https://doi.org/10.7236/JIIBC.2023.23.4.107

JIIBC 2023-4-17

딥러닝 기술을 이용한 영상에서 흡연행위 검출

Detection of Smoking Behavior in Images Using Deep Learning Technology

김동준**, 최유진**, 박경민**, 박지현**, 이재문*, 황기태**, 정인환**

Dong Jun Kim**, Yu Jin Choi**, Kyung Min Park**, Ji Hyun Park**, Jae-Moon Lee*, Kitae Hwang**, In Hwan Jung**

요 약 본 논문은 인공지능 기술을 활용하여 영상에서 흡연 행위를 검출하는 방법을 제안한다. 흡연은 정적 현상이 아니라 행위에 해당하기 때문에 객체 탐지 기술에 행위를 탐지할 수 있는 자세 추정 기술을 접목하였다. 이미지에서 흡연자를 검출하기 위하여 흡연자 검출 학습 모델을 개발하였으며, 영상에서 흡연행위를 검출하기 위하여 흡연행위의 특성을 자세 추정 기술에 적용하였다. 객체 탐지를 위하여 YOLOv8을 사용하였으며, 자세 추정을 위하여 OpenPose를 이용하였다. 또한, 영상에 흡연자 및 비흡연자가 포함되어 있는 경우 사람들만 분리하는 방법도 적용하였다. 제안된 방법은 파이선으로 Google Colab NVIDEA Tesla T4 GPU를 사용구현 하였고, 테스트 결과 주어진 영상에서 흡연 행위를 완벽하게 검출함을 알 수 있었다.

Abstract This paper proposes a method for detecting smoking behavior in images using artificial intelligence technology. Since smoking is not a static phenomenon but an action, the object detection technology was combined with the posture estimation technology that can detect the action. A smoker detection learning model was developed to detect smokers in images, and the characteristics of smoking behaviors were applied to posture estimation technology to detect smoking behaviors in images. YOLOv8 was used for object detection, and OpenPose was used for posture estimation. In addition, when smokers and non-smokers are included in the image, a method of separating only people was applied. The proposed method was implemented using Google Colab NVIDEA Tesla T4 GPU in Python, and it was found that the smoking behavior was perfectly detected in the given video as a result of the test.

Key Words: Smoking, Deep Learning, Object Detection, Pose Estimation, SORT Tracking Algorithm

1. 서 론

금연 구역에서의 흡연, 담배꽁초 무단투기 등 흡연으로 인한 문제는 사회적으로 많은 문제를 일으키며 그 심 각성이 심화되어 가고 있다. 담배꽁초로 인해 발생하는 화재 건수는 해를 거듭하며 증가하고, 무단투기로 인해 배수시설이 막혀 2022년 강남 지역 도로가 침수되는 원인이 되기도 하였다. 이에 대응하기 위해 지능형 CCTV 기술을 활용하여 무단 흡연을 억제하고 있지만 기존의 CCTV는 단순히 사람이 인식되는 경우 경고음을 내는

*정회원, 한성대학교 컴퓨터공학부(교신저자)
**정회원, 한성대학교 컴퓨터공학부 접수일자 2023년 7월 4일, 수정완료 2023년 7월 30일 게재확정일자 2023년 8월 4일

Received: 4 July, 2023 / Revised: 30 July, 2023 / Accepted: 4 August, 2023
*Corresponding Author: jmlee@hansung.ac.kr

Dept. of Computer Engineering, Hansung University, Korea

정도로 그쳐 완전한 대응에 한계가 있다.

최근 컴퓨터 비전 분야에서의 인공지능 기술 발전으로 인해, 이미지 및 영상 분석에 뛰어난 성능을 보여주고 있 다. 특히 이미지 인식, 객체 탐지 등의 분야에서 인공지능 은 인간의 개입을 줄이면서도 높은 정확도를 보여주고 있다.

본 논문은 이러한 인공지능 기술 중 객체 탐지(Object Detection) 모델인 YOLO^[1, 2]와 자세 추정(Pose Estimation) 모델인 OpenPose^[3]를 이용하여 흡연자의 행동 패턴을 식별하고 영상에서 흡연 행위를 검출하는 방법에 대해 제안 한다. 객체 탐지 기술을 이용하여 이미지내 흡연자를 검출하며, 자세 추정 기술을 이용하여 흡연행위를 검출한다. 영상에서 흡연자와 흡연행위가 동시에 검출되는 경우 흡연 영상으로 판정하는 것이다.

2장에서는 객체 탐지 기술을 이용하여 흡연자 검출하기 위한 학습 모델 개발에 대하여 설명하고, 3장에서는 흡연 영상 검출 방법의 구현을 설명한다. 4장에서는 개발된 모델의 성능을 논의하고, 마지막으로 5장에서 결론 및 향후 연구 방향에 대해 논의 한다.

II. 흡연자 검출을 위한 학습 모델 개발

객체 탐지 기술은 이미지 속에서 특정 객체를 탐지하여 위치와 크기를 추출하는 인공지능 기술이다^[1, 2, 4]. 이기술은 흡연자와 비흡연자를 학습시켜 흡연자를 인식하는데 활용될 수 있다. 본 논문에서는 이미지 내에 흡연자를 검출하기 위하여 YOLOv8n 모델^[1]을 사용하였다.

1. 학습 데이터 준비

이미지에서 흡연지점(smoke_head)만을 검출하는 객체 탐지 학습 모델 생성을 위해, AIHUB^[5]에서 제공하는 '공원 주요 시설 및 불법행위 감시 CCTV 영상 데이터'와 '일상생활 영상 데이터' 두 개의 영상 데이터를 활용하였다. 파이선 코드를 이용하여 각 영상 데이터마다 3개의 프레임을 추출하고, 그중 흡연지점 이미지만을 선별하였다. 더불어, Kaggle이 제공하는 'Cigarette Smoker Detection' 데이터^[6]의 흡연지점 이미지를 추가하여 학습을 위한 데이터를 구축하였다. 이렇게 준비한 학습 이미지는 194개이다.

Roboflow의 Annotation 툴^[7]을 이용하여 이미지 데 이터에서 흡연지점에 해당하는 부분을 바운당 박스로 표 시하는 라벨링 작업을 수행하였다. 흡연 도중의 자세 변 화와 입력 이미지 품질의 다양성을 반영하기 위하여 회 전(Rotation), 전단(Shear), 흐림(Blur), 노이즈(Noise) 총 4개의 데이터 증강(Data Augmentation) 기법을 적용하였다. 이를 통해 원본 데이터의 다양성을 확장하였고 최종적으로 582장의 흡연지점 이미지 데이터를 생성하였다. 그림 1은 원본 이미지(a)에 대한 회전(b), 전단(c) 및 흐림(d)에 대한 예이다.



그림 1. 이미지 증폭에 대한 예

Fig. 1. Example of Image Augmentation

2. 학습 및 학습 결과

Google Colab 환경의 NVIDEA Tesla T4 GPU를 사용해 학습하였다. Roboflow를 통해 구축한 데이터를 8:2 비율로 훈련 세트와 검증 세트로 나누어 16 Batch, 100 Epoch 설정으로 전이학습 진행하였다. Pre-trained 모델인 YOLOv8n 모델^[1]을 사용하였다. 흡연자 검출은 영상의 여러 프레임을 처리해야 하기에 검출 속도와 성능의 균형이 중요하다. 이러한 이유로 YOLO 시리즈 경량화 모델 중 가장 좋은 성능을 보이는 YOLOv8n 모델을 채택하였다.

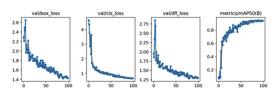


그림 2. 학습 결과 그래프(x축: epoch, y축: 손실-왼쪽 3그래프, 정확도-마지막 그래프)

Fig. 2. Learning Result Graphs(x-axis: epoch, y-axis: loss-left 3 graphs, accuracy-last graph)

그림 2는 학습 모델에 대한 성능을 보인다. 그림 2의학습 그래프를 보면 검증 데이터에 대한 박스 손실 (box_loss), 클래스 손실(cls_loss), 클래스 불균형 손실 (DFL: Dual Focal Loss)이 Epoch이 증가함에 따라 감소하는 경향을 보였다. 이는 모델이 각 Epoch 마다 흡연지점의 위치와 크기를 더욱 정확하게 예측하고, 클래스를 더욱 정확하게 분류하며, 클래스 불균형 문제를 더욱효과적으로 처리하고 있음을 의미한다. 객체 탐지에서객체의 박스와 객체의 카테고리(클래스) 탐지에 대한 결합 지표인 mAP50(Mean Average Performance 50)¹⁸¹ 지표는 Epoch이 증가함에 따라 증가하는 추세를 보인다. 이는 모델의 흡연지점 감지 능력이 점차 향상되고 있음을 의미한다.

학습 결과, 정밀도(Precision) 0.919, 재현율(Recall) 0.901, mAP50에서 0.942의 성능을 보였다. 이는 모델이 대부분의 테스트 이미지에서 흡연지점을 성공적으로 감지하였음을 의미한다. 특히, mAP50의 높은 성능은 모델이 물체의 위치와 클래스를 동시에 정확하게 예측하였음을 의미한다. 그림 3은 완성된 모델이 흡연자의 비흡연지점과 흡연지점을 구별하여, 흡연지점(smoke_head)을 검출하는 예시이다.



smoke head 0.8/

smoking frame

그림 3. 흡연지점 검출 예시

Fig. 3. Example of Detection of Smoke_head

Ⅲ. 딥러닝을 이용한 흡연행위 검출

1. 흡연행위 검출의 핵심 요소

흡연은 흡연자의 상태가 아닌 행위이기 때문에 하나의 이미지에 대한 물체탐지 인공지능 기술적용만으로 흡연 행위를 검출하지 못할 수도 있다. 이것은 다수의 연속된 프레임에서 흡연자의 행동 패턴을 검출해야 한다는 것을 의미한다.

이러한 관점에서 본 논문에서는 그림 4와 같이 이중

모형을 제안한다. YOLO을 이용하여 흡연자를 탐지함과 동시에 OpenPose 기술을 이용하여 행위 탐지를 하여 최종 흡연행위를 탐지하도록 하였다.

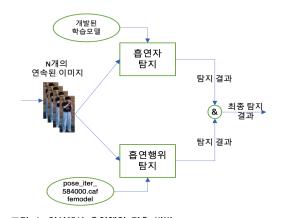


그림 4. 영상에서 흡연행위 검출 방법 Fig. 4. Method for detecting smoking behavior in video

2. SORT를 이용한 사람 분리

이미지 내에 다수의 사람 객체가 존재하는 경우, YOLO 및 OpenPose 모델을 사용하여 얻은 결과가 객체간 영향을 끼치게 된다. 이 문제를 해결하기 위해, SORT(Simple Online and Realtime Tracking)^[9] 알고리즘을 적용하여 각 프레임 내의 사람 객체를 독립적으로 분리하였다. 사람 객체 검출에는 YOLO 모델을 이용하고 검출된 각 객체에는 고유한 ID를 부여하여 개별적으로 관리하여 각 객체의 추적을 위해 칼만 필터 (Kalman-Filter)^[10]를 적용하였다.



그림 5. 사람 객체 분리 및 추적 예시 Fig. 5. Example of human object separation and tracking

객체 추적 과정에서 각 프레임별로 객체를 바운당 박 스(bounding box) 크기에 맞춰 잘라 저장하였다. 하지 만, 이 과정에서 사람의 전신이 일부 잘리는 현상이 발생 하였고, 이는 OpenPose를 통한 자세 추정에 문제를 일 으킬 수 있다는 것을 야기하였다. 이를 해결하기 위해 바운당 박스의 크기에 추가적인 패당을 반영하여 이미지를 자르는 방식을 적용하였다. 그림 5는 SORT 알고리즘을 이용하여 프레임별 사람 객체를 분리 및 추적하여 저장한 예시이다.

3. 물체탐지에 의한 흡연자 검출

그림 4에서 N을 100으로 하여 흡연자가 포함된 100개의 프레임으로 구성된 영상을 학습된 YOLOv8 모델에 적용하여 'smoke_head' 검출빈도를 측정한 결과 평균 30번 검출을 하였다. 이는 흡연 영상에서 평균 3.3 프레임당 흡연자를 검출한다는 것을 의미한다.

그림 6은 하나의 프레임에 대하여 흡연자 검출 여부를 저장하고 최종적으로 100개의 프레임에 흡연자 유무를 판단하는 함수이다. detect_smoker() 함수는 영상 프레임 순서대로 100번 호출되고 object_detector.find_smoker()함수를 이용하여 매개변수 frame에 흡연자가 있는지를 판별하여 이에 대한 컨피던스를 리턴한다. 컨피던스가 0.8 이상이면 흡연자가 있는 것으로 판단하여 smokers에 1을 저장한다. 여기서 0.8, 0.3 등의 상수는 다양한 실험에 의한 찾은 수치이다.

```
def detect_smoker(frame, smokers, criticalPoint=0.3):
# 흡연자 여부를 탐색한다.
confidence = object_detection.find_smoker(frame)

# 흡연자 가능성을 판정한다.
if confidence > 0.8:
    smokers.push(1)
else:
    smokers.push(0)
# 흡연자임을 판정한다.
if smkoer.count(1) / len(smokers) > criticalPoint:
    return True
return False
```

그림 6. Smoke_head 검출빈도를 통한 흡연자 판단 알고리즘 코드

Fig. 6. Algorithm for determining smoker by detecting Smoke_head frequency

4. 물체행위에 의한 흡연자 검출

앞서 물체탐지에 의한 흡연자 검출의 결과에서 볼 수 있듯이 흡연자가 포함된 영상에서 흡연자 검출은 겨우 30%에 지나지 않는다. 본 논문에서는 이를 보완하기 위하여 사람의 팔 각도의 변화를 탐지하여 흡연자 검출을 보완한다.

오픈소스 Pose Estimation 모델인 OpenPose은 이 미지에서 인체의 관절을 인식하는 인공지능 소프트웨어 이다. 25개의 관절(키포인트)을 추적하는 기존 OpenPose 모델을 수정하여, 팔과 관련된 'RShoulder', 'RElbow', 'RWrist', 'LShoulder', 'LElbow', 'LWrist' 6개의 키포 인트만 인식하도록 설정하였다. 각 키포인트는 PCM(Part Confidence Maps)을 통해 정의되며, 이는 각 픽셀 (Pixel)이 해당 키포인트에 속할 확률을 나타내는 2D에 서의 좌표이다. 이 PCM을 통해 모델은 입력 이미지에서 팔의 주요 관절에 해당하는 지점을 식별한다. 또한 키포인 트 간의 연결 관계를 정의하기 위해 4개의 POSE PAIRS 를 사용한다. 이는 각각 'RShoulder-RElbow', 'RElbow-RWrist', 'LShoulder-LElbow', 'LElbow-LWrist'를 연결이다. 이 연결 관계를 통해 모델은 이미지에서 팔의 구조를 인식 및 표현할 수 있다. 각 POSE PAIRS에 해 당하는 PAF(Part Affinity Fields)는 mapIdx를 통해 추 출된다. PAF는 키포인트 간의 연결과 방향을 나타내는 2D벡터 필드로, 이를 통해 모델은 각 관절의 위치뿐만 아니라 그들 사이의 연결 구조와 방향성을 알 수 있다. 이 정보는 팔의 움직임과 각도를 정확하게 추정하는 데 사용된다.

본 논문에서는 흡연자 탐지를 위한 보완책으로 OpenPose로부터 추출되는 키포인드를 기반으로 팔의 각도를 계산하고 그 각도가 흡연시 사람의 팔 각도와 유사한 경우 흡연으로 판정하는 하도록 하였다. 팔굽의 각도는

'RShoulder-RElbow, RElbow-RWrist' 두 벡터 또는 'LShoulder-LElbow, LElbow-LWrist' 두 벡터를 구한 후 벡터의 내적(Dot Product)를 이용하여 구하였다.

그림 7. 팔의 각도 변화를 인지하여 흡연행위 판단 알고리즘 Fig. 7. Algorithm for determining smoking behavior by recognizing changes in the angle of the arm

그림 7은 흡연행위를 판정하는 코드이다. 그림 7에서 pose_estimation.calculate_angle() 함수는 매개변수인 프레임에서 OpenPose를 이용하여 팔굽에 대한 벡터를 찾아 팔굽의 각도를 계산하여 리턴하는 함수이다. 100도 이상이었던 팔굽의 각도가 50이하로 줄어드는 경우 흡연행위의 가능성으로 판단하였고 이것이 100프레임에 대하여 12번 이상 발생하면 흡연행위로 판정하였다. 여기서 100, 50, 0.12 등의 상수는 다양한 실험에의한 찾은 수치이다.

5. 통합 알고리즘

최종적인 흡연자 판별은 YOLOv8 모델에 기반한 detect_smoker() 함수와 OpenPose 모델에 기반한 detect_smoking() 함수의 결과를 종합하여 판정하였다. 두 함수에서 모두 흡연자로 판정되었을 때 한해서 최종적으로 해당 대상을 흡연자가 있는 영상으로 판정한다. 두 모델 모두에서 흡연자로 판정이 내려질 때만 최종적으로 흡연자 판단을 내리기 때문에, 각 모델의 단점 및 오류를 최소화하는 효과가 있다. 또한, 각 모델이 각각 다른 특성(흡연지점 검출 빈도 및 팔의 각도 변화)을 중점으로 살피기 때문에, 더 다양한 흡연 행동 패턴을 인식할 수 있다는 이점이 있다. 그림 8는 다중 모델 결합방식을 포함한, 흡연자 검출 알고리즘이다.

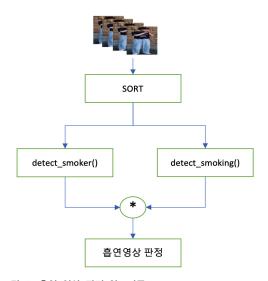


그림 8. 흡연 영상 판정 알고리즘 Fig. 8. Smoking image judgment algorithm

IV. 모델 검증

시험은 100 프레임(약2분)으로 구성된 10개의 영상을 이용하여 측정되었다. 7개의 영상은 2인이 나타나는 영상이고, 3개의 영상은 3인이 나타나는 영상이다. SORT 알고리즘이 적용되어 총 23개의 사람 객체에 대한 영상을 시험한 것이라 할 수 있다. 시험 영상과 결과의 예는 그림 9와 같다. 그림에서 볼 수 있듯이 사람의 자세와는 관계없이 판정하는 것을 알 수 있다.





(a) 왼쪽 흡연, 오른쪽 비흡연으로 (b) 왼쪽 비흡연, 오른쪽 흡연으로 판정 판정

그림 9. 시험 영상에서 캡처한 샘플 프레임들 Fig. 9. Sample frames captured from the test image

검증은 Google Colab 환경의 NVIDEA Tesla T4 GPU를 사용해 진행하였다. 검증 방식으론 모델의 흡연 자와 비흡연자 구별에 대한 성능을 판단하기 위해 혼돈 행렬(confusion matrix)^[111]을 사용하였다. 표 1은 검증 결과이다.

표 1. 검증 결과 Table 1. Verification Results

	예측 흡연자 (P)	예측 비흡연자 (N)
실제 흡연자 (P)	10(TP)	3(FN)
실제 비흡연자(N)	O(FP)	10(TN)

True Positives(TP)와 True Negatives(TN) 값이 각각 10으로, False Negative가 3으로 분류되었다. 이것은 23명 중에서 실제 흡연자가 13명인데 판정은 10명으로하였다는 것이다. 혼돈행렬에 기초하여 정밀도(Precision), 재현율(Recall) 및 F1-score^[11]는 각각 1, 0.77, 0.87이므로 제안된 방법은 영상에서 비교적 정확하게 흡연자를 검출하며, 흡연자와 비흡연자를 구분한다는 것을 알 수있다.

V. 결 론

본 논문은 YOLOv8을 전이학습 시킨 Object Detection 모델과 Pose Estimation 모델인 OpenPose를 이용해 '흡연자 검출 딥러닝 모델'을 제작하였다. 각각의 결과 값을 종합하여 최종 판단을 내리는 다중 모델 결합방식을 통해, 모델의 결정에 대한 신뢰성을 높이고 각 모델의 단점과 오류를 보완하였다. 또한 다양한 흡연 행동 패턴을 인식하며, 이를 통해 흡연자를 효과적으로 검출하는 모델을 구축하였다. 이러한 방식은 영상에서 흡연자를 효과적으로 검출해냈다.

그러나 본 모델에는 몇 가지 한계점이 존재한다. 첫째 로, 흡연을 하지 않지만 얼굴에 손을 가져다 대는 행동을 반복하는 경우, 이는 흡연 행동으로 오검출될 가능성이 있다. 이러한 문제를 해결하기 위해, 담배 자체를 검출하거나, 연기를 검출하는 등의 여러 추가적인 변수를 도입하여 모델의 정밀도를 높이는 것이 향후 목표이다. 둘째로, 프레임에서 신체 일부가 잘려 보이는 경우 모델의 검출 능력이 제한된다. 이러한 점들은 앞으로의 연구에서 개선될 필요성이 있다.

최종적으로 이 연구는 영상 속 흡연자 검출에 대한 새로운 접근 방식을 제시함으로써, 공공장소에서의 흡연 감시 및 행동 모니터링 시스템의 효율성을 높일 수 있는 발판을 마련하였다고 볼 수 있다.

References

- [1] Tamang, Sanjog, et al. "Enhancing COVID-19 Safety: Exploring YOLOv8 Object Detection for Accurate Face Mask Classification." International Journal of Intelligent Systems and Applications in Engineering, Vol. 11, No. 2, pp. 892-897, 2023.
- [2] Jae Moon Lee, Kitae Hwang, Inhwan Jung, Development of a Hole Cup Recognition Model on Golf Green Using Object Detection Technology, The Journal of The Institute of Internet, Broadcasting and Communication (IIBC), Vol. 23, No. 1, pp. 15-21, 2023.
 - DOI: http://doi.org/10.7236/JIIBC.2023.23.1.15
- [3] Fujiwara, K., and K. Yokomitsu. "Video-based tracking approach for nonverbal synchrony: a comparison of Motion Energy Analysis and OpenPose." Behavior Research Methods, Vol. 53, pp. 2700-2711, 2021.
- [4] Choi, Nak-Hun, Oh, Jong-Seok, Ahn, Jong-Rok and Kim, Key-Sun, "A Development of Defeat Prediction Model Using Machine Learning in Polyurethane

Foaming Process for Automotive Seat", Journal of the Korea Academia-Industrial cooperation Society, Vol. 22, No. 6, pp. 36-42, 2021.

DOI: http://doi.org/10.5762/KAIS.2021.22.6.36

- [5] AIHUB: https://aihub.or.kr/
- [6] Cigarette Smoker Detection: https://www.kaggle.com/datasets/vitaminc/cigarette-s moker-detection
- [7] Roboflow: https://app.roboflow.com/object-detection-nr1m4
- [8] Horváth, Dániel, et al. "Object detection using sim2real domain randomization for robotic applications." IEEE Transactions on Robotics, Vol. 39, No. 2, pp. 1225-1243, 2022.
- [9] Bewley, Alex, et al. "Simple online and realtime tracking." 2016 IEEE international conference on image processing (ICIP). IEEE, pp. 3464-3468, 2016.
- [10] Patel, Hitesh A., and Darshak G. Thakore. "Moving object tracking using kalman filter." International Journal of Computer Science and Mobile Computing, Vol. 2, No. 4, pp. 326-332, 2013.
- [11] Luque, Amalia, et al. "The impact of class imbalance in classification performance metrics based on the binary confusion matrix", Pattern Recognition, 91, pp 216–231, 2019).

저 자 소 개

김 동 준(비회원)



- 한성대학교 컴퓨터공학부 재학
- 딥러닝, 기계학습, 컴퓨터 비전

최 유 진(비회원)



- 한성대학교 컴퓨터공학부 재학
- 딥러닝, iOS

박 경 민(비회원)



- 한성대학교 컴퓨터공학부 재학
- 딥러닝, 게임 프로그래밍, iOS

정 인 환(정회원)



- KAIST 정보및통신공학과 박사
- 삼성전자 시스템사업부 수석연구원
- 한성대학교 컴퓨터공학과 교수
- 주관심분야 : 망관리, 멀티미디어신, IoT

박 지 현(비회원)



- 한성대학교 컴퓨터공학부 재학
- 인공지능, 모바일 시스템

이 재 문(정회원)



- 한양대학교 전자공학과(학사) 한국과학기술원 전기및전자공학과(석사) 한국과학기술원 전기및전자공학과 (박사)
- 경력 : 한국통신 연구개발단
- 관심분야 : 기계학습, 게임프로그래 밍, 감성컴퓨팅

황 기 태(정회원)



- 서울대학교 컴퓨터공학과 학사 서울대학교 컴퓨터공학과 석사 서울대학교 컴퓨터공학과 박사
- 경력 University of Callifornia, Irvine 방문교수

University of Florida 방문 교수

• 주관심분야 : 모바일 시스템

※ 본 연구는 한성대학교 교내 학술 연구비를 지원받았음