기 제2 클래스 내에 속할 것으로 예측하는,

방법.

청구항 4

제2항 또는 제3항에 있어서,

상기 분류 모델을 제공하는 단계는 제1 예측 임계치 및 제2 예측 임계치를 결정하는 단계를 포함하고,

상기 제1 예측 임계치 및 상기 제2 예측 임계치를 결정하는 단계는,

상기 제1 예측 임계치의 값들 및 상기 제2 예측 임계치의 값들의 복수의 가능한 조합들을 획득하는 단계, 상기 가능한 조합들 각각을 이용하여 분류 모델의 정확도를 결정하는 단계, 및 미리 결정된 정확도 값보다 더 높은 정확도를 가지는 조합을 선택하는 단계를 포함하는,

방법.

청구항 5

제1 클래스에 대응하는 제1 입력 데이터를 포함하는 제1 데이터세트를 획득하는 단계;

상기 제1 클래스에 대응하는 제2 입력 데이터를 포함하는 제2 데이터세트를 획득하는 단계;

상기 제1 데이터세트를 이용하여 분류기를 학습시키는 단계;

상기 제2 데이터세트의 상기 제2 입력 데이터를 적어도 하나의 분류기에 입력시키는 단계, 및 제1 분류를 포함하는 분류 모델을 제공하는 단계를 포함하고,

상기 제1 분류를 포함하는 상기 분류 모델을 제공하는 단계는 상기 제1 클래스에 대응하는 예측 임계치를 결정하는 단계를 포함하는,

컴퓨터로 구현되는 방법.

청구항 6

제1항 내지 제5항 중 어느 항에 있어서,

상기 제1 데이터세트는 제1 모집단에 대응하는 제1 입력 데이터를 포함하고, 상기 제2 데이터세트는 제2 모집단에 대응하는 제2 입력 데이터를 포함하는,

방법.

청구항 7

입력 데이터를 획득하는 단계;

상기 입력 데이터를 제1 분류 및 제2 분류를 적용하는 분류 모델에 입력시키는 단계 -상기 제1 분류는 상기 제2 분류보다 제1 클래스에 대응하는 입력 데이터의 더 큰 비율을 정확하게 예측하도록 구성됨-;

상기 제1 분류 및 상기 제2 분류의 출력에 기초하여 상기 입력 데이터에 대한 분류 예측을 결정하는 단계를 포함하는,

컴퓨터로 구현되는 방법.

청구항 8

제1항 또는 제7항에 있어서,

상기 제1 분류는 제1 분류기를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고,

상기 제2 분류는 상기 제1 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함하고,

상기 제1 예측 임계치는 상기 제2 예측 임계치보다 높은 것인,

방법.

청구항 9

제7항에 있어서,

상기 제1 분류는 제1 분류기를 적용하는 단계를 포함하고, 상기 제2 분류는 제2 분류기를 적용하는 단계를 포함하는,

방법.

청구항 10

제9항에 있어서,

상기 제1 분류는 제1 예측 임계치를 이용하고 상기 제2 분류는 제2 예측 임계치를 이용하며,

상기 제1 예측 임계치는 상기 제2 예측 임계치보다 높은 것인,

방법.

청구항 11

제7항 내지 제9항 중 어느 항에 있어서,

상기 입력 데이터는 조직의 이미지를 포함하고,

상기 분류 예측을 결정하는 단계는, 의료 진단과 관련된 정보를 결정하는 단계를 포함하는,

방법.

청구항 12

입력 데이터를 획득하는 단계;

상기 입력 데이터를 제1 분류를 적용하는 분류 모델에 입력시키는 단계 -상기 제1 분류는 제1 분류기를 적용하는 단계 및 제1 클래스에 대응하는 예측 임계치를 이용하는 단계를 포함함-;

상기 제1 분류의 출력에 기초하여 상기 제1 클래스에 대한 분류 예측을 결정하는 단계를 포함하는,

컴퓨터로 구현되는 분류의 방법.

청구항 13

제7항 내지 제12항 중 어느 항의 상기 방법을 수행하도록 구성된 하나 이상의 프로세서를 포함하는,

분류 시스템.

청구항 14

제1항 내지 제7항 중 어느 항의 상기 방법에 따라 학습된 분류 모델을 포함하는,

분류 시스템.

청구항 15

컴퓨터가 제1항 내지 제12항 중 어느 항의 상기 방법을 수행하도록 구성된 컴퓨터 판독 가능한 코드를 포함하는,
전달 매체.

발명의 설명

기술 분야

[0001] 본 개시는 분류의 방법, 분류를 위한 시스템, 및 학습시키는 방법에 관한 것이다.

배경 기술

[0002] 분류기 모델(classifier model)들은 넓은 범위의 애플리케이션들에서 응용될 수 있다. 예를 들어, 분류기 모델들은 다양한 의료 애플리케이션들에서 이용될 수 있다. 분류기 모델은 예를 들어 생리학적 측정값들을 입력으로 취하여 의료 진단을 제공하는 것에 이용될 수 있다. 다른 예에서, 분류기 모델은 예를 들어 입력 조직 이미지가 암세포들을 포함하는지 여부를 예측하기 위해 또는 특정 바이오마커가 조직 이미지와 연관되어 있는지 여부를 예측하기 위해 의료 이미지 데이터를 분석하는 것에 이용될 수 있다.

[0003] 이러한 분류기 모델들은 학습 데이터세트(training dataset)를 이용하여 학습될 수 있다. 예를 들어, 학습 데이터세트는 레이블링된 분류 예측들(labelled classification predictions)과 함께 수많은 예시 입력들을 포함할 수 있다. 그러나 학습 데이터세트를 통해 모델링되어 온 분류기 모델을 이용하는 분류 방법은 새로운 입력 데이터에 포될 때 일정 수준의 부정확성(inaccuracy)을 가질 것이다. 분류 방법의 정확도는 분류 방법에 의해 수행된 전체 예측 수에 대한 정확한 예측 수의 비율로 측정될 수 있다. 이것은 테스트 데이터세트(test dataset), 예를 들어 데이터세트 내의 각 입력에 대해 정확한 예측을 알 수 있는 테스트 데이터세트를 통해 평가될 수 있다.

[0004] 예를 들어 의료 이용 사례들과 같은 일부 이용 사례들에서는, 높은 정확도가 요구된다. 분류 방법의 정확도는 분류기 모델 성능을 향상시킴으로써 향상될 수 있다. 이러한 향상은 예를 들어 새로운 학습 데이터세트로 분류기 모델을 추가로 학습시킴으로써 달성될 수 있다. 그러나, 일부 사례들에서는, 예를 들어 추가적인 학습 데이터가 쉽게 획득될 수 없는 경우, 분류기 모델 성능의 추가적인 향상은 가능하지 않다.

발명의 내용

도면의 간단한 설명

[0005] 비제한적인 실시예들에 따른 시스템들 및 방법들은 이제 다음과 같은 첨부된 도면들을 참조하여 설명될 것이다:

도 1은 실시예에 따른 분류의 방법의 개략도를 도시한다;

도 2(a)는 데이터세트에 대한 분류기 모델 출력의 개략도이다;

도 2(b)는 제1 예시에 따른 방법의 개략도를 도시한다;

도 3은 제1 예시에 따른 방법의 개략도를 도시한다;

도 4는 컨볼루션 신경망의 예시를 도시한다;

도 5는 제1 예시에 따른 학습된 분류 시스템의 학습 파이프라인 및 배포를 도시하는 개략도이다;

도 6은 제1 예측 임계치를 도시하는 개략도이다;

도 7은 제2 예측 임계치를 도시하는 개략도이다;

도 8(a)는 제1 예시에 따른 분류를 위한 시스템을 학습시키는 방법의 개략도이다;

도 8(b)는 제1 예시에 따른 분류를 위한 시스템을 학습시키는 방법의 개략도이다;

- 도 9는 최고 민감도로부터 최고 특이도로의 전환의 개략도를 도시한다;
- 도 10(a)는 제1 예시의 변형에 따른 분류를 위한 시스템을 학습시키는 방법의 개략도이다;
- 도 10(b)는 예시적인 최적화 프로세스를 도시한다;
- 도 10(c)는 임계치 조합들에 대해 계산된 정확도 값들을 도시한다;
- 도 11 (a)는 제2 예시에 따른 분류의 방법을 도시한다;
- 도 11 (b)는 입력 데이터를 포함하는 방법의 개략도를 도시한다;
- 도 12(a)는 제2 예시에 따른 분류 시스템을 학습시키는 방법의 개략도이다;
- 도 12(b)는 제2 예시에 따른 학습된 분류 시스템의 학습 파이프라인, 및 배포를 도시하는 개략도이다;
- 도 12(c)는 제2 예시의 변형에 따른 분류기를 학습시키는 방법의 개략도이다;
- 도 13은 제3 예시에 따른 분류의 방법의 개략도이다;
- 도 14는 예시적인 분류의 방법의 개략도이다;
- 도 15는 예시적인 분류 시스템의 개략도를 도시한다;
- 도 16은 제4 예시에 따른 분류의 방법을 도시한다;

발명을 실시하기 위한 구체적인 내용

- [0006] 제1 측면에 따르면,
- [0007] 제1 클래스에 대응하는 제1 입력 데이터를 포함하는 제1 데이터셋을 획득하는 단계;
- [0008] 제1 클래스에 대응하는 제2 입력 데이터를 포함하는 제2 데이터셋을 획득하는 단계;
- [0009] 제1 데이터셋을 이용하여 적어도 하나의 분류기(classifier)를 학습시키는 단계;
- [0010] 제2 데이터셋의 제2 입력 데이터를 적어도 하나의 분류기에 입력시키는 단계, 및 제1 분류 및 제2 분류를 포함하는 분류 모델(classification model)을 제공하는 단계를 포함하고, 제1 분류는 제2 분류보다 제1 클래스에 대응하는 제2 입력 데이터의 더 큰 비율(greater proportion)이 제1 클래스 내에 속할 것으로 예측하는, 컴퓨터로 구현되는 방법이 제공된다.
- [0011] 예를 들어, 제2 분류는 제1 분류보다 제2 클래스에 대응하는 제2 입력 데이터의 더 큰 비율이 제2 클래스 내에 속할 것으로 예측한다.
- [0012] 예를 들어, 제1 클래스는 음성 클래스(negative class)이고 제2 클래스는 양성 클래스(positive class)이다.
- [0013] 예를 들어, 제1 분류는 제1 분류기를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고, 제2 분류는 제1 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함한다. 제1 예측 임계치는 제2 예측 임계치보다 높다.
- [0014] 예를 들어, 제1 분류기는 이진 분류기(binary classifier)이다. 제1 예측 임계치를 결정하는 단계는 제1 클래스에 대응하는 제2 데이터셋의 입력에 대한 이진 분류기의 최고 출력 값을 결정하는 단계 및 제1 예측 임계치를 최고 출력 값(highest output value)보다 작은 값 또는 최고 출력 값과 같은 값으로 설정하는 단계를 포함할 수 있고, 제2 예측 임계치를 결정하는 단계는 제2 클래스에 대응하는 제2 데이터셋의 입력에 대한 최저 출력 값(lowest output value)을 결정하는 단계 및 제2 예측 임계치를 최저 출력 값보다 큰 값 또는 최저 출력 값과 같은 값으로 설정하는 단계를 포함할 수 있다.
- [0015] 예를 들어, 적어도 하나의 분류기는 적어도 두 개의 분류기들을 포함하고, 분류 모델을 제공하는 단계는 적어도 두 개의 분류기들 중 제1 분류기를 결정하는 단계를 포함하고, 제1 분류기는 제1 클래스에 대응하는 제2 입력 데이터의 더 큰 비율이 적어도 두 개의 분류기들 중에서 제1 클래스 내에 속할 것으로 예측하고, 제2 분류기는 제2 클래스에 대응하는 제2 입력 데이터의 더 큰 비율이 적어도 두 개의 분류기들 중에서 제2 클래스 내에 속할 것으로 예측한다.
- [0016] 예를 들어, 분류 모델을 제공하는 단계는 제1 예측 임계치 및 제2 예측 임계치를 결정하는 단계를 포함하고, 제

1 예측 임계치 및 제2 예측 임계치를 결정하는 단계는 제1 예측 임계치의 값들 및 제2 예측 임계치의 값들의 복수의 가능한 조합들을 획득하는 단계, 가능한 조합들 각각을 이용하여 분류 모델의 정확도를 결정하는 단계, 및 미리 결정된 정확도 값보다 더 높은 정확도를 가지는 조합을 선택하는 단계를 포함한다.

- [0017] 예를 들어, 제1 데이터세트는 제1 입력 데이터에 대한 분류 레이블들(classification labels)을 포함하고 제2 데이터세트는 제2 입력 데이터에 대한 분류 레이블들을 포함한다.
- [0018] 예를 들어, 제1 분류기 및 제2 분류기는 서로 다른 모델 아키텍처들(model architectures), 및/또는 서로 다른 입력 데이터에 기초하여 업데이트된 파라미터들 및/또는 서로 다른 하이퍼-파라미터들(hyper-parameters)을 이용하여 업데이트된 파라미터들에 대응한다.
- [0019] 예를 들어, 제1 데이터세트는 제1 모집단(first population)에 대응하는 제1 입력 데이터를 포함하고, 제2 데이터세트는 제2 모집단(second population)에 대응하는 제2 입력 데이터를 포함한다.
- [0020] 다른 측면에 따르면,
- [0021] 제1 클래스에 대응하는 제1 입력 데이터를 포함하는 제1 데이터세트를 획득하는 단계;
- [0022] 제1 클래스에 대응하는 제2 입력 데이터를 포함하는 제2 데이터세트를 획득하는 단계;
- [0023] 제1 데이터세트를 이용하여 분류기를 학습시키는 단계;
- [0024] 제2 데이터세트의 제2 입력 데이터를 적어도 하나의 분류기에 입력시키는 단계, 및 제1 분류를 포함하는 분류 모델을 제공하는 단계를 포함하고, 제1 분류를 포함하는 분류 모델을 제공하는 단계는 제1 클래스에 대응하는 예측 임계치를 결정하는 단계를 포함하는, 컴퓨터로 구현되는 방법이 제공된다.
- [0025] 예를 들어, 예측 임계치를 결정하는 단계는 제1 클래스에 대응하는 제2 데이터세트로부터의 입력에 대한 분류기의 최고 출력 값을 결정하는 단계 및 예측 임계치를 최고 출력 값보다 작은 값 또는 최고 출력 값과 같은 값으로 설정하는 단계를 포함한다.
- [0026] 예를 들어, 예측 임계치를 결정하는 단계는 제1 클래스에 대응하는 제2 데이터세트로부터의 입력에 대한 분류기의 최저 출력 값을 결정하는 단계 및 예측 임계치를 최저 출력 값보다 큰 값 또는 최저 출력 값과 같은 값으로 설정하는 단계를 포함한다.
- [0027] 예를 들어, 예측 임계치를 결정하는 단계는 예측 임계치에 대한 복수의 가능한 값들(plurality of possible values)을 획득하는 단계, 각각의 가능한 값들을 이용하여 분류 모델에 대한 스코어를 결정하는 단계, 및 미리 결정된 스코어보다 더 높은 스코어를 가지는 값을 선택하는 단계를 포함한다. 스코어는 제1 클래스 내에 있을 것으로 분류되는 제1 클래스에 해당하지 않는 데이터의 비율의 수치(measure)일 수 있다. 복수의 가능한 값들은 제1 클래스에 대응하는 제2 데이터세트로부터의 입력에 대한 분류기의 최고 출력 값과 제1 클래스에 대응하는 제2 데이터세트로부터의 입력에 대한 분류기의 최저 출력 값 사이의 값들의 세트를 포함할 수 있다.
- [0028] 다른 측면에 따르면,
- [0029] 입력 데이터를 획득하는 단계;
- [0030] 입력 데이터를 제1 분류 및 제2 분류를 적용하는 분류 모델에 입력시키는 단계 -제1 분류는 제2 분류보다 제1 클래스에 대응하는 입력 데이터의 더 큰 비율을 정확하게 예측하도록 구성됨-;
- [0031] 제1 분류 및 제2 분류의 출력에 기초하여 입력 데이터에 대한 분류 예측을 결정하는 단계를 포함하는, 컴퓨터로 구현되는 방법이 제공된다.
- [0032] 다른 측면에 따르면,
- [0033] 입력 데이터를 획득하는 단계;
- [0034] 입력 데이터를 제1 분류 및 제2 분류를 적용하는 분류 모델에 입력시키는 단계 -제1 분류는 제1 예측 임계치를 이용하고 제2 분류는 제2 예측 임계치를 이용하고, 제1 예측 임계치는 제2 예측 임계치보다 높음-;
- [0035] 제1 분류 및 제2 분류의 출력에 기초하여 입력 데이터에 대한 분류 예측을 결정하는 단계를 포함하는, 컴퓨터로 구현되는 방법이 제공된다.
- [0036] 예를 들어, 제1 분류는 제1 분류기를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고, 제2 분류

는 제1 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함하고, 제1 예측 임계치는 제2 예측 임계치보다 높다.

- [0037] 예를 들어, 제1 분류는 제1 분류기를 적용하는 단계를 포함하고 제2 분류는 제2 분류기를 적용하는 단계를 포함한다.
- [0038] 예를 들어, 입력 데이터는 조직(tissue)의 이미지를 포함하고, 분류 예측을 결정하는 단계는 의료 진단(medical diagnosis)과 관련되는 정보를 결정하는 단계를 포함한다.
- [0039] 예를 들어, 제1 분류 및 제2 분류의 출력에 기초하여 입력 데이터에 대한 분류 예측을 결정하는 단계는,
- [0040] 제1 분류 및 제2 분류의 출력이 같은 경우, 제1 분류 또는 제2 분류의 출력인 분류 예측을 출력하는 단계; 및
- [0041] 분류 예측이 없다고 결정하는 단계를 포함한다.
- [0042] 다른 측면에 따르면,
- [0043] 입력 데이터를 획득하는 단계;
- [0044] 입력 데이터를 제1 분류를 적용하는 분류 모델에 입력하는 단계 -제1 분류는 제1 분류기를 적용하는 단계 및 제1 클래스에 대응하는 예측 임계치를 이용하는 단계를 포함함-;
- [0045] 제1 분류의 출력에 기초하여 제1 클래스에 대한 분류 예측을 결정하는 단계를 포함하는, 컴퓨터로 구현되는 분류의 방법(method of classification)이 제공된다.
- [0046] 예를 들어, 분류 예측을 결정하는 단계는,
- [0047] 제1 분류의 출력이 제1 클래스인 경우, 제1 클래스인 분류 예측을 출력하는 단계; 및
- [0048] 제1 분류의 출력이 제1 클래스가 아닌 경우, 분류 예측이 없다고 결정하는 단계를 포함한다.
- [0049] 다른 측면에 따르면, 상기의 방법들을 수행하도록 구성된, 하나 이상의 프로세서를 포함하는 분류 시스템(classification system)이 제공된다.
- [0050] 다른 측면에 따르면, 상기의 방법들에 따라 학습된 분류 모델을 포함하는 분류 시스템이 제공된다.
- [0051] 다른 측면에 따르면, 컴퓨터가 상기의 방법들을 수행하도록 구성된 컴퓨터 판독 가능한 코드(computer readable code)를 포함하는 전달 매체(carrier medium)가 제공된다. 방법들은 컴퓨터로 구현되는 방법들이다. 실시예들에 따른 일부 방법들은 소프트웨어에 의해 구현될 수 있기 때문에, 일부 실시예들은 임의의 적합한 전달 매체 상의 범용 컴퓨터에 제공되는 컴퓨터 코드를 포함한다. 전달 매체는 플로피 디스크, CD ROM, 자기 장치 또는 프로그램 작동이 가능한 메모리 장치와 같은 임의의 저장 매체(storage medium), 또는 임의의 신호, 예를 들어, 전기적, 광학적 또는 마이크로파 신호와 같은 임의의 일시적인 매체(transient medium)를 포함할 수 있다. 전달 매체는 비일시적(non-transitory) 컴퓨터 판독 가능한 저장 매체를 포함할 수 있다.
- [0052] 분류기 모델들은 넓은 분야들에서 애플리케이션을 가질 수 있다. 예를 들어, 분류기 모델들은 다양한 의료 애플리케이션들에서 이용될 수 있다. 분류기 모델은 예를 들어 생리학적 측정값들을 입력으로 취하여 의료 진단을 제공하는 것에 이용될 수 있다. 또 다른 예에서, 분류기 모델은 예를 들어, 입력 조직 이미지가 암세포들을 포함하는지 여부를 예측하기 위해 또는 특정 바이오마커가 조직 이미지와 연관되어 있는지 여부를 예측하기 위해 의료 이미지 데이터를 분석하는 것에 이용될 수 있다.
- [0053] 이러한 분류기 모델들은 학습 데이터셋을 이용하여 학습될 수 있다. 예를 들어, 학습 데이터셋은 레이블링된 분류 출력들과 함께 수많은 예시 입력들을 포함할 수 있다. 그러나 학습 데이터셋을 통해 모델링되어 온 분류기 모델을 이용하는 분류 방법은 새로운 입력 데이터에 배포될 때 일정 수준의 부정확성을 가질 것이다.
- [0054] 예를 들어, 이진 분류기 모델은 예측 임계치에 기초하여 분류(양성 또는 음성)로 변환되는 각 데이터 포인트에 대해 0과 1 사이의 스코어를 출력한다. 스코어가 임계치 위이면, 출력 예측은 양성이다. 스코어가 임계치 아래이면 출력 예측은 음성이다. 임계치는 예를 들어 0.5일 수 있다. 이진 분류기 모델은 예를 들어 전체 슬라이드 이미지가 암세포들을 포함하는지 여부를 예측하는 것에 이용될 수 있다. 입력 전체 슬라이드 이미지는 분류기 모델에 의해 제공된 스코어가 0.5 이상인 경우에만 암세포들을 포함하는 것으로 예측된다.
- [0055] 분류기 모델 출력은 분류 방법의 출력(양성 또는 음성)을 생성하기 위해 예측 임계치와 비교된다. 분류 방법의 정확도는 전체 분류 수에 대한 정확한 분류 수의 비율로 측정될 수 있다. 이는 예를 들어, 데이터셋 내의 각

입력에 대해 정확한 분류(양성 또는 음성)를 알 수 있는 테스트 데이터셋을 통해 평가될 수 있다.

- [0056] 예를 들어 의료 이용 사례들과 같은 일부 이용 사례들에서는, 높은 정확도가 요구된다. 분류 방법의 정확도는 분류기 모델 성능을 향상시킴으로써 향상될 수 있고, 분류기 모델 성능은 (아래에서 더 상세하게 기술되는) AUC에 의해 측정될 수 있다. 이러한 향상은 예를 들어 새로운 학습 데이터셋으로 분류기 모델을 추가로 학습시킴으로써 달성될 수 있다. 그러나, 일부 사례들에서는, 예를 들어 추가적인 학습 데이터가 쉽게 획득될 수 없는 경우, 분류기 모델 성능의 추가적인 향상이 가능하지 않다.
- [0057] 도 1은 실시예에 따른 분류의 방법의 개략도를 도시한다.
- [0058] 단계(S101)에서는, 입력 데이터가 획득된다. 입력 데이터는 예를 들어 생리학적 측정치들(physiological measurements) 또는 입력 의료 이미지일 수 있다.
- [0059] 단계(S102)에서는, 제1 분류(first classification) 및 제2 분류(second classification)가 적용된다. 제1 분류 및 제2 분류는 각각 분류를 양성 또는 음성 클래스로 출력한다. 양성 클래스(positive class)는 제2 클래스(second class)라고도 지칭되고 음성 클래스(negative class)는 제1 클래스(first class)라고도 지칭된다. 양성 클래스는 제1 분류 및 제2 분류에 대해 동일하다. 예를 들어, 양성 클래스는 의료 진단, 예를 들어 질병의 진단일 수 있다. 또는, 양성 클래스는 예를 들어 조직의 입력 이미지에 대한 특정 바이오마커의 표시(indication) 또는 암세포들의 검출일 수 있다. 제1 분류는 제2 분류보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된다. 제2 분류는 제1 분류보다 더 많은 양성 사례들을 정확하게 예측하도록 구성된다. 제2 분류는 제1 분류보다 더 높은 진양성률(true positive rate)을 가지도록 구성되고, 여기서 진양성률(TPR)은 방법이 양성으로 정확하게 예측하는 양성 사례들의 비율이다.
- [0060] 단계(S103)에서, 분류 예측은 이어서 제1 분류 및 제2 분류의 출력에 기초하여 결정된다. 특히, 제1 분류 및 제2 분류의 출력이 동일한 경우, 분류 예측은 제1 분류 (또는 제2 분류)의 출력으로 출력된다. 제1 분류 및 제2 분류의 출력이 서로 다른 경우, 분류 예측은 출력되지 않는다. 이어서 예를 들어, 수동 분류(manual classification)와 같은 일부 추가적인 단계가 수행될 수 있다. 예측이 더 불확실한 사례들이 배제되기 때문에, 분류 예측의 정확도는 향상된다.
- [0061] 이제 예시적인 분류의 방법들이 기술될 것이다.
- [0062] 제1 예시
- [0063] 분류 방법의 성능의 다양한 지표들(indicators)이 있다. 분류 방법의 민감도(sensitivity)는 진양성률(True Positive Rate, TPR)이고, 방법이 양성으로 정확하게 예측하는 양성 사례들의 비율이다. 분류 방법의 특이도(specificity)는 진음성률(True Negative Rate, TNR)이고, 방법이 음성으로 정확하게 예측하는 음성 사례들의 비율이다. 방법의 특이도는 $1 - \text{FPR}$ 과 같다. 위양성률(False Positive Rate, FPR)은 방법이 양성으로 잘못 예측하는 음성 사례들의 비율이다. 위음성률(False Negative Rate, FNR)은 방법이 음성으로 잘못 예측하는 양성 사례들의 비율이고 $1 - \text{TPR}$ 과 같다.
- [0064] 전술된 대로, 이진 분류기 모델은 예측 임계치에 기초하여 분류(양성 또는 음성)로 변환되는 각 데이터 포인트에 대해 0과 1 사이의 스코어를 출력한다. 스코어가 임계치 위(above)이면, 출력 예측은 양성이다. 스코어가 임계치 아래(below)이면 출력 예측은 음성이다. 예를 들어 심층 신경망들(deep neural networks), 서포트 벡터 머신들(support vector machines), 랜덤 포레스트들(random forests) 및 다른 것들과 같은 많은 분류기 모델들에서, 분류 방법의 민감도 및 특이도는 예측 임계치를 단순히 이동시킴으로써 분류기 모델을 학습시킨 이후 조정될 수 있다. 도 2(a)는 데이터셋에 대한 분류기 모델 출력의 개략도이다. 양성 클래스에 대한 분류기 모델 출력(0에서 1 사이의 확률 값)은 x축을 따라 0.05만큼의 이산적인 폭(discrete bins)으로 나타난다. 데이터셋에 대한 각 분류기 모델 출력의 빈도(frequency)는 y축에 대응한다. 참의 값(true value)이 양성 (1)인 사례들에 대응하는 출력들은 더 어두운 음영(darker shading)으로 도시되는 반면, 참의 값이 음성 (0)인 사례들에 대응하는 출력들은 더 밝은 음영(lighter shading)으로 도시된다.
- [0065] 예측 임계치가 높을수록 높은 특이도(higher specificity)를 가진 분류 방법을 초래하여, 더 많은 음성 사례들을 정확하게 예측할 수 있지만 더 많은 양성 사례들을 정확하게 예측하는 것에 실패할 것이다. 도 2(a)의 "특이도 임계치(specificity threshold)" 라인은 데이터셋 내의 모든 진음성 사례들(true negative cases)이 정확하게 예측되는 임계치를 도시한다.
- [0066] 동일한 분류기 모델이 낮은 예측 임계치와 결합되면, 높은 민감도(higher sensitivity)를 가진 분류 방법을 초

래하여, 더 많은 양성 사례들을 정확하게 예측할 수 있지만 더 많은 음성 사례들을 정확하게 예측하는 것에 실패할 것이다. 그림 2(a)의 "민감도 임계치(sensitivity threshold)" 라인은 데이터세트 내의 모든 양성 사례들이 정확하게 예측되는 임계치를 도시한다. "특이도" 임계치는 "민감도" 임계치보다 높다.

[0067] 분류기 모델과 높은 임계치를 이용하는 분류 방법은 분류기 모델과 낮은 임계치를 이용하는 분류 방법보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된다. 또한, 분류기 모델의 민감도 임계치에 기초한 분류 방법은 분류기 모델의 특이도 임계치에 기초한 분류 방법보다 더 많은 양성 사례들을 정확하게 예측하도록 구성된다.

[0068] 도 2(b)는 제1 예시에 따른 방법을 도시한다. 분류는 조직 이미지에 대응하는 입력 이미지 데이터에 대해 수행된다. 입력 이미지 데이터(41)를 포함하는 프로세스의 개략도가 도 3에 도시되어 있다. 그러나, 다른 예시들에서, 예를 들어, 다른 유형의 이미지(예를 들어 비조직 의료 이미지들), 환자 데이터(예를 들어 나이, 성별, 건강 상태), 액체 생검 데이터(liquid biopsy data) (예를 들어 DNA 시퀀싱 데이터), 또는 혈액 검사 데이터(예를 들어 혈액 세포의 수) 중 하나 이상을 포함하는 다른 유형의 입력 데이터가 이용될 수 있다는 것이 이해될 것이다.

[0069] 이 예시에서, 입력 이미지는 특정 바이오마커, ER 바이오마커(ER biomarker)에 대응하는지 여부에 따라 분류된다. 입력 이미지는 ER 바이오마커에 대응하는 경우 양성 클래스로 분류된다. 그러나, 다른 예시들에서, 다른 유형의 분류, 예를 들어, 이미지가 다른 바이오마커에 대응하는지 여부, 또는 예를 들어 암세포들에 대응하는지 여부를 분류하기 위해 수행될 수 있다는 것이 이해될 것이다.

[0070] 제1 예시에 따른 방법은 제1 분류 및 제2 분류를 이용한다. 제1 분류와 제2 분류는 동일한 분류기를 이용한다. 그러나, 예측 임계치는 제2 분류보다 제1 분류에서 더 높다. 따라서 제1 분류는 제2 분류보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된다. 제2 분류는 제1 분류보다 더 많은 양성 사례들을 정확하게 예측하도록 구성되며, 즉, 제2 분류는 제1 분류보다 더 높은 진양성률을 가지도록 구성된다.

[0071] 단계(S201)에서, 입력 데이터가 획득된다. 이 예시에서, 입력 데이터는 복수의 픽셀들(pixels)을 포함하는 입력 이미지 데이터(41)이다. 이 예시에서 입력 이미지 데이터(41)는 도 3에 도시된 바와 같이, 조직의 전체 슬라이드 이미지(whole slide image)로부터의 이미지 타일(image tile)이다. 아래의 설명에서, 픽셀 데이터는 레드(red)-그린(green)-블루(blue) (치수 높이 x 너비 x 3)이지만, 픽셀 데이터는 또는 예를 들어 그레이스케일 (치수 높이 x 너비 x 1)일 수 있다. 입력 이미지 데이터는 제1 수(number)의 픽셀들을 포함하고, 여기서 제1 수는 높이 x 너비의 값과 동일하다. 이미지 데이터는 조직(조직학적 단면(histological section)이라고도 함)의 이미지들을 캡처하는 디지털 카메라가 장착된 현미경을 이용하여 획득될 수 있다.

[0072] 본 명세서에서 기술되는 구체적인 예시에서, 입력 이미지 데이터(41)는 헤마톡실린(hematoxylin) 및 에오신(eosin) 염색으로 염색된 조직학적 단면의 이미지의 타일을 포함한다. 헤마톡실린 및 에오신 염색으로 염색된 조직학적 단면의 전체 슬라이드 이미지(WSI)의 예가 도 3(WSI)에 도시되어 있다. 그리드(grid)는 WSI상에 오버레이(overlay)되며, 각 그리드 정사각형은 타일에 대응한다. 전체 슬라이드 이미지 스캐너는 전체 조직 슬라이드를 스캔하여 예를 들어 약 높이 60,000픽셀, 너비 60,000픽셀을 포함하는 헤마톡실린 및 에오신 염색으로 염색된 조직학적 단면의 이미지를 초래한다. 이어서 전체 슬라이드 이미지는 인접하는(contiguous) 부분들 또는 타일들로 분할된다. 이미지 부분들(타일들)은 고정된 높이와 너비를 가진다. 부분들은 이미지 내부에서 오버랩(overlap)되거나 인접할 수 있다. 예를 들어, 이미지 타일 크기는 512 x 512 픽셀일 수 있다. 전체 슬라이드 이미지는 타일들로 분할되고, 각 타일은 단계(S201)에서 개별적으로 입력으로 취해진다. 물론 다른 타일 크기가 이용될 수 있다. 예를 들어, 2의 거듭제곱에 대응하는 타일 크기, 예를 들어, 128 x 128, 256 x 256, 512 x 512 또는 1024 x 1024 픽셀이 이용될 수 있다. 각 WSI 이미지는 서로 다른 크기일 수 있으며, 따라서 입력 WSI의 크기에 따라 입력 WSI로부터 서로 다른 수의 부분들이 추출될 수 있다. 원본 이미지 데이터는 입력, 즉 원본 픽셀 값들로 취해질 수 있다. 또는, 일부 전처리(pre-processing)는 예를 들어 컬러 정규화(colour normalization)를 위해 원본 픽셀 값들을 통해 수행될 수 있다. 셀 분할 단계(cell segmentation step)가 초기에(initially) 수행될 수 있고, 암이 없는 조직들만을 포함하는 타일들은 폐기될 수 있다. 사전 학습된 모델은 셀 분할을 위해 이용될 수 있다.

[0073] 단계(S202)에서, 입력 이미지 데이터(41)(이미지 타일)는 분류 모델에 대한 입력으로 취해진다. 분류 모델은 제1 분류기를 포함한다. 이 예시에서, 제1 분류기는 제1 컨볼루션 신경망(CNN)(40)이다. 예시적인 제1컨볼루션 신경망(convolutional neural network)(40)이 도 4에 도시되어 있다.

[0074] 제1 컨볼루션 신경망(first convolutional neural network)(40)은 적어도 하나의 컨볼루션 레이어

(convolutional layer)를 포함하는 제1 파트(46) 및 1차원 벡터를 입력으로 취하는 분류 파트인 제2 파트(47)를 포함한다. 제2 파트(47)는 적어도 하나의 완전 연결 레이어(fully connected layer)를 포함한다. 제1 CNN(40)은 다양한 시야(field-of-view) 및 깊이의 레이어들에서 이미지들에 적용되는 컨볼루션 필터들(convolutional filters)을 포함하는 제1 파트(46)를 포함하고, 이어서 데이터 감소를 위해 완전 연결 조밀 레이어들(fully connected dense layers) 및/또는 풀링 레이어들(pooling layers)을 포함하는 제2 파트(47)를 포함한다. 필터 가중치들(filter weights)은 학습 단계 동안 학습되는 학습 가능한 파라미터들(trainable parameters)이다. 하위 레벨 필터들(lower level filters)은 가장자리들(edges) 및 블롭들(blobs)과 같은 거친 구조들(coarse structures)을 감지하는 반면, 심층 레벨들(deeper levels)은 모양 및 질감과 같은 더 복잡한 속성들을 캡처하고 마지막으로 상위 레이어들(top layers)은 바이오마커의 식별과 관련하여 관심 대상들에 대해 일반화하는 방법(generalize)을 학습한다(learn).

- [0075] 제1 CNN(40)은 이진 분류기이다. 제1 CNN(40)은 입력 이미지 타일이 특정 분자 바이오마커(specific molecular biomarker)와 연관되어 있는지 여부를 결정하는 것에 이용된다. 타일들은 제1 CNN(40)에 전달된다. 픽셀 단위 데이터(per-pixel data)는 제1 CNN(40)에 직접 입력된다. 각 타일에 대해, CNN은 타일이 양성 클래스에 할당될 확률, 즉 타일이 바이오마커와 연관되어 있는 확률을 출력한다.
- [0076] 도 4에 도시된 제1 CNN(40)은 잔차 네트워크 아키텍처(residual network architecture)에 기초한다. 잔차 신경망(residual neural network)은 하나 이상의 스킵 연결(skip connection)을 포함한다. 도면은 단순화를 위해 적은 수의 레이어들을 도시하지만, 제1 CNN(40)은 예를 들어 100개 이상의 레이어들을 포함할 수 있다.
- [0077] CNN의 제1 레이어는 컨볼루션 레이어로, 도면에 "컨볼루션 레이어 1"이라고 레이블링되어 있다. 제1 레이어의 각 필터는 입력 데이터의 깊이와 일치(match)하는 깊이를 가진다. 예를 들어, 입력 데이터가 RGB인 경우, 제1 레이어의 필터 깊이는 3이다. 단순화를 위해, 도 4에 도시된 CNN은 1의 입력 데이터 깊이(즉, 그레이스케일 입력 데이터)를 가진다.
- [0078] 제1 레이어의 출력 볼륨(output volume)은 여러 요인들에 의해 결정된다.
- [0079] 제1 레이어의 출력 볼륨의 깊이는 필터의 수에 대응한다. 예를 들어, 제1 레이어 내에 32개의 필터가 있다면, 제1 레이어의 출력은 32의 깊이를 가진다. 따라서 이후 레이어(subsequent layer) 내의 필터들은 32의 깊이를 가진다.
- [0080] 출력 볼륨의 높이와 너비는 입력의 높이와 너비, 필터들의 수용 필드 크기(receptive field size)(높이와 너비 모두) 및 필터 스트라이드(filter stride)에 의해 결정된다. 스트라이드가 1이면 필터들은 한 번에 한 픽셀씩 슬라이딩한다. 스트라이드가 2이면 필터들은 한 번에 두 픽셀씩 슬라이딩하여 더 작은 출력 볼륨을 생성한다. 경계들(borders)에서 이용되는 임의의 제로 패딩(zero padding)도 또한 출력 크기에 영향을 미친다.
- [0081] 각 필터는 입력의 너비 및 높이를 따라 이동된다. 각 위치에서, 입력 값들 및 가중치 값들의 요소별 곱셈(element-wise multiplication)이 취해지고, 각 위치에 대해 단일 출력 값을 제공하기 위해 결과들이 합산된다. 하나의 필터의 출력 값들은 2D 어레이를 형성한다. 레이어 내의 모든 필터들로부터의 출력 어레이들은 깊이 차원을 따라 누적되고, 그 결과로 생성되는 볼륨은 다음 레이어에 입력된다.
- [0082] 각 컨볼루션 레이어에는 활성화 레이어(activation layer)가 뒤따른다. 활성화 레이어는 크기가 변경되지 않고 요소별 활성화 함수(elementwise activation function)를 적용한다. 단순화를 위해 도면에는 활성화 레이어들이 도시되어 있지 않다. 예를 들어, 모델은 요소별 활성화 함수를 적용하는 하나 이상의 ReLU 레이어들(Rectified Linear Unit layers)을 포함할 수 있다. 배치 정규화 레이어(batch normalization layer)는 또한 각 컨볼루션 레이어 이후 구현되며, 활성화 레이어는 배치 정규화 레이어 이후 구현된다. 모델은 컨볼루션 레이어, 배치 정규화 레이어 및 활성화 레이어를 포함하는 유닛들 또는 예를 들어 제1 컨볼루션 레이어, 제1 배치 정규화 레이어, 제2 컨볼루션 레이어, 제2 배치 정규화 레이어 및 활성화 레이어를 포함하는 유닛들을 포함한다.
- [0083] 제1 CNN(40)은 출력이 입력보다 작은 차원을 가지는, 예를 들어, 높이 및/또는 너비가 레이어에 대한 입력보다 작은 복수의 레이어들을 포함한다. 이 방식으로, 깊이가 증가하는 반면, 높이 및 너비는 여러 레이어들을 통하면서 감소한다. 제1 CNN(40)은 "인코더/디코더" 구조를 가지며, 이로 인해 레이어들은 (예를 들어, 스트라이드 크기와 같은 필터 하이퍼-파라미터들(filter hyper-parameters)을 통해) 먼저 깊이를 증가시키면서 높이 및 너비를 감소시키고, 이후 (예를 들어, 풀링 레이어들 및/또는 이중 선형 업-샘플링 레이어들(bilinear up-sampling layers)을 통해) 깊이를 감소시키면서 높이 및 너비를 증가시킨다. 이는 레이어들의 출력 크기들을 도시는 도 4에 도시되어 있다.

- [0084] 제1 CNN(40)은 공간적 크기(spatial size)를 변화시키기 위해 포함되는 하나 이상의 풀링 레이어들을 더 포함한다. 풀링 레이어들은 출력의 너비 및/또는 높이를 증가시키고 깊이를 감소시키기 위해 이용된다. 풀링 레이어들은 "평균 풀링(average pooling)" 레이어들일 수 있다. 평균 풀링 레이어는 각 위치에서 평균값을 취하면서 입력에 따라 이동되는 공간적 범위(spatial extent) 및 스트라이드를 가지는 필터를 포함한다. 그러나, 평균 외의 함수들, 예를 들어, 맥스 풀링(max pooling)이 이용될 수 있다. 업-샘플링 레이어들, 예를 들어, 하나 이상의 이중 선형 업-샘플링 레이어는 높이 및/또는 너비를 증가시키기 위해 추가적으로 또는 대안적으로 포함될 수 있다.
- [0085] 제1 CNN(40)은 적어도 하나의 스킵 연결을 더 포함한다. 실제로, 모델은 다수의 스킵 연결들을 포함할 수 있지만, 단순화를 위해 적은 수의 레이어들 및 단일 스킵 연결이 도 4에 도시되어 있다. 제2 레이어 "컨볼루션 레이어 2"는 출력 m이라고 지칭되는 출력을 생성한다. 제4 레이어 "컨볼루션 레이어 4"는 출력 m과 동일한 차원을 가지는 출력 o를 생성한다. "컨볼루션 레이어 5"에 대한 입력은 제4 레이어 o의 출력뿐만 아니라 제1 레이어 m의 출력으로부터도 생성된다. 이전 레이어의 출력을 나중 레이어에 직접 입력하는 것이 "스킵 연결"이다. 이 예시에서 출력들은 픽셀별 덧셈에 의해 결합된다. 순차(concatenation)은 대안적으로 예를 들어 출력들이 다른 크기인 경우 이용될 수 있다. 하나 이상의 스킵 연결을 이용하여, 다운스트림(downstream)의 정보가 업스트림(upstream)으로 직접 공급된다. 이는 네트워크 전체에 걸쳐 높은 수준의 전역(global) 및 지역(regional) 시각적 특징들을 유지한다. 이전 레이어의 특징들을 나중 레이어에 직접 입력하는 것은 하나 이상의 중간 계층을 스킵하면서 컨텍스트(context)를 제공한다.
- [0086] 평탄화 레이어(flattening layer)는 최종 컨볼루션 레이어 이후 포함된다. 평탄화 레이어는 최종 컨볼루션 레이어의 출력 데이터를 다음 레이어에 입력하기 위한 1차원 벡터 x로 변환한다. 이 예시에서 평탄화 레이어 이전의 레이어들은 CNN의 제1 파트(46)를 형성한다.
- [0087] 평탄화 레이어 이후 하나 이상의 완전 연결 레이어들(fully connected layers)이 포함된다. 최종 완전 연결 레이어는 양성 클래스에 대응하는 값을 출력한다. 활성화 함수(activation function), 예를 들어 시그모이드 함수(sigmoid function)는 출력에서 확률 값을 제공하기 위해 적용된다. 활성화 함수는 최종 완전 연결 레이어의 출력을 입력으로 취해 확률로 정규화한다. 따라서 활성화 함수는 양성 클래스에 대응하는 1과 0 사이의 값을 출력한다. 완전 연결 레이어(들) 및 활성화 함수는 제1 CNN(40)의 제2 파트(47)를 형성한다.
- [0088] 각 타일에 대해, CNN(40)은 타일이 양성 클래스에 할당될 확률, 즉, 타일이 바이오마커와 연관되어 있는 확률을 출력한다. 바이오마커는 특정한 병리학적(pathological) 또는 생리학적(physiological) 프로세스, 질병, 진단법, 치료법 또는 예후가 식별될 수 있는 자연적으로 발생하는 분자, 유전자 또는 특성이다. 본 명세서에서 기술되는 특정한 예시에서, 바이오마커는 암 바이오마커, 즉 특정한 유형의 암, 또는 특히 효과적인 암 치료법이 식별될 수 있는 자연적으로 발생하는 분자, 유전자 또는 특성이다. 또한, 본 명세서에 기술된 예시에서, 바이오마커는 분자 바이오마커(molecular biomarker)이다. 이 예시에서, 특정 바이오마커는 ER 바이오마커이다. 그러나, 다른 바이오마커들의 예들이 아래에서 더 자세히 기술될 것이다.
- [0089] 단계(S203)에서, 제1 분류기(40)의 출력에 제1 임계치가 적용된다. 즉, 제1 CNN(40)으로부터 출력된 확률 값은 제1 임계치에 대해 비교된다. 도 3에 도시된 예시에서, 제1 분류기로부터 출력된 확률 값은 0.764이다. 제1 임계치는 높은 예측 임계치(higher predictive threshold)(T1)이다. 이 예시에서, 제1 임계치(T1)는 0.6이다. 제1 분류기의 출력이 높은 예측 임계치보다 높은 경우, 제1 분류 출력은 1이고, 양성 클래스에 대응한다. 제1 분류기의 출력이 높은 예측 임계치보다 낮은 경우, 제1 분류 출력은 0이다. 도 6은 분류기 모델 예측 스코어에 대한 빈도의 플롯(plot) 상의 라인(T1)으로 표시되는 제1 예측 임계치를 도시하는 개략도이다.
- [0090] 단계(S203)에서, 제2 임계치는 또한 제1 분류기의 출력에 적용된다. 즉, 제1 CNN(40)으로부터 출력된 확률 값은 제2 임계치에 대해 개별적으로 비교된다. 제2 임계치는 낮은 예측 임계치(lower predictive threshold)(T2)이다. 이 예시에서, 제2 임계치(T2)는 0.45이다. 제1 분류기의 출력이 낮은 예측 임계치보다 높은 경우, 제2 분류 출력은 1이고, 양성 클래스에 대응한다. 제2 분류기의 출력이 낮은 예측 임계치보다 낮은 경우, 제2 분류 출력은 0이다. 제2 예측 임계치는 도 7에 라인(T2)으로 표시된다.
- [0091] 제1 분류(T1)에 대한 예측 임계치는 제2 분류(T2)에 대한 예측 임계치보다 높다. 따라서, 제1 분류는 제2 분류보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된다. 제2 분류는 제1 분류보다 더 많은 양성 사례들을 정확하게 예측한다.
- [0092] 제1 분류(T1)에 대한 제1 예측 임계치는 제2 분류(T2)에 대한 제2 예측 임계치보다 높다. 임계치들의 값들은 시

시스템의 이용 이전에 수행되는 학습 단계(training stage) 동안 결정된다. 학습 단계는 아래의 도 8(a)와 관련하여 기술될 것이다.

- [0093] 단계(S204)에서, 최종 분류 예측(final classification prediction)은 제1 분류 출력 및 제2 분류 출력에 기초하여 결정된다. 이 단계에서, 제1 분류가 1을 출력하고, 제2 분류가 1을 출력하면, 분류 예측은 1이 된다. 제1 분류가 0을 출력하고, 제2 분류가 0을 출력하면, 분류 예측은 0이 된다. 제1 분류와 제2 분류가 서로 다른 분류들을 출력하면, 분류 예측은 만들어지지 않는다. 최종 분류를 결정하기 위해 일부 추가적인 단계가 수행될 수 있다. 예를 들어, 이미지는 전문가에 의해 수동으로 분류될 수 있다.
- [0094] 이 방법에서, 두 개의 서로 다른 예측 임계치들을 이용함으로써, 두 개의 서로 다른 분류들이 단일 학습 프로세스를 가지는 단일 학습된 분류기를 통하여 수행될 수 있다. 제2 분류(M2)는 낮은 예측 임계치(T2)를 이용하고, 따라서 더 신뢰성 있는 양성 예측들을 생성하면서, 높은 민감도를 가지는 분류를 제공한다. 낮은 예측 임계치는 여기서 또한 음성 임계치(negative threshold)로도 지칭된다. 제1 분류(M1)는 높은 예측 임계치(T1)를 이용하고, 따라서 더 신뢰성 있는 음성 예측들을 생성하면서, 높은 특이도를 가지는 분류를 제공한다. 높은 예측 임계치는 여기서 또한 양성 임계치(positive threshold)로도 지칭된다. 두 개의 분류들을 동시에 배포함으로써, 0.5의 임계치를 가지는 제1 분류기를 이용하여 단일 분류를 적용하는 것보다, 분류 예측들에서 더 높은 정도의 확실성(certainty)이 달성될 수 있다.
- [0095] 도 8(a)는 제1 예시에 따른 분류를 위한 시스템을 학습시키는 방법의 개략도이다. 방법은 전술된 제1 예시에 따른 분류 방법에 이용될 수 있는 시스템을 학습시키는 것에 이용된다.
- [0096] 복수의 입력 이미지들을 각각 포함하는 학습 데이터세트(training dataset) 및 검증 데이터세트(validation dataset)는 학습 방법에 이용된다. 이미지들은 분류 방법 동안 이용하기 위한 의도된 유형의 입력 이미지들에 대응한다. 여기에 기술된 예시에서, 입력 이미지들은 전술된 바와 같이 헤마톡실린 및 예오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 파일들이다. 따라서 학습 데이터세트 및 검증 데이터세트는 헤마톡실린 및 예오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 파일들을 각각 포함한다.
- [0097] 각 이미지 파일은 모델이 검출할 특정 바이오마커에 대응하는지 여부에 따라 레이블링된다. 전술된 바와 같이, 이 예시의 특정 바이오마커는 ER 바이오마커이다. 각 이미지 파일은 ER 바이오마커에 대응하면 1, 대응하지 않으면 0으로 레이블링된다.
- [0098] 레이블들을 생성하기 위해, 예를 들어 IHC 염색 프로세스(IHC staining process)로부터의 정보가 이용될 수 있다. 일부 데이터세트들의 경우, 전문가는 IHC-염색된 이미지들(IHC-stained images)을 검토할 수 있고 예를 들어 메타데이터(metadata)로서 아직 이용 가능하지 않은지 등의 타겟 이미지들(target images)의 ER 상태(ER status)를 결정할 수 있다. 이들은 이어서 학습 동안 대응하는 H&E 이미지들에 대한 그라운드-트루스 레이블들(ground-truth labels)로서 이용된다. 유전자, 전사체학(transcriptomics) 및/또는 면역학적 검정들(immunological assays)을 통한 환자로부터의 인간 샘플들의 다양한 테스트(testing)이 추가적으로 또는 대안적으로 이용될 수 있다. 이러한 테스트들은 액체 및/또는 고체 형태의 생검들(biopsies)이라고 하는 인간 샘플들을 대상으로 수행되며, 이후 샘플의 분자 상태를 알려주는 절차를 거친다. 그 결과들은 이후 대응하는 H&E 이미지들에 대해 1 또는 0의 레이블을 생성하기 위해 전문가들 -조직 생검을 위한 전문 병리학자, 액체 생검을 위한 혈액학자, 세포학 샘플들을 위한 세포병리학자, 유전자/전사체학적 검정을 위한 유전학자 등-에 의해 분석된다. 어노테이션(annotation)은 훈련된 병리학자에 의해 수행될 수 있다.
- [0099] 데이터세트들을 준비하기 위해, 입력 이미지들 상에 수행되도록 의도된 것과 동일한 이미지 전처리가 수행된다. 따라서, 각 전체 슬라이드 이미지에 대해, 복수의 이미지 부분들(파일들)이 획득된다. 셀 분할은 암이 아닌 조직만을 포함하는 파일들을 데이터세트들로부터 폐기하기 위해 이용될 수 있다.
- [0100] 이미지 어노테이션은 전체 슬라이드 이미지의 각 파일에 대해 개별적으로 수행될 수 있다. 또는, 어노테이션은 전체 슬라이드 이미지에 대해 수행될 수 있고, 이미지로부터의 각 파일은 이후 전체 슬라이드 이미지에 제공된 어노테이션으로 레이블링된다. 따라서 이 경우에는 각 파일을 그들의 대응하는 슬라이드의 레이블과 페어링하는 단계가 수행된다.
- [0101] 따라서 학습 데이터세트 및 검증 데이터세트는 복수의 이미지 파일들을 각각 포함하며, 각 파일은 모델이 검출할 특정 바이오마커에 대응하는지 여부에 따라 레이블링된다.
- [0102] 단계(S801)에서, 학습 데이터세트 내의 이미지 파일들은 추론(배포) 단계(inference(deployment) stage)와 관련하여 전술된 바와 동일한 방식으로, 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기, 제1

CNN(40)에 전달된다.

[0103] 제1 CNN(40)은 연관 파라미터(associated parameter) θ_1 을 가진다. 파라미터들은 제1 CNN(40)의 제2 파트에 대한 가중치들뿐만 아니라 제1 CNN(40)의 제1 파트 내의 모든 컨볼루션 레이어들에 대한 필터 가중치들을 포함한다. 단계(S801) 내의 학습 프로세스의 목표는 학습 데이터세트 내의 이미지들의 어노테이션들과 이미지들에 대한 제1 분류기의 출력들 사이의 차이가 최소화되도록 파라미터 θ_1' 을 찾는 것이다.

[0104] 최적의 파라미터들은, 초기에 랜덤 값들을 θ_1 로 할당시키고, 그 다음에 손실 $\frac{\partial D1}{\partial \theta_1}$ 의 그라디언트(gradient)를 계산하는 것에 의해 θ_1 을 순차적으로 업데이트 하며, 계산된 그라디언트를 이용하여 θ_1 을 업데이트함으로써 계산된다. D1은 “타일당” 손실인 손실 함수(loss function)를 나타낸다. 이 예시에서는 교차 엔트로피 손실(cross entropy loss)이 이용된다:

수학식 1

$$-y_i \log(p_i) - (1 - y_i) \log(1 - p_i)$$

[0105]

[0106] 여기서, y_i 는 어노테이션 값 (1 또는 0)이고, p_i 는 학습 데이터세트 내의 예시 i 에 대한 분류기 출력 값이다. 분류기 모델의 학습 가능 파라미터들 각각에 대한 손실의 그라디언트는 역전파(back-propagation)를 통해 결정된다. 그라디언트들은 이후에 최적화 함수(optimizer function)를 이용하여, 업데이트된 파라미터들을 결정하는 것에 이용된다. 이 업데이트 방법들의 패밀리(family)는 일반적으로 다음과 같이 반복적으로 정의되는 경사하강법(gradient descent, GD)으로 알려져 있다:

수학식 2

$$\theta_1' = \theta_1 - \mu_1 \frac{\partial D1}{\partial \theta_1}$$

[0107]

[0108] 여기서 μ_1 은 학습률(learning rate)로서, 파라미터들이 얼마나 빨리 업데이트되는지를 정의한다. 업데이트는 배치 평균(batch average)에 기초하여 수행될 수 있다. 8개 타일의 배치 크기(batch size)가 예를 들어 이용될 수 있다. 8개의 타일에 걸친 평균 값 손실(average value loss)이 계산되고, 모델 파라미터들이 평균 손실에 기초하여 업데이트된다. 따라서 분류기는 파라미터들이 업데이트되기 전에 8개의 타일에 적용된다.

[0109] Adam 최적화 알고리즘(Adam optimization algorithm)이 이용될 수 있다. 선택된 최적화 전략은 이용 사례에 대한 각 전략의 성능에 의존할 수 있다. 예를 들어, 다음의 최적화 방법들 중 하나가 선택될 수 있다:

[0110] · Stochastic Gradient Descent (SGD)

[0111] · AdaDelta

[0112] · Adam

[0113] · AdaMax

[0114] · Nesterov Adam Optimizers

[0115] · RMSProp

[0116] 학습은 복수의 에포크들(epochs)에 걸쳐 수행될 수 있다. 에포크의 수는 분류기가 학습 동안 전체 학습 데이터 세트에 적용되는 횟수이다. 각 학습 에포크 동안, 학습 데이터세트 내의 학습 예시들은 무작위의 순서로 모델에

입력될 수 있어서, 학습 예시들은 각 에포크에 대해 서로 다른 순서로 입력될 수 있다.

[0117] 분류기 모델 성능은 검증 데이터셋을 이용하여 각 에포크 이후 평가될 수 있다. 검증 데이터셋으로부터의 데이터의 서로 다른 서브셋(sub-set)이 각 에포크 이후 평가를 위해 이용될 수 있다. 각 학습 에포크 이후, 0.5의 임계치에 기초한 분류기 모델 정확도가 결정될 수 있다. 검증 데이터셋으로부터의 각 이미지 타일은 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기에 제출된다. 분류기 모델 출력은 이후에 이미지에 대한 분류를 제공하기 위해 0.5의 임계치와 비교된다. 그리고 이미지에 대한 어노테이션 값(레이블)과 비교함으로써 분류가 정확한지 여부가 결정된다. 정확도는 정확한 분류들의 수를 검증 데이터에 대해 수행된 분류들의 수로 나누어 결정된다. 정확도 값은 분류기 모델 성능의 표시(indication)이다. 일단 분류기 모델이 원하는 성능을 달성하면, 학습은 중지되고(즉, 더 이상의 에포크는 수행되지 않는다) 분류기 모델 파라미터들은 고정된다. 원하는 성능은 최소의 정확도(예를 들어, 0.8 또는 0.9)를 충족함으로써, 또는 현재 에포크와 이전 에포크 사이의 정확도의 변화를 미리 결정된 값보다 작게 함으로써 결정될 수 있다. 정확도를 모니터링하는 대신, 예를 들어 동일한 방식으로 수행되는 모니터링을 위해 AUC 또는 손실이 이용될 수 있다.

[0118] 단계(S802)에서, 학습된 모델이 주어지면, 이어서 검증 데이터셋을 이용하여 예측 임계치들(T1 및 T2)이 결정된다. 이 단계에서, 검증 세트는 테스트 세트와 대부분 유사하다고 가정된다. 단계(S802)에서, 검증 데이터셋 내의 이미지 타일들은, 주론(배포) 단계와 관련하여 전술된 바와 동일한 방식으로 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기, 제1 CNN(40)에 제출된다. 분류기 모델 예측 스코어들, 즉 분류기 모델로부터 출력된 확률 값들은 이후에 이미지들의 레이블들과 비교된다. 이 예시에서, 검증 데이터셋 결과들로부터, (음성 클래스 내의) 레이블 0을 가지는 이미지에 대한 최고 확률 값(highest probability value)이 제1 예측 임계치(T1)로 취해진다. (양성 클래스 내의) 레이블 1을 가지는 이미지에 대한 최저 확률 값(lowest probability value)은 제2 예측 임계치(T2)로 취해진다. 따라서, 검증 데이터셋을 통해 제1 분류기 모델을 실행한 결과들로부터 결정된 바와 같이, 제1 예측 임계치(T1)는 특이도 임계치로 설정되고 제2 예측 임계치(T2)는 민감도 임계치로 설정된다. 제1 예측 임계치(T1)를 특이도 임계치로, 제2 예측 임계치(T2)를 민감도 임계치로 설정함으로써, 분류 방법은 양성 및 음성 클래스 모두에 대해 신뢰성 있는 예측들을 생성할 수 있다.

[0119] 상기 도 2(b)와 관련하여 기술된 바와 같이 제1 분류는 제1 임계치(T1)를 이용하여 적용되고 제2 분류는 제2 임계치(T2)를 이용하여 적용되는 완전 분류법(full classification method)의 성능은 추가적인 테스트 데이터셋을 이용하여 평가될 수 있다. 예를 들어, 완전 분류법의 정확도는 테스트 데이터셋으로부터 결정될 수 있다.

[0120] 이 예시에서, 제1 예측 임계치(T1)는 특이도 임계치로 설정되고 제2 예측 임계치(T2)는 민감도 임계치로 설정된다. 그러나, 대안적인 예시들에서, 제1 예측 임계치(T1)는 특이도 임계치보다 낮게 설정되고, 제2 예측 임계치(T2)는 민감도 임계치보다 높게 설정되는 동시에, 제1 예측 임계치(T1)는 제2 예측 임계치(T2)보다 높아지도록 보장한다. 이는 분류 방법을 이용하여 예측될 수 있는 사례들의 수를 증가시킨다. 예를 들어, 의료 진단 시스템들에서 불확실성 또는 오류의 비용들보다 이익들이 더 큰 경우 허용 가능한 수준의 불확실성 또는 오류가 도입될 수 있는 것이 일반적이다. 민감도/특이도 컨텍스트에서, 이는 (M2에 의해 이용되는) 낮은 예측 임계치(T2)를 민감도 임계치보다 약간 높은 값으로 증가시키고, (M1에 의해 이용되는) 높은 예측 임계치(T1)를 특이도 임계치보다 약간 낮게 감소시킴으로써 달성될 수 있다. 이렇게 함으로써, 분류 방법이 예측할 수 있는 사례들의 수가 증가된다.

[0121] 비교적 우수한 성능을 가지는 학습된 분류기 모델이 주어지면, 예측 임계치들은 전술된 바와 같이 단계(S802)에서 결정된다. 검증 세트가 테스트 세트와 대부분 유사하다고 가정하는 전술된 예시에서, 두 클래스들에 대한 신뢰성 있는 예측들을 생성하기 위해, T1은 특이도 임계치로 설정되고 T2는 검증 데이터셋에 대한 민감도 임계치로 설정된다. 민감도 임계치로 설정된 임계치와 함께, 분류기 모델은 검증 데이터셋에서 100% PPV를 달성한다. PPV는 양성 예측 값(Positive Predictive Value)이며, $100 \times \text{TPR} / (\text{TPR} + \text{FPR})$ 로 주어진다. 특이도 임계치로 설정된 임계치와 함께, 분류기 모델은 검증 데이터셋에서 100% NPV를 달성하며, 여기서 NPV는 음성 예측 값(Negative Predictive Value)이고 $100 \times \text{TN} / (\text{FN} + \text{TN})$ 로 주어진다.

[0122] 도 5는 제1 예시에 따른 학습 파이프라인(training pipeline) 및 학습된 분류 시스템의 배포를 나타낸 개략도이다. 제1 분류기는 전술된 바와 같이, 학습 데이터셋을 이용하여 단계(S801)에서 학습된다. 이어서, 제1 분류기는 검증 데이터셋을 통해 실행되어, 도 5에 도시되어 있는 바와 같이 모델 예측 스코어들의 세트를 생성한다. 이어서, 모델 예측 스코어들의 세트에 기초하여, 단계(S802)에서 제1 임계치(T1) 및 제2 임계치(T2)가 선택된다. 제1 임계치(T1) 및 제2 임계치(T2)는 자동화된 방식으로 결정된다. 예를 들어, 제1 예측 임계치(T1)는 특이도 임계치로 설정되고 제2 예측 임계치(T2)는 민감도 임계치로 설정된다.

- [0123] 단계(S501)에서, 분류 방법은 이어서 분류 방법의 정확도를 평가하기 위해 테스트 데이터셋을 통해 수행된다. 단계(S502)에서, 제1 분류 및 제2 분류가 일치하는 테스트 사례들에만 기초하여 평가된 테스트 데이터셋에서의 분류 정확도가 의도된 목적에 충분한지 여부가 결정된다. 예를 들어, 의도된 애플리케이션에 따라, 0.90 또는 0.95의 정확도가 충분할 수 있다. 정확도가 충분하다면, 상기 도 2(b)와 관련하여 기술된 바와 같이, 이어서 시스템이 배포(deploy)된다. 그렇지 않으면, 시스템은 폐기(discard)된다.
- [0124] 배포 동안, 입력 이미지들은 단계(S201)에서 수신되고 단계(S202) 내의 제1 분류기에 의해 분석된다. 제1 및 제2 분류들의 예측은 단계(S203)에서 출력된다. 예측들이 일치하면, 이어서 최종 분류 예측이 단계(S204)에서만 들어진다. 그렇지 않으면, 이어서 수동 테스트(manual testing)와 같은 추가적인 단계가 수행될 수 있다.
- [0125] 예측 분류기 모델의 성능은 다음의 하나 또는 모두에 의해 측정될 수 있다:
- [0126] · 정확도(accuracy): 정확한 예측의 수 / 특정 예측 임계치 값 (예를 들어, 0.5)을 가지는 분류기 모델에 의해 생성된 예측의 수 및
- [0127] · AUC: 모든 예측 임계치들에 걸쳐 정확도를 측정하는 역할을 하는 수신기 작동 특성 곡선(receiver operating characteristic curve)(ROC 곡선) 아래의 영역
- [0128] ROC 곡선은 다양한 임계치 값들에 있어서 위양성률(FPR)에 대한 진양성률(TPR)의 플롯이다. 진양성률은 민감도라고도 하며, 분류기 모델이 임계치 값을 이용하여 양성으로 정확하게 예측하는 양성 사례들의 비율이다. 위양성률은 분류기 모델이 임계치 값을 이용하여 양성으로 잘못 예측하는 음성 사례들의 비율이다. 학습된 분류기 모델의 (예를 들어, 0.5와 같은 특정 임계치 값에 대한) 정확도 및 AUC는 테스트 데이터셋을 이용하여 평가될 수 있다. 의료 이용 사례들에 있어서, 임상 등급(clinical-grade) 의료 시스템은 예를 들어 0.94 내지 0.98의 정확도 및 AUC를 가지는 분류기 모델을 이용할 수 있다.
- [0129] 낮은 정확도 및 AUC를 가지는 분류기 모델은 의도된 이용이 일부 오차 한계(margin of error)를 허용하거나, 또는 참인 예측의 이익들이 거짓인 예측의 위험 및 영향보다 더 크거나, 또는 루프 내에 예측을 확인하는 사람이 있는 애플리케이션들에서 이용될 수 있다. 그러나 많은 이용 사례들에서, 이러한 분류기 모델을 가지는 것은, 예를 들어, 사람이 예측을 확인해야 하기 때문에 장치가 경제성이 높지 않다는 것(economically unfeasible)을 의미한다. 이는 장치가 임상 실습에 채택되는 것을 방해할 수 있다. 분류기 모델 성능은 예를 들어 새로운 학습 데이터셋으로 분류기 모델을 추가로 학습시킴으로써 향상될 수 있다. 그러나, 일부 사례들에는, 예를 들어 추가적인 학습 데이터가 쉽게 획득될 수 없는 경우, 분류기 모델 성능에 대한 추가적인 개선은 가능하지 않다.
- [0130] 그러나 분류 방법의 민감도와 특이도는 분류기 모델과 함께 이용되는 예측 임계치를 선택함으로써 조정될 수 있다. 도 9는 ROC 곡선을 오른쪽에서 왼쪽으로 가로지름으로써 관찰되는 완전한 민감도(perfect sensitivity) (위 음성 없음)에서 완전한 특이도(perfect specificity) (위양성 없음)로의 전이(transition)를 개략적으로 도시한 것이다.
- [0131] 전술된 분류 방법에서, 검증 데이터셋을 통해 평가된 대로 높은 민감도 또는 최고 민감도 (위음성이 거의 없거나 전혀 없음)를 가지는 제1 분류와, 높은 특이도 또는 최고 특이도 (위양성이 거의 없거나 전혀 없음)를 가지는 제2 분류가 이용된다. 따라서, 분류 방법의 정확도가 향상될 수 있다.
- [0132] 전술된 예시에서, 제1 예측 임계치 및 제2 예측 임계치는 검증 데이터셋을 이용하여 결정된다. 그러나, 학습, 검증 및 테스트 세트들을 이용하는 대안적인 예시에서는, 단계(S801)에서 학습 세트를 이용하여 학습이 수행되고, 검증 세트를 이용하여 학습의 성능이 검증된 다음, 테스트 세트를 이용하여 임계치들이 결정된다.
- [0133] 도 8(a)와 관련하여 기술된 학습 방법에서, 학습 데이터셋 및 검증 데이터셋은 동일한 모집단에 대응한다. 그러나, 대안적인 실시예들에서, 학습 데이터셋 및 검증 데이터셋은 서로 다른 모집단들에 대응할 수 있다. 예를 들어, 학습 데이터셋은 많은 양의 데이터가 이용 가능한 일반 모집단(general population)에 대응할 수 있다. 검증 데이터셋은 더 적은 양의 데이터가 이용 가능하고 의도된 이용 사례의 모집단에 더 밀접하게 대응하는 특정 모집단(specific population)에 대응할 수 있다. 예를 들어, 학습 데이터셋은 한 국가 또는 지역의 사람들에 대응할 수 있고, 검증 데이터셋은 다른 국가 또는 지역의 사람들에 대응할 수 있다. 검증 데이터셋은 특정 병원의 환자에 대응할 수 있다. 검증 데이터셋은 연령 범위와 같은 특정 특성(trait) 또는 특질을 가지는 환자에 대응할 수 있다. 이러한 방식으로, 제1 분류기는 큰 학습 데이터셋을 이용하여 학습되고, 즉, 제1 분류기의 파라미터들은 큰 학습 데이터셋을 이용하여 업데이트된다. 이어서, 검증 데이터셋은 제1 임계치 및 제2 임계치를 결정하는 것에 이용된다. 이는 이용이 의도된 모집단과 더 밀접한 모집단에 대응하는

검증 데이터세트에 기초하여 제1 임계치 및 제2 임계치가 결정되는 "캘리브레이션(calibration)" 단계로 간주될 수 있다.

[0134] 도 8(b)는 제1 예시에 따른 분류를 위한 시스템을 학습시키는 방법을 개략적으로 도시한 것으로, 검증 데이터세트는 학습 데이터세트와 다른 모집단에 대응한다. 방법은 전술된 제1 예시에 따른 분류 방법에 이용될 수 있는 시스템을 학습시키는 것에 이용된다.

[0135] 복수의 입력 이미지들을 각각 포함하는 학습 데이터세트 및 검증 데이터세트는 학습 방법에 이용된다. 이미지들은 분류 방법 동안 이용하기 위한 의도된 유형의 입력 이미지들에 대응한다. 여기에 기술된 예시에서, 입력 이미지들은 전술된 바와 같이 헤마톡실린 및 예오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 타일들이다. 따라서 학습 데이터세트 및 검증 데이터세트는 각각 헤마톡실린 및 예오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 타일들을 포함한다. 학습 데이터세트는 제1 모집단(P1)에 대응하는 이미지들을 포함하고 검증 데이터세트는 제2 모집단(P2)에 대응하는 이미지들을 포함한다. 제2 모집단(P2)은 타겟 모집단(target population), 즉 제1 모집단(P1)보다 의도된 이용 사례에 대한 모집단에 더 밀접하게 대응한다. 예를 들어, 학습 데이터세트는 모든 연령의 환자들로부터 얻은 조직의 이미지들을 포함할 수 있는 반면, 검증 데이터세트는 특정 연령 범위의 환자들로부터 얻은 조직의 이미지들을 포함한다.

[0136] 각 이미지 타일은 전술된 바와 같이, 모델이 검출할 하는 특정 바이오마커에 해당하는지 여부에 따라 레이블링된다. 전술된 바와 같이, 이 예시에서의 특정 바이오마커는 ER 바이오마커이다. 각 이미지 타일은 ER 바이오마커에 대응하면 1, 그렇지 않으면 0으로 레이블링된다. 따라서, 학습 데이터세트 및 검증 데이터세트는 각각 복수의 이미지 타일들을 포함하며, 각 타일은 검출될 특정 바이오마커에 대응하는지 여부에 따라 레이블링된다.

[0137] 단계(S801)에서, 학습 데이터세트 내의 이미지 타일들은 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기, 제1 CNN(40)에 제출된다. 전술된 바와 같이, 학습 데이터세트 내의 이미지들의 어노테이션들과 이미지들에 대한 제1 분류기의 출력들 간의 차이가 최소화되도록 업데이트된 파라미터 벡터 θ_1 가 결정된다. 학습되지 않은 제1 분류기는 제1 모집단(P1)에 대응하는 학습 데이터세트를 이용하여 학습된다. 단계(S801)의 출력은 제1 모집단(P1)을 통해 학습된, 학습된 분류기이다.

[0138] 전술된 바와 같이, 각 학습 에포크(training epoch) 후에, 0.5의 임계치에 기초한 분류기 정확도가 결정될 수 있다. 검증 데이터세트로부터의 각 이미지 타일은 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기로 제출된다. 분류기 출력은 이어서 이미지에 대한 분류를 제공하기 위해 0.5의 임계치와 비교된다. 결과들은 모델 예측 스코어들을 보여주는 도면의 "결과들 1"로 레이블링된 제1 그래프에 개략적으로 도시되어 있다. 0.5 위의 모든 스코어들은 양성으로 예측되고, 0.5 아래의 모든 스코어들은 음성으로 예측된다. 참인 분류들(true classifications)은 또한 제1 그래프에 도시되어 있는데, 어두운 음영(darker shading)은 실제 값(true value)이 양성임을 나타낸다.

[0139] 이미지에 대한 어노테이션 값(레이블)과 비교함으로써 분류가 정확한지 여부가 결정된다. 정확도는 검증 데이터에 대해 수행된 분류들의 수로 정확한 분류들의 수를 나누어 결정된다. 정확도 값은 분류기 모델 성능의 표시이다. 분류기 모델이 원하는 성능을 달성하면, 학습이 중단되고 (즉, 더 이상의 에포크가 수행되지 않음) 분류기 파라미터들은 고정된다.

[0140] 단계(S802)에서, 학습된 모델이 주어지면, 예측 임계치들(T1 및 T2)은 이후에 검증 데이터세트를 이용하여 결정된다. 검증 데이터세트는 제2 모집단(P2), 타겟 모집단에 대응한다. 따라서, 단계(S802)는 캘리브레이션 프로세스(calibration process)로서 고려될 수 있다. 학습된 모델은 입력으로 취해지고, 예측 임계치들이 결정된다.

[0141] 앞서 기술된 바와 같이, 단계(S802)에서, 검증 데이터세트 내의 이미지 타일들은, 추론(배포) 단계와 관련하여 전술된 바와 동일한 방식으로, 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기에 전달된다. 분류기 모델 예측 스코어들, 즉 분류기 모델로부터 출력된 확률 값들은 그런 다음 이미지들의 레이블들과 비교된다. 분류기 예측 스코어들은 그래프 "결과들 2"에 나타나 있다. 이 예시에서, 검증 데이터세트 결과들로부터, (음성 클래스에 있는) 레이블 0을 가지는 이미지에 대한 최고 확률 값이 제1 예측 임계치(T1)로 취해진다. (양성 클래스에 있는) 레이블 1을 가지는 이미지에 대한 최저 확률 값이 제2 예측 임계치(T2)로 취해진다. 따라서, 검증 데이터세트를 통해 제1 분류기 모델을 실행한 결과들로부터 결정된 바와 같이, 제1 예측 임계치(T1)는 특이도 임계치로 설정되고 제2 예측 임계치(T2)는 민감도 임계치로 설정된다. 제1 예측 임계치(T1)를 특이도 임계치로 설정하고 제2 예측 임계치(T2)를 민감도 임계치로 설정함으로써, 분류 방법은 양성 클래스 및 음성 클래스 모두에 대해 신뢰성 있는 예측들을 생성할 수 있다.

- [0142] 위의 도 2(b)와 관련하여 기술된 바와 같이 제1 임계치(T1)를 이용하여 제1 분류가 적용되고 제2 임계치(T1)를 이용하여 제2 분류가 적용되는 완전 분류법(full classification method)의 성능은 추가적인 테스트 데이터셋을 이용하여 평가될 수 있다. 예를 들어, 완전 분류법의 정확도는 테스트 데이터셋으로부터 결정될 수 있다. 이 예시에서 설정된 테스트 데이터셋은 또한 제2 모집단(P2)에 대응한다.
- [0143] 이 예시에서, 제1 예측 임계치(T1)은 특이도 임계치로 설정되고 제2 예측 임계치(T2)는 민감도 임계치로 설정된다. 그러나, 대안적인 예시들에서, 제1 예측 임계치(T1)은 특이도 임계치보다 낮게 설정되고 제2 예측 임계치(T2)는 민감도 임계치보다 높게 설정되는 동시에, 제1 예측 임계치(T1)는 제2 예측 임계치(T2)보다 높아지도록 보장한다.
- [0144] 단계(S802)에서의 캘리브레이션의 단계는 예를 들어 분류기 모델이 타겟 병원 모집단에 기초하여 조정, 즉 모집단 시프트(population shift)에 맞게 조정될 수 있도록 한다. 이는 모집단 변화(population change)에 따른 예측 스코어들의 분포 변화(distribution change)를 조정하기 위한 예측 임계치들(T1, T2)을 결정함으로써 수행된다. 모집단 변화를 조정하기 위해 모델의 가중치들(weights)을 조정하는 중간 단계가 또한 포함될 수 있다. 예를 들어, 제2 모집단(P2)에 대응하는 제2 학습 데이터셋이 제1 학습 데이터셋을 이용하여 수행된 학습 과정 이후 추가적인 학습 과정을 수행하기 위해 이용될 수 있다. 이 추가적인 학습 과정에서, 제1 분류기의 파라미터들은 타겟 모집단에 대응하는 데이터에 기초하여 더 조정된다. 이어서, 제1 및 제2 임계치들은 전술된 바와 같이 단계(S802)에서 결정된다.
- [0145] 전술된 예시에서, 제1 분류기가 학습되고, 이어서 제1 예측 임계치 및 제2 예측 임계치가 타겟 모집단에 따라 결정된다. 캘리브레이션 단계(S802) 동안, 입력 이미지들이 제1 분류기에 대한 입력으로 취해진다. 예측 이전의 최종 레이어로부터 출력된 값이 예측 스코어로서의 출력이다. 따라서, 타겟 모집단 예측 스코어 분포가 생성된다. 이어서, 제1 및 제2 임계치들이 결정된다.
- [0146] 표 1은 테스트 데이터셋으로부터의 이미지들을 ER 바이오마커에 대응하는 것으로 분류하는 제1 예시에 따른 제1 분류의 방법을 이용하여 생성된 결과들을 도시한다. 방법은 제1 및 제2 임계치를 가지는 전술된 바와 같은 분류기 모델을 이용한다. 표 1은 또한 테스트 데이터셋으로부터의 이미지들을 ER 바이오마커에 대응하는 것으로 분류하는 비교 예시(comparative example)에 따른 제1 분류의 방법을 이용하여 생성된 결과들을 도시한다. 방법은 0.5의 단일 임계치(single threshold)를 가지는 전술된 바와 같은 동일한 분류기 모델을 이용한다.
- [0147] 표 1은 또한 테스트 데이터셋으로부터의 이미지들을 PR 바이오마커에 대응하는 것으로 분류하는 제1 예시에 따른 제2 분류의 방법을 이용하여 생성된 결과들을 도시한다. 방법은 제1 및 제2 임계치를 가지는 전술된 바와 같은 분류기 모델을 이용한다. 표 1은 또한 테스트 데이터셋으로부터의 이미지들을 PR 바이오마커에 대응하는 것으로 분류하는 비교 예시에 따른 제2 분류의 방법을 이용하여 생성된 결과들을 도시한다. 방법은 0.5의 단일 임계치를 가지는 전술된 바와 같은 동일한 분류기 모델을 이용한다.

표 1

[0148]

	비교 예시		제1 예시	
	ER 바이오마커	PR 바이오마커	ER 바이오마커	PR 바이오마커
정확도	0.82234957	0.768361582	0.870090634	0.826086957
F1	0.89084507	0.836653386	0.926243568	0.889655172
AUC	0.825141878	0.756788921	0.823986358	0.721239607
FPR	0.484848485	0.47008547	0.689655172	0.530612245
FNR	0.106007067	0.113924051	0.010989011	0.044444444
PPV	0.887719298	0.79245283	0.870967742	0.832258065
NPV	0.53125	0.696629213	0.857142857	0.793103448
민감도	0.893992933	0.886075949	0.989010989	0.955555556

특이도	0.515151515	0.52991453	0.310344828	0.469387755
PLR	1.843860424	1.884925201	1.434065934	1.800854701
NLR	0.205778424	0.214985708	0.035409035	0.09468599

- [0149] 위의 표 1에서 PLR은 양성 우도비(positive likelihood ratio)이고 NLR은 음성 우도비(negative likelihood ratio)이다.
- [0150] 제1 예시의 변형
- [0151] 도 10(a)는 제1 예시의 변형에 따른 분류를 위한 시스템의 학습 방법의 개략도이고, 여기서 낮은 예측 임계치(T2)는 민감도 임계치 위의 값으로 설정될 수 있고 높은 예측 임계치(T1)는 특이도 임계치 아래의 값으로 설정될 수 있다. 이렇게 함으로써, 모델이 예측할 수 있는 경우의 수가 증가될 수 있다.
- [0152] 단계(S1001)에서, 제1 분류기(40)는 단계(S801)와 관련하여 전술된 바와 동일한 방식으로, 학습 데이터셋를 이용하여 학습된다.
- [0153] 단계(S1002)에서, 민감도 및 특이도 임계치들은 단계(S802)와 관련하여 전술된 바와 동일한 방식으로, 검증 데이터셋를 이용하여 결정된다.
- [0154] 단계(S1003)에서, 제1 및 제2 임계치들은 최적화 프로세스(optimization process)를 이용하여 결정된다. 제1 예시의 변형에 따른 방법에서, 제1 및 제2 임계치들을 각각 특이도 및 민감도 임계치들로 설정하는 대신, 제1 및 제2 임계치들을 결정하기 위해 최적화 프로세스가 수행될 수 있다.
- [0155] 단계(S1003)에서 수행될 수 있는 예시적인 최적화 프로세스가 이하 도10(b)와 관련하여 기술된다.
- [0156] 단계(S901)에서, 가능한 값들의 제1 세트(first set of possible values)가 제1 임계치에 대해 획득되고 가능한 값들의 제2 세트(second set of possible values)가 제2 임계치에 대해 획득된다. 제1 세트 및 제2 세트 내의 모든 값들은 민감도 임계치보다 크거나 같고 특이도 임계치보다 작거나 같다. 가능한 값들의 제1 세트는 특이도 임계치 값을 포함하지만 민감도 임계치 값을 포함하지 않는 민감도 임계치 값과 특이도 임계치 값 사이 일정한 간격의 값들 중의 고정된 수를 취함으로써 획득된다. 가능한 값들의 제2 세트는 민감도 임계치 값을 포함하지만 특이도 임계치 값을 포함하지 않는 민감도 임계치 값과 특이도 임계치 값 사이 일정한 간격의 값들 중의 고정된 수를 취함으로써 획득된다.
- [0157] 이 예시에서, 특이도 임계치는 0.65이고 민감도 임계치는 0.4이다. 0.025의 증분(increment)으로 일정한 간격의 값들의 세트가 값들의 제1 세트(first set of values) 및 값들의 제2 세트(second set of values)로 취해진다. 값들의 제1 세트는 {0.425, 0.450, 0.475, 0.500, 0.525, 0.550, 0.575, 0.600, 0.625, 0.650}의 값들을 포함한다. 값들의 제2 세트는 {0.400, 0.425, 0.450, 0.475, 0.500, 0.525, 0.550, 0.575, 0.600, 0.625}의 값들을 포함한다.
- [0158] 단계(S902)에서, 정확도 값은 가능한 제1 임계치 값들 및 가능한 제2 임계치 값들의 복수의 조합들 각각에 대해 계산된다. 정확도는 테스트 데이터셋로부터 결정된다. 이 단계에서, 제1 세트로부터의 각각의 값은 제2 세트로부터의 각각의 하위 값과 결합된다. 테스트 데이터셋로부터의 입력들은 제1 분류기를 통해 실행되고, 제1 분류기로부터의 출력 확률 값들은 분류 예측들을 제공하기 위해 조합 내의 제1 임계치 값 및 제2 임계치 값과 비교된다. 이어서, 조합에 대한 정확도는 정확한 분류 예측들의 수를 전체 분류 예측들의 수로 나눈 것으로서 결정된다. 이는 복수의 조합들 각각에 대해 반복된다.
- [0159] 이 단계의 결과들은 도 10(c)에 본 예시로 도시되어 있다. 제1 세트 {0.425, 0.450, 0.475, 0.500, 0.525, 0.550, 0.575, 0.600, 0.625, 0.650}의 각각의 값이 그리드의 맨 위를 따라 도시되어 있다. 제2 세트 {0.400, 0.425, 0.450, 0.475, 0.500, 0.525, 0.550, 0.575, 0.600, 0.625}의 각각의 값이 그리드의 왼쪽에 도시되어 있다. 그리드의 나머지 셀들은 (가능한 조합들에 대응하는 셀들에 대한) 정확도 값을 나타내거나 (제2 세트의 값이 제1 세트의 값보다 낮은 조합들에 대응하는 셀들에 대해) 비어 있다. 제1 세트의 각각의 값은 제2 세트의 각각의 낮은 값과 결합되어 대응하는 정확도 값을 가지는 조합들 (0.425, 0.400), (0.450, 0.400), (0.450, 0.425), ... (0.650, 0.600), (0.652, 0.600)을 초래한다. 각 조합에 대해서, 제1 임계치 값은 제2 임계치 값보

다 높다. 도10(c)는 조합들 각각에 대해 계산된 정확도 값들을 도시한다.

[0160] 단계(S903)에서, 의도된 애플리케이션에 있어서 원하는 정확도보다 크거나 같은 정확도 스코어를 초래하는 제1 및 제2 임계치 값들의 모든 조합들이 검색된다. 원하는 정확도는 미리 결정된 정확도 값이다. 이 예시에서, 원하는 정확도는 0.9이다. 이 예시에서, 조합들 (0.575, 0.400), (0.600, 0.400), (0.600, 0.425), (0.625, 0.400), (0.625, 0.425), (0.625, 0.425), (0.625, 0.450), (0.625, 0.475), (0.650, 0.400), (0.650, 0.425), (0.650, 0.450), (0.650, 0.475) 이 0.9보다 큰 정확도를 초래한다. 따라서, 이러한 조합들은 단계(S903)에서 검색된다.

[0161] 단계(S904)에서, 검색된 조합들로부터 하나의 조합이 선택된다. 조합은 커버리지(coverage)에 기초하여 선택된다. 예를 들어, 테스트 데이터세트로부터 가장 많은 결과들에 대한 분류를 예측했던 검색된 조합의 조합이 선택된다. 또는, 제1 임계치 값과 제2 임계치 값 사이의 차이가 가장 큰 조합이 선택된다. 이 예시에서, 조합 (0.625, 0.475)이 가장 큰 커버리지를 가지는 것으로 선택된다. 따라서, 단계(S1003)에서, 제1 임계치는 0.625로 설정되고, 제2 임계치는 0.475로 설정된다.

[0162] 제2 예시

[0163] 도 11(a)는 제2 예시에 따른 분류의 방법을 도시한다. 도 11(a)와 관련하여 기술되는 방법에서, 분류 방법은 제1 예시와 관련하여 기술된 것과 동일한 입력 이미지 데이터를 통해 수행된다. 도 11(b)는 입력 데이터를 포함하는 방법의 개략도를 도시한다.

[0164] 제2 예시에 따른 방법도 또한 제1 분류와 제2 분류를 이용한다. 그러나, 제1 분류 및 제2 분류는 서로 다른 분류기들을 이용한다. 제1 분류는 제2 분류에 이용되는 제2 분류기보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된 제1 분류기를 이용한다. 제2 분류기는 제1 분류기보다 더 높은 진양성률(true positive rate)을 가지도록 구성된다. 제2 실시예에서, 두 개의 별개의 분류기들이 학습되는데, 그 중 하나는 높은 특이도와 낮은 민감도를 가지고, 다른 하나는 높은 민감도와 낮은 특이도를 갖는다.

[0165] 단계(S1101)에서, 입력 데이터는 위의 단계(S201)와 관련하여 기술된 바와 동일한 방식으로 획득된다.

[0166] 단계(S1102)에서, 입력 이미지 데이터(input image data)(41)(이미지 타일)는 분류 모델에 대한 입력으로 취해진다. 분류 모델은 제1 분류기 및 제2 분류기를 포함한다. 따라서, 입력 이미지 데이터(41)(이미지 타일)는 제1 분류기에 대한 입력으로 취해지고, 또한 별개로 제2 분류기에 대한 입력으로 취해진다. 이 예시에서, 제1 분류기는 상기 도 4와 관련하여 기술된 바와 같은 컨볼루션 신경망(CNN)(40)이다. 제2 분류기도 상기 도 4와 관련하여 기술된 바와 같은 컨볼루션 신경망(CNN)(40)이다. 이 예시에서, 제1 분류기 및 제2 분류기는 동일한 모델 아키텍처를 가지지만, 서로 다른 학습 데이터세트들을 이용하여 학습되었으므로 서로 다른 파라미터 값들을 가질 수 있다. 상기 도 4와 관련하여 기술된 바와 같이, 분류기들의 구조는 서로 동일하다. 그러나, 대안적인 실시예들에서, 제1 분류기 및 제2 분류기는 서로 다른 모델 아키텍처들을 가진다. 예를 들어, 제1 분류기 및 제2 분류기는 서로 다른 수의 레이어들, 레이어들 내의 서로 다른 수의 필터들, 또는 서로 다른 스킵 연결들을 가질 수 있다.

[0167] 단계(S1103)에서, 제1 임계치는 제1 분류기(40a)의 출력에 적용된다. 즉, 제1 CNN(40a)으로부터 출력된 확률 값은 제1 임계치에 대해 비교된다. 도 11(b)에 도시된 예시에서, 제1 분류기(40a)로부터 출력된 확률 값은 0.764이다. 이 예시에서, 제1 분류기(40a)와 함께 이용되는 제1 예측 임계치는 검증 데이터세트로부터 결정된 제1 분류기(40a)에 대한 특이도 임계치이다. 제1 분류기의 출력이 제1 예측 임계치보다 높으면, 제1 분류 출력은 1이다. 제1 분류기의 출력이 제1 예측 임계치보다 낮으면, 제1 분류 출력은 0이다.

[0168] 단계(S1103)에서, 제2 임계치는 제2 분류기(40b)의 출력에 적용된다. 즉, 제2 CNN(40b)으로부터 출력된 확률 값은 제2 임계치에 대해 개별적으로 비교된다. 이 예시에서, 제2 분류기(40b)와 함께 이용되는 제2 예측 임계치는 검증 데이터세트로부터 결정된 제2 분류기(40b)에 대한 민감도 임계치이다. 제2 분류기(40b)의 출력이 제2 예측 임계치보다 높으면, 제2 분류 출력은 1이다. 제2 분류기의 출력이 제2 예측 임계치보다 낮으면, 제2 분류 출력은 0이다.

[0169] 이 예시에서, 제1 분류기(40a)와 함께 이용된 제1 예측 임계치는 검증 데이터세트로부터 결정된 제1 분류기(40a)에 대한 특이도 임계치이고, 제2 분류기(40b)와 함께 이용된 제2 예측 임계치는 검증 데이터세트로부터 결정된 제2 분류기(40b)에 대한 민감도 임계치이다. 이 임계치들이 어떻게 결정되는지는 아래의 도 12와 관련하여 기술될 것이다. 다만, 대안적인 예시들에서, 제1 및 제2 임계치들은 동일한, 미리 결정된 값으로 설정된다. 예

를 들어, 제1 및 제2 임계치들은 0.5로 설정될 수 있다.

- [0170] 제1 분류는 제2 분류보다 더 많은 음성 사례들을 정확하게 예측하도록 구성된다. 제2 분류는 제1 분류보다 더 많은 양성 사례들을 정확하게 예측하도록 구성된다. 더 많은 음성 사례들을 정확하게 예측하도록 구성된 제1 분류에 이용되는 제1 분류기(40a)는 배포 이전에 수행되는 학습 단계 동안 획득된다. 더 많은 양성 사례들을 정확하게 예측하도록 구성된 제2 분류에 이용되는 제2 분류기(40b)도 또한 학습 단계 동안 획득된다. 제2 분류기(40b)는 제1 분류기보다 더 높은 진양성률(true positive rate)을 가지도록 선택된다.
- [0171] 단계(S1104)에서, 분류 예측은 제1 분류 출력 및 제2 분류 출력에 기초하여 결정된다. 이 단계에서, 제1 분류가 1을 출력하고, 제2 분류가 1을 출력하면, 분류 예측은 1이다. 제1 분류가 0을 출력하고, 제2 분류가 0을 출력하면, 분류 예측은 0이다. 제1 분류와 제2 분류가 서로 다른 분류들을 출력하면, 분류 예측은 만들어지지 않는다. 이어서, 최종 분류를 결정하기 위해 일부 추가적인 단계가 수행될 수 있다. 예를 들어, 이미지는 전문가에 의해 수동으로 분류될 수 있다.
- [0172] 도 12의 (a)는 제2 예시에 따른 분류 시스템을 학습시키는 방법의 개략도이다. 방법은 전술된 제2 예시에 따른 분류 방법에 이용될 수 있는 분류 시스템을 학습시키는 것에 이용될 수 있다.
- [0173] 복수의 이미지들을 각각 포함하는 학습 데이터세트 및 검증 데이터세트가 이용된다. 이미지들은 의도된 유형의 입력 이미지들에 대응한다. 앞서 기술된 바와 같이, 학습 데이터세트 및 검증 데이터세트는 동일한 모집단에 대응하거나, 또는 학습 데이터세트는 제1 모집단에 대응하고, 검증 데이터세트는 제2, 타겟, 모집단에 대응할 수 있다.
- [0174] 여기서 기술된 예시에서, 입력 이미지들은 해마톡실린 및 에오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 타일들이다. 따라서, 학습 데이터세트 및 검증 데이터세트는 각각 해마톡실린 및 에오신 염색으로 염색된 조직학적 단면들의 이미지들로부터의 이미지 타일들을 포함한다. 제1 예시와 관련하여 기술된 바와 같이, 각 이미지 타일은 또한 모델이 검출할 특정 바이오마커에 대응하는지 여부에 따라 레이블링된다.
- [0175] 여기에 기술된 예시에서, 학습 데이터세트는 다섯 개의 개별 학습 데이터 서브세트들로 분할된다. 학습 데이터세트 내의 각 이미지는 다섯 개의 개별 학습 데이터 서브세트들 중 하나에 무작위로 할당된다. 각 학습 데이터 서브세트(training data subset)는 따라서 서로 다른 이미지들을 포함한다. 각 서브세트는 동일하거나, 또는 실질적으로 동일한 수의 이미지들을 가진다.
- [0176] 단계(S1201)에서, 다섯 개의 개별 분류기들, A, B, C, D 및 E가 학습된다. 각 분류기는 동일한 모델 아키텍처를 가지지만, 서로 다른 학습 데이터 서브세트를 이용한다.
- [0177] 제1 학습 데이터 서브세트로부터의 이미지 타일들은 양성 클래스에 할당될 확률을 생성하는 분류기 A에 제출된다. 분류기 A는 연관 파라미터 벡터(associated parameter vector) θ_A 를 갖는다. 최적의 파라미터들은, 랜덤 값들을 θ_A 로 할당하고, 그 다음 손실 $\frac{\partial DA}{\partial \theta_A}$ 의 그래디언트를 계산하고, 계산된 그래디언트를 이용하여 θ_A 를 업데이트하는 것에 의해 순차적으로 θ_A 를 업데이트함으로써 계산된다. DA는 "타일당(per-tile)" 손실인 손실 함수를 나타낸다. 교차 엔트로피 손실(cross entropy loss)은 앞서 기술된 바와 같이 이용된다. 제1 분류기의 학습 가능한 파라미터들 각각에 대한 손실의 그래디언트는 역전파(back-propagation)를 통해 결정된다. 그래디언트들은 이후에, 앞서 기술된 바와 같이, 최적화 함수(optimiser function)를 이용하여, 업데이트된 파라미터들을 결정하는 것에 이용된다.
- [0178] 제2 학습 데이터 서브세트로부터의 이미지 타일들은 양성 클래스에 할당될 확률을 생성하는 분류기 B에 제출된다. 분류기 B는 연관 파라미터 벡터 θ_B 를 갖는다. 최적의 파라미터들은, 랜덤 값들을 θ_B 로서 할당하고, 그 다음 손실 $\frac{\partial DB}{\partial \theta_B}$ 의 그래디언트를 계산하고, 계산된 그래디언트를 이용하여 θ_B 를 업데이트함으로써 계산된다. 이 예시에서, 동일한 형태의 손실 함수가 각 분류기를 학습시키기 위해 이용된다. 모델의 학습 가능한 파라미터들 각각에 대한 손실의 그래디언트는 역전파를 통해 결정된다. 그래디언트들은 이후에, 앞서 기술된 바와 같은 최적화 함수를 이용하여, 업데이트된 파라미터들을 결정하는 것에 이용된다. 이 예시에서, 각 분류기에 대해 동일한 학습률(learning rate) μ 가 이용된다.

- [0179] 제3, 제4 및 제5 학습 데이터 서브세트들은 분류기들 C, D 및 E를 각각 동일한 방식으로 학습시키기 위해 이용된다. 이 예시에서, 다섯 개의 분류기들 사이의 유일한 차이는 따라서 단계(S1201)에서의 학습 과정에 대해 이용된 학습 데이터이다.
- [0180] 각 분류기에 대해, 모델 성능은 앞서 기술된 바와 같이 검증 데이터세트로부터의 데이터를 이용하여 평가된다. 일단 학습이 원하는 성능을 달성하면, 모델 파라미터들은 각 분류기에 대해 고정된다.
- [0181] 단계(S1202)에서, 검증 데이터세트로부터의 동일한 데이터가 다섯 개의 분류기들 각각에 입력된다. 예측 스코어들, 즉 분류기들로부터 출력된 확률 값들은 이어서 이미지 레이블들과 비교된다. 특이도 임계치 및 민감도 임계치는 이어서 각 분류기에 대해 계산된다. 각 분류기에 대해, 검증 데이터세트 결과들은 (음성 클래스에 있는) 레이블 0을 가지는 이미지에 대한 최고 출력 확률 값(highest output probability value)을 결정하는 것에 이용된다 - 이는 특이도 임계치이다. 각 분류기에 대해, (양성 클래스에 있는) 레이블 1을 가지는 이미지에 대한 최저 출력 확률 값(lowest output probability value)이 민감도 임계치로서 결정된다. 최고 특이도 임계치에 대응하는 분류기가 다른 분류기들보다 더 많은 음성 사례들을 정확하게 예측할 수 있을 제1 분류기로서 선택된다. 최저 민감도 임계치를 가지는 분류기가 다른 분류기들보다 더 많은 양성 사례들을 정확하게 예측할 수 있을 제2 분류기로서 선택된다. 제1 분류기와 함께 이용될 제1 예측 임계치는 선택된 분류기의 특이도 임계치로 설정된다. 제2 분류기와 함께 이용될 제2 예측 임계치는 선택된 분류기의 민감도 임계치로 설정된다. 이 제2 예시에서, 두 개의 모델들, 즉 음성 사례들을 예측하기 위한 제1 모델 및 양성 사례들을 예측하기 위한 제2 모델이 선택된다. 특이도 임계치는 이후 제1 모델에 대해 선택되고 민감도 임계치는 제2 모델에 대해 선택된다. 이 접근법은 일부 모델들이 음성들을 잘 예측하지만 양성들을 잘 예측하지 못하는 반면, 다른 모델들이 양성들을 잘 예측하지만 음성들을 잘 예측하지 못하는 애플리케이션들에 이용될 수 있다.
- [0182] 이어서, 상기 도 11(a)와 관련하여 기술된 바와 같이 제1분류기를 이용하여 제1분류를 적용하고 제2분류기를 이용하여 제2분류를 적용하는 완전 분류법의 성능은 테스트 데이터세트를 이용하여 평가될 수 있다.
- [0183] 상기 예시에서, 최고 특이도 임계치(highest specificity threshold)에 대응하는 분류기가 제1 분류기로 선택되고, 최저 민감도 임계치(lowest sensitivity threshold)를 가지는 분류기가 제2 분류기로 선택되어, 다섯 개의 개별 분류기들이 학습된다. 그러나, 대안적인 예시들에서, 최고 특이도 임계치에 대응하는 분류기가 제1 분류기로 선택되고, 다른 분류기가 제2 분류기로 선택되어, 단지 두 개의 분류기들만이 학습된다. 대안적인 예시들에서, 최고 특이도 임계치에 대응하는 분류기가 제1 분류기로 선택되고, 최저 민감도 임계치를 가지는 분류기가 제2 분류기로 선택되어, 두 개 이상의 분류기들이 학습된다.
- [0184] 상기 예시에서, 최고 특이도 임계치에 대응하는 분류기가 제1 분류기로 선택되고 최저 민감도 임계치를 가지는 분류기가 제2 분류기로 선택된다. 대안적으로 그러나, 0.5의 임계치를 이용하는 최고 특이도에 대응하는 분류기가 제1 분류기로 선택되고 0.5의 임계치를 이용하는 최고 민감도를 가지는 분류기가 제2 분류기로 선택된다. 제1 예측 임계치 및 제2 예측 임계치는 이후에 모두 0.5로 설정된다.
- [0185] 상기 예시에서, 복수의 분류기들 각각은 학습 데이터세트의 개별 서브세트로 학습되고, 여기서 각 이미지는 서브세트들 중 하나에 무작위로 할당된다. 그러나, 대안적인 예시들에서, 이미지들은 이미지들의 하나 이상의 속성에 기초하여 할당될 수 있다. 예를 들어, 하나 이상의 서브세트는 다른 것들보다 더 많은 양성 예시들을 가질 수 있다. 예를 들어, 서브세트들 중 두 개는 80%의 양성 예시들을 가지고, 서브세트들 중 하나는 50%의 양성 예시들을 가지며, 서브세트들 중 두 개는 20%의 양성 예시들을 가지도록 이미지들이 할당될 수 있다. 상기 예시에서, 서브세트들은 실질적으로 동일한 수의 이미지들을 가지지만, 대안적으로 서브세트들 중 일부는 다른 것들보다 더 많은 이미지들을 가질 수 있다.
- [0186] 상기 예시에서, 복수의 분류기들 각각은 동일한 모델 아키텍처를 가진다. 그러나, 대안적인 예시들에서, 분류기들의 일부 또는 전부는 서로 다른 모델 아키텍처들을 가질 수 있고, 예를 들어, 분류기들은 서로 다른 수의 레이어들, 레이어들 내의 서로 다른 수의 필터들, 또는 서로 다른 스킵 연결들을 가질 수 있다.
- [0187] 상기 예시에서, 복수의 분류기들 각각은 동일한 하이퍼-파라미터들(hyper-parameters)을 이용하여 학습된다. 그러나, 대안적인 예시들에서, 분류기들 중 일부 또는 전부는 서로 다른 하이퍼-파라미터들을 이용하여 학습될 수 있으며, 예를 들어, 학습률(learning rate)은 서로 다를 수 있다.
- [0188] 도 12(b)는 제2 예시에 따른 학습 파이프라인 및 학습된 분류 시스템의 배포를 도시한 개략도이다. 전술된 바와 같이, 두 개의 분류기들은 단계(S1201)에서 학습 데이터세트를 이용하여 학습된다. 분류기들은 이어서 검증 데이터세트를 통해 실행되어, 도 12(b)에 도시된 바와 같이 각 분류기에 대한 모델 예측 스코어들의 세트를 생성

한다. 이후 모델 예측 스코어들의 세트에 기초하여, 제1 분류기 및 제2 분류기가 선택된다. 제1 분류기 및 제2 분류기는 자동화된 방식(automated manner)으로 결정된다. 예를 들어, 검증 데이터세트 결과들에 기초하여 결정된 바와 같이, 제1 분류기는 최고 특이도 임계치에 대응하는 분류기로 선택되고, 제2 분류기는 최저 민감도 임계치를 가지는 분류기로 선택된다.

[0189] 단계(S501)에서, 방법의 정확도를 평가하기 위해, 완전 분류법(full classification method)이 이후 테스트 데이터세트를 통해 수행된다. 단계(S502)에서, 확실하다고 주장하는 테스트 사례들(즉, 제1 및 제2 분류가 일치하는지 여부)로만 평가된 테스트 데이터세트에서의 정확도가 의도된 목적에 충분한지 여부가 결정된다. 정확도가 충분하다면, 상기 도 11(a)에 관련하여 기술된 바와 같이, 이어서 시스템이 배포된다. 그렇지 않으면, 시스템은 폐기된다.

[0190] 배포 동안, 입력 이미지들은 단계(S1101)에서 수신되고 단계(S1102)에서 제1 및 제2 분류기에 의해 분석된다. 제1 및 제2 분류들의 예측은 단계(S1103)에서 출력된다. 예측들이 일치하면, 최종 분류 예측이 단계(S1104)에서 만들어진다. 그렇지 않으면, 그 다음 수동 테스트(manual testing)과 같은 추가적인 단계가 수행될 수 있다.

[0191] 제2 예시의 변형

[0192] 도 12(c)는 제2 예시의 변형에 따른 분류기를 학습시키는 방법의 개략도이다.

[0193] 단계(S1401)에서, 복수의 분류기들은 상기 단계(S1201)와 관련하여 기술된 바와 동일한 방식으로 학습된다.

[0194] 단계(S1402)에서, 제1 분류기 및 제2 분류기들은 상기 단계(S1202)와 관련하여 기술된 바와 동일한 방식으로 선택된다. 단계(S1202)와 관련하여 기술된 바와 같이, 이 단계의 일부로서, 제2 분류기의 민감도 임계치 및 제1 분류기의 특이도 임계치는 검증 데이터세트를 이용하여 결정된다.

[0195] 전술된 제2 예시에서, 제1 분류기(40a)와 함께 이용되는 제1 예측 임계치는 검증 데이터세트로부터 결정된 제1 분류기(40a)에 대한 특이도 임계치로 설정되고, 제2 분류기(40b)와 함께 이용되는 제2 예측 임계치는 검증 데이터세트로부터 결정된 제2 분류기(40b)에 대한 민감도 임계치로 설정된다. 그러나, 제2 예시의 변형에 따른 방법에서, 제1 및 제2 임계치들을 결정하기 위해 최적화 프로세스가 수행된다.

[0196] 단계(S1403)에서, 제1 및 제2 임계치들은 최적화 프로세스를 이용하여 결정된다. 상기 도 10(b)와 관련하여 기술된 예시적인 최적화 프로세스는 이 단계에서 수행될 수 있다.

[0197] 가능한 값들의 제1 세트는 제1 임계치에 대해 획득되고, 가능한 값들의 제2 세트는 제2 임계치에 대해 획득된다. 이 예시에서, 제1 세트의 모든 값들은 제1 분류기의 민감도 임계치보다 크거나 같고 제1 분류기의 특이도 임계치보다 작거나 같은 반면, 제2 세트의 모든 값들은 제2 분류기의 민감도 임계치보다 크거나 같고 제2 분류기의 특이도 임계치보다 작거나 같다. 예를 들어, 가능한 값들의 제1 세트는 특이도 임계치 값 및 민감도 임계치 값을 포함하는 제1 분류기의 민감도 임계치 값으로부터 특이도 임계치 값까지 일정한 간격의 값들 중의 고정된 수의 값들을 취함으로써 획득될 수 있다. 가능한 값들의 제2 세트는 민감도 임계치 값 및 특이도 임계치 값을 포함하는 제2 분류기의 민감도 임계치 값으로부터 특이도 임계치 값까지 일정한 간격의 고정된 수의 값들을 취함으로써 획득될 수 있다.

[0198] 정확도는 이어서 가능한 제1 임계치 값들 및 가능한 제2 임계치 값들의 복수의 조합들 각각에 대해 결정된다. 이 단계에서, 제1 세트로부터의 각각의 값은 제2 세트로부터의 각각의 값과 결합된다. 테스트 데이터세트로부터의 입력들은 제1 및 제2 분류기들을 통해 실행된다. 분류 예측들을 제공하기 위해 제1 분류기 모델 출력들은 조합 내의 제1 임계치 값에 대해 비교되고 제2 분류기 모델 출력들은 조합 내의 제2 임계치 값에 대해 비교된다. 정확도는 이어서 조합들 각각에 대해 결정된다.

[0199] 애플리케이션에 있어서 원하는 정확도보다 크거나 같은 정확도 스코어를 초래하는 제1 및 제2 임계치 값들의 모든 조합들이 검색되고, 검색된 조합들로부터 조합은 커버리지(coverage)에 기초하여 선택된다. 예를 들어, 검색된 조합들 중에서 테스트 데이터세트로부터 가장 많은 결과들을 분류한 조합이 선택된다.

[0200] 제3 예시

[0201] 제1 예시에서, 분류 방법은 타일 단위 기초(per-tile basis)로 적용된다. 이미지 타일은 전체 슬라이드 이미지로부터 취해지고, 제1 분류기에 입력으로 취해진다. 제1 임계치를 이용하는 제1 분류 및 제2 임계치를 이용하는 제2 분류가 적용된다. 이미지 타일이 ER 바이오마커에 대응하는지 여부에 대한 분류 예측이 이어서 결정된다.

[0202] 도 13은 제3 예시에 따른 분류의 방법의 개략도이고, 여기서 분류 방법은 이미지 단위 기초(per-image basis)로

적용된다. 즉, 전체 슬라이드 이미지가 ER 바이오마커에 대응하는지 여부에 대한 분류 예측이 획득된다.

- [0203] 단계(S1501)에서, 입력 데이터가 획득된다. 이 예시의 입력 이미지 데이터는 앞서 기술된 바와 같이 조직의 전체 슬라이드 이미지로부터의 이미지 타일이다. 이 예시에서, 도 13과 관련하여 도시되어 있는 바와 같이, 전체 슬라이드 이미지로부터 다수의 이미지 타일들이 획득된다. 셀 분할은 암이 아닌 조직만을 포함하는 타일들을 폐기하기 위해 이용될 수 있다.
- [0204] 단계(S1502)에서, 각 입력 이미지 타일은 이 예시에서 제1 예시와 관련하여 기술된 바와 같은 제1 CNN(40)인 제1 분류기를 포함하는 분류 모델에 대한 입력으로 개별적으로 취해진다. 제1 분류기는 각 이미지 타일에 대해, 이미지 타일이 ER 바이오마커와 연관되어 있는 확률을 출력한다.
- [0205] 이어서, 집계 단계(aggregation step)가 수행된다. 이 예시에서, 집계 단계는 풀링 연산자(pooling operator)를 적용하는 단계를 포함한다. 최고 확률 값에 대응하는 이미지 타일이 식별된다. 이 이미지 타일은 "상위 랭크된(top-ranked)" 이미지 타일로서 지칭될 것이다. 이 "상위 랭크된" 이미지 타일은 이어서 분류 방법의 이후 단계들(subsequent steps)에서 이미지를 분류하는 것에 이용된다.
- [0206] 단계(S1503)에서, 제1 임계치는 상위 랭크된 이미지 타일에 대한 제1 분류기(40)의 출력에 적용된다. 즉, 상위 랭크된 이미지 타일에 대한 제1 CNN(40)으로부터 출력된 확률 값은 제1 분류의 출력을 제공하기 위해 제1 임계치에 대해 비교된다. 제2 임계치는 상위 랭크된 이미지 타일에 대한 제1 분류기(40)의 출력에 적용된다. 즉, 상위 랭크된 이미지 타일에 대한 제1 CNN(40)으로부터 출력된 확률 값은 제2 분류를 제공하기 위해 제2 임계치에 대해 별도로 비교된다. 제1 및 제2 임계치들은 앞서 제1 예시와 관련하여 기술된 바와 같이, 학습 단계 동안 결정된다.
- [0207] 단계(S1504)에서, 전체 슬라이드 이미지에 대한 최종 분류 예측이 제1 분류 출력 및 제2 분류 출력에 기초하여 결정된다. 이 단계에서, 제1 분류가 1을 출력하고, 제2 분류가 1을 출력하면, 분류 예측은 1이다. 제1 분류가 0을 출력하고, 제2 분류가 0을 출력하면, 분류 예측은 0이다. 제1 분류 및 제2 분류가 서로 다른 분류들을 출력하면, 분류 예측은 만들어지지 않는다. 최종 분류를 결정하기 위해 일부 추가적인 단계가 수행될 수 있다. 예를 들어, 이미지는 전문가에 의해 수동으로 분류될 수 있다.
- [0208] 제3 예시에서, 배포 동안, 제1 CNN 분류기(40)는 타일 단위 기초로 확률들을 반환하고, 이 개별 스코어들은 최대 연산자(max operator)를 통해 집계된다. 평균화(averaging)와 같은 다른 학습 불가능한 집계 함수들(non-trainable aggregation functions)이 대안적으로 이용될 수 있다. 제1 분류기(40)는 제1 예시와 동일한 방식으로 학습된다. 전체 슬라이드 이미지에 대한 분류 예측은 단일 "대표" 타일에 대한 분류 예측으로 취해진다.
- [0209] 대안적인 예시들에서, 학습 가능한 집계 함수가 이용된다. 입력 전체 슬라이드 이미지가 바이오마커와 연관되어 있는지 여부의 표시는 복수의 타일들로부터의 데이터를 집계 모델에 입력함으로써 결정된다. 집계 모델은 순환 신경망(recurrent neural network), 예를 들어, 장단기 메모리(LSTM) 네트워크(long-short term memory network)를 포함할 수 있다. 제1 분류기(40)의 출력은 감소된 이미지 타일들의 세트, 예를 들어 제1 분류기(40)로부터 출력된 확률들에 기초한 상위 k개의 타일들을 선택하는 것에 이용될 수 있고 여기서 k는 2보다 큰 정수이다. 감소된 이미지 타일들의 세트는 분류 모델에 대한 입력으로 취해진다. 특징 벡터(feature vector)가 제1 학습된 모델(first trained model)을 이용하여 이미지 타일들 각각에 대응하여 먼저 생성된다. 이들 특징 벡터들은 이어서, 예를 들어 LSTM을 포함할 수 있는 제2 학습된 모델(second trained model)에 입력으로서 순서대로 취해진다. 순서 내의 최종 입력에 대응하는 LSTM의 출력은 입력 이미지가 (양성 클래스에 대응하는) ER 바이오마커와 연관되어 있는 확률을 출력하는 완전 연결 신경망 레이어(fully connected neural network layer)에 입력된다. 이 출력 값은 이어서 제1 분류의 출력을 제공하기 위해 제1 임계치와 비교되고, 제2 분류의 출력을 제공하기 위해 제2 임계치와 비교된다. 전체 슬라이드 이미지에 대한 분류 예측은 이어서 제1 분류 및 제2 분류의 출력에 기초하여 결정된다. 학습 동안, 학습 데이터세트로부터의 이미지 타일들은 추론 동안과 동일한 방식으로 분류 모델에 입력으로 취해진다. 출력은 예를 들어 앞서 기술된 바와 같은 교차 엔트로피 손실 함수 및 최적화 함수를 이용하여, 제1 및 제2 모델들을 학습시키기 위해 전체 슬라이드 이미지에 대한 레이블과 비교된다. 대안적인 예시들에서, 분류 모델은 제1 임계치를 가지는 제1 학습된 모델과 제2 학습된 모델을 포함하는 제1 분류 및 제2 임계치를 가지는 서로 다른 제1 학습된 모델과 서로 다른 제2 학습된 모델을 포함하는 제2 분류를 적용하는 단계를 포함할 수 있다.
- [0210] 제4 예시
- [0211] 전술된 예시들에서, 예측 커버리지 즉, 분류 모델이 예측할 수 있는 사례들의 수의 감소로 인해 모든 클래스들

에 대해 향상된 정확도가 획득된다. 그러나, 일부 애플리케이션들에서, 분류의 방법은 일부 클래스들, 예를 들어, 음성 클래스만 또는 양성 클래스만을 예측하기 위해 이용된다. 예를 들어, 분류의 방법은 이미지가 HER2 바이오마커에 대응하는지 여부를 분류하기 위해 이용될 수 있고, 여기서 모델은 음성 클래스를 예측하기 위해서만 이용된다.

- [0212] 이러한 예시들에서, 분류의 방법은 단일 분류기 및 단일 예측 임계치를 이용할 수 있다. 특정 클래스들에 대해 향상된 정확도가 획득된다. 예를 들어, 분류기는 제1 클래스를 예측하기 위해서만 이용되며, 여기서 제1 클래스에 대해 향상된 정확도가 획득된다. 제1 클래스에 대응하는 임계치가 방법에 이용된다. 분류가 제1 클래스를 예측하면, 제1 클래스의 분류 예측이 만들어진다. 분류 예측이 제1 클래스를 예측하지 않으면, 분류 예측이 만들어지지 않는다.
- [0213] 앞서 기술된 바와 같이, 높은 예측 임계치는 더 많은 음성 사례들을 정확하게 예측할 수 있지만 더 많은 양성 사례들을 정확하게 예측하는 것에 실패하는 높은 특이도를 가지는 분류 방법을 초래한다. 예를 들어, 도 2(a)의 "특이도 임계치" 라인은 데이터세트 내의 모든 진음성 사례들이 정확하게 예측되는 임계치를 도시한다. 낮은 예측 임계치와 결합된 동일한 분류기는 더 많은 양성 사례들을 정확하게 예측할 수 있지만 더 많은 음성 사례들을 정확하게 예측하는 것에 실패하는 높은 민감도를 가지는 분류 방법을 초래한다. 도 2(a)의 "민감도 임계치" 라인은 데이터세트 내의 모든 양성 사례들이 정확하게 예측되는 임계치를 도시한다. "특이도" 임계치는 "민감도" 임계치보다 높다.
- [0214] 도 16은 제4 예시에 따른 방법을 도시한다. 제1 예시와 관련하여 기술된 바와 같이, 분류는 조직의 이미지에 대응하는 입력 이미지 데이터를 통해 수행된다. 입력 이미지는 HER2 바이오마커에 대응하지 않으면 음성 클래스로 분류된다.
- [0215] 제4 예시에 따른 방법은 제1 분류를 이용한다.
- [0216] 단계(S1601)에서, 입력 데이터는 제1 예시와 관련하여 단계(S201)에 기술된 바와 동일한 방식으로 획득된다.
- [0217] 단계(S1602)에서, 입력 데이터는 분류 모델에 입력으로 취해진다. 분류 모델은 제1 분류기를 적용한다. 제1 분류기는 제1 예시의 단계(S202)와 관련하여 기술된 바와 같이, CNN일 수 있다.
- [0218] 단계(S1603)에서, 임계치는 제1 분류기의 출력에 적용된다. 즉, 제1 분류기로부터 출력된 확률 값은 임계치에 대해 비교된다. 임계치는 이 예시에서 음성 클래스에 대한 임계치에 대응한다. 예를 들어, 임계치는 데이터세트를 이용하여 결정된 바와 같은 민감도 임계치에 대응할 수 있다.
- [0219] 단계(S1604)에서, 음성 클래스에 대한 최종 분류 예측은 단계(S1603)에서 출력된 제1 분류에 기초하여 결정된다. 이 단계에서, 제1 분류가 0을 출력하면, 분류 예측은 0이 된다. 제1 분류가 1을 출력하면, 분류 예측은 만들어지지 않는다. 최종 분류를 결정하기 위해 일부 추가적인 단계가 수행될 수 있다. 예를 들어, 이미지는 전문가에 의해 수동으로 분류될 수 있다.
- [0220] 이 예시에서, 제1 분류는 음성 클래스를 예측하기 위해 이용된다. 대안적인 예시에서, 제1 분류는 양성 클래스를 예측하기 위해 이용된다. 이 경우에, 임계치가 양성 클래스에 대한 임계치에 대응하도록, 높은 예측 임계치가 이용된다. 예를 들어, 임계치는 데이터세트를 이용하여 결정된 바와 같은 특이도 임계치에 대응할 수 있다.
- [0221] 도 17은 제4 예시에 따른 분류를 위한 시스템을 학습시키는 방법의 개략도이다. 방법은 전술된 제1 예시에 따른 분류 방법에 이용될 수 있는 시스템을 학습시키는 것에 이용된다.
- [0222] 복수의 입력 데이터를 각각 포함하는 학습 데이터세트 및 검증 데이터세트는 학습 방법에 이용된다. 전술된 바와 같이, 입력 데이터는 분류 방법 동안 사용하기 위한 의도된 유형의 입력 데이터에 대응한다. 전술된 바와 같이, 학습 데이터세트는 제1 모집단에 대응할 수 있고, 검증 데이터세트는 타겟 모집단에 대응할 수 있다. 이 예시에서, 제1 예시와 관련하여 기술된 바와 같이, 입력 데이터는 모델이 검출할 특정 바이오마커에 대응하는지 여부에 따라 레이블링된 조직의 이미지들을 포함한다. 전술된 바와 같이, 이 예시의 특정 바이오마커는 HER2 바이오마커이다. 각 이미지 타일은 바이오마커에 대응하면 1로, 그렇지 않으면 0으로 레이블링된다.
- [0223] 단계(S1701)에서, 학습 데이터세트 내의 이미지 타일들은 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기, 제1 CNN(40)에 제출된다. 전술된 바와 같이, 학습 데이터세트 내의 이미지들의 어노테이션들 및 이미지들에 대한 제1 분류기의 출력들 사이의 차이가 최소화되도록 업데이트된 파라미터 벡터 θ_1 이 결정된다.
- [0224] 단계(S1702)에서, 학습된 모델이 주어지면, 예측 임계치는 이어서 검증 데이터세트를 이용하여 결정된다. 검증

데이터세트 내의 이미지 타일들은 추론(배치) 단계와 관련하여 전술된 바와 동일한 방식으로, 이미지가 양성 클래스에 할당될 확률을 생성하는 제1 분류기, 제1 CNN(40)에 제출된다. 분류기 모델 예측 스코어들, 즉 분류기 모델로부터 출력된 확률 값들은 이어서 이미지들의 레이블들과 비교된다. 이 예시에서, 검증 데이터세트 결과들로부터, (양성 클래스 내의) 레이블 1을 가지는 이미지에 대한 최저 확률 값이 예측 임계치로 취해진다. 예측 임계치는 따라서, 검증 데이터세트를 통해 제1 분류기 모델을 실행한 결과들로부터 결정된 바와 같이, 민감도 임계치로 설정된다. 예측 임계치를 민감도 임계치로 설정함으로써, 분류 방법은 음성 클래스에 대한 정확한 예측들을 생성할 수 있고, 즉, 모델이 음성으로 잘못 예측하는 양성 사례들의 비율인 위음성률이 감소된다.

[0225] 전술된 예시에서, 예측 임계치는 검증 데이터세트로부터 결정된 바와 같이 민감도 임계치로 설정된다. 그러나, 일부 예시들에서, 예측 임계치는 민감도 임계치보다 높은 값으로 설정된다. 예를 들어, 예측 임계치는 최적화 프로세스를 이용하여 결정될 수 있다. 임계치에 대한 가능한 값들의 세트가 획득된다. 이 예시에서, 세트 내의 모든 값들은, 제1 예시의 변형과 관련하여 앞서 기술된 바와 같이, 제1 분류기의 민감도 임계치보다 크거나 같고 제1 분류기의 특이도 임계치보다 작거나 같다.

[0226] 이어서, 스코어가 각 가능한 임계치 값에 대해 결정된다. 스코어는 예를 들어, FNR, 또는 NPV일 수 있다. 테스트 데이터세트로부터의 입력들은 제1 분류기를 통해 실행된다. 제1 분류기 모델 출력들은 분류 예측들을 제공하기 위해 임계치 값에 대해 비교된다. 스코어는 이어서 예를 들어, 예측들의 세트에 대한 FNR을 결정함으로써 임계치에 대해 결정된다. 이는 각 가능한 임계치에 대해 반복된다.

[0227] 애플리케이션에 있어서 원하는 스코어보다 크거나 같은 스코어를 초래하는 모든 임계치들이 검색되고, 임계치는 검색된 임계치들로부터 커버리지에 기초하여 선택된다. 예를 들어, 검색된 임계치들 중에서 테스트 데이터세트로부터 가장 많은 결과를 분류한 임계치가 선택되거나, 또는 최고 임계치가 선택된다.

[0228] 제1 분류가 양성 클래스를 예측하기 위해 이용되는 예시들에서, 예측 임계치는 위양성률을 감소시키기 위해, 검증 데이터세트로부터 결정된 바와 같이 특이도 임계치로 설정될 수 있다. 또는, 임계치는 전술된 바와 유사한 방식으로, 최적화 프로세스에 의해 설정될 수 있다.

[0229] 제4 예시에서, 분류의 방법은 단일 분류기를 이용하고, 예를 들어 타겟 모집단에 대응하는 검증 데이터세트를 이용하여 타겟 모집단에 따라 임계치를 조정할 수 있다. 예시는 하나의 임계치를 가지는 하나의 분류기를 이용하고, 여기서 분류의 방법은 예를 들어 음성 클래스만 예측하거나 양성 클래스만 예측하도록 의도된다. 일부 예시들에서, 분류기는 학습 세트로 학습되고, 학습의 성능은 검증 세트를 통해 검증되고 이어서 임계치는 테스트 세트를 통해 설정된다.

[0230] 표 2는 테스트 데이터세트로부터의 이미지들을 HER2 바이오마커에 대응하지 않는 것으로 분류하는 제4 예시에 따른 제1 분류의 방법을 이용하여 생성된 결과들을 도시한다. 방법은 음성 클래스에 대응하는 단일 임계치를 가지는 전술된 바와 같은 분류기 모델을 이용한다. 표 2는 테스트 데이터세트로부터의 이미지들을 HER2 바이오마커에 대응하는 것으로 분류하는 비교 예시에 따른 제1 분류의 방법을 이용하여 생성된 결과들도 또한 도시한다. 방법은 0.5의 단일 임계치를 가지는 전술된 동일한 분류기 모델을 이용한다.

표 2

	비교 예시	제4 예시
정확도	0.647482014	0.867521368
F1	0.319444444	0.114285714
AUC	0.612370114	0.558557127
FPR	0.331914894	0.009852217
FNR	0.465116279	0.935483871
PPV	0.227722772	0.5
NPV	0.88700565	0.873913043

민감도	0.534883721	0.064516129
특이도	0.668085106	0.990147783
PLR	1.611508646	6.548387097
NLR	0.696193157	0.944792168

[0232] 일반의 환자 및/또는 암의 특정 분자 프로파일을 이해함으로써, 그 중에서도 호르몬 치료법, 면역 요법 또는 표적 약물 치료법을 포함하여 암에 대해 수행된 다양한 절차들을 알 수 있다. 환자가 특정 치료들에 얼마나 잘 반응하는지를 결정하기 위해 임의의 진단법(diagnostics), 치료법(therapeutics) 또는 예후(prognostics) 마커들을 포함하는, 그 중에서도 돌연변이 상태(mutation status), 수용체 상태(receptor status) 또는 유전자 복제수 변이들(copy number variations)를 포함하여 다양한 의학적으로 관련된 바이오마커들이 식별될 수 있다. 돌연변이 상태, 수용체 상태 또는 유전자 복제수 변이들은 분자 바이오마커들의 예시들이다. 예를 들어, 일부 사례들에서 분자 바이오마커(molecular biomarker)는 단백질 발현 수준(protein expression level)일 수 있다. 예를 들어, 특정 바이오마커는 에스트로겐 수용체(ER), 프로게스테론 수용체(PR) 또는 인간 표피 성장 인자 수용체(HER2)일 수 있다. 이러한 필러 바이오마커들(pillar biomarkers)은 유방암(breast cancer)에 특이적이다. 이들은 유방암의 예후(prognosis)에 가장 중요한 바이오마커들이며 표적 치료법(targeted therapies)의 기초에 있다. ER과 HER2는 각각 암 치료제인 타목시펜 및 허셉틴과 가장 일반적으로 연관되어 있다. 환자는 이러한 치료들에 대한 적합성(suitability)을 결정하기 위해 이 두 개의 바이오마커에 대해 테스트될 수 있다.

[0233] 전술된 예시들에서, 입력 이미지는 ER 바이오마커와 연관되어 있는지 여부에 따라 분류된다. 또는, 방법은 입력 이미지 타일이 HER2 바이오마커, PR 바이오마커, (폐 선암종(lung adenocarcinoma)과 연관되어 있는) EGFR 바이오마커, 또는 (대장 선암종(colon adenocarcinoma)과 연관되어 있는) MSI와 연관되어 있는지 여부를 결정하기 위해 이용될 수 있다. 특정 암들을 유방암 또는 결장암(colorectal)과 같은 카테고리들로 분류하기 위해 다양한 분자 바이오마커들이 이용될 수 있다. 예를 들어, 유방암은 다섯 개의 서로 다른 분자 "아형들(subtypes)"을 가지며, 각각은 ER, PR 및 HER2의 상태들에 기초하여 결정된다. 예를 들어, ER, PR 및 HER2가 모두 음성인 경우, 분자 아형은 "기저 세포(basal-like)"이다. 따라서, 다수의 분자 바이오마커들의 존재 유무를 결정함으로써, 분자 아형이 예측될 수 있다. "분자 아형(molecular subtype)"은 특정 유형의 암을, 바이오마커들의 존재 유무, 또는 경우에 따라 하나 또는 일련의 바이오마커들의 레벨에 기초하여 분류하는 방법이다.

[0234] 분류 방법은 다양한 다른 바이오마커들을 검출하기 위해 이용될 수 있다. 예를 들어, 항원 Ki-67은 또한 암의 공격성을 나타내는 세포 증식(cell proliferation)에 대한 마커로서 점점 더 테스트되고 있다. 따라서, 특정 바이오마커는 대안적으로 Ki-67일 수 있다. Ki67 핵 항원(nuclear antigen)의 IHC-염색에 기초한 레이블링 인덱스(labeling index)는 예를 들어, HER2- 및 ER+ 유방암의 종양 증식(tumour proliferation)을 평가할 때 등급 체계들(grading schemes)에서 유사 분열수(mitotic counts)에 대한 대안으로서 다른 IHC 마커들과 함께 이용될 수 있다. 이는 보조항암화학요법(adjuvant chemotherapy)에 대한 임의의 요구사항(requirement)과 같은 치료법 결정들을 위한 추가적인 정보를 제공할 수 있다. 다양한 연구에서, 이는 생존의 강력한 예측 인자(predictor)임을 보여주었다. 예를 들어, PREDICT는 초기 침습성 유방암에 대한 서로 다른 치료법들이 수술 후 생존율을 어떻게 향상시킬 수 있는지를 보여주는 온라인 툴이다. PREDICT 모델 성능은 예측 마커로서의 Ki67의 관여와 함께 향상되었다. IHC-염색된 Ki67 슬라이드들을 해석하기 위한 수동 배점 방법(manual scoring method)은 종양 주변부와 같이 무작위로 선택된 관심 영역에서 침습성 세포들의 수를 세는 단계 및 모든 침습성 종양 세포들에 대한 Ki67 염색의 퍼센트를 결정하는 단계를 포함한다. 전술된 종래의 분자 프로파일링 기술들과 유사하게, 이 프로세스는 노동 집약적이고, 인간 오류들이 발생하기 쉽고, 인터 /인트라 옵저버(inter/-intra observer)에서 개방적이다. 예를 들어, H&E 이미지들로부터의 Ki67 인덱스를 예측함으로써, 이러한 프로세스가 더 짧아지고, 잠재적으로 정확도가 향상될 수 있다.

[0235] 분류 방법은 암의 진단법, 치료법 및/또는 예후와 관련된 특정 바이오마커의 자동 프로파일링(automatic profiling)을 제공할 수 있다. 특정 바이오마커는 다른 예들 중에서도 돌연변이 상태, 수용체 상태 또는 유전자 복제수 변이들일 수 있다. 프로파일링은 다른 이미지들 또는 입력들이 이용될 수 있지만, 기술된 예시들의 전체 슬라이드 H&E 이미지들로부터 수행된다.

- [0236] 바이오마커는 예를 들어 분자, 또는 특정 분자의 양(amount)과 같은 하나 이상의 분자 중 하나와 연관된 특징일 수 있다. 일부 사례에서, 바이오마커는 특정 암 치료와 연관된 분자이다. 바이오마커는 임상적 의미를 갖는 유전자 변이(clinically actionable genetic alteration)일 수 있다. 이 예시에서, 특정 바이오마커는 ER 바이오마커이다. 다른 바이오마커들의 예시들은 이하에서 더 상세히 기술될 것이다.
- [0237] 또는, 분류 예측은 다른 속성들에 기초할 수도 있다. 일부 예시들에서, 분류 방법은 예를 들어, 이미지가 암세포들을 포함하는지 여부를 분류한다.
- [0238] 상술된 예시에서는, 헤마톡실린 및 에오신 염색으로 염색된 조직학적 단면의 이미지가 이용되지만, 다양한 방법들을 이용하여 획득된 다양한 유형의 조직 이미지들은 기술된 방법을 이용하여 처리될 수 있다. 예를 들어, 대안적으로, IHC(Immunohistochemistry) 염색을 한 조직학적 단면의 이미지가 입력으로 취해질 수 있다. IHC 염색은 조직 단면의 세포들에서 항원들을 선택적으로 식별하는 단계를 포함한다. 항체들은 생물학적 조직들 내의 항원들에 특이적으로 결합한다. 염색은 항체-항원 상호작용(antibody-antigen interaction)을 시각화(visualisation)하도록 한다. 예컨대, CIH(chromogenic immunohistochemistry)를 이용하여, 항체는 발색 반응(colour-producing reaction)을 촉매(catalyse)할 수 있는 효소(enzyme)에 접합(conjugate)된다.
- [0239] 전술된 예시들에서, 이진 분류 방법이 수행된다. 그러나, 대안적인 예시들에서, 분류의 방법은 다원 분류(multi-way classification)를 위해 이용될 수 있다. 다원 분류에서, 분류 모델은 다수의 클래스들의 각 클래스에 대해 제1 분류 및 제2 분류를 적용하고, 여기서 제2 분류는 제1 분류보다 클래스에 대해 더 높은 진양성률을 가지도록 구성된다. 검증 데이터셋으로부터 결과들이 주어지면, 클래스에 대한 진양성률은 방법이 클래스 내에 속할 것으로 정확하게 예측하는 클래스 내의 사례들의 비율로서 결정된다.
- [0240] 세 개의 클래스들로 분류가 수행되는 예시가 이하에서 기술될 것이다. 삼원 분류기 모델(three-way classifier)이 입력에 적용된다. 제1 클래스에 대한 제1 분류는 제1 클래스에 대한 출력을 제1 예측 임계치와 비교하는 단계를 포함한다. 제1 클래스에 대한 제2 분류는 제1 클래스에 대한 출력을 제2 예측 임계치와 비교하는 단계를 포함한다. 제1 예측 임계치는 제2 예측 임계치보다 높다. 제2 클래스에 대한 제1 분류는 제2 클래스에 대한 출력을 제3 예측 임계치와 비교하는 단계를 포함한다. 제2 클래스에 대한 제2 분류는 제2 클래스에 대한 출력을 제4 예측 임계치와 비교하는 단계를 포함한다. 제3 예측 임계치는 제4 예측 임계치보다 높다. 제3 클래스에 대한 제1 분류는 제3 클래스에 대한 출력을 제5 예측 임계치와 비교하는 단계를 포함한다. 제3 클래스에 대한 제2 분류는 제3 클래스에 대한 출력을 제6 예측 임계치와 비교하는 단계를 포함한다. 제5 예측 임계치는 제6 예측 임계치보다 높다. 제1 예측 임계치는 제1 클래스에 있지 않은 입력에 대해(즉, 입력은 제2 또는 제3 클래스에 있음), 검증 데이터셋 결과들로부터 제1 클래스에 대한 최고 출력 확률 값으로 설정된다. 제2 예측 임계치는 제1 클래스의 입력에 대해, 검증 데이터셋 결과들로부터 제1 클래스에 대한 최저 확률 값으로 설정된다. 따라서 검증 데이터셋을 통해 분류기를 실행한 결과들로부터 결정된 바와 같이, 제1 예측 임계치(T1)는 제1 클래스에 대해 특이도 임계치로 설정되고, 제2 예측 임계치(T2)는 제1 클래스에 대해 민감도 임계치로 설정된다. 제3 예측 임계치는 제2 클래스에 있지 않은 입력에 대해, 검증 데이터셋 결과들로부터 제2 클래스에 대한 최고 출력 확률 값으로 설정된다. 제4 예측 임계치는 제2 클래스의 입력에 대해, 검증 데이터셋 결과들로부터 제2 클래스에 대한 최저 확률 값으로 설정된다. 제5 예측 임계치는 제3 클래스에 있지 않은 입력에 대해, 검증 데이터셋 결과들로부터 제3 클래스에 대한 최고 출력 확률 값으로 설정된다. 제6 예측 임계치는 제3 클래스의 입력에 대해, 검증 데이터셋 결과들로부터 제3 클래스에 대한 최저 확률 값으로 설정된다.
- [0241] 세 개의 클래스들로 분류가 수행되는 대안적인 예시가 이하에서 기술될 것이다. 제1 클래스에 대한 제1 분류는 제1 이진 분류기(first binary classifier)를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고, 제1 클래스에 대한 제2 분류는 제1 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함하며, 여기서 제1 예측 임계치는 제2 예측 임계치보다 높다. 제1 이진 분류기는 입력이 제1 클래스에 속하는지(양성 클래스) 또는 제1 클래스에 속하지 않는지(음성 클래스) 여부를 결정하도록 학습된다. 제2 클래스에 대한 제1 분류는 제2 이진 분류기(second binary classifier)를 적용하는 단계 및 제3 예측 임계치를 이용하는 단계를 포함하고, 제2 클래스에 대한 제2 분류는 제2 분류기를 적용하는 단계 및 제4 예측 임계치를 이용하는 단계를 포함하며, 제3 예측 임계치는 제4 예측 임계치보다 높다. 제2 이진 분류기는 입력이 제2 클래스에 속하는지(양성 클래스) 또는 제2 클래스에 속하지 않는지(음성 클래스) 여부를 결정한다. 제3 클래스에 대한 제1 분류는 제3 이진 분류기(third binary classifier)를 적용하는 단계 및 제5 예측 임계치를 이용하는 단계를 포함하고, 제3 클래스에 대한 제2 분류는 제3 분류기를 적용하는 단계 및 제6 예측 임계치를 이용하는 단계를 포함하며, 제5 예측 임계치는 제6 예측 임계치보다 높다. 제3 이진 분류기는 입력이 제3 클래스에 속하는지(양성 클래스) 또는 제3 클래스에 속하지 않는지(음성 클래스) 여부를 결정한다. 제1 예시와 관련하여 기술된 바와 같이 임계치

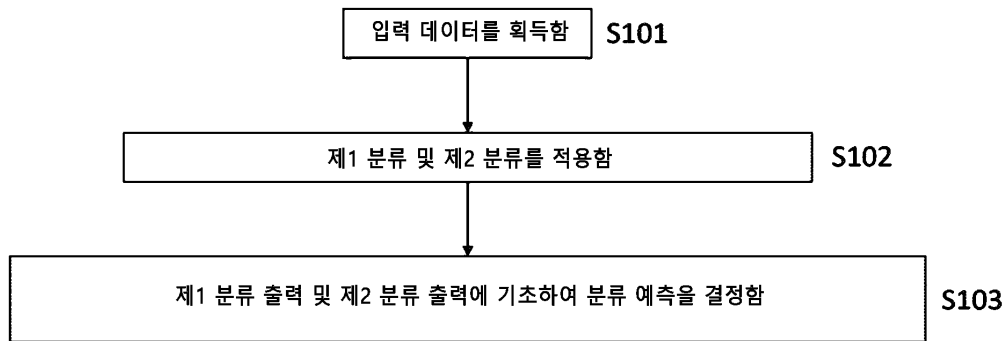
들이 선택되고, 각 분류기는 학습 데이터셋으로 개별적으로 학습된다.

- [0242] 세 개의 클래스들로 분류가 수행되는 대안적인 예시가 이하에서 기술될 것이다. 제1 클래스에 대한 제1 분류는 제1 이진 분류기를 적용하는 단계 및 제1 예측 임계치를 이용하는 단계를 포함하고, 제1 클래스에 대한 제2 분류는 제2 이진 분류기를 적용하는 단계 및 제2 예측 임계치를 이용하는 단계를 포함한다. 제1 이진 분류기 및 제2 이진 분류기는 입력이 제1 클래스에 속하는지 (양성 클래스) 또는 제1 클래스에 속하지 않는지 (음성 클래스) 여부를 결정하도록 학습된다. 제2 클래스에 대한 제1 분류는 제3 이진 분류기를 적용하는 단계 제3 예측 임계치를 이용하는 단계를 포함하고, 제2 클래스에 대한 제2 분류는 제4 이진 분류기를 적용하는 단계 및 제4 예측 임계치를 이용하는 단계를 포함한다. 제3 이진 분류기 및 제4 이진 분류기는 입력이 제2 클래스에 속하는지 (양성 클래스) 또는 제2 클래스에 속하지 않는지 (음성 클래스) 여부를 결정하도록 학습된다. 제3 클래스에 대한 제1 분류는 제5 이진 분류기를 적용하는 단계 및 제5 예측 임계치를 이용하는 단계를 포함하고, 제3 클래스에 대한 제2 분류는 제6 이진 분류기를 적용하는 단계 및 제6 예측 임계치를 이용하는 단계를 포함한다. 제5 이진 분류기 및 제6 이진 분류기는 입력이 제3 클래스에 속하는지 (양성 클래스) 또는 제3 클래스에 속하지 않는지 (음성 클래스) 여부를 결정하도록 학습된다. 제2 예시와 관련하여 전술된 바와 같이 제1 및 제2 분류에 이용되는 분류기들은 각 클래스에 대해 선택되고, 각 클래스에 대응하는 세 개 이상의 분류기들은 학습 데이터셋으로 개별적으로 학습된다. 각 클래스에 있어서, 양성 클래스는 두 개의 분류기들에 대해 동일하다.
- [0243] 도 14는 분류의 방법의 개략도이다. 해당 방법은 테스트 데이터셋에 대한 이전 성능에 기초하여, 이진 예측 값뿐만 아니라 예측이 정확한지 여부의 지표를 출력하는 예측 의료 장치(predictive medical device, PMD)를 사용(deploy)하는 단계를 포함한다. PMD가 확실한 사례들에 대해서는, 해당 사례에 대한 예측이 제공된다. PMD가 확실하지 않은 사례들에 대해서는, 예측이 제공되지 않고 해당 사례는 표준 치료(standard of care)로 계속될 것이다.
- [0244] 단계(S1301)에서, 분석될 입력 의료 샘플 또는 데이터가 획득된다. 단계(S1302)에서, 입력은 PMD에 입력된다. 단계(S1303)에서, 확실성(certainty)과 함께, 예측 값(1 또는 0)이 출력된다. 제1 분류 및 제2 분류가 일치하면, 확실성은 1이고, 방법은 단계(S1305)로 이동한다. 분석 (예측 값)은, 표준치료 분석을 바이패싱(bypassing)하면서 제공된다. 제1 분류 및 제2 분류가 일치하지 않으면, 확실성은 0이고, 방법은 단계(S1304)로 이동한다. 이어서 표준치료 분석이 수행되고, 이는 수동 검사를 포함할 수 있다. 이 분석은 시간이 소요되거나 또는 비쌀 수 있다. 일단 분석이 제공되면, 단계(S1305) 또는 단계(S1034) 중 어느 하나에서, 방법은 단계(S1306)으로 이동하고 이후 프로세스(subsequent process)를 수행한다.
- [0245] 도 15는 일 실시예에 따른 분류 시스템(1)의 개략도를 도시한다. 시스템(1)은 입력(11), 프로세서(3), 워킹 메모리(working memory)(9), 출력(13), 및 스토리지(7)를 포함한다. 시스템(1)은 입력 이미지 데이터를 취하여 출력을 생성한다. 출력은 진단 정보를 포함할 수 있다. 특히, 출력은 입력 이미지가 바이오마커와 연관되어 있는지 여부를 표시할 수 있다.
- [0246] 시스템(1)은 컴퓨팅 시스템, 예를 들어 범용 컴퓨터, 또는 서버와 같은 엔드 유저 시스템(end-user system)이다.
- [0247] 프로세서(3)는 스토리지(7)에 커핑되고 RAM(9)을 포함하는 워킹 메모리에 액세스한다. 프로세서(3)는 워킹 메모리에 저장된 코드의 명령어들에 응답하고 이를 처리하는 로직 회로(logic circuitry)를 포함할 수 있다.
- [0248] 컴퓨터 프로그램(5)은 비휘발성 메모리(7)에 저장된다. 비휘발성 메모리(7)는 프로세서(3)에 의해 액세스되고 저장된 코드(5)는 프로세서(3)에 의해 검색되고 실행된다. 특히, 실행될 때, 본 명세서에 기술된 방법들을 구현하는 컴퓨터 프로그램 코드(5)는 워킹 메모리(9)에 저장된 소프트웨어 제품으로서 대표(represent)된다. 프로세서(3)에 의한 코드(5)의 실행은 본 명세서에 기술된 바와 같은 실시예들이 구현되게 할 것이다.
- [0249] 프로세서(3)는 또한 입력 모듈(11) 및 출력 모듈(13)에 액세스한다. 입력 모듈들 또는 인터페이스들(11) 및 출력 모듈들 또는 인터페이스들(13)은 단일 컴포넌트일 수 있거나, 또는 별도의 입력 인터페이스(11) 및 별도의 출력 인터페이스(13)로 나뉘어질 수 있다.
- [0250] 입력 모듈(11)은 데이터를 수신하기 위해 입력(15)에 연결된다. 입력(15)은 외부 저장 매체로부터 또는 통신 네트워크를 통해 데이터를 수신하기 위한 수신기(receiver)일 수 있다. 또는, 입력(15)은 이미지 캡처 장치와 같은 하드웨어를 포함할 수 있다. 또는, 입력(15)은 시스템에 또는 플로피 디스크, CD ROM, 자기 디바이스 또는 프로그래머블 메모리 디바이스와 같은 별도의 저장 매체에 저장될 수 있는 저장된 이미지 파일로부터 데이터를 읽을 수 있다.

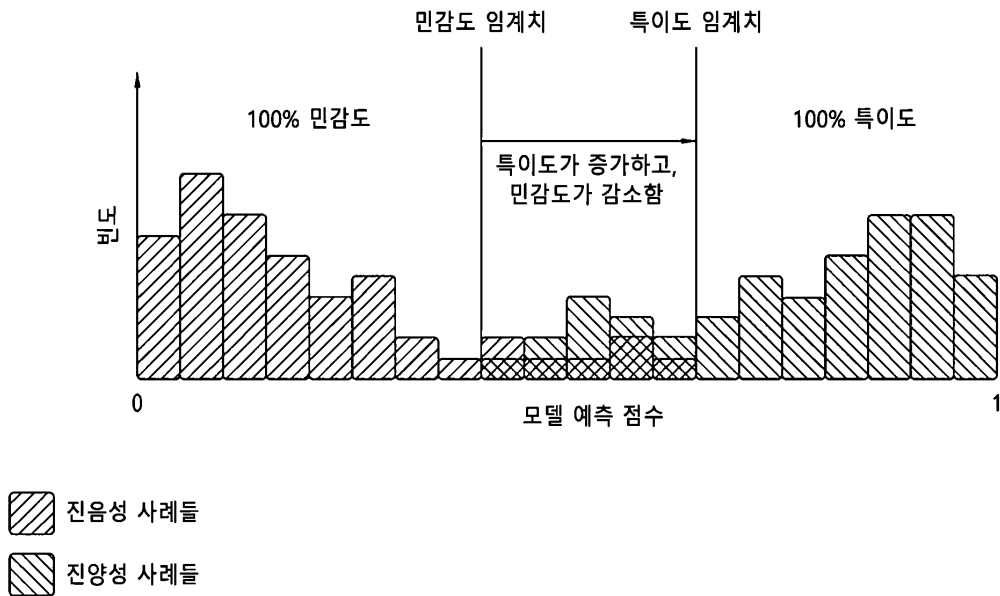
- [0251] 출력 모듈(13)에는 출력(17)이 연결된다. 출력(17)은 시각적인 디스플레이와 같은 하드웨어를 포함할 수 있다. 또는, 출력은 외부 저장 매체로 또는 통신 네트워크를 통해 데이터를 송신하기 위한 송신기(transmitter)일 수 있다. 또는, 출력(17)은 시스템에 또는 플로피 디스크, CD ROM, 자기 디바이스 또는 프로그래머블 메모리 디바이스와 같은 별도의 저장 매체에 저장될 수 있는 저장된 이미지 파일에 데이터를 기록할 수 있다.
- [0252] 스토리지(7)는 프로세서(3)와 통신 방식으로(communicatively) 커플링된다. 스토리지(7)는 프로세서(3)에 의해 실행될 때 코드(5)에 의해 이용되는 데이터를 포함할 수 있다. 도시된 바와 같이, 스토리지(7)는 디바이스 내에 포함되는 로컬 메모리이다. 대안적으로, 그러나, 스토리지(7)는 예를 들어 (인터넷과 같은) 통신 네트워크를 통해 원격으로 액세스될 수 있는 클라우드 기반 메모리를 이용하여 전체적으로 또는 부분적으로 원격지에 (remotely) 위치할 수 있다. 코드(5)는 또한 스토리지(7)에 저장된다. 코드(5)는 실행될 때 워킹 메모리(9)에 배치된다.
- [0253] 시스템(1)은 데이터를 입력 및 출력하는 하드웨어와 함께 공통 시스템(common system)에 위치할 수 있다. 또는, 시스템(1)은 별도의 유닛(예를 들어, 이미지 캡처 장치)으로부터 전송된 이미지 데이터를 수신하고, 출력 데이터를 다른 별도의 유닛(예를 들어, 화면을 포함하는 이용자 컴퓨터)으로 전송하는 원격 시스템(1)일 수 있다. 예를 들어, 시스템은 데이터를 수신하고 전송하는 클라우드 컴퓨팅 시스템 상에 구현될 수 있다.
- [0254] 기술된 시스템에서는, 디바이스에 위치한 단일 프로세서(3)가 이용되지만, 시스템은 동일한 시스템에 위치되거나 원격지에 위치할 수 있는, 프로세싱(processing)의 서로 다른 부분들을 수행하고 그들 사이에서 데이터를 전송하도록 구성된 두 개 이상의 프로세서들을 포함할 수 있다. 예를 들어, 시스템은 그래픽 처리 장치(GPU) 및 일반적인 중앙 처리 장치(CPU)를 포함할 수 있다. 다양한 연산들이 GPU에 의해 구현되는 반면, 다른 연산들은 CPU에 의해 구현된다. 예를 들어, 매트릭스 연산들이 GPU에 의해 수행될 수 있다.
- [0255] 소프트웨어를 메모리에 로딩하고 데이터를 스토리지 유닛(7)에 저장하는 일반적인 절차들이 적용된다. 코드(5)는 기존 장비에 내장될 수 있거나, 제조 후에 전체적으로 또는 부분적으로 제공될 수 있다. 예를 들어, 코드는 전체적으로 다운로드의 형태일 수 있는 컴퓨터 프로그램 제품으로서 도입될 수 있거나, 또는 광 디스크와 같은 컴퓨터 프로그램 스토리지 매체를 통해 도입될 수 있다. 또는, 기존 대화 관리자 소프트웨어(dialogue manager software)에 대한 수정들은 기술된 실시예들의 특징들을 제공하기 위해 업데이트, 또는 플러그인(plugin)에 의해 이루어질 수 있다.
- [0256] 분류 모델을 학습시키는 방법 및 학습된 분류 모델을 이용한 분류의 방법은 동일한 시스템 상에서 수행될 수 있다. 예를 들어, 분류 모델을 학습시키는 방법 및 학습된 분류 모델을 이용한 분류의 방법은 도 15와 관련하여 기술된 바와 같은 시스템 상에서 수행될 수 있으며, 여기서 학습된 분류 모델은 일단 학습되면 메모리(7)에 저장된다. 또는, 분류 모델을 학습시키는 방법은 제1 시스템에서 수행될 수 있고, 분류의 방법이 제2 시스템에서 수행되도록 상기 학습된 분류 모델은 제2 시스템 내에 저장될 수 있다.
- [0257] 기술된 실시예들이 임의의 컴퓨팅 시스템에 적용 가능하다는 것은 인지하겠지만, 도 15에 도시된 예시적인 컴퓨팅 시스템은 본 명세서에 기술된 바와 같은 실시예를 실행에 옮길 수 있는 수단들을 제공한다. 이용 시에, 시스템(1)은 데이터 입력(11)을 통해 데이터를 수신한다. 프로세서(3) 상에서 실행되는 프로그램(5)은 상기 도면들을 참조하여 기술된 방식으로 데이터를 출력(13)을 통해 출력한다.
- [0258] 특정 실시예들이 기술되었지만, 실시예들은 단지 예시적으로 제시된 것일 뿐이고, 발명들의 범위를 제한하기 위한 것은 아니다. 실제로 본 명세서에 기술된 신규한 방법들 및 장치들은 다양한 다른 형태들로 구현될 수 있다; 나아가, 본 명세서에 기술된 방법들 및 장치들의 형태에서 다양한 생략, 치환 및 변경이 이루어질 수 있다.

도면

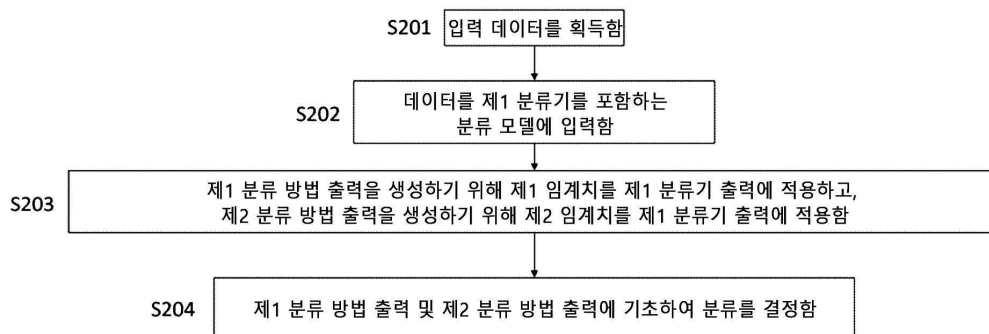
도면1



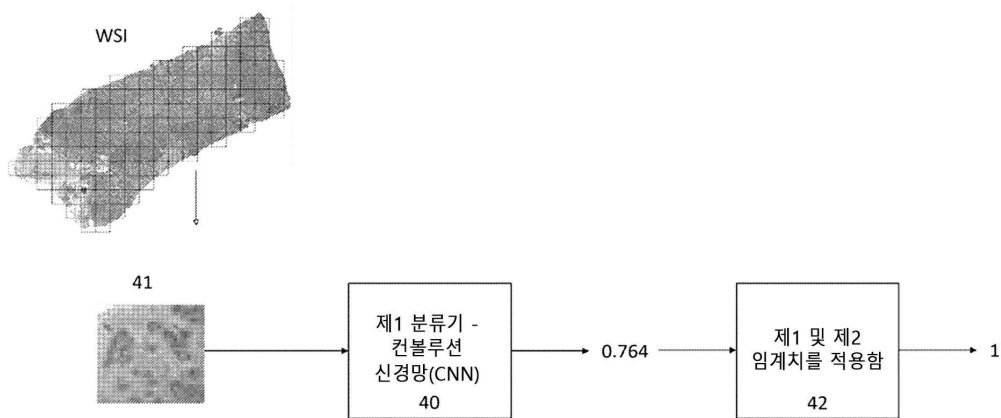
도면2a



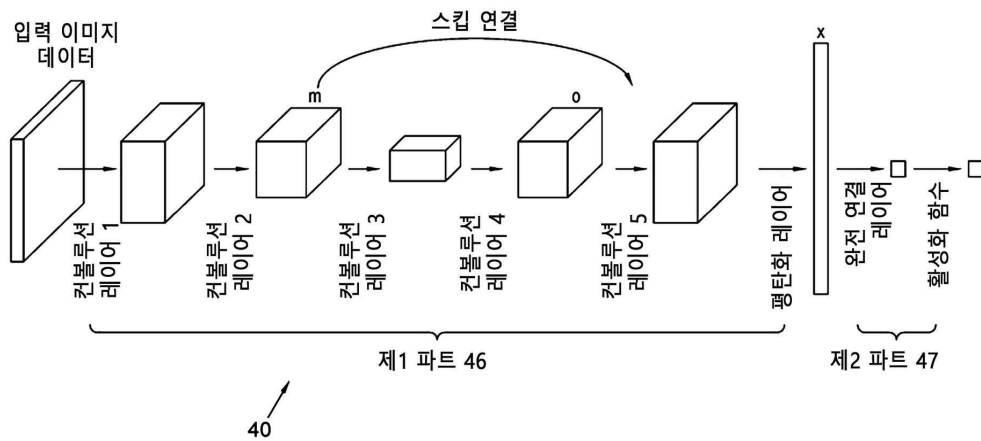
도면2b



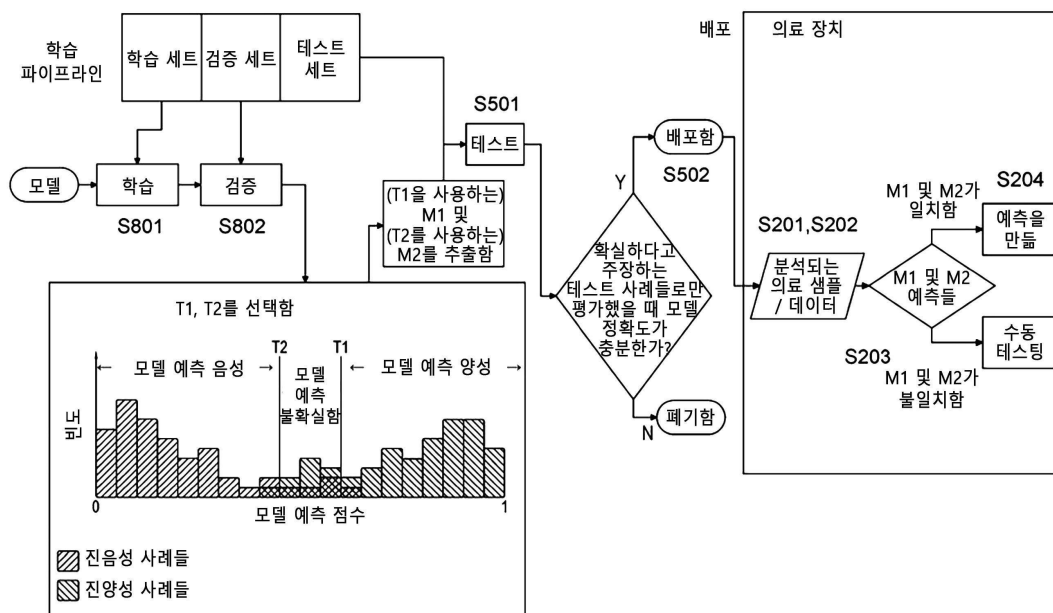
도면3



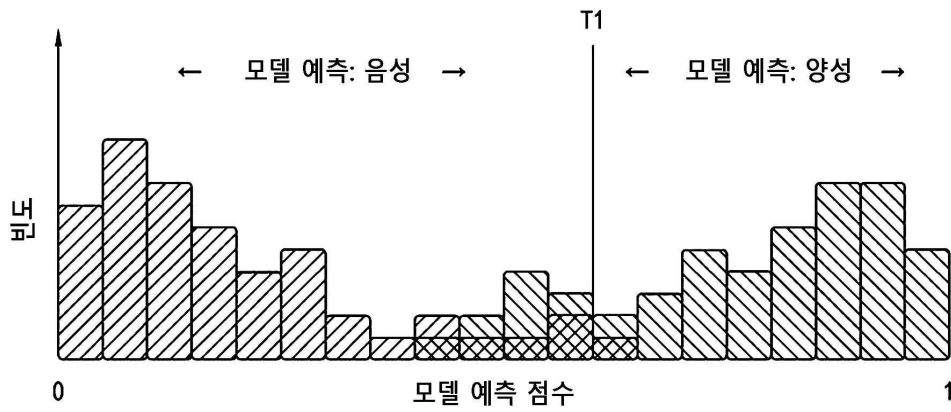
도면4



도면5

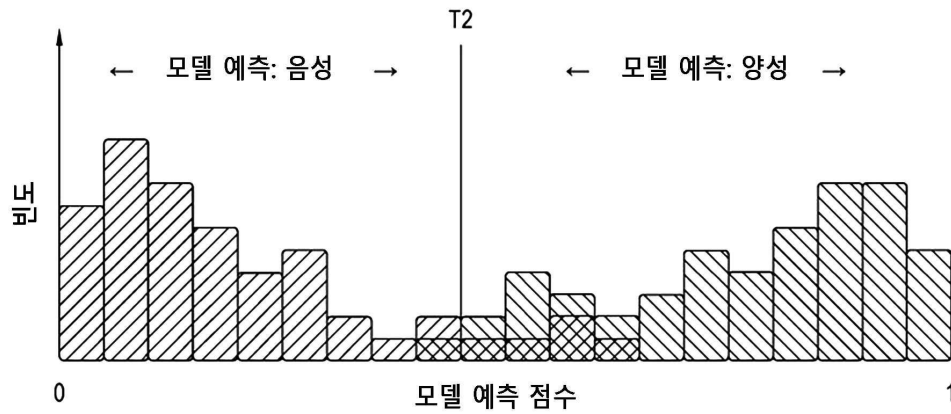


도면6



-  진음성 사례들
-  진양성 사례들

도면7

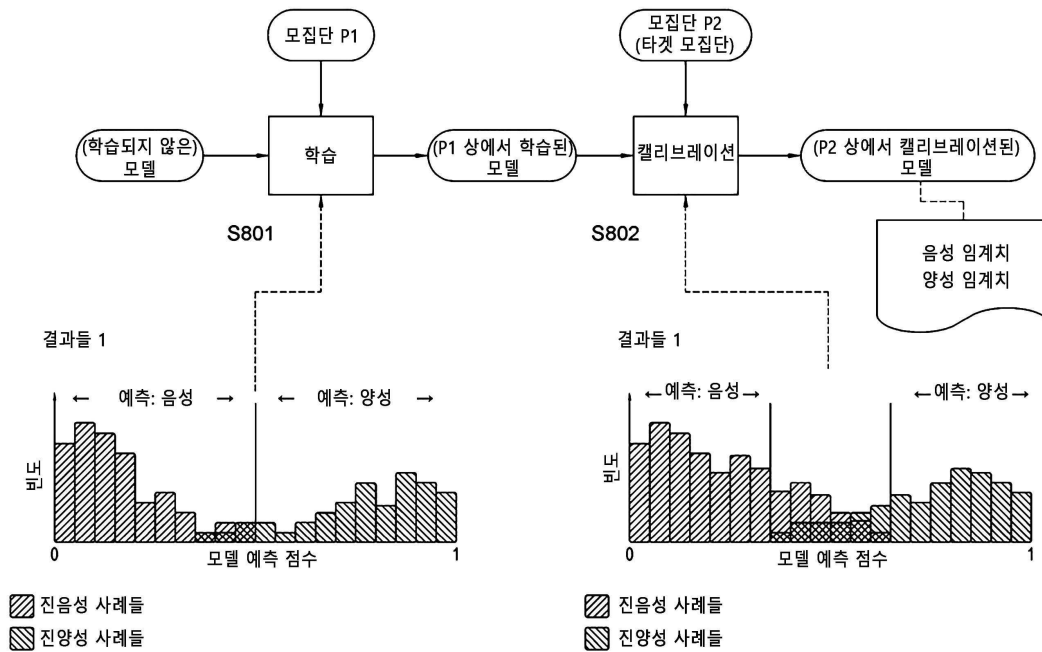


-  진음성 사례들
-  진양성 사례들

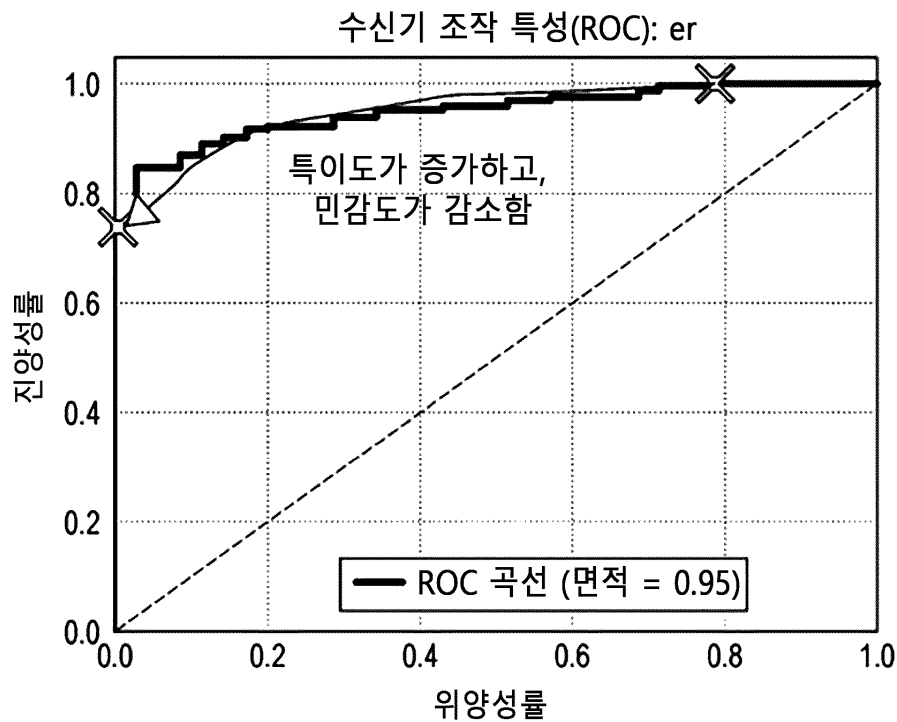
도면8a



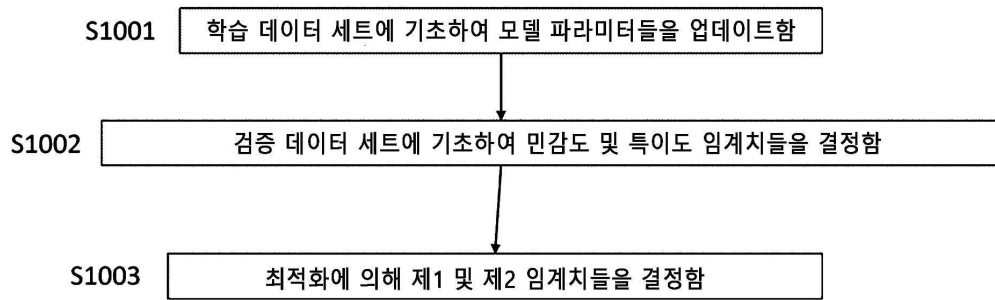
도면8b



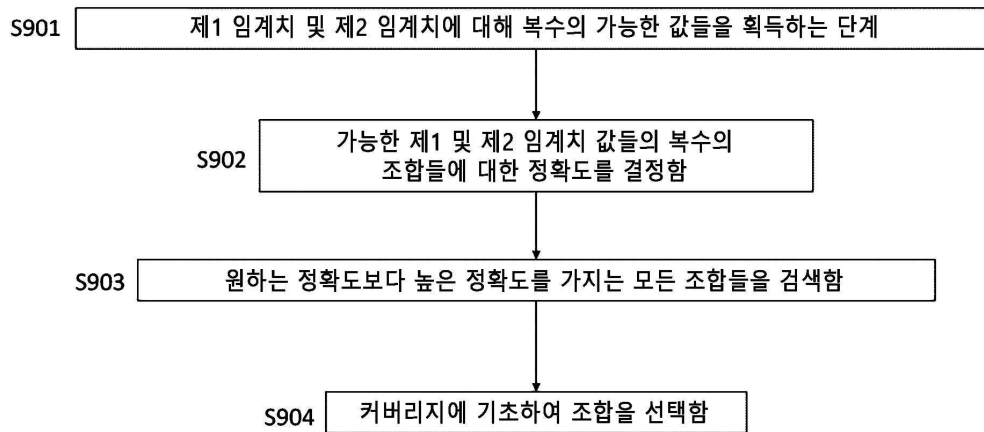
도면9



도면10a



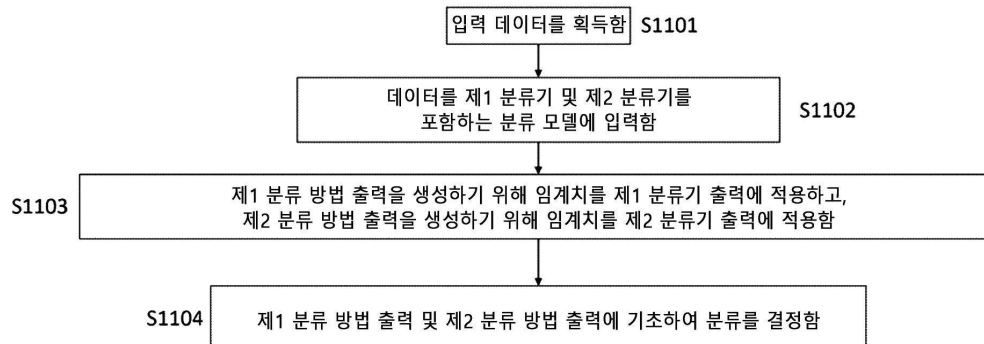
도면10b



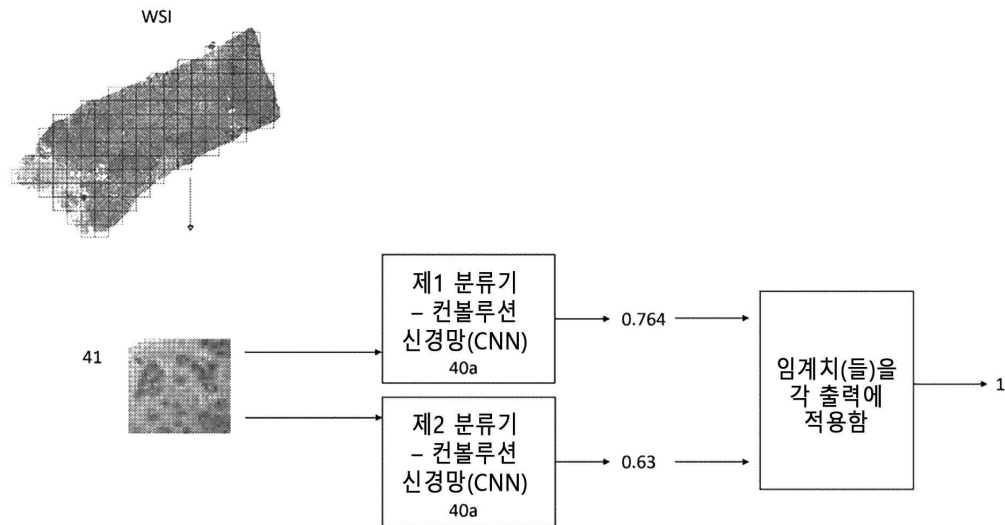
도면10c

T1 T2	0.425	0.455	0.475	0.500	0.525	0.550	0.575	0.600	0.625	0.650
0.400	0.34	0.34	0.38	0.61	0.72	0.85	0.91	0.96	0.98	0.99
0.425		0.31	0.37	0.58	0.74	0.83	0.89	0.94	0.96	0.97
0.450			0.33	0.53	0.72	0.81	0.87	0.87	0.95	0.95
0.475				0.47	0.69	0.75	0.80	0.83	0.92	0.93
0.500					0.60	0.65	0.75	0.82	0.82	0.84
0.525						0.42	0.61	0.63	0.63	0.66
0.550							0.35	0.37	0.39	0.41
0.575								0.37	0.40	0.40
0.600									0.37	0.38
0.625										0.32

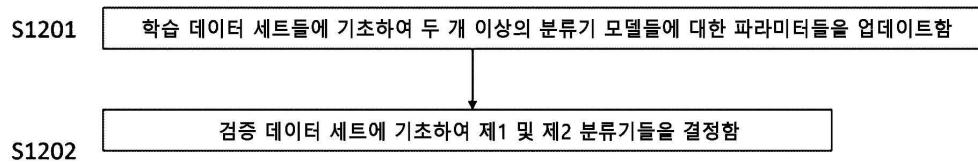
도면11a



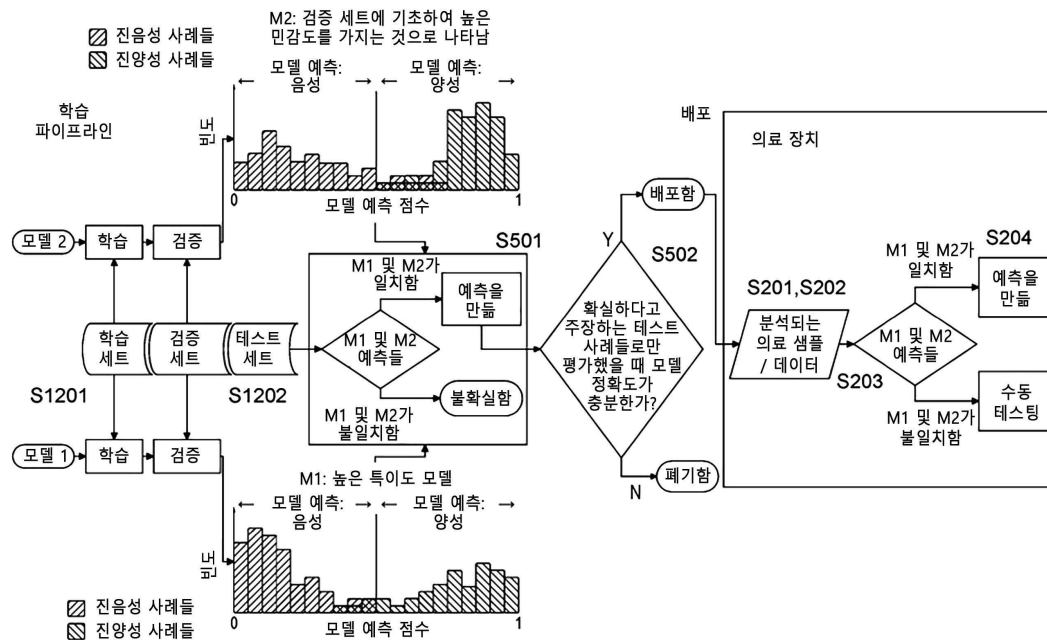
도면11b



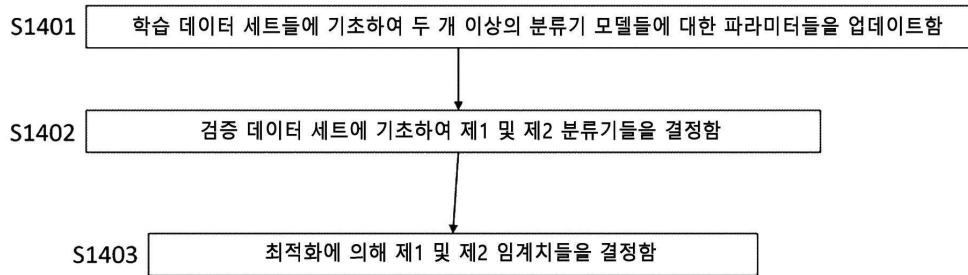
도면12a



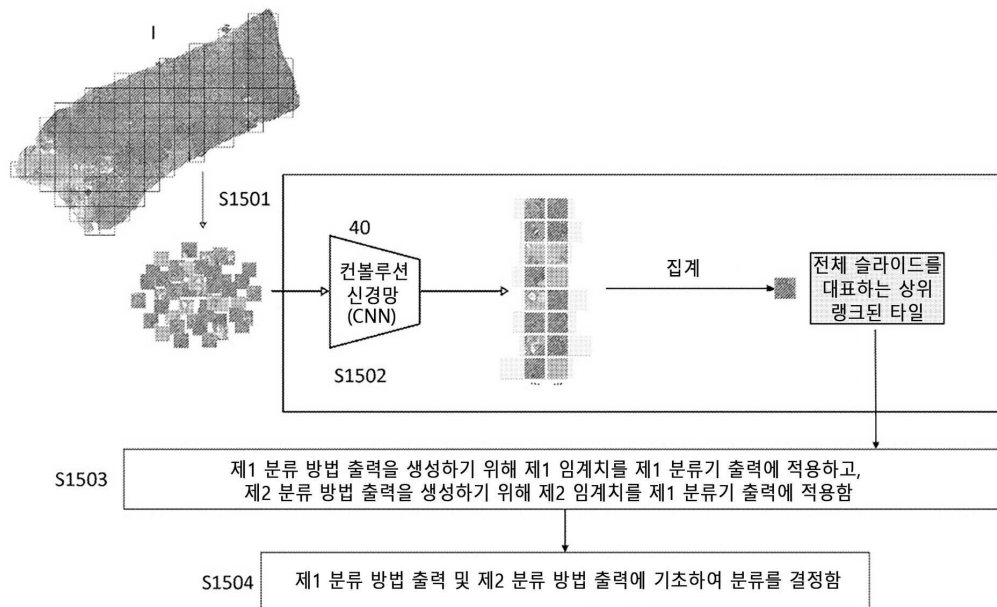
도면12b



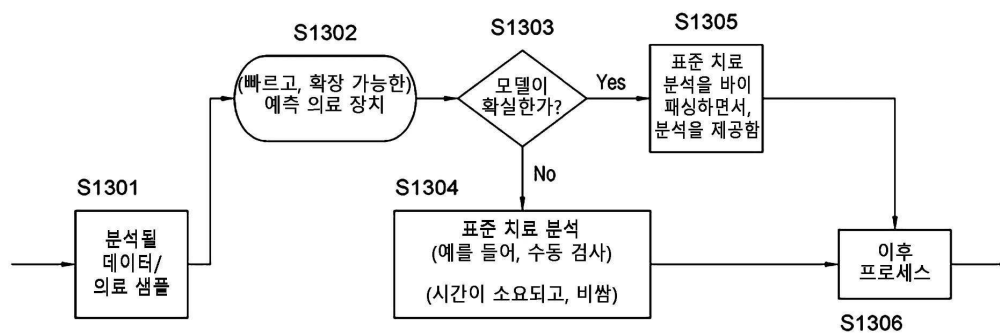
도면12c



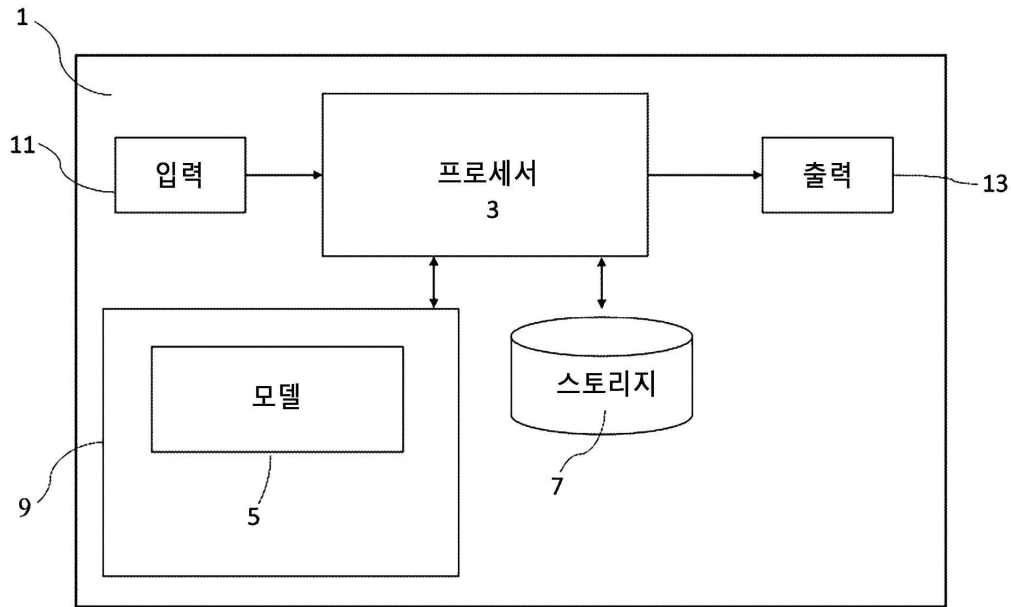
도면13



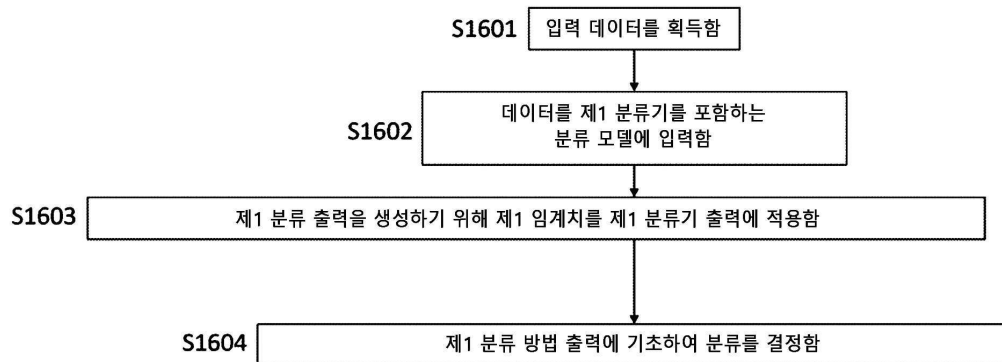
도면14



도면15



도면16



도면17

