

아동 외현화 장애 조기발견을 위한 인공지능 기반의 영상 분석 시스템 개발

안시현*, 박윤하**, 문병인***

Developing an AI-based Image Analysis System for Children's Externalization Disorder

Shi-Hyun Ahn*, Yun Ha Park**, and Byungin Moon***

본 성과물은 산업통상자원부에서 지원하는 2023년도 지식서비스산업기술개발사업(No. 20014682)의 연구 수행으로 인한 결과물임을 밝힙니다

요 약

아동 주의력 결핍/과잉행동 장애(ADHD) 및 적대적 반항장애는 주의산만, 적대적 반항행동, 충동성을 주 증상으로 보이는 정신질환이다. 본 논문에서는 이러한 장애를 조기발견 할 수 있는 실시간 다중센서 데이터 기반 영상 동시 입력 기술, 집중도 분석 및 딥러닝 기반 아동 이상행동 분석기술들을 제안한다. 다채널 실시간 데이터 수신을 위해 Codec SDK와 연동하여 입력되는 프레임 RAW 용량을 줄여서, 기존 연구 20개 동시 입력에서 30개로 채널수를 확장하였다. 아동 집중도 분석을 위해 영상 흔들림에 강인한 시선 추정 모델을 적용하여 평균 0.968 Accuracy를 확보하였다. 이상행동 분석은 아동 특징 맵 추출과 시공간 특징 동시 분석을 통해 기존 연구 대비 f1-Score가 80%에서 83% 이상의 높은 성능개선 결과를 도출하였다. 향후 비식별화 한 15,000건 이상의 학습데이터 및 강인한 행동 분석알고리즘 추가 적용으로 행동 분석 정확도를 90%이상으로 개선할 예정이다.

Abstract

Children's Attention-Deficit Hyperactivity Disorder and Oppositional Defiant Disorder are mental diseases that have main symptoms of distraction, oppositional defiant behavior, and impulsivity. In this paper, we propose technologies(multiple image processing, concentration and deep learning-based behavior analysis) for early detection. In conjunction with the Codec SDK, RAW capacity in input frames was reduced to expand the number of channels from 20 to 30 in real time. To analyze children's concentration, an average accuracy of 0.968 was achieved by applying VSL2G-Net, a gaze estimation model with image stabilization. The f1-Score abnormal behavior recognition resulted in a high performance improvement of 80%(previous studies) to 83% or more with extraction feature maps and simultaneous analysis of spatio-temporal features. In the future, we plan to improve more than 90% of behavioral analysis accuracy by adding more than 15,000 learning data with de-identification and applying a robust behavioral analysis algorithm.

Keywords

externalization disorder, abnormal behavior, AI, concentration, multiple image processing

* ㈜ 우경정보기술 AI/DX융합개발부문 이사

- ORCID: <https://orcid.org/0000-0002-9992-4216>

** ㈜ 우경정보기술 대표이사

- ORCID: <https://orcid.org/0000-0002-9247-2587>

*** 경북대학교 전자공학부/대학원 전자전기공학부 교수
(교신저자)

- ORCID: <https://orcid.org/0000-0002-8102-4818>

· Received: Jul. 26, 2023, Revised: Aug. 07, 2023, Accepted: Aug. 10, 2023

· Corresponding Author: Byungin Moon

School of Electronics Engineering, Graduate School of Electronic and

Electrical Engineering, Kyungpook National University, 80 Daehakro, Bukgu,

Daegu 41566, Korea

Tel.: +82-53-950-7580, Email: bihmoon@knu.ac.kr

I. 서 론

아동 주의력 결핍/과잉행동 장애(ADHD, Attention-Deficit Hyperactivity Disorder)는 신경발달질환으로 초창기에는 아동기인 4세 이전에 시작하며 12세 이전에는 확실히 나타난다. 만성적으로 경과하여 성인으로까지 증상이 이어지는 경우가 빈번하다. 적절한 치료를 하지 않으면 학업, 직업 사회생활, 대인관계 등에 문제가 발생할 수 있다. 적대적 반항장애(Oppositional defiant disorder)는 부모가 자신의 필요에 의해 아동의 행동을 힘이나 권위로 억제하려 하면 투쟁이 일어나서 발병하게 된다. 생물학적 요인으로는 유전적 요인과 DHEA-S의 증가 등이 거론되고 있다. 거부적, 적대적 반항적 행동이 주 증상으로 행실장애(Conductive disorder)나 반사회적 인격장애(Antisocial personality disorder)처럼 사회적 규범을 위반하거나 타인의 권리를 침해하는 반사회적, 공격적인 행동은 많지 않다. 빠르면 3세부터 시작될 수 있으며, 전형적으로는 8세 이전에 시작한다. 하은혜 등은 복지시설 아동을 대상으로 ADHD증상인 주의력결핍 및 과잉행동 충동과 굵은 인지적 템포(SCT, Sluggish Cognitive Tempo)가 내재화-외현화 문제와 학업적-사회적 기능에 미치는 영향력을 검증하였다[1]. 본 과제에서는 우선 외현화 장애의 표적 행동을 선별하여 목록화 하였다. 9가지 이상행동에 대해서 비슷한 유형으로 크게 4가지 Class로 구분하여 이상행동을 재정의하였고 딥러닝 기반의 이상행동 분석에 적용하였다. 또한 외현화 장애의 대표적인 증상인 주의력 결핍 발견을 위해 얼굴 탐지 및 딥러닝 기반의 시선추정 모델을 적용하여 집중 유무를 판단하였다.

최종적으로 주의력 결핍/과잉행동 장애 및 적대적 반항장애를 조기발견 할 수 있는 다채널 실시간 데이터 기반 영상 동시 입력 기술, 집중도 분석 및 딥러닝 기반 아동 이상행동 분석기술을 제안한다. 이러한 기술들을 활용한다면 아동의 외현화 장애를 발견하여 조기치료를 받을 수 있다. 아동의 이상행동 및 집중도를 조기발견하기 위한 다양한 연구들이 진행되었다. D. Ni et al.은 TGAM 칩에서 알고리즘 최적화를 통해 iPad 기반 모바일 학습자 그룹의 EEG 실험을 수행하였다[2]. X. Xu et al.은 특징 선택에 기반한 다단계 주의 인식 방법을 제안하였

다. 10개의 뇌파도(EEG) 채널에서 시간 영역 측정, 샘플 엔트로피 및 주파수 대역 에너지 비율을 포함하여 총 10개의 특징 추출 방법을 제안하였다[3]. 추상현 등은 2가지 유형의 감정 이미지(사회적/비사회적 공포, 사회적/비사회적 슬픔, 사회적/비사회적 중립적)를 볼 때 뇌 신호를 수집, 보조 컨텍스트 정보를 사용한 감정 인식에서 원시 EEG 기반 컨볼루션 신경망(CNN)에 대한 다중 작업 학습(MTL)의 효과를 연구하였다[4]. B. Meriem et al.은 e-learning 환경에서 전이 학습 및 데이터 증강 기술을 사용하여 학습자의 표정에서 실시간으로 집중 수준을 결정하는 방법을 제안하였다[5]. X. Miao et al.은 학생들의 수업 집중력 평가를 위해 집중력 분석 및 평가 모델을 사용하여 표정을 인식하고 해당 가중치를 부여하여 표정 점수를 계산하였다[6]. 이런 기존 연구들 대부분은 뇌파 측정 장비를 착용하여 집중도를 분석하였다. 정확도는 높으나 고가의 뇌파 측정 장비가 반드시 필요하다.

본 논문에서는 카메라로 촬영한 영상을 사용하여 아동 집중도를 분석하였다. 아동 얼굴을 검출하고 시선 추정을 통해 집중력을 분석하기에 별도의 장비 착용 없이 다수의 공간과 아동들을 대상으로 집중도를 분석할 수 있다. T. Nose et al.은 아동 추락 사고를 방지하기 위해 빠르게 탐색하는 랜덤 트리를 기반으로 어린이의 낙상을 측정하는 시스템을 개발하였다[7]. H. Duan et al.은 스켈레톤 기반 동작 인식 방법 GCN을 채택하여 인간 골격 톤의 특징을 추출하였고 골격 기반 동작 인식에 대한 새로운 접근 방식인 PoseConv3D를 제안하였다[8]. T. Ahmad et al.은 스켈레톤 기반 동작 인식에 널리 사용되고 있는 그래프 컨볼루션 네트워크를 통한 동작 인식 기술을 설명하고 소스코드 및 사용 가능한 리소스를 제공하였다[9]. Y. Zhang et al.은 정형화된 행동의 인식 과정을 가속화하기 위해 골격 데이터와 장단기 기억(LSTM)을 기반으로 한 행동인식 방법을 제안하였다[10]. K. Hwang et al.은 딥러닝 기반의 얼굴 인식과 손 제스처 인식을 활용하는 원격 제어 시스템을 제안하였다[11]. B. Hwang et al.은 동적인 운동 자세를 분석하여 운동 자세 정확도를 보여주는 딥러닝 모델을 제안하였다[12]. 기존 연구에서 제안한 방법은, 급속탐색 랜덤트리(Rapidly explore

random tree)의 관절 데이터들이 충분치 않다면 자세 예측은 불가능하다. 영상에서 많은 객체가 존재할 경우 시간 흐름과 함께 관절 좌표값은 서로 다른 객체 간 겹치는 문제가 발생할수 있다.

본 논문에서 제안한 행동 분석 방법은 현재 프레임에 데이터가 손실되더라도 일정한 다른 데이터들을 기본으로 학습이 가능하다. 또한 한 화면에 다양한 행동이 다수 존재하여도 동시에 다중 행동 분류가 가능하다.

II. 외현화 장애 조기발견 시스템

아동 외현화 장애를 조기에 발견하여 적절한 치료를 위한 다중채널 실시간 데이터용 영상 동시 입력 기술, 집중도 분석 및 딥러닝 기반 아동 이상행동 분석기술들을 제안드린다. 맞춤형 아동 케어를 지원하는 통합서비스 플랫폼 핵심기술로 사용된다. 그림 1은 본 논문에서 제안하는 아동 외현화 장애 조기 발견 시스템의 전체구조를 나타낸다. 처음으로 실증 어린이집을 통해 외현화 장애 기반의 행동에 대해 진단검사를 수행하였다. 대학병원 정신의학과 도움으로 아동의 이상행동(외현화 및 내현화) 식별을 위한 표적행동 목록을 정의하였고, ADHD, 적대적반항장애(ODD)등 외현화 장애의 표적 행동 목록을 구체화하였다. 그 결과 대표적인 이상 행동 9가지를 정의하였다. 이상행동 분석을 위한 데이터 추출 및 생성을 위해 집중도 및 행동 분석 데이터셋 구축, 학습모델 설계를 하였다. 아동 주요 '이상행동(Abnormal behavior)' 및 집중도 분석을 위해 정상적인 행동을 제외한 이상행동들을 규정한 후 영상기반 행동분석을 수행하

였다. 그 결과 이상 행동 인식 f1-Score 83% 이상 정확도를 도출하였다. 또한 영상에서 집중해야 하는 영역 지정을 통해 아동의 집중도를 확인한 결과 Accuracy 0.968 결과를 도출하였다.

기존 실내 아동 안전을 위한 딥러닝 기반의 아동 행동분석기술 개발 연구[13]에서 적용한 DeepStream의 multi-stream으로 다중 영상을 받은 경우 모델 예측에서 속도 저하가 발생하였다. 이러한 문제점을 개선하기 위해 Codec SDK와 연동하여 입력되는 프레임 RAW 용량을 줄였다. 데이터수신부는 딥스트림에서 다중센서로부터 MultiURI 및 RTSP를 통해 영상 데이터를 수신하였다. Streammux 연결부 코드 변경하였고 nvstreammux에 있는 입력 GstBuffer의 GstMeta는 출력 배치 버퍼의 NvDsBatchMeta에 복사되고 CUDA Encoder 코드와 연동하였다. 그 결과, 실시간 다중 영상데이터 입력을 위해 20개에서 30개로 채널 수가 확장되었다. 그림 2는 다양한 멀티 채널 데이터 스트리밍 결과이다.



그림 2. 30채널 데이터(스트리밍)
Fig. 2. 30-channel data(streaming)

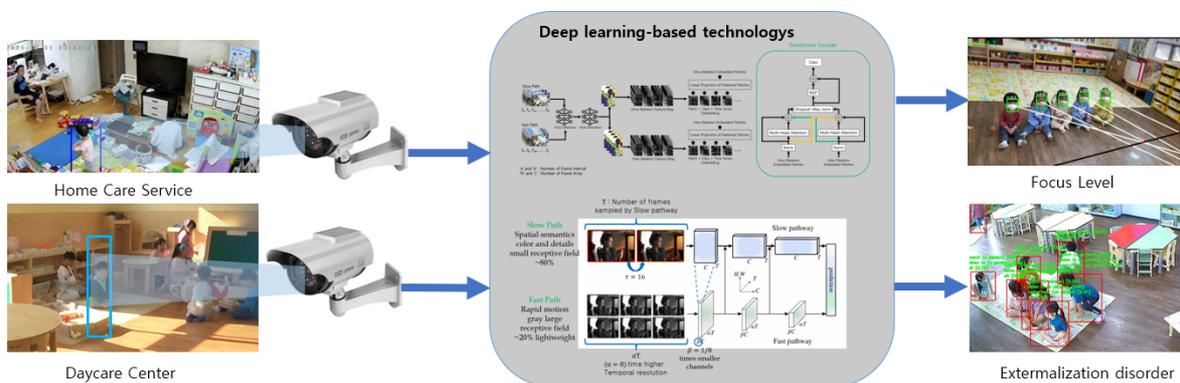


그림 1. 아동 외현화 장애 조기발견 시스템
Fig. 1. Early detection system for child with externalization disorder

2.1 아동 집중도 분석 시스템 개발

시선 추정(Gaze estimation) 기술을 이용한 아동 집중도 분석 시스템 설계하였다. 오재광 등은 특징 기반으로 단일 눈 이미지를 통해 눈 영역의 랜드마크를 감지하여 시선을 추정하는 방법을 제안하였다 [14]. X. Zhou et al.은 Kinect센서를 활용하여 3D 눈 모델을 통해 피사체의 시선을 자유로운 움직임의 허용 범위로 추정하는 효과적인 방법을 제시하였다 [15]. A. A. Abdelrahman et al.은 시선추정 예측 정확도를 향상시키기 위해 개별적으로 각도를 조절할 수 있는 강인한 CNN기반 모델을 제시하였다[16]. 본 논문에서는 아동의 집중도를 분석할 영상이 입력되고 아동의 시선이 향해야 하는 영역인 ROI(Region of Interest)를 우선 지정한 후, 영상 안정화 딥러닝 모델로 흔들리는 영상을 보정하고 객체 탐지 모델로 몸, 얼굴을 탐지하였다. 탐지된 몸, 얼굴 결과로 IOU(Intersection Over Union)를 계산하

여 한 명씩 몸, 얼굴을 매칭시키고, 몸만 탐지됐다면 비 집중 상태로, 몸, 얼굴 모두 탐지됐다면 다음 단계로 이동한다.

탐지된 얼굴 이미지를 사용하여 그림 3과 같은 시선 추정 모델로 아동의 시선을 추정한다. 추정된 시선 좌표가 설정한 ROI영역을 지나가는지 여부를 확인하여 지나간다면 집중 상태로, 지나가지 않으면 비 집중 상태로 분석한다. 그림 4는 아동 집중도 분석 시스템 구성도를 나타낸다.

어린이집에서 수업 시간에 아동을 정면으로 바라보는 각도로 영상 수집을 진행하였다. 아동 집중도 분석 데이터셋 구축 과정은 다음과 같다. 우선 데이터를 정의한 후 수집을 진행하였다. 그 후 데이터 가공을 위해 라벨링 툴을 제작하고 오토 라벨링을 통해 데이터 라벨링을 진행하였고, 최종적으로 인공지능 학습에 적절한 데이터인지 시스템 테스트 과정을 통해 검증하였다.

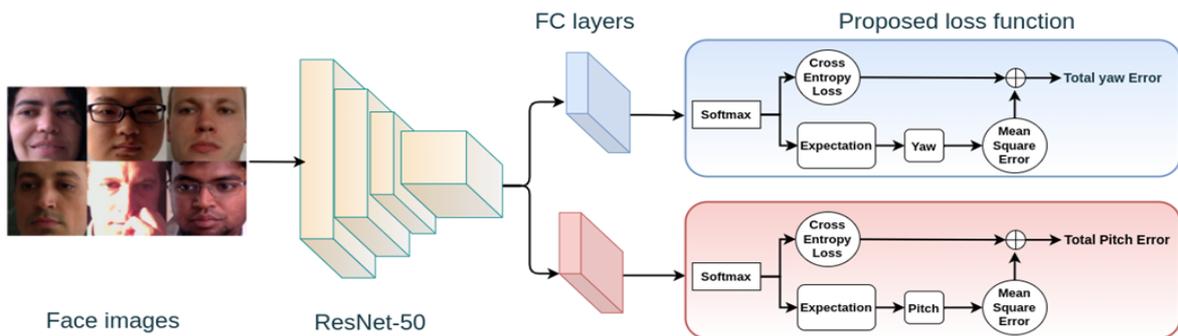


그림 3. 시선 추정 모델 구조
Fig. 3. Structure of the gaze-estimation model

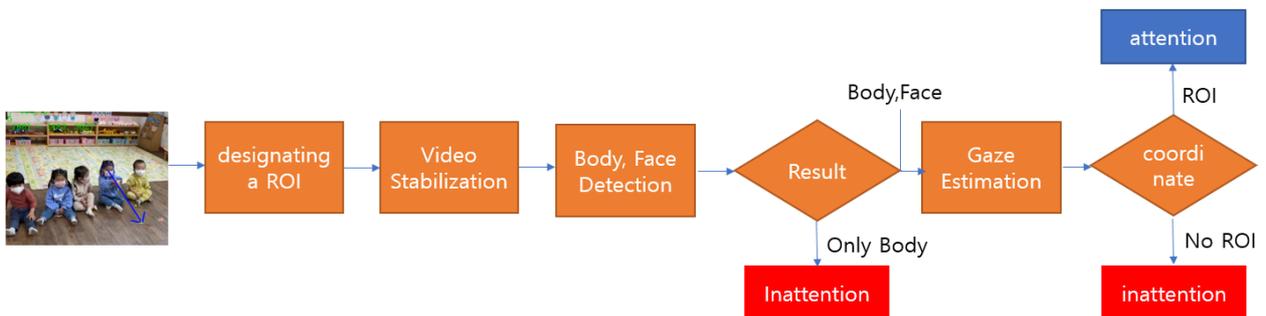


그림 4. 아동 집중도 분석 시스템 구성도
Fig. 4. Structure of the children's visual attention analysis system

아동 집중도 분석 데이터셋 설계 기준은 다음과 같다. 이미지 샘플링은 OPENCV를 통해 영상을 이미지로 변환 및 프레임 수를 조절하고 이미지 확장자는 PNG로 지정한다. 프레임 수는 30프레임당 1장씩 영상에서 이미지를 저장한다. 저장한 이미지들은 객체 탐지 모델로 몸과 얼굴을 오토 라벨링 하여 1차 결과 확인한다. 라벨링 작업자는 1차 결과를 검수하고 몸과 얼굴이 탐지되지 않은 부분은 직접 라벨링한다. 제작한 라벨링 툴을 사용하여 이전 탐지 결과의 이미지에서 탐지된 얼굴의 시선 방향을 추정하여 직접 라벨링한다.

최종적으로 영상 및 촬영정보와 객체의 정보는 라벨링 정보를 입력한다. 그 결과 아동들의 개인정보가 보호되는 범위에서 500건의 집중도 데이터셋을 구축하였다.

J. Redmon et al.은 Yolov5 기반 아동 탐지 및 IOU 기반 아동의 몸, 얼굴 매칭 기술을 구현하였다[17]. 얼굴 검출기를 적용하여 실증 어린이집에서 확보된 아동의 얼굴 대상 테스트 진행하였을 때, 오검출 또는 미검출 결과가 확인되었다. 모델의 임계값에 변화를 주어 오검출과 미검출을 감소시키는 것은 영상마다 임계값을 조절해야 하고 정확하지 않았다. 몸 검출기를 추가로 적용한 결과 오검출이 줄었다. 또한 얼굴을 추가로 추적한 결과 미검출이 되는 경우가 줄어들었다. 추가적으로 유치원 수업시간에 촬영된 영향으로 시선 추정을 위한 영상 흔들림 보정 기술 개발이 필요하였다. 기존 영상 안정화 방법들은 왜곡을 생성하거나 aggressive cropping이 동반되지만 영상 안정화 딥러닝 모델인 Hybrid Neural Fusion을 적용함으로써 이러한 문제점들을 감소시켰다. 그림 5 처럼 아동 집중도 분석을 위해서 집중해야 할 영역을 OPENCV를 통해 영상에서 직접 지정하였다.

프레임에서 탐지된 몸 bbox와 얼굴 bbox로 IOU 계산 결과, 몸만 탐지됐을 때 해당 아동은 얼굴을 숙였다거나 뒤를 돌아보는 행위로 인해 얼굴이 탐지되지 않았다고 판단하여 비 집중 상태로 판단한다. 시선 추정 결과 해당 시선 방향이 지정된 집중 영역을 통과한다면 해당 아동은 집중 한다고 판단한다. 시선 추정 결과 해당 시선 방향이 지정된 집중 영역을 통과하지 않는다면 해당 아동은 비 집중 상태로 계산한다.



그림 5. 집중 영역 설정
Fig. 5. Attention area setting

L2CS-Net, VSL2G-Net 두 가지 시선추정 딥러닝 모델을 적용하여 테스트를 진행하였다. L2CS-Net 테스트 결과, 영상이 흔들릴 때 추정 결과와 실제 아동의 시선에 차이가 있었다. 영상 흔들림 보정이 가능한 VSL2G-Net 시선 추정 모델을 사용하여 테스트 진행했을 때는 같은 프레임에서 올바른 시선 추정 결과를 확인할 수 있었다. 아동의 아동 집중도 분석 결과 지표로 Accuracy를 사용하였다. 그 결과, 평균 Accuracy는 표 1처럼 0.960에서 0.968로 0.008 향상된 정확도를 확인할 수 있었다.

표 1. 데이터 세트별 아동 집중도 분석 정확도
Table 1. Children' visual attention assessment accuracy for each dataset

Data set (number of children)	L2CS-Net accuracy	VSL2G-Net accuracy
A(5)	0.977	0.980
B(7)	0.970	0.978
C(8)	0.948	0.957
D(9)	0.947	0.956
Average	0.960	0.968

2.2 이상행동 분석 시스템 개발

아동의 이상행동(외현화 및 내현화) 식별을 위한 표적행동목록을 정의하였다. DSM-5의 진단기준과 진단을 위해 사용하는 척도 질문지를 사용해 과학적 근거를 기반으로 외현화 장애의 표적 행동을 선별하여 목록화 하였다. ADHD, 적대적반항장애(ODD)등 외현화 장애의 표적 행동 목록을 구체화 하였다. 임상 현장에서 사용하는 심리검사의 문항들의 내용을 분석하였다. 그 결과는 표 2와 같다.

표 2. 아동 이상행동 목록

Table 2. List of children' abnormal behavior

	Externalization disorder		Internalized emotional disorder	
	ADHD	Oppositional defiant disorder	Depressive disorder	Anxiety disorder
fall	fall			
lie down	lie down			
hit		hit		
push		push	push	
steel	steel	steel		
pull	pull	pull		pull
hit	hit	hit		
throw	throw	throw		
swing	swing	swing		

9가지 이상행동에 대해서 비슷한 유형을 크게 4 가지 Class로 구분하여 이상행동을 재정의하였다. Class 1 : 아동이 넘어지는 경우, 아동이 바닥 등에 등을대고 누워있는 상태. Class 2: (성인, 아동등) 손을 사용하여 상대를 폭행한 경우, 아동이 물건 등을 던지는 경우, 사물이나 주먹 등을 허공에 휘두르는 행위. Class 3: 아동이 손, 몸 등을 사용해 상대를 미는 경우, 상대방 물건을 강제로 뺏는 경우, 손, 물건 등을 사용해 상대를 잡아 당기는 경우. Class 4: 타인 또는 사물 등에 신체가 부딪히는 경우이다.

아동의 행동 및 학습활동에 대한 이상행동 인식을 위한 데이터셋을 구축하였다. 총 4가지 이상 행동 class에 대해 3,000개의 영상 데이터를 수집 하였다. 원본 비디오 영상으로부터 아동의 특이 행동이 발생하는 시간 영역을 CSV로 작성하였다. CSV 파일로부터 Clip Video Frame을 자동 추출하였다. 생성된 모든 Clip Video Frame에 대하여 Openpose 기반의 Key-point 추출하였다. 표 3과 같은 Key-point 정보로부터 Feature Map 생성하고 생성된 Feature Map은 Stream Embedding 통해 분할, 분할된 Patch는 Transformer 기반의 병렬 학습 진행하였다.

그림 6과 같은 아동 특징 맵 추출을 위해 방향성이 있거나(Directed) 없는(Undirected) 엣지(Edge)로 연결된 노드(Nodes=verticles)들의 집합인 Graph를 생성하였다. Openpose를 활용하여 Key-point 추출하여 25개 Node와 23개 Edge로 구성하였다. 연결된 각 Node들의 Edge 길이를 Matrix의 값으로 지정하고

지정된 Matrix 값을 L2 Normalization 하여 정규화 처리(0~1) 값으로 변경하였다. 한 프레임 내의 객체마다 Adjacency Matrix를 구성한 다음, 프레임 간의 Adjacency Matrix를 구성하였다. 프레임 간 구성시 같은 Key-point끼리 사잇각을 연산하는데 입력 단에서 프레임 스트림의 재사용이 있으므로 사잇각에서 중복은 제외한다. 기존 2D Convolution으로는 시간에 따른 특징은 분석할 수 없어 3D Convolution으로 변경하였다. 3D Convolution은 x, y의 RGB 픽셀 특징과 시간에 따른 변화인 z 특징까지 같이 Convolution 연산하였다.

표 3. 아동 키포인트 정보

Table 3. Child key-point information

ID	Name	ID	Name
0	Nose	13	LKnee
1	Neck	14	LAnkle
2	RShoulder	15	REye
3	RElbow	16	LEye
4	RWrist	17	REar
5	LShoulder	18	LEar
6	LElbow	19	LBigToe
7	LWrist	20	LSmallToe
8	MidHip	21	LHeel
9	RHip	22	RBigToe
10	RKnee	23	RSmallToe
11	RAnkle	24	RHeel
12	LHip	25	Bbox

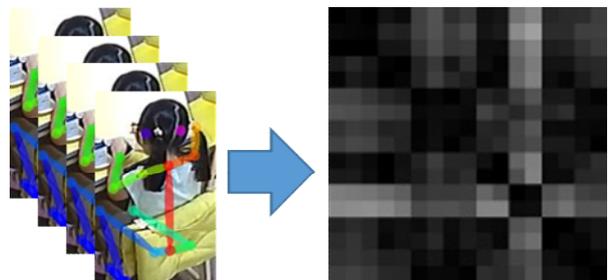


그림 6. 내부 스켈레톤 특징맵
Fig. 6. Inter-skeleton feature map

Transformer는 최근 텍스트뿐만 아니라 이미지, 동영상 등 다양한 데이터에도 적용하고 있다[18]. 기존 아동 행동 분석 모델의 경우 자세 추정 모델을 3D 차원 좌표계로 변환하여 급속탐색 랜덤트리 알고리즘 기반의 분류 기법을 사용한다.

하지만 초기 입력값인 3D 좌표가 부족할 경우 최적의 분류 결과를 얻을 수 없다. 그 결과 자세 추정과 분류 모델을 단일 스트림으로 처리하여 다중 인원 분석 시 딜레이가 생성된다. 본 논문에서 제안하는 방법은 공간적 특징과 시간적 특징을 분리하여 특징 정보를 학습 할 수 있다. Transformer Encoder 구조 적용을 통한 시공간 특징 동시 분석을 진행하였다.

우선 두 개의 서로 다른 특징을 학습하는 모델의 결합된 구조인 SlowFast 네트워크 모델을 활용한 딥러닝 학습 병렬화를 적용하였다. 공간적 영역에서 객체가 무엇인지 인지하는 Slow-path와 시간적 영역에서 객체의 움직임을 인지하는 Fast-path로 구성된다. 두 path 입력되는 Frame의 시간 간격이 다르다는 차이점이 있다.

Slow-path의 경우 공간적 정보인 색상, 모양 및 크기와 같은 세부 정보를 학습하기 위해 16개 프레임 중 하나를 선택하고 Fast-path의 경우 모션 정보만 얻기 위해 높은 프레임 속도로 작동되어 단 2개의 하나를 선택한다. 사용되는 Layer는 모두 3D Resnet이며, Fast-path에서 모션 정보만 보기 때문에 채널 수는 Slow-path보다 8배 작게 들어간다. 채널 수가 다르므로 전체 계산의 20%는 Fast-path에 할당되고 나머지 80%는 Slow-path에 할당된다. 공간적 시간적 특성을 모두 적용하기 위해 같은 입력에 대해 서로 다른 Patch의 정보를 합쳐서 학습한다.

그림 7처럼 아동행동분석 SW를 실행하여 아동 행동 패턴을 확인하였다.

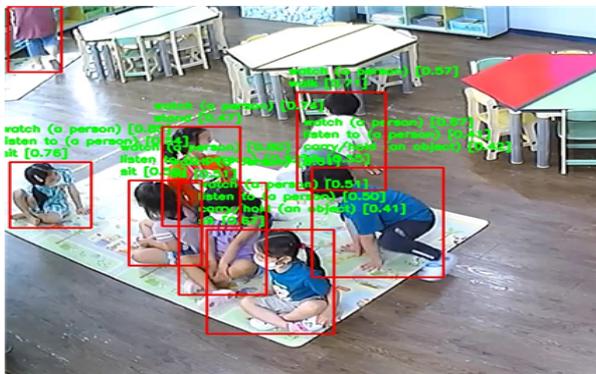


그림 7. 행동분석 결과
Fig. 7. Result of a image analysis

Mean Average Precision 계산을 사용하여 성공률을 계산 (모든 영상에서 Precision 값과 Recall 값을 구하여 Precision-Recall 그래프의 면적을 계산 후 백분율로 변환)하였다. 표 4는 본 과제에서 제안한 모델과 Tsubasa 모델의 f1-Score를 비교한 것이다. 제안한 모델은 83% 이상이고 2차원의 카-포인트를 3차원으로 사영하여 급속탐색 랜덤트리 알고리즘 기반으로 행동 예측하는 Tsubasa 모델은 74%로 F1-Score[19]는 9% 차이를 보인다.

표 4. 테스트 결과
Table 4. Result of test

	Test 1		Test 1	
	Tsubasa scheme	Proposed scheme	Tsubasa scheme	Proposed scheme
TP	19,846	21,927	15,546	20,338
FP	4,412	2,923	6,941	2,754
FN	5,844	5,847	7,150	5,142
Precision	0.818	0.882	0.691	0.880
Recall	0.772	0.789	0.684	0.798
F1-Score	0.794	0.833	0.688	0.837

III. 통합서비스 플랫폼

그림 8처럼 이상행동 감지 및 집중도 분석 기술을 통한 아동 외현화 장애 표출을 위한 통합서비스 플랫폼을 개발하였다. 공통 언어인 HTML/CSS 웹접근성 및 W3C 웹 표준 기반 구현하였고 웹호환성을 위한 서브메뉴 네비, 아웃라인 알고리즘 활용 및 웹폼/CSS 적용하였다. 현재 활성화된 기기에 한하여 실시간 이벤트 감지 기능을 제공, 집중 관찰이 요구되는 학습에 대한 관리자 모니터링 용이하도록 설계하였다.

스크롤 방식을 활용한 인터페이스 디자인으로 시스템에 대한 별도의 교육 없이 직관적으로 사용 가능하도록 개발하였다. 그림 9처럼 관리자가 모바일 환경에서도 모니터링이 가능하도록 웹 기반 시스템을 활용하여 업로드한 영상을 모바일에 최적화하여 기존 가로 레이아웃에서 세로 레이아웃으로 시안성을 개선하였다.



그림 8. 아동 외현화 장애 서비스 플랫폼
Fig. 8. Service platform for child with externalization disorder

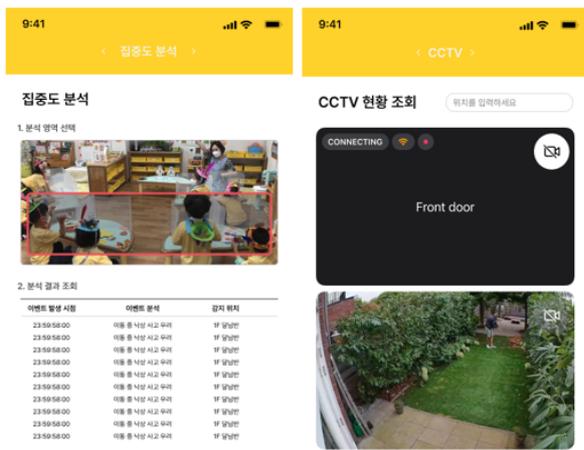


그림 9. 아동 외현화 장애 서비스 모바일플랫폼
Fig. 9. Service platform(mobile) for child with externalization disorder

IV. 결론 및 향후 과제

첫째 실시간 다중 센서 데이터 동시 입력 및 분석 가능 채널 수는 기존 연구 결과인 20개 채널에서 30개로 확장되었다. 그 결과 초등학교와 같은 대규모 CCTV가 설치된 장소에서 한 개의 통합시스템으로 운영할 수 있게 되었다. 둘째 아동 외현화 장애 증상 중 하나인 주의력 장애를 조기발견할 수 있는 시스템을 추가 개발하였다. 아동 집중도 분석을 위해 500건의 데이터셋을 구축하였고, 영상 흔들림 보정 모델의 구조를 적용한 시선 추정 모델인 VSL2G-Net을 적용하여 평균 0.968 Accuracy를 확보하였다. 셋째 아동 이상행동 분석을 위해 DSM-5의 진단기준과 척도 질문지를 사용해 표적행동목록을 크게 4가지 Class로 정의하였다. 3,000개 이상의 학습 데이터를 구축하였

고 아동 특징 맵 추출과 시공간 특징 동시 분석을 통해 기존 연구 대비 f1-Score가 80%에서 83% 이상의 정확도 성능개선 결과를 도출하였다.

마지막으로 이상행동 감지 및 집중도 분석 기술을 통한 아동 외현화 장애의 보호자 모니터링 및 알람을 위한 통합서비스 플랫폼을 개발하였다. 향후 아동 외현화 장애 맞춤형 교육 및 조기치료를 위해 시선 추정 오차 개선을 통한 집중력 분석 정확도 향상 및 학습데이터 셋을 10,000건 이상 추가확보, 알고리즘 추가 개발을 통해 이상행동 분석 정확도 90%이상 달성할 계획이다.

References

- [1] Y. K. Cho and E. H. Ha, "Effects of ADHD symptoms and SCT in children on internalizing and externalizing problems and academic and social functions", The Korean Society of Emotional and Behavioral Disorders, Vol. 35, No. 2, pp. 311-335, Jun. 2019. <https://doi.org/10.33770/JEBD.35.2.15>.
- [2] D. Ni, S. Wang, and G.g Liu, "The EEG-Based Attention Analysis in Multimedia m-Learning", Computational and Mathematical Methods in Medicine, Vol. 2020, pp. 10, Jun. 2020. <https://doi.org/10.1155/2020/4837291>
- [3] X. Xu, X. Nie, J. Zhang, and T. Xu, "Multi-Level Attention Recognition of EEG Based on Feature Selection", Int J Environ Res Public Health, Vol. 20, No. 4, Feb. 2023. <https://doi.org/10.3390/ijerph20043487>.
- [4] S. Choo, et al., "Effectiveness of multi-task deep learning framework for EEG-based emotion and context recognition", Expert Systems with Applications, Vol. 227, Oct. 2023. <https://doi.org/10.1016/j.eswa.2023.120348>.
- [5] B. Meriem, et al., "Determine the Level of Concentration of Students in Real Time from their Facial Expressions", International Journal of Advanced Computer Science and Applications,

- Vol. 13, No. 1, 2022. <https://doi.org/10.14569/IJACSA.2022.0130119>.
- [6] X. Miao, Z. Yu, and M. Liu, "Using Partial Differential Equation Face Recognition Model to Evaluate Students' Attention in a College Chinese Classroom", *Journal of Institute of Control Advances in Mathematical Physics*, Vol. 2021, pp. 10, Oct. 2021. <https://doi.org/10.1155/2021/3950445>.
- [7] T. Nose, K. Kitamura, M. Oono, Y. Nishida, and M. Ohkura, "Data-driven child behavior prediction system based on posture database for fall accident prevention in a daily living space", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11, No. 12, pp. 5845-5855, May 2020. <https://doi.org/10.1007/s12652-020-02097-6>.
- [8] H. Duan, Y. Zhao, K. Chen, D. Lin, and B. Dai, "Revisiting Skeleton-based Action Recognition", *arXiv:2104.13586*, pp. 2969-2978, Apr. 2022. <https://doi.org/10.48550/arXiv.2104.13586>.
- [9] T. Ahmad, L. Jin, X. Zhang, S. Lai, G. Tang, and L. Lin, "Graph Convolutional Neural Network for Human Action Recognition: A Comprehensive Survey", *IEEE Transactions on Artificial Intelligence*, Vol. 2, No. 2, pp. 128-145, Apr. 2021. <https://doi.org/10.1109/TAI.2021.3076974>.
- [10] Y. Zhang, Y. Tian, P. Wu, and D. Chen, "Application of Skeleton Data and Long Short-Term Memory in Action Recognition of Children with Autism Spectrum Disorder", *Sensors*, Vol. 21, No. 2, pp. 411-427, Jan. 2021. <https://doi.org/10.3390/s21020411>.
- [11] K. Hwang, J.-M. Lee, and I. Jung, "Remote Control System using Face and Gesture Recognition based on Deep Learning", *The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)*, Vol. 20, No. 6, pp. 115-121, Dec. 2020. <https://doi.org/10.7236/IIBC.2020.20.6.115>.
- [12] B. Hwang, J. Kim, Y.-R. Lee, C. Kyeong, J. Seon, Y.-G. Sun, and J.-Y. Kim, "Performance of Exercise Posture Correction System Based on Deep Learning", *The Journal of the Institute of Internet, Broadcasting and Communication*, Vol. 22, No. 5, pp. 177-183, Oct. 2022. <https://doi.org/10.7236/IIBC.2022.22.5.177>.
- [13] S.-H. Ahn and B. Moon, "A Development of Deep Learning-based Analysis of Behavior for Child Safety Indoors", *The Journal of Korean Institute of Information Technology*, Vol. 21, No. 1, pp. 91-100, Jan. 2023. <https://doi.org/10.14801/jkiit.2023.21.1.91>
- [14] J. Oh, Y. Lee, J. Yoo, and S. Kwon, "Improved Feature-Based Gaze Estimation Using Self-Attention Module and Synthetic Eye Images", *Sensors*, Vol. 22, No. 11, pp. 4026, May 2022. <https://doi.org/10.3390/s22114026>.
- [15] X. Zhou, H. Cai, Z. Shao, H. Yu, and H. Liu, "3D eye model-based gaze estimation from a depth sensor", 2016 IEEE international conference on robotics and biomimetics, Qingdao, China, pp. 369-374, Mar. 2017. <https://doi.org/10.1109/ROBIO.2016.7866350>.
- [16] A. A. Abdelrahman, T. Hempel, A. Khalifa, and A. Al-Hamadi, "L2CS-Net: Fine-Grained Gaze Estimation in Unconstrained Environments", *arXiv:2203.03339*, Mar. 2022. <https://doi.org/10.48550/arXiv.2203.03339>.
- [17] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement.", *arXiv:1804.02767*, Apr. 2018. <https://doi.org/10.48550/arXiv.1804.02767>.
- [18] A. Dosovitskiy, et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", *arXiv:2010.11929*, Jun. 2021. <https://doi.org/10.48550/arXiv.2010.11929>.
- [19] Z. C. Lipton, C. Elkan, and B. Naryanaswamy, "Optimal Thresholding of Classifiers to Maximize f1 Measure", *Machine Learning and Knowledge Discovery in Databases*, Vol. 8725, pp. 225-239, Sep. 2014. https://doi.org/10.1007/978-3-662-44851-9_15.

저자소개

안 시 현 (Shi-Hyun Ahn)



2016년 2월 : 경북대학교:
전자공학부(박사수료)
2022년 5월 ~ 현재 :
(주)우경정보기술 이사
관심분야 : 인공지능, 컴퓨터비전,
임베디드

박 윤 하 (Yun Ha Park)



1997년 2월 : 영남대학교:
전자공학과(공학사)
1997년 ~ 2008년 :
(주)엔트릭커 본부장
2008년 ~ 현재 :
(주)우경정보기술 대표이사
관심분야 : BigData, 스마트시티,
AI컴퓨터 비전, 딥러닝

문 병 인 (Byungin Moon)



1995년 : 연세대학교:
전자공학과(학사)
1997년 : 연세대학교:
전자공학과(석사)
2002년 : 연세대학교:
전기전자공학과(박사)

2002년 ~ 2004년 : 하이닉스반도체, 선임연구원

2004년 ~ 2005년 : 연세대학교, 연구교수

2005년 ~ 현재 : 경북대학교, 전자공학부 교수

관심분야 : SOC, 컴퓨터 구조, 비전 프로세서