



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0126440
(43) 공개일자 2023년08월30일

(51) 국제특허분류(Int. Cl.)

A61B 6/00 (2006.01) A61B 5/00 (2021.01)
A61B 5/08 (2006.01) G06T 7/00 (2017.01)
G16H 50/20 (2018.01)

(52) CPC특허분류

A61B 6/5217 (2020.08)
A61B 5/08 (2013.01)

(21) 출원번호 10-2022-0023596

(22) 출원일자 2022년02월23일

심사청구일자 2022년02월23일

(71) 출원인

인하대학교 산학협력단

인천광역시 미추홀구 인하로 100(용현동, 인하대학교)

(72) 발명자

이상철

인천광역시 연수구 아트센터대로97번길 76, 1402동 805호(송도동, 송도 더샵 하버뷰 14단지)

남주현

전라남도 순천시 팔마1길 27, 104동 302호(연향동, 블루시아파트)

(74) 대리인

양성보

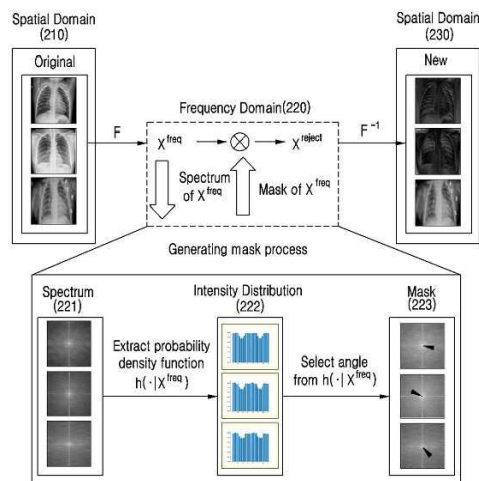
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 폐렴 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법 및 장치

(57) 요약

폐렴 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법 및 장치가 제시된다. 본 발명에서 제안하는 폐렴 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법은 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계, 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계 및 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계를 포함한다.

대표도 - 도2



(52) CPC특허분류

A61B 5/7257 (2013.01)

A61B 5/7264 (2023.08)

A61B 6/50 (2013.01)

G06T 7/0012 (2013.01)

G16H 50/20 (2018.01)

G06T 2207/10116 (2013.01)

G06T 2207/20056 (2013.01)

G06T 2207/30061 (2013.01)

명세서

청구범위

청구항 1

푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계;

마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계; 및

역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계

를 포함하는 데이터 증강 방법.

청구항 2

제1항에 있어서,

상기 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계는,

푸리에 변환된 영상의 푸리에 스펙트럼에서 획득한 확률 밀도 함수를 정의하고, 상기 확률 밀도 함수에 기초하여 상기 푸리에 변환된 영상에 대해 랜덤으로 각도를 샘플링하여 주파수 영역을 마스킹하는

데이터 증강 방법.

청구항 3

제1항에 있어서,

상기 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계는,

상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하기 위해 원점과 마스크 사이의 최단 거리, 원점으로부터 가장 긴 거리와 가장 짧은 거리의 차이, 마스크의 직선 에지 사이의 각도를 포함하는 파라미터를 이용하여 마스크를 극좌표계에서 지시 함수로 정의하고, 생성된 마스크를 이용하여 주파수 영역에서 기초 단위 (elementary-wise)로 곱하여 해당 주파수 패턴을 제거하는

데이터 증강 방법.

청구항 4

제1항에 있어서,

상기 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계는,

제거된 주파수 패턴의 영상에 2D 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하고, 변환된 공간 영역을 새로운 데이터셋으로 추가하여 영상 분류를 위한 학습에 이용하는

데이터 증강 방법.

청구항 5

입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 푸리에 변환부;

선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서

상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 마스크 생성부; 및

상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 역 푸리에 변환부

를 포함하는 데이터 증강 장치.

청구항 6

제5항에 있어서,

상기 푸리에 변환부는,

푸리에 변환된 영상의 푸리에 스펙트럼에서 획득한 확률 밀도 함수를 정의하고, 상기 확률 밀도 함수에 기초하여 상기 푸리에 변환된 영상에 대해 랜덤으로 각도를 샘플링하여 주파수 영역을 마스킹하는

데이터 증강 장치.

청구항 7

제5항에 있어서,

상기 마스크 생성부는,

상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하기 위해 원점과 마스크 사이의 최단 거리, 원점으로부터 가장 긴 거리와 가장 짧은 거리의 차이, 마스크의 직선 에지 사이의 각도를 포함하는 파라미터를 이용하여 마스크를 극좌표계에서 지시 함수로 정의하고, 생성된 마스크를 이용하여 주파수 영역에서 기초 단위(elementary-wise)로 곱하여 해당 주파수 패턴을 제거하는

데이터 증강 장치.

청구항 8

제5항에 있어서,

상기 역 푸리에 변환부는,

제거된 주파수 패턴의 영상에 2D 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하고, 변환된 공간 영역을 새로운 데이터셋으로 추가하여 영상 분류를 위한 학습에 이용하는

데이터 증강 장치.

발명의 설명

기술 분야

[0001] 본 발명은 페렘 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법 및 장치에 관한 것이다.

배경 기술

[0002] 심층 신경망 모델은 특히 대규모 데이터셋을 사용할 수 있는 경우 뛰어난 성능으로 작동하는 머신러닝 시스템에 필수적인 도구이다. 데이터 증강은 증강 방법을 랜덤으로 적용하여 영상의 수와 다양성을 증가시키는 효과적인 방법이다. 공간 영역에서 표준 증강 방법에는 공간 영역을 랜덤으로 자르는 것, 영상을 몇 픽셀씩 변환하는 것, 영상을 수평 또는 수직으로 뒤집는 것, 영상 회전이 포함된다. 직관적으로 데이터 증강은 영상을 인식할 때 학습 속성의 불변성이나 위치와 크기의 패턴을 개선하는 것으로 생각할 수 있다. 게다가, 정규화 방법은 심층 신경망 모델의 일반화를 개선하기 위해 사용된다. 그러나 이것은 여전히 데이터셋 부족 문제에 대한 근본적인 해결책은 아니다. 따라서 최근 연구는 이 문제를 직접 해결할 수 있는 데이터 증강 방법에 초점을 맞추고 있다.

[0003] 최근에는 여러 가지 데이터 증강 방법이 제안되어 다양한 과제와 심층 신경망 모델에 사용되고 있다. 종래기술에서는 영상에서 직사각형 영역을 랜덤으로 선택하고 객체 감지 또는 영상 분류를 위해 임의의 값 또는 ImageNet 평균 픽셀 값으로 픽셀을 지우는 방법을 제안했다. 데이터 영역에 특정한 확률 분포에 따라 다른 클래스의 데이터와 레이블을 가중치와 결합함으로써 혼합 방법이 제안되었다. 적대적 생성망은 데이터 증강에 사용

할 수 있으며, 데이터셋에 따라 추가 데이터를 생성할 수 있다. 다른 딥러닝 기반 방법에는 사전 정의된 정책 경사(policy gradient)를 통해 각 파라미터와 데이터 증강 연산자를 결합하는 최적화 방법을 제시하는 강화 학습이 포함된다.

선행기술문헌

비특허문헌

- [0004] (비특허문헌 0001) [1] H. Wang, X. Wu, Z. Huang, and E. P. Xing, "High-frequency component helps explain the generalization of convolutional neural networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- (비특허문헌 0002) [2] H. Chen, Y. Wang, H. Shu, Y. Tang, C. Xu, B. Shi, and C. Xu, "Frequency domain compact 3d convolutional neural networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- (비특허문헌 0003) [3] S. Bian, T. Wang, M. Hiromoto, Y. Shi, and T. Sato, "Ensei: Efficient secure inference via frequency-domain homomorphic convolution for privacy-preserving visual recognition," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
- (비특허문헌 0004) [4] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in Advances in Neural Information Processing Systems 25, pp. 2960-2968, 2012.

발명의 내용

해결하려는 과제

- [0005] 본 발명이 이루고자 하는 기술적 과제는 컨볼루션 신경망 기반 접근법에서 영상 분류에 사용할 수 있는 새로운 데이터 증강 방법 및 장치를 제공하는데 있다. 본 발명은 공간 영역보다 고주파 패턴이 많은 주파수 영역에서 주어진 영상의 주파수 영역 패턴을 임의로 변경하여 보다 다양한 패턴을 획득하여, 주파수 영역에서 의미 있는 패턴을 선택하고 특정 패턴을 제거한다.

과제의 해결 수단

- [0006] 일 측면에 있어서, 본 발명에서 제안하는 패림 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법은 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계, 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계 및 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계를 포함한다.
- [0007] 상기 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계는 푸리에 변환된 영상의 푸리에 스펙트럼에서 획득한 확률 밀도 함수를 정의하고, 상기 확률 밀도 함수에 기초하여 상기 푸리에 변환된 영상에 대해 랜덤으로 각도를 샘플링하여 주파수 영역을 마스킹한다.
- [0008] 상기 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계는 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하기 위해 원점과 마스크 사이의 최단 거리, 원점으로부터 가장 긴 거리와 가장 짧은 거리의 차이, 마스크의 직선 에지 사이의 각도를 포함하는 파라미터를 이용하여 마스크를 극좌표계에서 지시 함수로 정의하고, 생성된 마스크를 이용하여 주파수 영역에서 기초 단위(elementary-wise)로 곱하여 해당 주파수 패턴을 제거한다.
- [0009] 상기 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계는 제거된 주파수 패턴의 영상에 2D 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환

하고, 변환된 공간 영역을 새로운 데이터셋으로 추가하여 영상 분류를 위한 학습에 이용한다.

[0010] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 패텀 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 장치는 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 푸리에 변환부, 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 마스크 생성부 및 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 역 푸리에 변환부를 포함한다.

발명의 효과

[0011] 본 발명의 실시예들에 따르면 컨볼루션 신경망 기반 접근법에서 영상 분류에 사용할 수 있는 패텀 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법 및 장치를 통해 공간 영역보다 고유 패턴이 많은 주파수 영역에서 주어진 영상의 주파수 영역 패턴을 임의로 변경하여 보다 다양한 패턴을 획득할 수 있다.

[0012] 본 발명의 실시예에 따르면, 주파수 영역에서 의미 있는 패턴을 선택하고 특정 패턴을 제거하며, 활성화 가중 맵의 시각화를 이용하여 X선 기반 COVID-19 진단에 사용될 때 공간적 중요성에 대한 중요한 단서를 제공할 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일 실시예에 따른 패텀 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법을 설명하기 위한 흐름도이다.

도 2는 본 발명의 일 실시예에 따른 데이터 증강 과정의 전체 개념을 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시예에 따른 주파수 필터링기반 데이터 증강 알고리즘을 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 패텀 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 장치의 구성을 나타내는 도면이다.

도 5는 본 발명의 일 실시예에 따른 질병별 시각화 결과의 예시를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0014] 딥러닝을 기반으로 과제를 해결하려면 충분히 큰 데이터셋이 필요하다. 그러나 딥러닝에 의료 데이터셋을 사용하는 것은 접근이 제한적이기 때문에 어려움이 있다.

[0015] 본 발명에서는 주파수 영역의 새로운 데이터 증강 방법을 제안한다. 영상 분류를 위해 CNN 기반 심층 신경망 모델의 주파수 영역에서 데이터 증강을 수행하는 기술에 관한 종래기술은 없었다. 본 발명과 종래기술과의 주요 차이점은 종래기술에서는 직사각형 영역을 임의로 삭제하는 반면, 본 발명은 주파수 영역에서 의미 있는 패턴을 선택하고 특정 패턴을 제거한다는 것이다.

[0016] 본 발명은 컨볼루션 신경망 기반 접근법에서 영상 분류에 사용할 수 있는 새로운 데이터 증강 패러다임을 제안한다. 주파수 영역은 공간 영역보다 고유 패턴이 많아 주어진 영상의 주파수 영역 패턴을 임의로 변경하여 보다 다양한 패턴을 획득할 수 있다. 먼저, 주파수 영역의 의미 있는 패턴은 일반적으로 푸리에 스펙트럼(Fourier spectrum)의 고강도 영역(High-Intensity Regions)에 분포한다고 가정한다. 다음으로, 푸리에 스펙트럼에서 의미 있는 패턴을 제거하는 각도를 선택한다. 이후, 새로운 마스크 패턴을 생성하고 고속 푸리에 변환(Fast Fourier transform) 영상에서 특정 주파수 패턴을 제거한 다음, 역 고속 푸리에 변환(Inverse Fast Fourier transform)을 제거된 주파수 패턴의 영상에 적용하여 그것을 공간 영역으로 다시 변환한다.

[0017] 본 발명의 실시예에 따르면, 입력 영상의 영역 특성을 기반으로 주파수 영역에서 새로운 마스크 패턴을 이용한 마스크 기법을 제안한다. 제안하는 방법이 영상 분류를 위해 다른 메트릭으로 평가할 때 CNN 기반 신경망에 적용된 기존 증강 방법을 능가한다는 것을 실험적으로 입증했다. 또한, 활성화 가중 맵의 시각화는 X선 기반 COVID-19 진단에 사용될 때 공간적 중요성에 대한 중요한 단서를 제공할 수 있다.

[0018] 제안하는 방법은 다양한 백본을 사용하여 X선 영상에서 일관된 성능 향상을 달성한다. 평균적으로 종래기술보다 정확도, 정밀도, 리콜, F1-점수, AUC에서 3.30%, 7.09%, 7.75%, 8.14%, 4.91% 향상되었다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

- [0020] 도 1은 본 발명의 일 실시예에 따른 패턴 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법을 설명하기 위한 흐름도이다.
- [0021] 제안하는 패턴 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 방법은 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택하는 단계(110), 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거하는 단계(120) 및 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하는 단계(130)를 포함한다.
- [0022] 단계(110)에서, 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택한다. 이때, 푸리에 변환된 영상의 푸리에 스펙트럼에서 획득한 확률 밀도 함수를 정의하고, 상기 확률 밀도 함수에 기초하여 상기 푸리에 변환된 영상에 대해 랜덤으로 각도를 샘플링하여 주파수 영역을 마스킹한다.
- [0023] 단계(120)에서, 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거한다.
- [0024] 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하기 위해 원점과 마스크 사이의 최단 거리, 원점으로부터 가장 긴 거리와 가장 짧은 거리의 차이, 마스크의 직선 에지 사이의 각도를 포함하는 파라미터를 이용하여 마스크를 극좌표계에서 지시 함수로 정의하고, 생성된 마스크를 이용하여 주파수 영역에서 기초 단위(elementary-wise)로 곱하여 해당 주파수 패턴을 제거한다.
- [0025] 단계(130)에서, 역 푸리에 변환부가 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환한다.
- [0026] 본 발명의 실시예에 따르면, 제거된 주파수 패턴의 영상에 2D 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하고, 변환된 공간 영역을 새로운 데이터셋으로 추가하여 영상 분류를 위한 학습에 이용한다.
- [0028] 도 2는 본 발명의 일 실시예에 따른 데이터 증강 과정의 전체 개념을 설명하기 위한 도면이다.
- [0029] 지도 학습 전략을 사용하여 CNN 기반 모델 $f(\cdot)$ 를 학습시키기 위해 입력 데이터 $\mathcal{X} \in \mathcal{R}^{N \times C \times C}$ ($\mathcal{X}$ 가 다중 채널을 가질 때 $\mathcal{X} \in \mathcal{R}^{N \times C \times C \times 3}$)가 필요하다. 여기서 $N, C, k, \mathcal{Y}$ 각각은 영상 수, 영상 폭/높이, 클래스 수, 해당하는 k차원 이진 원핫 벡터 레이블 $\mathcal{Y} \in \{0, 1\}^{N \times k}$ 을 나타낸다. 학습 중에 CNN 기반 모델은 각 레이블의 확률을 나타내는 k차원 실수값 벡터 $\hat{\mathcal{Y}} \in [0, 1]^{N \times k}$ 를 출력한다. 다중 클래스 문제에서 손실 함수 $\mathcal{L}$ 은 다음과 같이 정의된 범주형 교차 엔트로피이다:
- $$\mathcal{L}(\mathcal{Y}, \hat{\mathcal{Y}}) = - \sum_{i=1}^N \hat{y}_i \log y_i \quad (1)$$
- [0030]
- [0031] 여기서 모든 $i \in \mathcal{N}$ 에 대해 $\hat{y}_i \in \hat{\mathcal{Y}}, y_i \in \mathcal{Y}$ 이다. 데이터 증강 방법 간의 성능을 비교하기 위해 Dropout, DropConnect, Batch Normalization과 같은 다른 정규화를 사용하지 않는다.
- [0032] 본 발명의 실시예에 따르면, 공간 영역으로 나타난 입력 영상을 주파수 영역으로 변환하면 공간 영역과 비교하여 패턴을 보다 명확하게 표현할 수 있다. CNN 기반 모델은 저주파 및 고주파 성분을 모두 고려하므로 주파수 영역에서 고주파 성분을 다루는 것이 중요한 역할을 한다[1]. 공간 영역 대신 주파수 영역의 다양한 문제를 해결하기 위한 여러 최신 기술이 개시되었다[2, 3].
- [0033] 본 발명에서는 주파수 영역의 특정 패턴을 필터링하는 새로운 데이터 증강 방법을 제안한다. 편의를 위해 $\mathcal{X}$ 를 $\mathcal{X}$ 의 샘플이라고 한다.

[0034] 본 발명의 실시예에 따른 첫 번째 단계는 주어진 공간 영역의 영상 $X(210)$ 를 $X^{\text{freq}} = F(X)$ 를 나타내는 주파수 영역(220)으로 변환하는 것이다. 여기서 F 는 2D 고속 푸리에 변환(Fast Fourier Transform; FFT)이다. 주파수 영역에서 패턴 구조(221)가 중심에서 바깥쪽으로 확장되는 것을 관찰할 수 있다. 따라서 부분 섹터 모양을 마스크 모양으로 선택한다. 두 번째 단계는 0에서 359 사이의 각도를 선택하여 주파수 영역에서 의미 있는 위치를 선택하는 것이다.

[0035] X^{freq} 의 푸리에 스펙트럼(221)에서 얻은 확률 밀도 함수 $h(\cdot)$ 는 다음과 같이 정의된다:

$$h(\theta; X^{\text{freq}}) = \sum_{r=1}^{c/2} |X^{\text{freq}}|_{\lfloor r \cos \theta \rfloor, \lfloor r \sin \theta \rfloor} \quad (2)$$

[0037] 여기서 $\lfloor \cdot \rfloor$ 는 플로워 함수(floor function)를 나타낸다. 주어진 X^{freq} 에 대한 θ 의 각도 확률 밀도 함수 h 를 기반으로(222), 랜덤으로 각도를 샘플링하여 영역을 마스킹한다. 마스크는 부분 섹터 형태이기 때문에, 각각 원점과 마스크 사이의 최단 거리인 r_0 , 원점으로부터 가장 길고 가장 짧은 거리의 차이 l , 마스크의 직선 가장자리 사이의 각도 α 등 세 가지 추가 파라미터가 필요하다. 이와 같은 파라미터를 이용하여 마스크(223)를 극좌표계에서 지시 함수 $M(\cdot)$ 로 정의한다:

$$M(\theta, \alpha, l, r) = \begin{cases} 1 & \text{for } r \in [r_l, r_h], \theta \in [\theta_l, \theta_h] \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

[0039] 여기서 $r_l = r_0$, $r_h = r_0 + l$, $\theta_l = \theta_0 - \frac{\alpha}{2}$, $\theta_h = \theta_0 + \frac{\alpha}{2}$ 이고, θ_0 은 확률밀도함수 $h(\theta; X^{\text{freq}})$ 에서 랜덤으로 선택된다. 마지막 단계로, 생성된 마스크(223)는 주파수 영역에서 기초 단위(elementary-wise)로 곱하여 특정 패턴을 제거한다.

[0041] 도 3은 본 발명의 일 실시예에 따른 주파수 필터링기반 데이터 증강 알고리즘을 나타내는 도면이다.

[0042] 제거된 주파수 영상에 2D 역 고속 변환(Inverse Fast Fourier Transform; IFFT)을 적용하여 공간 영역(230)(도 2 참조)으로 다시 변환하고 X^{new} 를 나타내는 새로운 데이터셋으로 추가한다:

$$X^{\text{new}} = \mathcal{F}^{-1} (M(\theta, \alpha, l, r) \otimes X^{\text{freq}}) \quad (4)$$

[0044] 여기서 $\mathcal{F}^{-1}$ 은 IFFT이고, $\otimes$ 는 기초 단위 곱(element-wise product)이다. 전반적인 절차는 도 3의 알고리즘 1과 도 2에 도시되어 있다. r_0 , α , l 은 $[0, 10]$, $[50, 90]$, $[10, 30]$ 의 범위에서 베이지안 최적화 기술[4]을 사용하여 각각 5, 70, 20으로 선택된다.

[0046] 도 4는 본 발명의 일 실시예에 따른 페럼 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 장치의 구성을 나타내는 도면이다.

[0047] 제안하는 페럼 엑스레이 영상 분류를 위한 주파수 필터링기반 데이터 증강 장치는 푸리에 변환부(410), 마스크 생성부(420) 및 역 푸리에 변환부(430)를 포함한다.

[0048] 본 발명의 실시예에 따른 푸리에 변환부(410)는 푸리에 변환부가 입력 영상을 고속 푸리에 변환을 통해 주파수 영역으로 변환하고, 주파수 영역에서의 특정 주파수 패턴을 제거하기 위한 각도를 선택한다.

[0049] 본 발명의 실시예에 따른 푸리에 변환부(410)는 푸리에 변환된 영상의 푸리에 스펙트럼에서 획득한 확률 밀도 함수를 정의하고, 상기 확률 밀도 함수에 기초하여 상기 푸리에 변환된 영상에 대해 랜덤으로 각도를 샘플링하여 주파수 영역을 마스킹한다.

[0050] 본 발명의 실시예에 따른 마스크 생성부(420)는 마스크 생성부가 선택된 각도를 이용하여 상기 주파수 패턴을

제거하기 위한 마스크 패턴을 생성하고, 푸리에 변환된 영상에서 상기 생성된 마스크 패턴을 이용하여 해당 주파수 패턴을 제거한다.

- [0051] 본 발명의 실시예에 따른 마스크 생성부(420)는 상기 주파수 패턴을 제거하기 위한 마스크 패턴을 생성하기 위해 원점과 마스크 사이의 최단 거리, 원점으로부터 가장 긴 거리와 가장 짧은 거리의 차이, 마스크의 직선 에지 사이의 각도를 포함하는 파라미터를 이용하여 마스크를 극좌표계에서 지시 함수로 정의하고, 생성된 마스크를 이용하여 주파수 영역에서 기초 단위(elementary-wise)로 곱하여 해당 주파수 패턴을 제거한다.
- [0052] 본 발명의 실시예에 따른 역 푸리에 변환부(430)는 상기 제거된 주파수 패턴의 영상에 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환한다.
- [0053] 본 발명의 실시예에 따른 역 푸리에 변환부(430)는 제거된 주파수 패턴의 영상에 2D 역 고속 푸리에 변환을 적용하여 공간 영역으로 다시 변환하고, 변환된 공간 영역을 새로운 데이터셋으로 추가하여 영상 분류를 위한 학습에 이용한다.
- [0055] 도 5는 본 발명의 일 실시예에 따른 질병별 시각화 결과의 예시를 나타내는 도면이다.
- [0056] 본 발명의 실시예에 따라 하나의 GPU, TITAN RTX 2080TI를 사용하여 Ubuntu 18.04에 대한 실험을 수행했다. 제안하는 방법은 Pytorch, Python3에서 구현된다. 본 발명의 실시예에 따른 데이터셋은 영상 분류를 위한 정상(Normal), 결핵(Tuberculosis), 세균성 폐렴(Bacterial Pneumonia) 및 바이러스(Virus) 등 다양한 유형의 폐렴과 같은 환자의 질병 레이블을 포함하는 데이터셋으로 다양한 출처의 폐 X선 영상을 구성했다.
- [0057] VGG, ResNet, DenseNet과 같은 CNN 기반 모델로 다섯 가지 데이터 증강 방법을 적용하여 제안하는 방법을 평가했다. 본 발명의 실시예에 따르면, 확률적 경사 강하 최적화 도구(stochastic gradient descent optimizer)를 사용하여 학습 속도 0.01, 배치 크기 256 및 30 epoch의 모델을 훈련하고 각 신경망의 마지막 세 개의 컨볼루션 계층의 파라미터를 미세 조정했다. 평가를 위해 각 증강 방법에 대해 5-fold 교차 검증을 사용한다.
- [0058] 정확도, 정밀도, 리콜, F1 점수, AUC 등 5가지 지표를 사용하여 결과를 평가했다. F1 점수는 평균 정밀도와 평균 리콜의 조화 평균으로 계산되었다. 정밀도, 리콜과 마찬가지로, 본 발명의 실시예에 따른 각 클래스에 대한 ROC 곡선을 구한 후 평균 AUC를 계산했다.
- [0059] 본 발명의 실시예에 따르면, 평균적으로 여러 백본에서 정확도, 정밀도, 리콜, F1-점수, AUC에서 3.30%, 7.09%, 7.75%, 8.14%, 4.91% 개선을 달성했다. 특히 ResNet18과 ResNet50에서는 다른 증강 방법과 비교하여 정밀도, 리콜, F1-점수가 9.90%, 10.24%, 11.10%로 가장 크게 향상되었다. 회전 기반 방법과 비교하여, 제안하는 방법이 일관된 성능 향상을 보인다는 것을 관찰했다. 주파수 영역에서 연산을 적용하면 모델을 훈련하는 동안 공간 영역에서 간단한 회전 연산보다 더 다양한 영상이 생성되기 때문이다.
- [0060] 본 발명의 실시예에 따르면, 종래기술과 비교하여 2.37%, 1.09%, 1.45%, 2.21%, 1.97%의 개선을 달성했다. 랜덤 소거 절차에 의해 제거될 때, 신경망 모델은 관심 대상을 포함하는 중요한 영역을 무시한다는 것을 이해할 수 있다. 제안하는 방법은 주파수 공간에서 영역별 패턴을 제거하기 때문에 객체에 대한 정보를 보존하여 신경망을 통한 학습을 방해하지 않는다. 공간 드롭 기반 방법과 달리, 제안하는 방법은 전체 영상에서 반복되는 주파수 패턴을 랜덤으로 제거하므로 X선 영상에서 문제의 특성이 다른 영역에 대하여 더 풍부한 정보를 제공할 수 있다.
- [0061] 제안하는 방법의 합리성을 추가로 평가하기 위해, 신경망의 활성화된 가중 맵을 시각화했다. 본 발명에서는 정상(Normal), 결핵(Tuberculosis), 세균성 폐렴(Bacterial Pneumonia) 및 바이러스(Virus)를 예측하기 위해 Grad-CAM을 사용하여 가중 활성화 맵을 시각화했다. 그 결과는 도 5와 같다.
- [0062] 전반적으로, 제안하는 방법이 종래기술보다 폐 캡처 능력이 더 높고 CNN 기반 신경망에서 객체를 찾는 능력을 더욱 향상시킨다는 것을 관찰할 수 있다. 특히, 제안하는 방법은 ResNet에서 우수한 객체 지역화(localization)를 개선하는 데 도움이 된다. 예를 들어 ResNet18과 ResNet50은 폐 영역의 가중치가 제안하는 방법으로 더 활성화되었음을 확인함으로써 모든 범주의 객체의 위치를 강조 표시할 수 있다. CAM은 의료 영상 진단에 널리 사용되며 제안하는 방법의 CAM은 다른 증강 방법에 비해 상대적으로 더 많은 관심 영역, 예를 들어 폐 영역이 활성화된다는 것을 명확히 보여준다.
- [0063] 본 발명의 실시예에 따르면, 영상 분류를 위해 주파수 영역에서 데이터 증강 방법을 제안한다. 제안하는 방법은

영상 분류뿐만 아니라 객체 지역화(localization)가 중요한 다양한 과제에서도 널리 사용될 수 있다. 특히 의료 데이터셋에 Grad-CAM을 사용하는 것은 제안하는 방법이 기존의 지역화에 비해 더 합리적임을 증명한다. 다양한 CNN 기반 신경망 데이터 증강 방법을 사용한 본 발명의 실시예에 따른 실험은 다른 종래기술에 따른 증강 방법에 비해 일관된 성능 향상을 보인다. 평균적으로 종래기술보다 정확도, 정밀도, 리콜, F1-점수, AUC에서 3.30%, 7.09%, 7.75%, 8.14%, 4.91% 향상되었다.

[0065] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0066] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

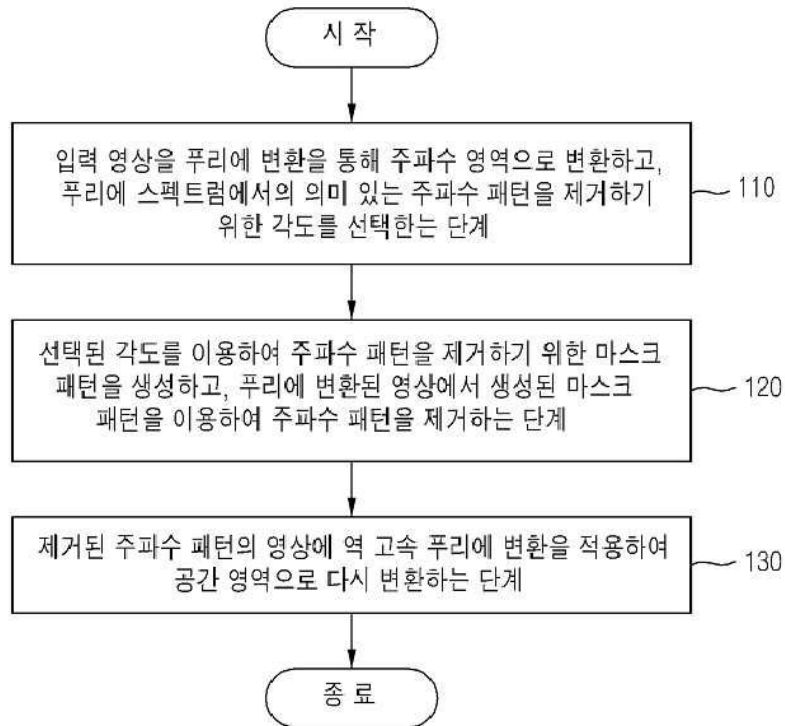
[0067] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0068] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

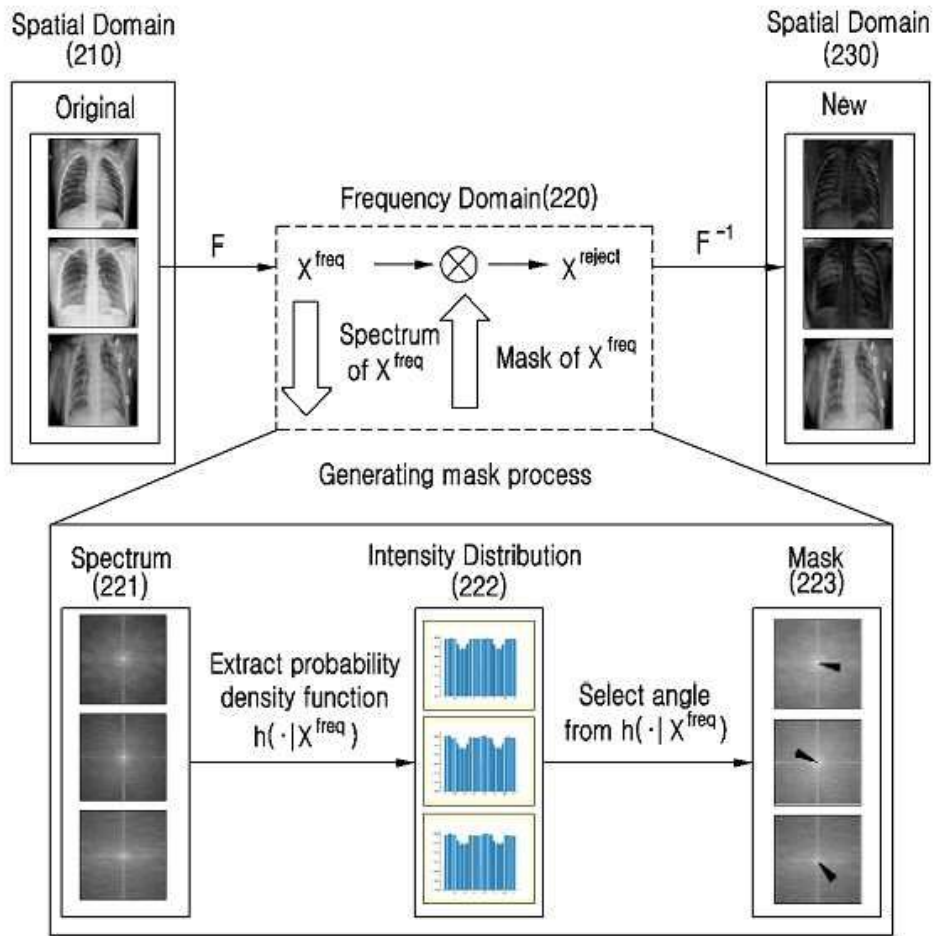
[0069] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

도면1



도면2



도면3

Algorithm 1 Frequency domain data augmentation

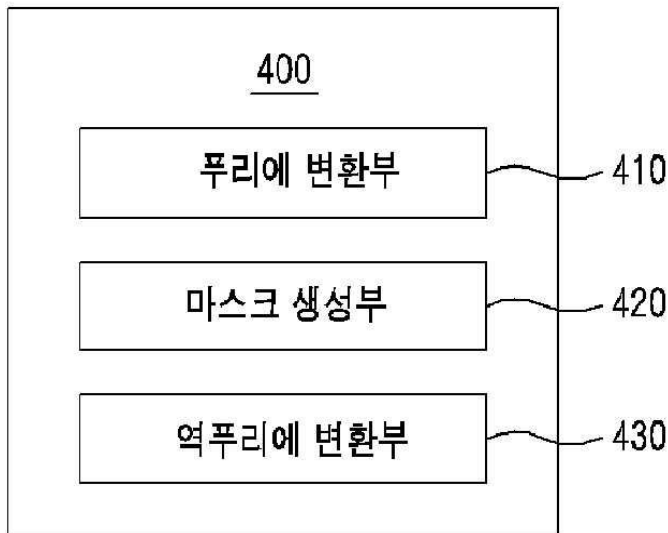
Input: original dataset $(X, y) \in \{\mathcal{X}, \mathcal{Y}\}$, and other mask parameters r_0, l , and α

if dimension of $X = 3$ **then**
 $X^{\text{temp}} \leftarrow X[:, :, 0]$
end if

if dimension of $X = 1$ **then**
 $X^{\text{temp}} \leftarrow X$
end if

$X^{\text{freq}} \leftarrow \mathcal{F}(X^{\text{temp}})$
 $\theta_0 \leftarrow$ randomly sampling from $h(\theta; X^{\text{freq}})$
 $r_l, r_h, \theta_l, \theta_h \leftarrow r_0, r_0 + l, \theta_0 - \frac{\alpha}{2}, \theta_0 + \frac{\alpha}{2}$
 $X^{\text{reject}} \leftarrow M(\theta, \alpha, l, r) \otimes X^{\text{freq}}$
 $X^{\text{new}} \leftarrow \mathcal{F}^{-1}(X^{\text{reject}})$
 $X^{\text{new}}[:, :, i] \leftarrow X^{\text{new}}$ ▷ for $i = 0, 1, 2$
 Adding (X^{new}, y) to $\{\mathcal{X}, \mathcal{Y}\}$
return new images $X^{\text{new}} \in \mathcal{R}^{C \times C \times 3}$

도면4



도면5

