

영유아/아동의 발달장애 조기선별을 위한 제한적/상동적 행동 검출에 관한 연구

유철환*, 유장희, 김호원, 한병옥
한국전자통신연구원

{ch.yoo*, jhy, hw_kim, byungok.han}@etri.re.kr

A study on the detection of restricted and repetitive behavior for early screening of developmental disorders in infants and children

Yoo Cheol Hwan*, Yoo Jang Hee, Kim Ho Won, Han Byung Ok
Electronics and Telecommunications Research Institute (ETRI)

요 약

제한적/상동적 행동은 자폐 스펙트럼 장애 (Autism Spectrum Disorder, ASD)를 가진 아동과 전형적 발달 (Typical Development, TD) 아동을 구분 짓는 중요한 행동 지표 중 하나이다. 딥러닝 기반의 제한적/상동적 행동 검출 기술은 현재의 의료전문가에 의한 시간 소모적인 검사를 완화하고 치료와 향후 예후에 중요한 ASD 아동의 조기 선별에 도움을 줄 수 있다. 본 논문에서는 반복적 행동 구간 제안, 휴먼 포즈 예측 모듈, 그리고 단위 행위 인식 모듈로 구성된 제한적/상동적 행동 검출 통합 파이프라인에 대한 설계와 실험결과에 관한 연구를 제시한다. 제안하는 검출 파이프라인은 자체 제작한 RBC 데이터셋 뿐만 아니라 PERTUBE, QUVA 등의 다양한 크로스 데이터셋에서 평균 정확도 71.3%의 우수한 검출 성능을 확보하였으며, 이를 통해 제안하는 알고리즘의 ASD 아동 조기선별 분야에서의 높은 활용 가능성을 증명한다.

I. 서 론

자폐 스펙트럼 장애 (Autism Spectrum Disorder, ASD)는 미국정신의학회에서 발행하는 정신장애 분류체계인 정신장애의 진단 및 통계 편람 (Diagnostic and Statistical Manual of Mental Disorder, DSM)에 의해 크게 사회적 의사소통 및 상호작용의 지속적인 결함, 그리고 제한적이고 반복적인 행동과 같은 특징으로 정의된다. 이러한 ASD의 유병률은 매년 지속적으로 증가하고 있으며, 따라서 ASD 아동에 대한 조기 진단은 아동의 뇌에 대한 가소성이 높은 시기에 정상적인 형태로 변화할 수 있는 기회를 제공할 뿐만 아니라 이차적인 신경학적 손상과 행동문제가 축적되는 것을 방지할 수 있다는 측면에서 매우 중요하다.

본 논문에서는 이러한 ASD 아동의 조기선별을 보조하기 위해 우리의 선행 연구인 논문[1,2]를 확장하여 딥러닝 기반의 제한적/상동적 행동 검출 통합 파이프라인 설계 및 실험결과에 관한 연구를 제시한다.

II. 본론

제한적/상동적 행동 검출 파이프라인의 대략적인 구조도는 그림 1과 같다. 제한적/상동적 행동 검출 파이프라인의 경우 크게 3가지의 주요 부분으로 구성되며, 제한적/상동적 행동이 자극에 의한 강제적인 반응 유도가 힘들고 시계열 데이터 상의 임의의 구간에서 발생한다는 점에 대응하기 위한 반복적 행동 구간 제안 모듈, 멀티모달 데이터 기반의 성능향상을 위한 휴먼 포즈 예측 모듈, 최종적으로 반복적 행동 발생 유무를 판단하기 위한 단위 행위 인식 모듈로 구성된다.

먼저 반복적 행동 구간 제안 모듈의 경우 시간적 멀티 스케일 특징(temporal multi-scale features)을 활용하기 위해 논문[1]에서와 같이 멀티 스케일 U-Net과 멀티 스케일 특징 앙상블 모듈로 구성된 네트워크를 채택하였다. 또한 인식 성능 향상을 위한 멀티모달 데이터의 경우 그림 1에서 볼 수 있듯 3차원 휴먼 포즈 정보를 활용하였으며, 이를 위해 단일 RGB 영상 기반 어피어런스 정보를 바탕으로 3차원 신체 좌표 추론이 가능한 3DMPPE[3] 네트워크를 채택하였다. 마지막으로 단위 행위 인식 네트워크의 경우 반복적 행동이 비반복적 행동에 비해 시계열 데이터상의 시간 축을 기준으로 행동의 유사도가 높다는 관찰을 토대로 Siamese network 기반의 contrastive learning 기법을 제안한 SiRepNet[2]를 활용하였으며, backbone으로는 I3D[4] 네트워크를 채택하였다.

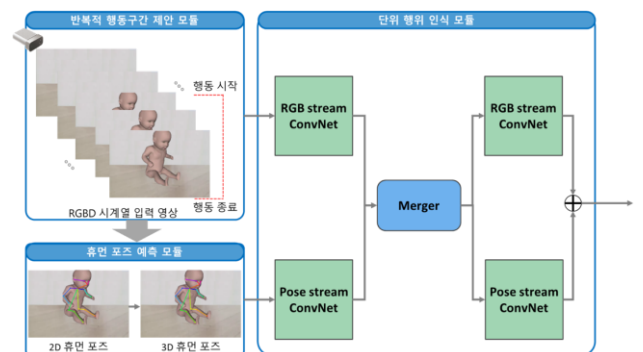


그림 1. 제한적/상동적 행동 검출 파이프라인

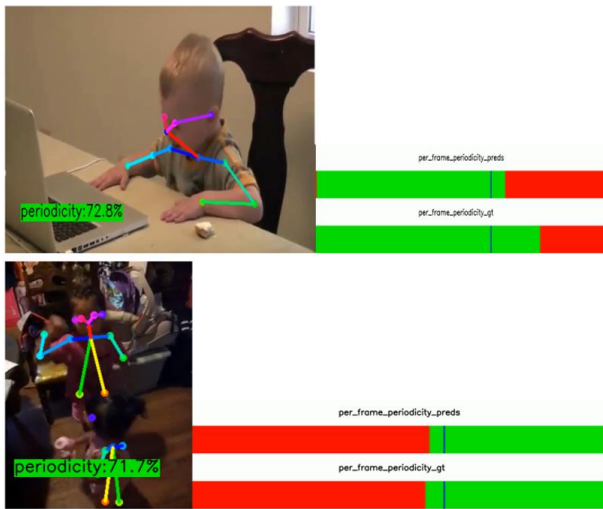


그림 2. 제한적/상동적 행동 검출 결과 예시

따라서 최종적으로 통합된 검출 파이프라인의 경우 그림 2의 예시와 같이 입력 비디오내의 임의의 시간대에 발생하는 행동 시작, 행동 종료에 대한 구간을 제안하고, 제안된 구간에 대한 RGB 및 3D 휴먼 포즈의 멀티모달 데이터를 활용하여 반복적 행동 발생 확률을 추론하는 통합 알고리즘으로 구성된다.

III. 실험

먼저 이진 주석으로 반복적 행동 분류 태스크를 처리하는 공개 비디오 데이터셋이 없기 때문에 네트워크 학습 및 성능 평가를 위해 그림 3과 같이 기존의 행동인식에 널리 쓰이는 Kinetics[4]와 반복구간 검출을 위한 Countix[5] 데이터셋을 목적에 맞게 재구성하였다. 우리는 Countix 데이터셋을 positive 샘플로 활용하였으며, 이와 같은 숫자의 샘플을 Kinetics 데이터셋에서 임의로 선택하여 총 6,574 개의 학습 비디오 데이터 및 4,672 개의 테스트 비디오 데이터로 구성된 RBC 데이터셋을 구성하였다. 이 밖에도 기존의 반복 구간 검출 혹은 반복횟수 추론 태스크에 활용되는 PERTUBE[6]와 QUVA[7] 데이터셋을 성능 평가를 위해 활용하였다. 상동적 행동 검출 성능 측정을 위해 accuracy(Acc), recall(R), precision(P), 그리고 F1-score(F1)의 평가지표를 활용하였으며, 실험 결과는 표 1과 같다. 표 1에서 제시된 것과 같이 제안하는 파이프라인을 활용하였을 때 학습에 활용된 도메인의 데이터셋 이외에 PERTUBE, QUVA 등의 크로스 도메인 데이터셋에서도 우수한 검출 성능을 보여주었으며, 평균 검출률 71.3%의 제한적/상동적 행동 검출률을 확보하였다. 또한 그림 2의 결과 예시와 같이 임의의 입력 동영상에 대해서 높은 정밀도의 검출 성능을 보임을 확인하였다.

표 1. 제한적/상동적 행동 검출 정확도

학습 DB	테스트 DB	Acc(%)	R(%)	P(%)	F1(%)
RBC	RBC	75.9	62.6	85.4	72.2
RBC	PERTUBE	76.0	76.0	100.	86.4
RBC	QUVA	62.0	62.0	100.	76.5
평균		71.3	66.9	95.1	78.4

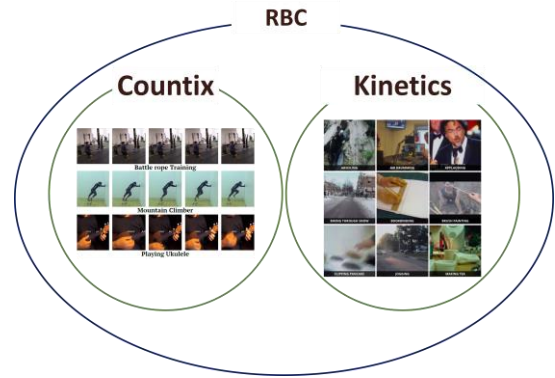


그림 3. RBC 데이터셋 구성도

IV. 결론

본 논문에서는 자폐 스펙트럼 장애를 가진 아동의 주요한 특징인 제한적/상동적 행동을 효율적으로 검출하기 위한 통합 검출 파이프라인 설계와 실험결과에 대해 제시하였다. 제안하는 파이프라인은 반복적 행동 구간 제안 모듈을 통해 임의의 시간대에 발생하는 아동의 반복적 행동 구간을 제안하고 제안된 구간에 대한 멀티 모달 데이터 기반의 반복적 행동 인식 모듈을 통해 최종적으로 제한적/상동적 행동 발생 확률을 추론한다. 제안하는 검출 파이프라인을 통해 다양한 테스트 데이터셋에 대해 평균 정확도 71.3%의 검출 성능을 확보하였으며, 이를 통해 제안하는 알고리즘이 실제 자폐 스펙트럼 장애를 가진 아동을 조기 선별하는데 있어 핵심 요소 기술로 사용될 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2019-0-00330, 영유아/아동의 발달장애 조기선별을 위한 행동·반응 심리인지 AI 기술 개발)

참 고 문 헌

- [1] 유철환, et al, "주기적 행동 검출을 위한 멀티스케일 U-Net," 전자공학회논문지, 2021.
- [2] Yoo, C. et al, "Simple Yet Effective Approach to Repetitive Behavior Classification based on Siamese Network," *ICPR*, 2022.
- [3] Moon, G. et al, "Camera distance-aware top-down approach for 3d multi-person pose estimation from a single rgb image," *ICCV*, 2019.
- [4] Carreira, J. et al, "Quo vadis, action recognition? a new model and the kinetics dataset," *CVPR*, 2017.
- [5] Dwibedi, D. et al, "Counting out time: Class agnostic video repetition counting in the wild," *CVPR*, 2020.
- [6] Panagiotakis, C. et al, "Unsupervised detection of periodic segments in videos," *ICIP*, 2018.
- [7] Runia, T. F. et al, "Real-world repetition estimation by div, grad and curl," *CVPR*, 2018.