

다중 융합 네트워크 기반 이동 객체 행동 인식

김진아, 문남미
호서대학교 컴퓨터공학과
kkim.jinah00@gmail.com, nammee.moon@gmail.com

Behavior Recognition of Moving Object based on Multi-Fusion Network

Jinah Kim, Nammee Moon
Dept. of Computer Science and Engineering, Hoseo University

요 약

단일 데이터로부터의 이동 객체에 대한 행동 인식 연구는 데이터 수집 과정에서 발생하는 노이즈의 영향을 크게 받는다. 본 논문은 영상 데이터와 센서 데이터를 이용하여 다중 융합 네트워크 기반 이동 객체 행동 인식 방법을 제안한다. 영상으로부터 객체가 감지된 영역의 추출과 센서 데이터의 이상치 제거 및 결측치 보간을 통해 전처리된 데이터들을 융합하여 시퀀스를 생성한다. 생성된 시퀀스는 CNN(Convolutional Neural Networks)과 LSTM(Long Short Term Memory)기반 다중 융합 네트워크 모델을 통해 시계열에 따른 행동 특징들을 추출하고, 깊은 FC(Fully Connected) 계층을 통해 특징들을 융합하여 행동을 예측한다. 본 연구에서 제시된 방법은 사람을 포함한 동물, 로봇 등의 다양한 객체에 적용될 수 있다.

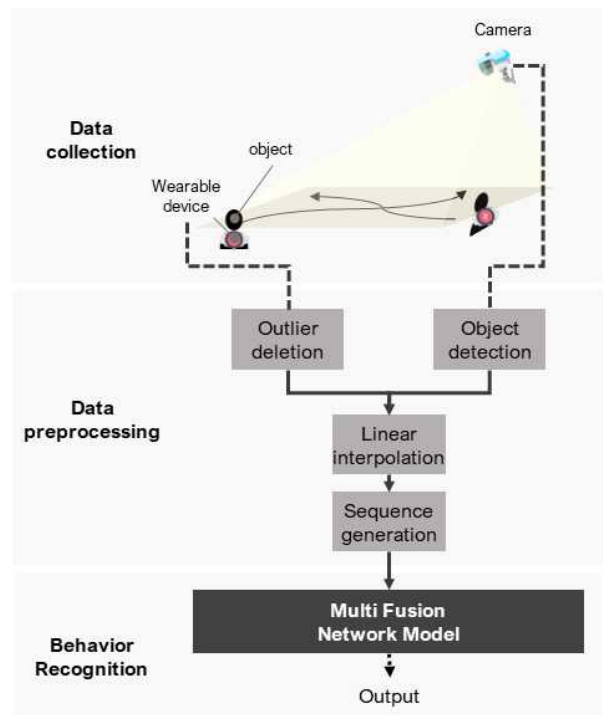
1. 서론

이동 객체에 대한 행동 인식 연구는 영상 중심의 연구에서 센서를 활용한 연구로 확장되고 있다[1, 2, 3]. 한 시점에 대하여 앉다, 서다 등 자세 위주의 단편적인 행동을 예측하거나 시계열에 따른 객체의 이동 방향과 속도를 고려하여 걷다, 뛰다 등의 동적인 행동을 예측할 수 있다.

그러나 데이터 수집 과정에서 발생하는 노이즈로 인하여 행동 인식에 대한 어려움이 있다[2]. 영상의 경우 객체가 가려져 보이지 않을 수 있으며, 센서의 경우 센서가 부착된 위치나 방향에 따른 영향이 크다[4, 5]. 이를 보완하기 위해 다른 유형의 데이터를 융합할 수 있으며, 행동 인식의 정확도를 개선할 수 있다[6].

본 논문에서는 영상 데이터와 센서 데이터를 이용하여 다중 융합 네트워크 기반의 이동 객체 행동 인식에 대하여 다룬다. 영상으로부터 객체의 움직임 영역과 센서로부터 객체의 움직임을 살피고, 융합 네트워크 기반의 모델을 통해 행동을 인식하고자 한다. 이때, 객체는 사람뿐만 아니라 동물 등 주체적으로 움직일 수 있는 객체로 한정한다.

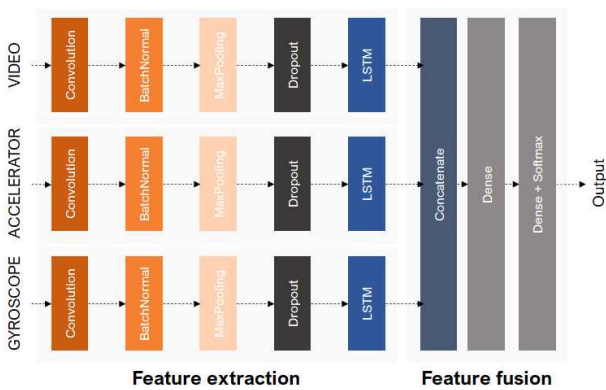
2. 이동 객체 행동 인식 프로세스



(그림 1) 다중 융합 네트워크 기반 이동 객체 행동 인식 프로세스

본 논문에서 제안하는 다중 융합 네트워크 기반 이동 객체 행동 인식은 (그림 1)과 같이 진행된다. 영상 데이터는 제한된 공간에서 RGB 카메라를 이용해 객체의 움직임에 대해 수집된다. 이때, 객체는 가속도와 자이로 센서가 있는 웨어러블 기기를 착용한 상태로 센서 데이터가 같이 수집된다. 데이터 전처리 과정은 영상과 센서 데이터의 특징에 맞게 개별로 진행된다. 영상 데이터에서는 프레임마다 감지된 영역(중심좌표 x , 중심좌표 y , 너비, 높이)을 추출하고, 객체가 감지되지 않은 프레임에 대하여 앞과 뒤 프레임의 감지 영역으로 선형 보간을 통해 감지 영역을 추정한다. 센서 데이터에서는 z-score 정규화를 통해 이상치 제거를 진행하며, 데이터 결측이 발생한 부분은 마찬가지로 선형 보간을 통해 예측된 값을 대체한다. 이후 전처리된 모든 데이터를 시간 기준으로 동기화하여 시퀀스를 생성한다. 최종적으로, 생성된 시퀀스는 다중 융합 네트워크 모델에 입력되어 서다, 걷다, 뛰다 등의 객체 행동들을 예측한다.

3. 이동 객체 행동 인식을 위한 다중 융합 네트워크 모델



(그림 2) 다중 융합 네트워크 모델 구조

이동 객체의 행동 인식을 위한 다중 융합 네트워크 모델은 각 데이터로부터 시계열에 따른 특징들을 추출하는 부분과 추출된 특징들을 융합하는 부분으로 구성된다. (그림 2)는 Keras 환경에서 모델 구조를 시각화한 그림이다. CNN(Convolutional Neural Networks)과 LSTM(Long Short Term Memory)을 하이브리드하여 시계열에 따른 특징들을 추출하고, 추출된 모든 특징들을 깊은 FC(Fully Connected) 계층을 통해 융합한다. 최종적으로 Softmax 활성화 함수를 통해 객체의 행동을 예측한다.

4. 결론

본 연구는 영상 데이터와 센서 데이터를 이용하여 다중 융합 네트워크 기반의 이동 객체 행동 인식 방법을 제안하였다. 제시된 방법은 사람을 포함한 동물, 로봇 등의 다양한 객체에 적용될 수 있다. 다만, 객체가 외부 영향을 적게 받기 위해 적합한 위치의 센서 부착이 필요하며, 시퀀스의 길이는 객체의 평균 행동 시간을 기준으로 설정되어야 한다. 향후 연구에서는 일부 데이터의 결측이 발생하더라도 객체의 행동 특성들을 반영하여 행동 인식률을 높이고자 한다.

ACKNOWLEDGEMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2021R1A2C2011966).

참고문헌

- [1] Dang, L. M., Min, K., Wang, H., Piran, M. J., Lee, C. H., Moon, H. "Sensor-based and vision-based human activity recognition: A comprehensive survey" Pattern Recognition, 108, 107561, 2020.
- [2] Ehatisham-Ul-Haq, M., Javed, A., Azam, M. A., Malik, H. M., Irtaza, A., Lee, I. H., Mahmood, M. T. "Robust human activity recognition using multimodal feature-level fusion" IEEE Access, 7, 60736-60751, 2019.
- [3] Uddin, M. Z., Hassan, M. M., Alsanad, A., Savaglio, C. "A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare" Information Fusion, 55, 105-115, 2020.
- [4] Kim, J., Moon, N. "Dog Behavior Recognition Based on Multimodal Data from a Camera and Wearable Device" Applied Sciences, 12, 6, 3199, 2022.
- [5] 김정훈, 최중혁, 박영호 "객체 탐지와 행동인식을 이용한 영상내의 비정상적인 상황 탐지 네트워크" 멀티미디어학회논문지, 24, 2, 186-198, 2021.
- [6] Ranieri, C. M., MacLeod, S., Dragone, M., Vargas, P. A., Romero, R. A. F. "Activity recognition for ambient assisted living with videos, inertial units and ambient sensors", Sensors, 21, 3, 2021.