

PA 흉부 X-선 영상 패치 분할에 의한 지역 특수성 이상 탐지 방법[☆]

A Method for Region-Specific Anomaly Detection on Patch-wise Segmented PA Chest Radiograph

김 현 빈¹ 진 준 철^{2*}
Hyun-bin Kim Jun-Chul Chun

요 약

COVID-19로 대표되는 팬데믹 상황에서 의료 인력 부족으로 인한 문제가 대두되고 있다. 본 논문에서는 진단 업무를 지원하기 위한 컴퓨터 비전 솔루션으로 PA 흉부 X-선 영상에 대한 병변 유무 진단 방법에 대해 제시한다. 디지털 영상에 대한 특징 비교 방식의 이상 탐지 기법을 X-선 영상에 적용하여 비정상적인 영역을 예측할 수 있다. 정렬된 PA 흉부 X-선 영상으로부터 특징 벡터를 추출하고 패치 단위로 분할하여 지역적으로 등장하는 비정상을 포착한다. 사전 실험으로 다중 객체를 포함하는 시뮬레이션 데이터 세트를 생성하고 이에 대한 비교 실험 결과를 제시한다. 정렬된 영상에 대해 적용 가능한 패치 특징 하드마asking을 통해 프로세스의 효율성 및 성능을 향상하는 방법을 제시한다. 지역 특수성 및 전역 이상 탐지 결과를 합산하여 기존 연구 대비 6.9%p AUROC 향상된 성능을 보인다.

☞ 주제어 : 이상 탐지, 흉부 X-선, 컴퓨터 비전, 반지도 학습, 영상 특징 모델링

ABSTRACT

Recently, attention to the pandemic situation represented by COVID-19 emerged problems caused by unexpected shortage of medical personnel. In this paper, we present a method for diagnosing the presence or absence of lesional sign on PA chest X-ray images as computer vision solution to support diagnosis tasks. Method for visual anomaly detection based on feature modeling can be also applied to X-ray images. With extracting feature vectors from PA chest X-ray images and divide to patch unit, region-specific abnormality can be detected. As preliminary experiment, we created simulation data set containing multiple objects and present results of the comparative experiments in this paper. We present method to improve both efficiency and performance of the process through hard masking of patch features to aligned images. By summing up regional specificity and global anomaly detection results, it shows improved performance by 0.069 AUROC compared to previous studies. By aggregating region-specific and global anomaly detection results, it shows improved performance by 0.069 AUROC compared to our last study.

☞ keyword : Anomaly Detection, Chest Radiograph, Computer Vision, Semi-supervised Learning, Image Feature Modeling

1. 서 론

최근의 심층 신경망 기반의 컴퓨터 비전 성능의 급격

한 발전은 의료 영상 진단 분야에서의 연구 활성화에도 그 영향을 미치고 있다. 이는 지난 수년간의 COVID-19 팬데믹(Pandemic)으로 인한 의료 인력의 예상치 못한 부족 문제가 대두된 것에도 기인하는 것으로 보인다. 의료 영상에 대한 컴퓨터 비전 기반의 진단 자동화는 직접 또는 간접적으로 활용될 수 있다. 직접적인 활용은 의사를 대체하여 솔루션이 최종 진단을 내리는 것으로 설명된다. 간접적인 활용은 의료 인력을 지원하는 것으로, 다양한 세부 목적을 위해 활용될 수 있다.

본 연구에서는 의료 인력을 지원하기 위해서 이진 중증도 분류를 통해 의료 워크플로우를 개선하고자 한다. PA 흉부 X-선 영상은 가장 대중적인 의료 영상 매체로, 환자의 접근이 용이하며 다양한 질환에 대한 진단의 초

¹ Department of Computer Science, Kyonggi University, Suwon 16227, South Korea

² Division of Computer Science and Engineering, Kyonggi University, Suwon 16227, South Korea

* Corresponding author (jcchun@kyonggi.ac.kr)

[Received 27 December 2022, Reviewed 6 January 2023, Accepted 31 January 2023]

☆ This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

☆ This work was supported by Kyonggi University's Graduate Research Assistantship 2023.

석이 된다. 이에 이상 영상 탐지 방법을 적용하여 중증도가 높은 환자를 선별하고 우선적으로 의사의 진단을 받을 수 있도록 할 수 있다.

기존 연구[1]에서 사전 훈련된 모델로부터 추출된 패치 특징을 비교하는 이상 영상 탐지 방법을 PA 흉부 X-선 영상에 적용한 실험 결과를 제시했다. 기존 연구의 이상 탐지 방법은 비정상이 등장하는 위치를 특정하지 않는다. 이 방법은 영상 전역에서 등장할 가능성이 있는 병변을 효과적으로 탐지할 수 있는 한편, 특정 지역에서만 등장하는 병변에 대한 탐지 중요성을 간과할 수 있다.

본 연구를 수행하기 전 지역적인 비정상을 탐지하는 성능에 대한 문제를 실험적으로 확인하고자 했다. 의료데이터에 대한 직접 실험은 다소 부정확할 수 있을 것이라 판단하였다. 지역적인 비정상을 탐지하는 성능에 대한 문제를 실험적으로 확인하기 위해 MVTEC-AD 데이터 세트 [2]를 활용하여 예상되는 문제 상황을 시뮬레이션 하였다. MVTEC-AD 데이터세트는 산업 불량 검사 목적으로 생성된 공개 데이터 세트로, 평가 데이터 세트에 비정상 범주 및 이진 마스크 형태의 라벨이 포함되어 있다. 따라서 보다 예측 결과에 대한 육안 식별이 용이하고 성능 편차를 측정할 수 있다.

본 논문에서는 사전 연구로 수행한 지역 특수성 이상 탐지를 위한 시뮬레이션 데이터 세트에 대한 이상 탐지 방법 비교 실험 결과를 제시한다. 이를 통해 지역 특수성 이상 탐지 방법의 성능을 검증하였다. 또, 지역 특수성 이상 탐지를 위해 영상 정렬을 수행하였다. 영상 정렬을 수행함으로써 추론에 불필요한 영역을 제거하는 하드 마스킹을 적용할 수 있었다. 결론적으로 지역 특수성 및 전역적 이상 탐지 결과를 합산하여 기존 연구 결과인 0.705 분류 AUROC 보다 6.9%p 성능을 향상하여 0.774 AUROC를 달성하였다.

2. 관련 연구

2.1 이상 영상 탐지

2.1.1 학습 기법에 따른 이상 영상 탐지 방법

이상 탐지는 정상 범주로 보기 어려운 데이터를 검출하는 작업이다. 디지털 영상 및 음성 데이터와 같이 개별 데이터가 정형적으로 샘플링(Sampling)되어 있는 경우, 비정상 데이터 포인트를 검출하는 것 또한 이상 탐지에 속한다. 디지털 영상에 대한 이상 탐지는 영상에서의 비

정상적인 영역을 탐지 또는 분할하거나, 개별 영상 자체의 비정상 여부를 판별함으로써 수행된다. 결론적인 이상 탐지의 작업 내용은 영상 또는 화소 단위의 이진 분류로, 훈련 데이터 및 목표하는 작업의 성격에 따른 적합한 훈련 방법을 다양하게 적용할 수 있다.

지도 학습(Supervised Learning) 기반의 이상 영상 탐지 방법은 훈련에 필요한 디지털 영상과 이에 대응되는 라벨을 요구한다. 수행하고자 하는 작업에 따라 비정상 클래스 라벨, 경계상자 또는 마스크 형태의 라벨이 요구된다. 일반적인 심층 신경망 학습 기반의 이진 분류 및 영역 탐지(Localization), 이진 분할(Binary Segmentation) 방법이 적용될 수 있다. 또한, 서포트 벡터 머신(Support Vector Machine; SVM)과 같은 전통적인 이진 분류 알고리즘을 활용할 수 있다.

반면, 이상 탐지가 적용되는 경우는 일반적으로 다음과 같은 문제를 수반한다. 현실에서 비정상 데이터 샘플을 획득하기 어려운 경우, 이는 지도 학습 시 클래스 불균형을 유발한다. 위 문제는 비정상 데이터 특징을 정의하기 어려운 문제로 이어진다. 알려지지 않은 비정상 데이터, 즉 콜드 클래스(Cold-class)의 등장을 대비해야 하는 경우에도 이상 탐지가 적용될 수 있다. 지도 학습 기반의 이상 탐지는 콜드 클래스가 등장했을 경우 모델을 재학습해야 하는 단점을 가지며, 이때 기존 클래스에 대한 추론 성능을 유지하기 위한 방법을 고려해야 한다.

비지도 학습(Unsupervised Learning)의 경우 훈련 데이터 세트가 분류되어 있지 않은 상태에서 수행되는 방법이다. 분류되지 않은 훈련 데이터 세트에는 정상 및 비정상 데이터 샘플이 무작위로 혼재되어 있으며, 정상 및 비정상 데이터에 대한 특징을 군집화(Clustering)하는 방법이다. 이는 전통적인 군집화 알고리즘을 적용하여 모수적(Parametric) 또는 비모수적(Non-parametric) 방법으로 수행될 수 있다.

대표적인 이상 탐지 데이터 세트인 MVTEC-AD[2]의 훈련 데이터 세트는 정상 데이터 샘플로만 구성되어 있다. 이외의 유사한 영상 데이터 세트들[3, 4] 또한 동일하게 훈련 데이터가 모두 정상 샘플로만 구성되어 있다. 이는 앞에서 언급한 비정상 데이터 샘플을 획득하기 어려운 상황을 상정하는 것에 기인한다. 이에 따라, 다수의 비지도 학습 기반 이상 탐지 방법은 모의 비정상 데이터에 대한 시뮬레이션 과정의 선행이 요구된다.

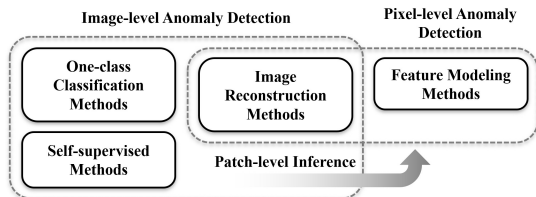
Cohen M, et al[5]은 트랜스포머(Transformer)기반의 자기지도 학습(Self-supervised Learning) 방법으로 이상 탐지를 수행하는 방법을 제시하였다. Yi J, et al[6]은 영상

패치(Patch) 단위의 심층 단일 클래스 이상 탐지 방법[7]을 패치 위치 손실 함수(Patch Position Loss)를 결합하여 자기 지도적으로 이상 탐지기를 훈련하는 방법을 제시하였다.

반지도 학습 기반의 이상 영상 탐지는 훈련 데이터 세트가 정상 데이터로만 구성되어 있다는 사실을 인지한 상태에서 훈련을 수행하는 방법들을 의미한다.

2.1.2 반-지도 학습 기반의 이상 영상 탐지

Yang, J. et al[8]에 의하면, 디지털 영상에 대한 이상 탐지 방법은 크게 영상 단위의 추론 방법과 화소 단위 추론 방법으로 나뉜다. 그림 1은 디지털 영상에 대한 이상 탐지 방법 분류를 보여준다.



(그림 1) 이상 탐지 방법 분류도

(Figure 1) Diagram about categories of anomaly detection methods

영상 단위의 이상 탐지를 수행하는 방법에는 대표적으로 이진 분류기 기반의 방법이 있다. SVM과 같은 이진 분류기를 활용하며, 정상 데이터로부터 추출한 특징이 벡터 공간에서 한 점으로 모이도록 유도하는 방법이다[7]. 영상 단위의 이상 탐지 방법은 패치 단위로 추론함으로써 화소 단위의 이상 탐지기와 근사하게 동작할 수 있다[6].

영상 재건 방법은 영상의 정상 특징을 복원하도록 생성적 모델을 학습시키는 방법이다. Bergmann P, et al[2]는 인코더를 통해 잠재 벡터(Latent Vector)를 생성하고 디코더를 통해 정상 특징만 복원하도록 하는 오토인코더(Auto Encoder) 모델을 영상에 적용하여 제시했다. 적대적 생성 신경망(Generative Adversarial Network;GAN)을 활용한 이상 탐지 연구 사례[9] 및 의료 영상에 적용한 사례[10]가 있다.

특징 모델링 또는 특징 비교 방식의 이상 탐지는 사전 훈련된 심층 신경망으로부터 추출된 특징 정보를 가공하여 정상 및 비정상 특징을 비교하는 방법이다. Bergman L, et al[11]은 추출된 영상 특징 간 유클리드 거리를 측정

하여 영상을 비교하는 방법을 제시했다. Cohen N, et al[12]은 k-최근접 이웃 탐색 알고리즘을 활용하여 다수 이웃 벡터와의 거리 평균을 상대적 이상 점수로 도출하는 방법을 제시했다. Defard T, et al[13]은 신경망의 각 블록으로부터 특징 영상을 추출 및 결합하여 지역적 패치 특징 벡터로 변형하는 방법을 제시했다. Roth K, et al[14]는 앞서의 연구 사례들을 기반으로 패치 특징 벡터에 대해 k-최근접 이웃 탐색을 적용하는 PatchCore를 제시했다. PatchCore는 추론에 필요한 중간 수준 벡터만 추출하며, 추론 시간 효율성을 위한 코어-셋트 서브샘플링(Sub-sampling) 적용 방법을 제시했다.

2.2 컴퓨터 비전 기반 PA 흉부 X-선 영상 진단

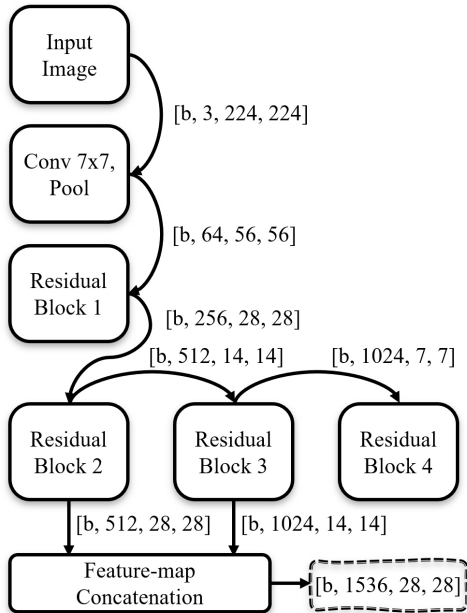
다수의 PA 흉부 X-선 영상 데이터 세트는 지도 학습 기반의 컴퓨터 비전 모델을 대상으로 한다. 이러한 데이터 세트를 활용하여 이상 탐지를 위한 데이터 세트로 변형할 수 있지만, 이러한 연구 사례는 드물다. 주로 비지도 학습 기반의 진단 모델 훈련 및 지도 학습 기반의 고성능 진단 모델에 대한 연구가 진행되고 있다.

Liu J, et al[15]은 대조 학습(Contrastive Learning) 기반의 병변 진단 모델을 제시했다. 패치 영역별 어텐션을 생성하여 정상 및 비정상 샘플의 어텐션을 대조 학습한다. 패치 영역별 어텐션의 정확성을 위한 전처리로 영상 정렬 방법을 제시하였으며, 본 논문에서는 해당 영상 정렬 기법을 기반으로 흉부 X-선 영상을 정렬하였다.

3. 사전 연구

3.1 이상 영상 탐지 방법

지도 학습을 통해 훈련된 심층 신경망 기반의 분류기는 콜드 스타트(Cold-start) 문제를 해결하기 어렵다. 이는 훈련에 활용된 데이터에 의존하는 데에 기인한다. 이상 탐지 분류는 훈련에 사용된 데이터 집합에 포함되지 않는 여집합을 구분함으로써 콜드 스타트 문제에 강건하게 동작할 수 있다. 이상 탐지는 세부 클래스 분류를 수행하기 위해서는 추가 분류 작업이 수행되어야 한다. 이러한 한계에도 불구하고 이상 탐지는 콜드 스타트 문제를 해결하는 방법으로 주목받고 있다. 본 논문에서는 디지털 영상에 대한 특징 비교 방식의 이상 탐지 방법[14]을 기반으로 다중 객체를 포함하는 영상에서 효과적으로 비정상 탐지하는 방법을 제시한다.



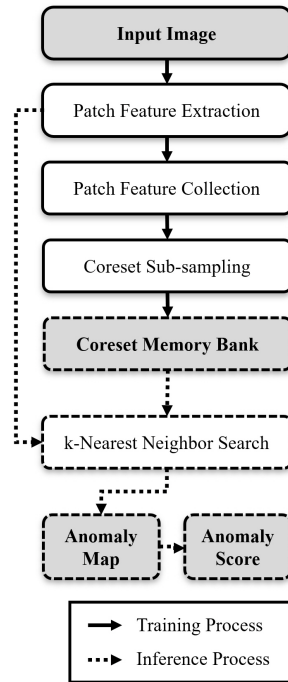
(그림 3) Wide-ResNet-50 모델로부터 중간 수준 특징을 추출 및 결합하는 과정

(Figure 3) Process of extracting and concatenating intermediate-level features from Wide-ResNet-50

본 연구는 사전 연구[1]에서 제시한 방법을 기반으로 한다. 특징 비교 방식의 이상탐지를 수행하기 위해 ImageNet 분류 데이터 세트[14]로 사전 훈련된 심층 신경망을 활용한다. 활용한 심층 신경망 모델은 Wide-ResNet-50[17] 모델로, 블록으로 구분될 수 있는 다른 절차적 신경망 모델과 비교하여 더 높은 이상 탐지 성능을 보인다. 그림 3은 입력 해상도 224×224의 Wide-ResNet-50 모델로부터 이상 탐지에 활용하기 위해 중간 수준 특징을 추출하는 과정을 보인다.

추출된 특징은 각 영상마다 28×28×1536의 크기를 가진다. Defard T, et al[13]에서 제안한 방법을 참고하면 이를 각 위치를 대표하는 정보를 가지는 768개의 1536차원 패치 임베딩 벡터로 여길 수 있다. 전체 훈련 데이터 세트로부터 추출한 임베딩 벡터를 취합하여 모든 벡터를 메모리에 적재하며, 이를 특징 메모리 뱅크(Feature Memory Bank)라 한다.

Cohen N, et al[12]의 방법에서는 주요 벡터 선정을 위한 거리 쌍 연산 및 k-최근접 이웃 탐색 알고리즘으로 인한 과도한 연산 비용 문제가 발생한다. 단순한 k-최근접



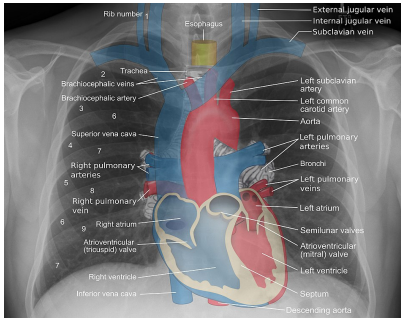
(그림 4) 패치 특징 비교 방식의 이상 영상 탐지 기본 프로세스
(Figure 4) Basic process of visual anomaly detection using patch feature modeling method

이웃 탐색 알고리즘은 NP-난해(NP-Hard)하기 때문에, 탐색 알고리즘을 활용하여 근사한 결과를 취한다. 그림 4는 사전 연구에서 활용한 이상 탐지 방법의 프로세스 흐름을 보인다.

3.2 다중 객체 포함 영상에 대한 분석

사전 연구에서는 PA(Posterior-Anterior) 흉부 X-선 영상에 대해 히스토그램 평활화(Equalization) 방법인 CLAHE (Contrast Limited Adaptive Histogram Equalization)를 적용하여 이상 탐지를 수행하는 방법과 그 실험 결과를 제시하였다[1].

PA 흉부 X-선 영상은 촬영 프레임 내에 여러 가지의 장기에 대해 표현하는 영역을 포함하고 있다. 영상이 정렬되어 있다는 가정 하에 대부분의 장기는 프레임 내의 일정한 위치에 등장한다. 그림 5는 정상 PA 흉부 X-선 영상에서 확인할 수 있는 정보와 해부학적 영역 구분을 보인다.

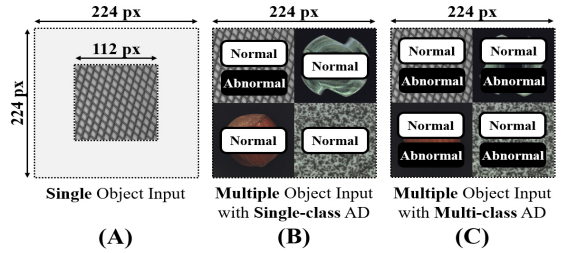


(그림 5) 정상적인 PA 흉부 X-선 영상에서 확인 가능한 장기의 해부학적 구성
(Figure 5) Anatomical composition of organs visible on normal PA chest X-ray images

Li C, et al[18]은 비지도 학습 기반의 이상 탐지를 수행하기 위해 비정상 영상을 시뮬레이션 하는 방법으로 CutPaste를 제시하였다. CutPaste는 한 영상의 일부 영역을 복사하여 다른 영역에 붙여 넣는 영상 처리 방법을 의미한다. CutPaste가 적용된 영상 영역의 특징이 실제 비정상 데이터와 같은 특징을 지니지는 않는다. 그러나 추출된 특징에 대한 t-SNE 분석 결과 정상 데이터 특징과는 명확히 다른 점을 확인할 수 있다. 이 실험 결과는 영상의 일부 영역이 정상일지라도 위치와 주변 영역 영상과의 관계 및 맥락에 따라 비정상으로 취급될 수 있다는 점을 보인다. 다양한 객체가 동시에 등장하는 복잡한 영상 데이터에 대한 이상 탐지를 수행할 경우, 위와 같은 문제가 부각될 수 있다.

정렬된 PA 흉부 X-선 영상에서는 다양한 장기가 일정한 위치에 중첩되어 있다. 의료 영상의 병변은 명확한 비정상적 특징을 보이는 경우 이외에도 장기의 크기나 위치가 달라지는 현상을 보일 수 있다. 또한 여러 객체가 한 영상에 동시에 등장함으로써 모든 클래스에 대한 전체적인 이상 탐지 성능이 저하될 수 있다. 본 연구에서는 이러한 가능성을 실험적으로 확인하기 위해 기존 MVTec-AD 데이터 세트를 활용하여 유사한 문제를 가지는 새로운 데이터 세트를 생성하였다.

그림 6은 MVTec-AD 데이터 세트를 활용하여 시뮬레이션 데이터 세트를 생성하는 예시를 보인다. A는 다른 비교군과 클래스 영상에 대한 해상도 기준을 동일하게 설정하기 위해 112×112 해상도로 영상 크기를 조정하였음을 나타낸다. B는 평가 데이터 세트에 격자(Grid) 클래스만 정상 및 비정상 데이터를 모두 포함하고, 나머지 클래스는 정상 데이터만 붙여 넣은 것을 나타낸다. C는 평



(그림 6) 다중 객체 포함 영상에 대한 이상탐지 성능 측정을 위한 시뮬레이션 데이터 생성 예시
(Figure 6) Example of generating simulation data to evaluate anomaly detection performance for images containing multiple objects

가 데이터 세트에 모든 클래스에 대해 정상 및 비정상 샘플을 포함시킨 것을 나타낸다. 이 경우 합쳐진 영상 단위의 정상 여부는 비정상 영상의 존재 유무로 결정된다. 하나의 포함 영상만 비정상이라도 조합 영상이 비정상으로 취급된다. 모든 평가 데이터 세트는 정상 및 비정상 샘플이 동일한 수로 구성되도록 설정하였으며, 무작위로 샘플이 조합되도록 하여 조합 영상 20,000장을 생성하였으며, 이를 활용하여 각 방법 간 이상 탐지 성능을 비교하였다.

성능 측정을 위해 다음 4장에서 제안하는 지역 특화 이상 탐지 방법 및 사전 연구에서 제안한 이상 탐지 방법을 활용하였으며, 5장 비교 실험 및 분석에서 실험 결과를 보인다.

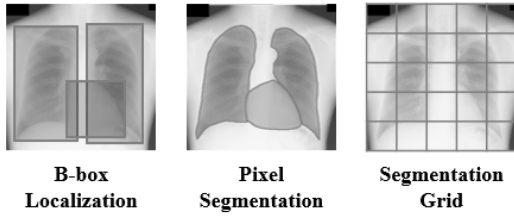
4. 지역 특수성 이상 탐지 방법

제안하는 지역 특수성(Region-Specific) 이상 탐지는 PA 흉부 X-선 영상이 정렬되어 있을 것을 전제한다. 본 논문에서는 Liu J, et al[15]에서 제시된 흉부 X-선 영상 정렬 방법을 활용하여, 전체 실험 데이터에 대한 전처리로 영상 정렬을 수행하였다.

4.1 지역 특수성 이상 탐지

PA 흉부 X-선 영상은 그림 6에서 보인 시뮬레이션 데이터의 예와 같이 다양한 객체가 한 영상 내에 혼합되어 있다. 이러한 데이터에 대한 정확한 이상 탐지를 위해서는 각 객체에서 고유적으로 등장하는 비정상을 정확하게 포착할 필요가 있다. 이를 위해 각 객체를 영역별로 분리하는 과정이 선행되어야 하며, 그림 7과 같은 다양한 방식으로 수행될 수 있다.

Examples of Cropping out Specific Regions from Chest X-Ray Image



(그림 7) PA 흉부 X-선 영상에 대한 객체 분리 방법 예시 (Figure 7) Example of Object Separation Method for PA Chest X-ray Image

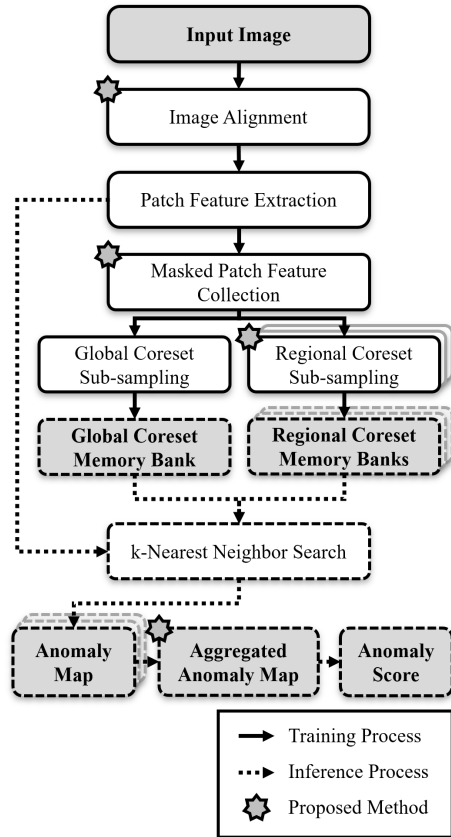
그림 7의 좌측은 객체 탐지를 통해 경계 상자 형태로 관심 영역을 제안하는 방법 예시를 보인다. 중앙은 픽셀 단위의 분류를 통해 정밀하게 영상 영역을 분할하는 방법을 보인다. 가장 우측은 단순히 분할 그리드 방식으로, 패치 단위로 영상을 분할하는 방식을 보인다. 분할 그리드 방식의 영상 분할은 단순히 영상 영역을 구분하여 패치 단위로 나눈다. 이 방법은 일반적인 자연 영상에 대해서는 객체 분할의 역할을 수행하지 못하지만 구도가 획일화된 즉, 정렬된 영상의 경우 객체 분할의 역할을 대체할 수 있다.

본 연구에서는 사전에 전체 입력 데이터에 대한 전처리로 영상 정렬을 수행하였으며, 추가적인 연산 소모 없이 단순히 분할 그리드 방식으로 영역을 분리할 수 있다. 앞서 방법들은 선행 프로세스(Cascaded Process)로 수행되며, 이 과정에서의 잘못된 예측은 전체 파이프라인의 성능 저하로 이어진다. 반면 영상 정렬은 영상의 변형 정도를 인위적으로 제어할 수 있다. 원본 영상이 프레임 밖으로 벗어나지 않도록 제어하여 전체 파이프라인 성능에 미치는 영향을 최소화할 수 있다.

그림 8은 제안하는 지역 특수성 이상 탐지 및 기존 이상 탐지 방법이 결합된 프로세스 흐름을 보인다. 분할 그리드 방식의 영상 영역 분할은 추출된 패치 특징에 대해 적용된다. 본 논문에서는 4×4 분할된 패치 특징을 16개의 메모리 बैं크에 적재하고, 이를 통해 지역적인 비정상 탐지하는 방법을 제시한다.

기존 이상 탐지 방법[1, 14]은 전체 데이터셋에서 추출된 특징 벡터를 모두 열거하여 핵심 정상 특징 벡터로만 구성된 집합으로 서브샘플링한다. 이 방법은 비정상 특징이 영상 전역에서 등장한다고 가정하며, 따라서 본 논문에서는 전역(Global) 이상 탐지 방법으로 기술하였다.

제안하는 방법에서 전역 이상 탐지 및 지역 특수성 이



(그림 8) 제안하는 영상 데이터에 대한 전역 및 지역 특수성 이상 탐지 프로세스 흐름도

(Figure 8) Flow chart of proposed global and region-specific visual anomaly detection process

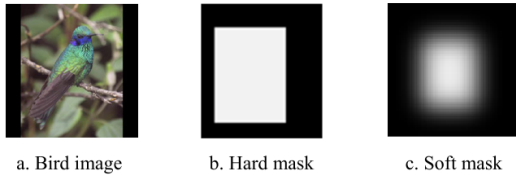
상 탐지는 병렬 수행된다. 각 이상 탐지 과정은 특징 추출 및 마스킹 이후에 분기되기 때문에 같은 해상도로 입력된 영상의 특징을 공유한다. 지역 특수성 이상 탐지에서 각 지역적 코어세트는 더 좁은 영상 영역의 패치 특징을 다루기 때문에 전역 이상 탐지에 비해 더 낮은 서브샘플링 비율을 설정해야 한다. 전역 이상 탐지와 특수성 이상 탐지에서 다루는 패치 특징의 수는 각 대응되는 영상의 면적에 비례하기 때문에, $n \times n$ 분할에 대해서 $1: \frac{1}{n^2}$ 의 서브샘플링 비율을 적용할 수 있다. 이상 탐지 결과는 공통적으로 히트맵 형태의 결과를 출력한다. 이를 이상도 영상(Anomaly Map)이라 하며, 해당 영상의 각 픽셀 밝기 값은 대응되는 영상 영역에 대한 이상도(Anomaly Score)

를 나타낸다. 각 영상 영역에 대해 수행된 지역 특수성 이상 탐지 결과는 분할된 영상 영역에 대응된다. 이를 다시 전역 이상 탐지 결과와 병합해야 하기 때문에 분할된 방식과 동일하게 결합시킨다.

전역 및 지역 특수성 이상 탐지를 통해 추론된 이상도 영상은 임의 비율에 따라 중첩되는 방식으로 합산된다. 이는 가중치에 따른 영상 합 방식으로 수행된다. 최적 가중치를 실험적으로 탐색하였으며, 5장에서 각 가중치에 따른 결합 이상 탐지 성능 평가 결과 비교를 제시한다.

4.2 패치 특징 하드 마스크

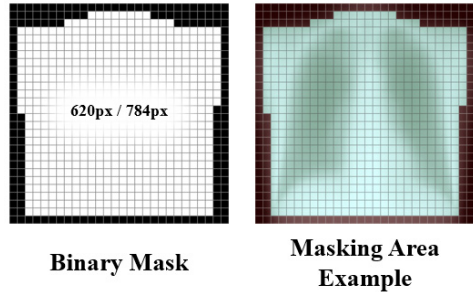
전역 및 지역 특수성 이상 탐지의 병렬적 실행은 기존 방법에 비해 많은 메모리 점유율 및 연산 처리 시간이 요구된다. 이는 코어셋트 서브샘플링 비율 및 입력 영상 해상도 설정에 따라 제한 없이 증가할 수 있다. 또한 일시적으로 전체 데이터 세트로부터 추출한 특징 벡터를 모두 메모리에 적재하며, 데이터 세트가 과도하게 큰 경우 프로세스 실행이 불가능할 수 있다. 이러한 문제를 일부 해소하는 방법으로 패치 특징에 대한 하드 마스크를 제안한다.



(그림 9) 하드 마스크 및 소프트 마스크 비교[19]
(Figure 9) Description

하드 마스크 및 소프트 마스크는 입력 영상의 관심 영역을 모델에 알리기 위해 인위적으로 생성되는 마스크이다. 그림 9의 소프트 마스크는 훈련 과정에서 지정한 영역에서만 손실 값을 계산하기 위해 활용된다. 관심 영역 (Region of Interest; ROI)을 이후 프로세스에 전달하기 위한 방법으로 활용된다.

본 논문에서는 정렬된 PA 흉부 X-선 영상에서 정확한 신체 영역만 이상 탐지를 위한 고려 대상으로 포함시키기 위해 패치 특징에 하드 마스크를 적용하는 방법을 제시한다. 그림 10은 정렬된 PA 흉부 X-선 영상 500장을 평균한 표준 영상을 활용하여, 평균 밝기가 10이하인 영역을 제외하도록 하는 이진 마스크를 보인다.



(그림 10) 하드 마스크 및 소프트 마스크 비교
(Figure 10) Comparison of hard and soft mask

추출된 패치 특징에 위와 같은 하드 마스크를 적용함으로써 불필요하게 처리해야 하는 특징 벡터의 수를 약 21% 감소시킬 수 있다. 이는 전체 파이프라인 처리 시간 및 일시적 최대 메모리 점유율을 감소시킨다. 동시에 실질적인 비정상성 등장하는 신체 영역에 해당되는 패치 특징만을 추출한다. 이로 인해 아날로그 라벨 및 클립 등 노이즈에 의한 비정상적 특징을 제거할 수 있으며, 이는 전체 파이프라인 성능을 향상하는 데에 기여한다.

5. 비교 실험 및 분석

5.1 실험 환경 및 데이터 세트

평가를 위한 훈련 및 평가 흉부 X-선 영상 데이터 세트는 PadChest 데이터 세트[20]로부터 생성되었다. PadChest는 10만장 이상의 대규모 데이터 세트로, PA를 포함하여 다양한 구도의 촬영 영상을 포함한다. 본 논문에서는 실제 의사가 진단한 데이터만을 활용하며, PA 방식으로 촬영된 영상만 추출하였다. 표 1은 본 연구에서 비교하는 모델의 성능 평가에 활용되는 데이터세트의 구성을 보인다. 실험에 사용된 하드웨어 및 소프트웨어 환경 구성은 표 2와 같다.

(표 1) PadChest 데이터 세트로부터 추출한 실험 데이터 세트 구성

(Table 1) Specification of experimental data set extracted from PadChest data set

분류	병변 유무	영상 수
평가	정상	1,182
	비정상	1,182

(표 2) 실험 환경 하드웨어 및 소프트웨어
(Table 2) Hardware and software environment for experiments

구분	내용
CPU	Intel i9-10980XE 18c36t
GPU	Nvidia Geforce RTX 3090
RAM	128GB
OS	Ubuntu 20.04
Language	Python 3.8
DL Framework	pytorch 1.8.1
CUDA	CUDA 11.1, cudnn 8.0
Other	faiss-gpu

5.2 다중 객체 영상 시뮬레이션 비교 실험

이 절에서 보이는 실험 결과는 다중 객체를 포함하는 PA 흉부 X-선 영상과 유사하게 생성된 시뮬레이션 데이터에 대한 이상 탐지 성능 비교이다. 시뮬레이션 영상의 예는 본 논문 3.2절의 그림 6에서 보인다.

(표 3) grid 클래스단일 데이터셋 A 및 grid, hazelnut, metal_nut, tile 클래스 조합 시뮬레이션 데이터 세트 B, C에 대한 성능 비교

(Table 3) Performance comparison for dataset A with only grid class and composed dataset B,C with grid, hazelnut, metal_nut and tile classes.

이상 탐지 방법	A	B	C
Global	0.951	0.932	0.896
Region-Specific	0.948	0.951	0.944

표 3은 MVTEC-AD 데이터셋의 Grid 클래스 및 Hazelnut, Metal Nut, Tile 클래스를 하나의 영상으로 조합하여 생성한 시뮬레이션 데이터 세트에 대한 각 방법의 성능 비교를 보인다. A 실험 데이터 세트의 경우 다중 클래스 영상을 합성한 시뮬레이션 데이터 세트와 해상도 조건을 일치시키기 위해 저해상도로 크기를 변형하여 입력한 결과다. B 실험 데이터 세트는 A 세트 클래스만 정상 및 비정상적으로 나뉘며, 기타 클래스는 정상 데이터만으로 생성한 데이터 세트다. C 실험 데이터 세트는 영상 생성에 사용된 모든 클래스 데이터가 정상 및 비정상적으로 나뉘어 무작위 구성되도록 한 데이터 세트다.

이 실험은 본 논문에서 제안하는 지역 특수성 이상 탐지 방법의 유효성을 사전에 확인하기 위해 수행되었다. B 실험 및 C 실험에서는 특정 위치에만 국한되어 등장하는 비정상 특징이 존재한다. 전역 이상 탐지 방법은 단일 클래스 영상에 대한 실험 A에서 가장 높은 성능을 보였으나, 지역적으로 등장하는 B 실험에서 성능이 소폭 하락하는 경향을 보인다. C 실험에서는 영상에 포함된 모든 클래스가 비정상 특징을 무작위적으로 가지며, 전역 이상 탐지 방법은 모든 경우에서 지역적 이상 탐지 방법보다 낮은 성능을 보였다.

아래 표 4 및 표 5는 각각 Pill 클래스, Screw 클래스를 다른 클래스 영상과 조합하여 실험한 결과를 보인다. 모든 경우의 조합 영상에 대한 이상 탐지 실험에서 지역 특수성 이상 탐지 방법이 전역 이상 탐지 방법의 성능을 상회하는 결과를 보인다.

(표 4) pill 클래스단일 데이터셋 A 및 grid, hazelnut, metal_nut, tile 클래스 조합 시뮬레이션 데이터 세트 B, C에 대한 성능 비교

(Table 4) Performance comparison for dataset A with only grid class and composed dataset B,C with pill, hazelnut, metal_nut and tile classes.

이상 탐지 방법	A	B	C
Global	0.935	0.918	0.880
Region-Specific	0.935	0.939	0.927

(표 5) screw 클래스단일 데이터셋 A 및 grid, hazelnut, metal_nut, tile 클래스 조합 시뮬레이션 데이터 세트 B, C에 대한 성능 비교

(Table 5) Performance comparison for dataset A with only grid class and composed dataset B,C with screw, hazelnut, metal_nut and tile classes.

이상 탐지 방법	A	B	C
Global	0.930	0.909	0.868
Region-Specific	0.926	0.932	0.906

이 실험 결과는 지역 특수성 이상 탐지 방법이 절대적으로 우수하다는 점을 의미하지는 않는다. 그러나 비정상성이 지역적으로만 등장한다는 특수 조건 하에서 제안하는 지역 특수성 이상 탐지를 적용하는 것이 효과적임을 보인다.

5.3 지역 특수성 이상 탐지 성능 평가

이 절에서는 표 1의 PadChest PA 흉부 X-선 데이터 세트 세트를 활용한 실험 결과를 보인다.

표 6은 영상 정렬 실행 여부에 따른 각 이상 탐지 방법의 성능 비교를 보인다. 정렬되지 않은 흉부 X-선 영상에 대해서는 지역 특수성 이상 탐지가 매우 낮은 성능을 보인다. 이는 제안하는 지역 특수성 이상 탐지가 정렬되어있다는 전제 조건에 크게 의존하는 점을 의미한다. 반면, 정렬 된 영상에 대한 지역 특수성 이상 탐지는 기존 전역 이상 탐지보다 동등하거나 오차 범위 내에서 더 높은 성능을 보인다.

(표 6) 영상 정렬 여부에 따른 이상 탐지 방법 성능 비교
(Table 6) Comparison of anomaly detection method performance according to either image alignment or not

이상 탐지 방법	정렬 여부	AUROC
Global	X	0.705
Global	O	0.711
Region-Specific	X	0.623
Region-Specific	O	0.719
Aggregated	X	0.650
Aggregated	O	0.755

(표 7) 패치 특징 하드마스킹 여부에 따른 이상 탐지 방법 성능 비교

(Table 7) Comparison of anomaly detection method performance according to either patch feature hard masking or not

이상 탐지 방법	하드 마스킹 적용 여부	AUROC
Global	X	0.711
Global	O	0.724
Region-Specific	X	0.719
Region-Specific	O	0.738
Aggregated	X	0.755
Aggregated	O	0.761

정렬된 PA 흉부 X-선 영상에 대한 패치 특징 하드마스킹은 모든 방법에 대해 표 7과 같은 성능 향상 효과를 보

인다. 표 8은 전역 및 지역 특수성 이상 탐지 결과 합산시 가중치를 4:6으로 설정했을 때 최적 비율에 가장 근사하다는 점을 실험적으로 보인다.

기존의 이상 영상 탐지 방법을 동일한 실험 데이터 세트에 적용한 비교 실험 결과는 표 9과 같다. 기존 방법에 비해 높은 성능을 달성하였다. 또한, 사전 연구[1]에서의 0.705 AUROC에 비해 최고 6.9%p 성능을 향상했다.

(표 8) 지역 특수성 및 전역 이상 탐지 결과 합산 가중치에 따른 결합 이상 탐지 성능 평가

(Table 8) Evaluation of aggregated anomaly detection performance according to weight ratio of region-specific and global.

이상 탐지 방법	Global	Region-specific	AUROC
Aggregated	0.9	0.1	0.723
	0.8	0.2	0.729
	0.7	0.3	0.737
	0.6	0.4	0.741
	0.5	0.5	0.758
	0.4	0.6	0.761
	0.3	0.7	0.753
	0.2	0.8	0.744
	0.1	0.9	0.739

(표 9) 동일 데이터 세트에 적용된 기존 방법들과의 비교

(Table 9) Comparison with existing methods applied to the same data set

이상 탐지 방법	AUROC	ACC@max
F-AnoGAN [10]	0.598	0.542
Patch-SVDD [6]	0.674	0.597
Patch-SVDD MR Ensembled [6]	0.746	0.731
Aggregated	0.761	0.732
Aggregated @Max	0.774	0.747

6. 결론 및 향후 연구

본 논문에서는 PA 흉부 X-선 영상에 대한 이상 탐지를 수행함으로써 병변 유무를 진단하는 방법 및 기존 연구

에 비해 성능을 향상한 방법을 제시하였다. 기존 연구된 이상 영상 탐지 방법을 적용하기 앞서 흉부 X-선 영상과 제조업 이상 탐지 데이터 세트의 데이터 차이를 분석하여 제시했다. 다중 객체를 포함하는 흉부 X-선 데이터 세트와 유사하도록 제조업 이상 탐지 데이터 세트를 활용하여 시뮬레이션 데이터 세트를 생성하였으며 이에 대한 사전 실험을 수행했다. 사전 실험을 통해 지역 특수성 이상 탐지 방법의 조건부 유효성을 확인하였으며, 흉부 X-선 영상 데이터 세트에 적용하였다. 정렬된 흉부 X-선 영상에 대한 패치 특징 하드마스킹을 통해 효율성 및 성능을 동시에 향상하는 방법을 제시하였다. 결론적으로 기존 연구에 비해 6.9%p AUROC 향상된 성능을 보였다. 본 논문에서 제시한 전역 및 지역 특수성 이상 탐지 결과의 상호 성능 보완은 AutoML을 비롯한 앙상블 방법에 의해 추가적으로 개선될 여지가 있다. 향후 본 논문에서 제안된 방법을 다중 객체를 포함하는 영상 데이터 세트에 적용하여 다른 작업에서의 유효성 및 범용성을 검증하려 한다.

참고문헌(Reference)

- [1] Hyunbin Kim and Chun-chul Chun, "Lesion Detection in Chest X-ray Images based on Coreset of Patch Feature," *Journal of Internet Computing and Services*, Vol.23, No.3, pp.35-45, 2022.
<https://doi.org/10.7472/jksii.2022.23.3.35>
- [2] Bergmann P. et al., "MVTec AD--A Comprehensive Real-world Dataset for Unsupervised Anomaly Detection," in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.9592-9600, 2019.
<https://doi.org/10.1109/CVPR.2019.00982>
- [3] Zhang Y. et al., "Single-image crowd counting via multi-column convolutional neural network," in *Proc. of the IEEE conference on computer vision and pattern recognition*, pp.589-597, 2016.
<https://doi.org/10.1109/CVPR.2016.70>
- [4] Lu C. et al., "Abnormal event detection at 150 fps in matlab," in *Proceedings of the IEEE international conference on computer vision*, pp.2720-2727, 2013.
<https://doi.org/10.1109/ICCV.2013.338>
- [5] Cohen M. J. et al., "Transformal-Two (Feature Spaces) Are Better Than One," in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.4059-4068, 2022.
<https://doi.ieeecomputersociety.org/10.1109/CVPRW56347.2022.00451>
- [6] Yi J. et al., "Patch SVDD: Patch-level SVDD for Anomaly Detection and Segmentation," in *Proc. of the Asian Conference on Computer Vision*, pp.375-390, 2020.
https://doi.org/10.1007/978-3-030-69544-6_23
- [7] Ruff L. et al., "Deep One-class Classification," in *Proc. of the 35th International Conference on Machine Learning*, PMLR, pp.4393-4402, 2018.
<https://proceedings.mlr.press/v80/ruff18a.html>
- [8] Yang J. et al., "Visual anomaly detection for images: A survey," *Computer Vision and Pattern Recognition*, 2021.
<https://doi.org/10.48550/arXiv.2109.13157>
- [9] Akcay S. et al., "Ganomaly: Semi-supervised anomaly detection via adversarial training," *Asian conference on computer vision*. Springer, Cham, pp.622-637, 2018.
https://doi.org/10.1007/978-3-030-20893-6_39
- [10] Schlegl T. et al., "f-AnoGAN: Fast Unsupervised Anomaly Detection with Generative Adversarial Networks," *Medical Image Analysis*, Vol. 54, pp.30-44, 2019.
<https://doi.org/10.1016/j.media.2019.01.010>
- [11] Bergman L. et al., "Deep Nearest Neighbor Anomaly Detection," *Machine Learning*, arXiv preprint, 2020.
<https://doi.org/10.48550/arXiv.2002.10445>
- [12] Cohen N. et al., "Sub-image Anomaly Detection with Deep Pyramid Correspondences," *Computer Vision and Pattern Recognition*, arXiv preprint, 2020.
<https://doi.org/10.48550/arXiv.2005.02357>
- [13] Defard T. et al., "PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization," *International Conference on Pattern Recognition*. Springer, Cham, pp.475-489, 2021.
https://doi.org/10.1007/978-3-030-68799-1_35
- [14] Roth K. et al., "Towards Total Recall in Industrial Anomaly Detection," in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.14318-14328, 2022.
<https://openaccess.thecvf.com/content/CVPR2022/pap>

- ers/Roth_Towards_Total_Recall_in_Industrial_Anomaly_Detection_CVPR_2022_paper.pdf
- [15] Liu J. et al., "Align, Attend and Locate: Chest X-ray Diagnosis via Contrast Induced Attention Network with Limited Supervision," in Proc. of 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp.10632-10641, 2019.
<https://doi.org/10.1109/ICCV.2019.01073>
- [16] Deng J. et al., "ImageNet: A Large-scale Hierarchical Image Database," in Proc. of 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp.248-255, 2009.
<https://doi.org/10.1109/CVPR.2009.5206848>
- [17] Zagoruyko S. et al., "Wide Residual Networks," Computer Vision and Pattern Recognition, arXiv preprint, 2016.
<https://doi.org/10.48550/arXiv.1605.07146>
- [18] Li C. L. et al., "Cutpaste: Self-supervised learning for anomaly detection and localization," in Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.9664-9674, 2021.
<https://doi.org/10.48550/arXiv.2104.04015>
- [19] Lin C. et al., "An Interpretable Visual Attention Plug-in for Convolutions," IEEE Access, Vol. 8, pp.136992-137003, 2020.
<https://doi.org/10.1109/ACCESS.2020.3011696>
- [20] Bustos A. et al., "Padchest: A Large Chest X-ray Image Dataset with Multi-label Annotated Reports," Medical Image Analysis, Vol. 66, pp.101797, 2020.
<https://doi.org/10.1016/j.media.2020.101797>

● 저 자 소 개 ●



김 현 빈(Hyun-bin Kim)
 2020년 경기대학교 컴퓨터과학과(이학사)
 2021년~현재 경기대학교 일반대학원 컴퓨터과학과(석사과정)
 관심분야 : 컴퓨터 비전, 딥러닝, 이상탐지
 E-mail : khb8813162@kyonggi.ac.kr



전 준 철(Jun-chul Chun)
 1984년 중앙대학교 전자계산학과(학사)
 1986년 중앙대학교 전자계산학과(석사)
 1992년 CONNECTICUT 대학교 컴퓨터공학과(공학석사)
 1995년 CONNECTICUT 대학 컴퓨터공학과(공학박사)
 1995년~현재 경기대학교 소프트웨어경영대학 AI컴퓨터공학부 교수
 관심분야 : 컴퓨터 그래픽스, 영상처리
 E-mail : jcchun@kyonggi.ac.kr