

# A Prediction Model for Detecting Developmental Disabilities in Preschool-Age Children Through Digital Biomarker-Driven Deep Learning in Serious Games: Development Study

---

Ho Heon Kim\* , RN, MS; Jae Il An\* , MS; Yu Rang Park, PhD

# 연구 개요

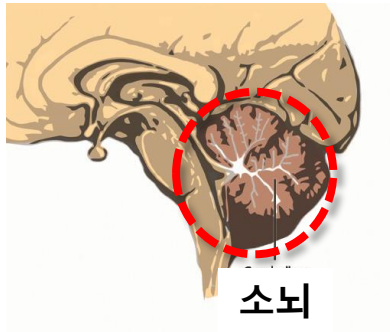
연구대상: 미취학 아동



조기 발견을 통한 조기 개입이 매우 중요  
**But**, 아동에게 검사의 진행이 어려움,  
부정확한 진단을 내릴 가능성이 높음

- 미취학 아동의 10~15%는 발달장애를 가지고 있음
- 뇌의 가소성이 뛰어나며, 두뇌가 빠르게 성장하는 시기

제안 근거: 운동 제어 능력과 지능의 연관성



- 운동제어 능력 관련된 소뇌는 높은 지능과 상관성이 존재
- 소뇌와 연관된 운동 제어 능력은 발달장애 진단에 중요한 요소



But, 현재 검사를 위해선 비싼  
기구가 필요로 제약이 많음



장소와 시간을 구애받지 않은  
새로운 측정 방법의 필요성

→ 게임 실행 데이터 바탕 발달장애 예측 딥러닝 모델

# 연구 방법

게임: DoBrain



- 지능개발을 위한 프로그램을 제공하는 모바일 기반 게임

타겟능력

공간  
인식

수리

창의성

추론

구성

기억

시각적  
식별

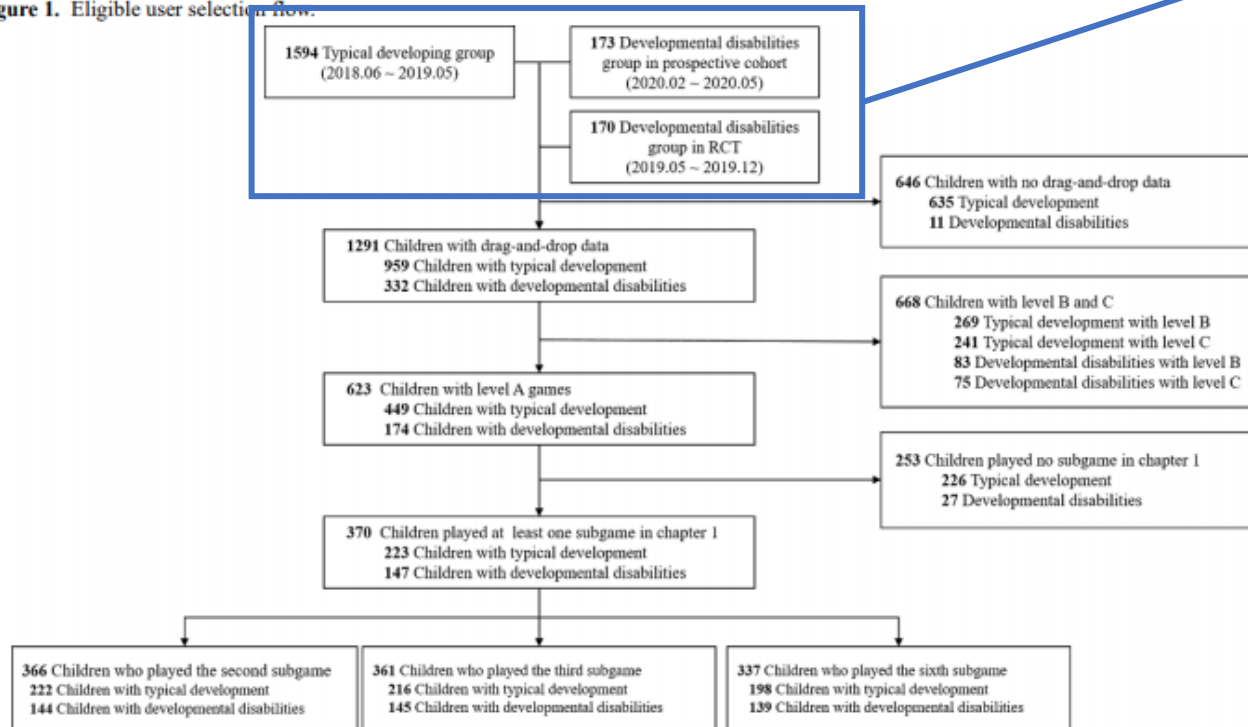
조작방식

답안 선택

드래그앤드롭

## 방법 | Methods – 연구설계 | Study Design

Figure 1. Eligible user selection flow.

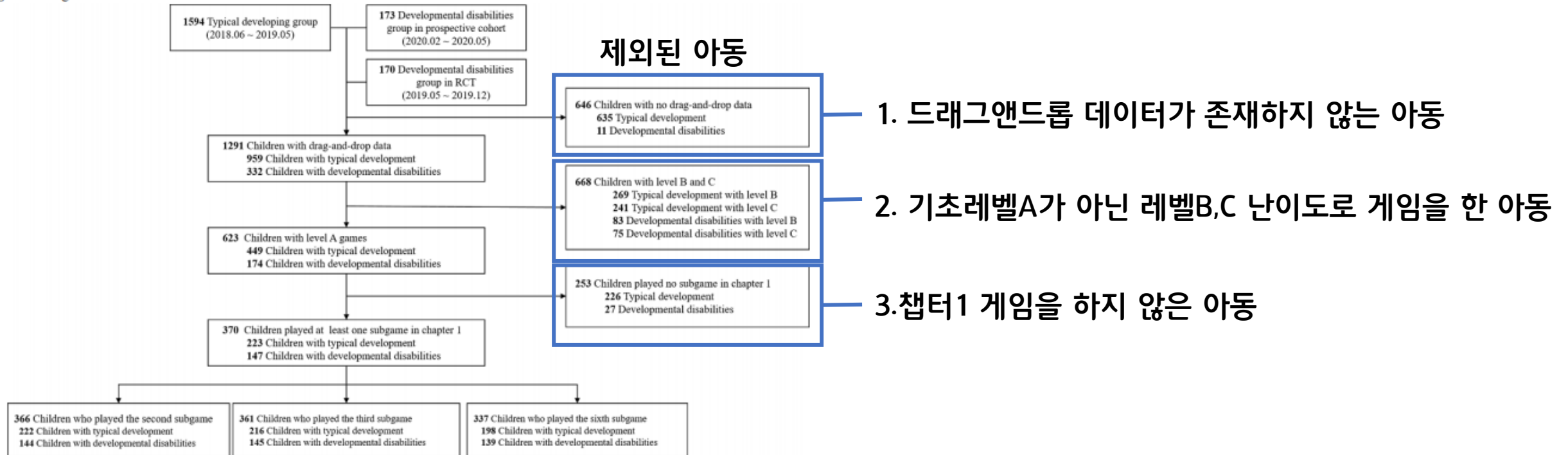


총 3개의 연구에서의 참가자들을 비식별화

- 2018.06 ~ 2020.06 retrospective study
- 2019.03 ~ 2019.12 randomized clinical trials
- 2020.02 prospective study

## 방법 | Methods – 연구설계 | Study Design

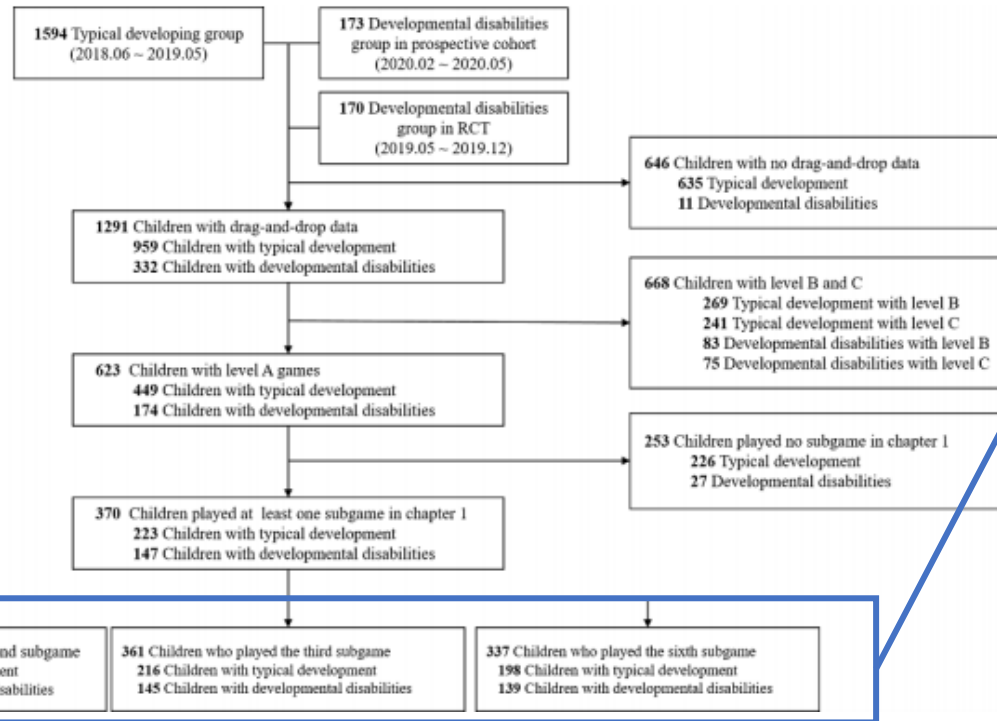
Figure 1. Eligible user selection flow.





## 방법 | Methods – 연구설계 | Study Design

Figure 1. Eligible user selection flow.



### 최종 활용한 데이터

1. 두번째 서브게임을 즐긴 366명
2. 세번째 서브게임을 즐긴 361명
3. 여섯번째 서브게임을 즐긴 337명

### 사용 변수

연속형 변수 - 나이, 기기 사이즈 등  
카테고리 변수 - 성별 등



손가락 움직임(스트로크) 데이터

### 분석 절차

정규성 검정

(시각적 분포 확인과 콜로모고로브-스미모브 검정)



이분산성 검정

(F검정)



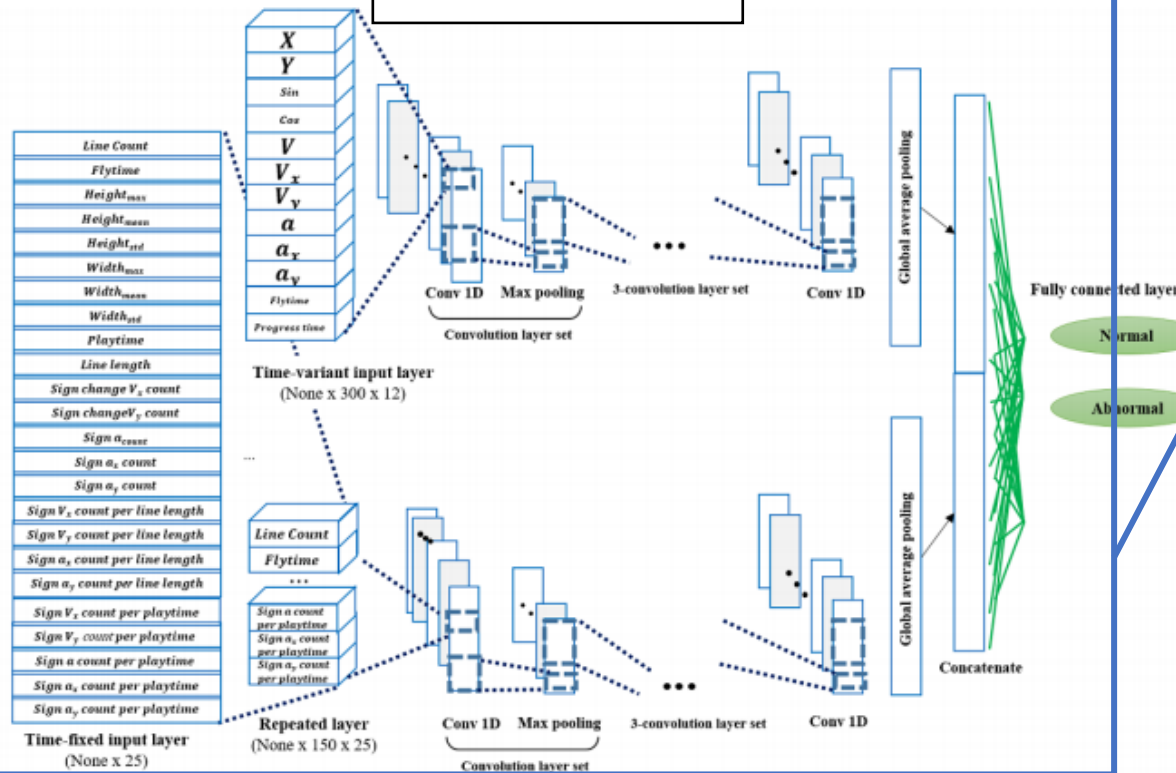
두 집단의 차이 유효성 검정

(정규형 - T검정, 비정규형 - 맨-휘트니검정)

(카테고리 변수 - 카이스퀘어 검정)

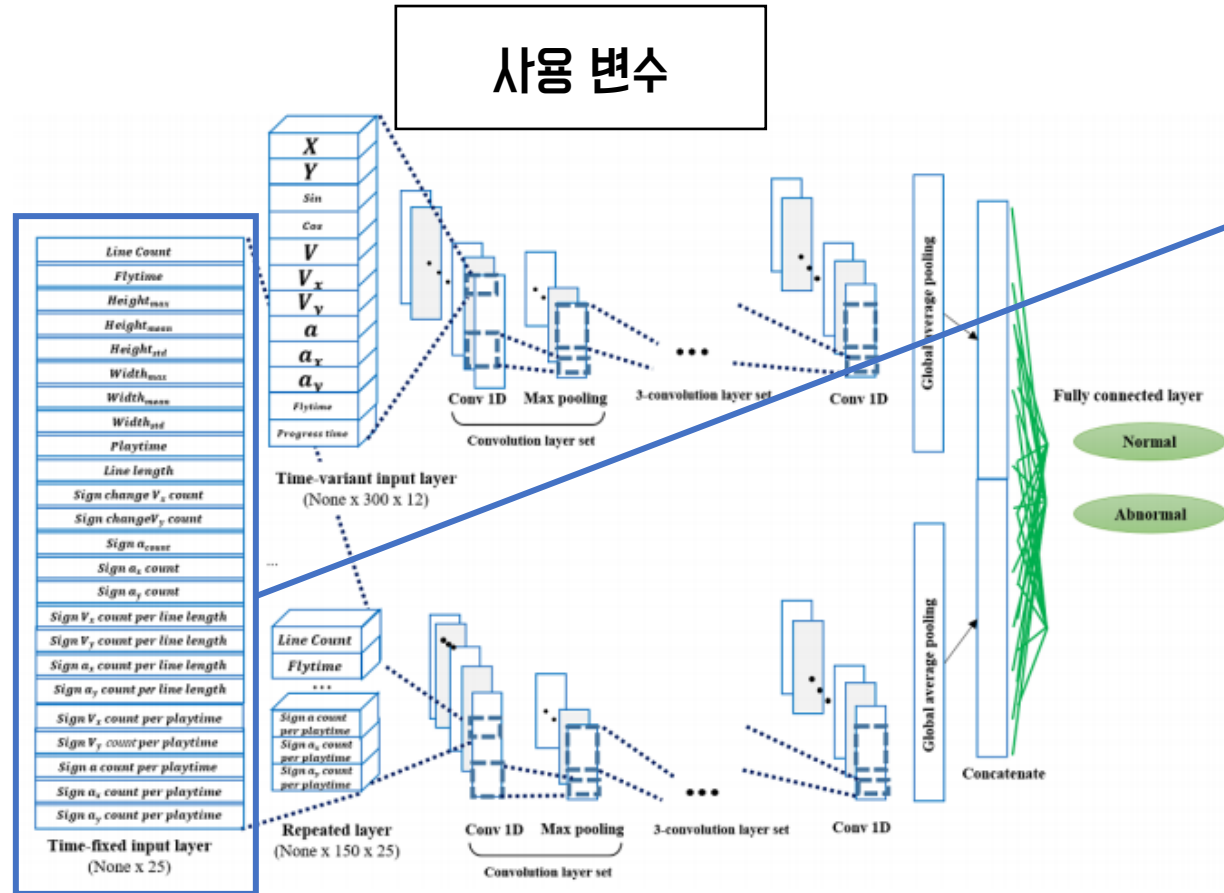
기타: effect size(CLES) 측정

## 사용 모델



사용한 딥러닝 모델:  
1D Convolution Neural Netowork(이하 'CNN')

## 방법 | Methods – 분석절차 | Analytical Procedure

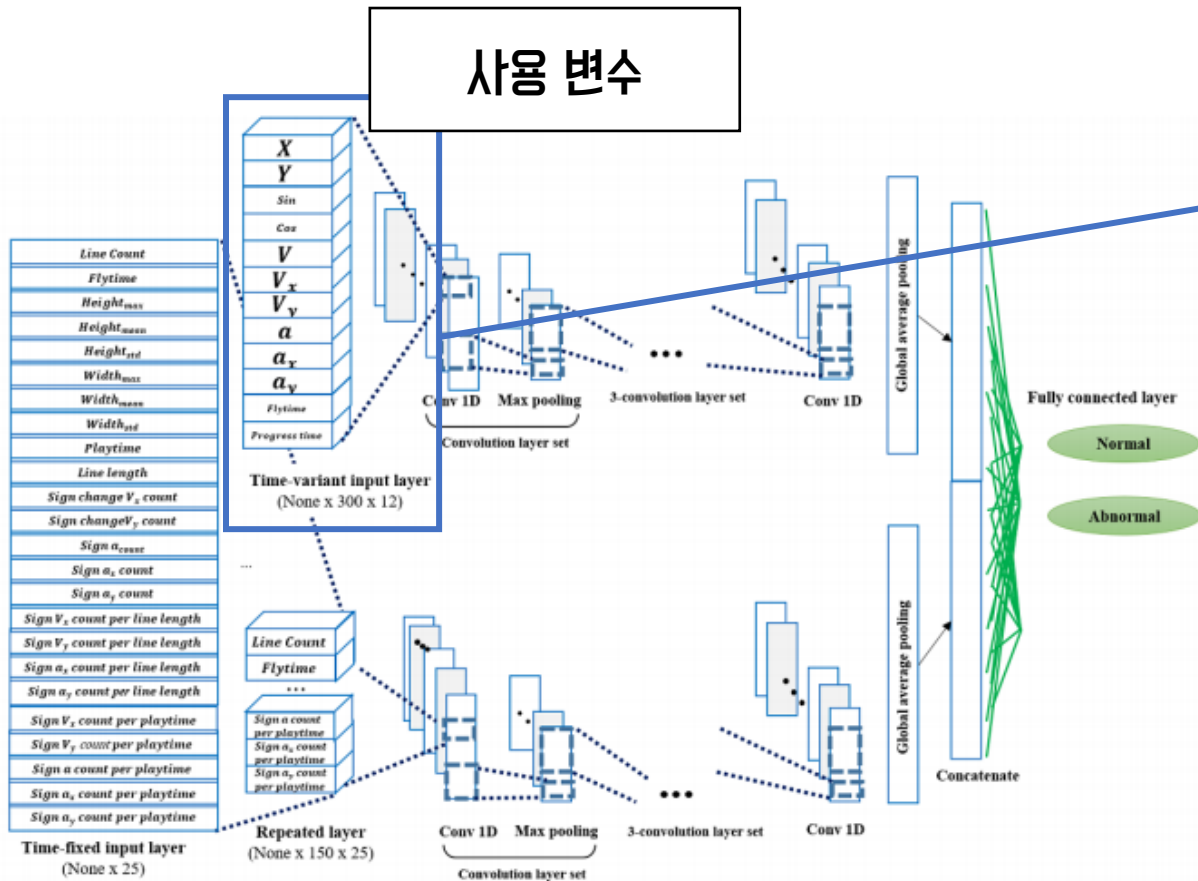


시간 고정 인풋 레이어

: 게임점수, 총 터치 면적, 인구통계학적 변수

## 방법 | Methods – 분석절차 | Analytical Procedure

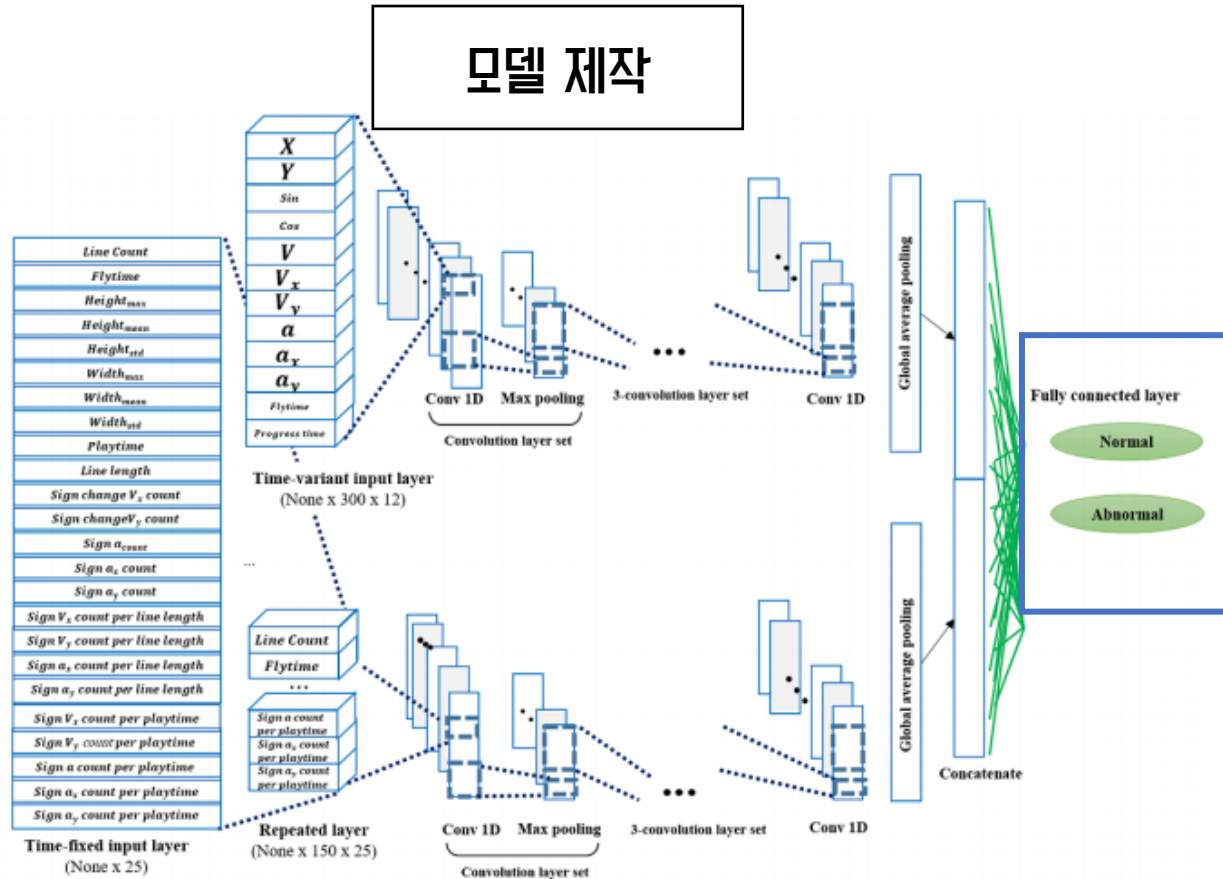
### 사용 변수



시간 변동 인풋 레이어: 터치 좌표 변수 및 **파생 변수**  
속도, 가속도

## 방법 | Methods – 분석절차 | Analytical Procedure

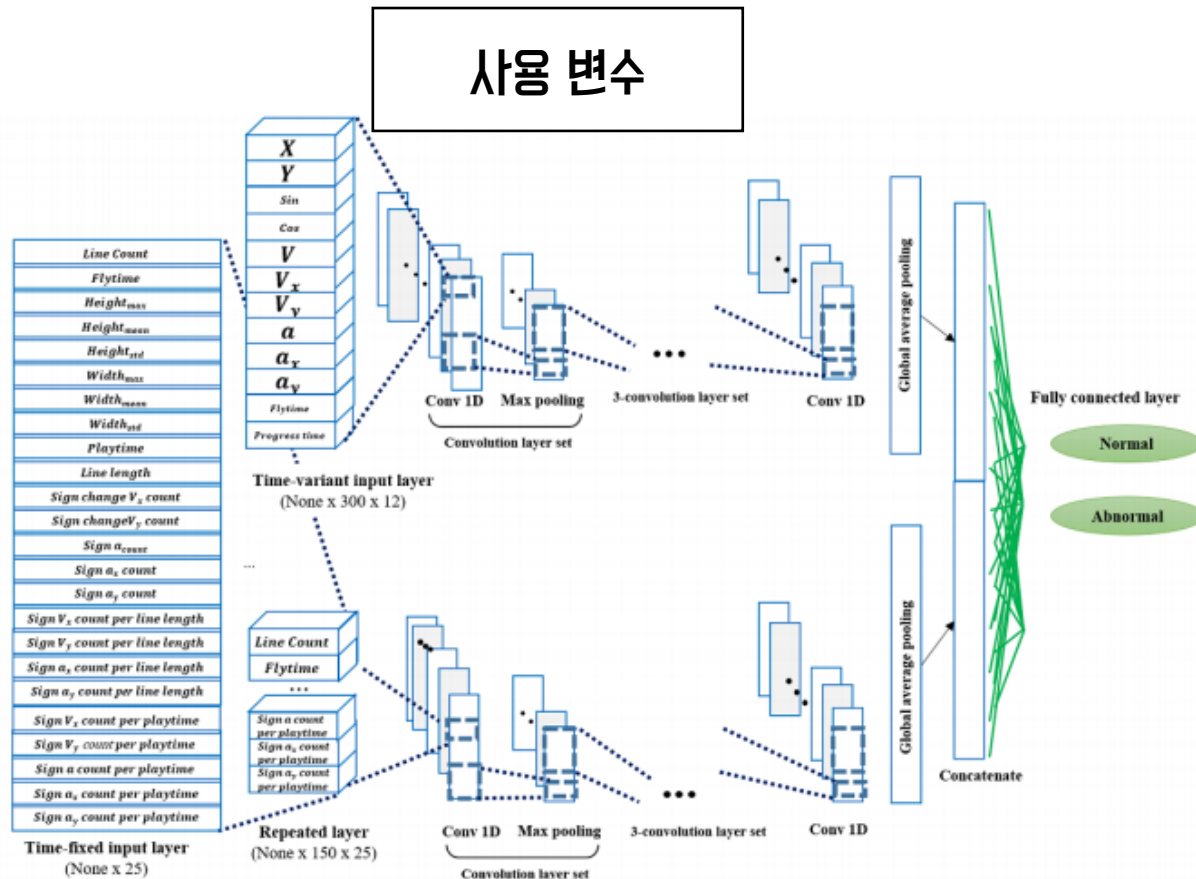
### 모델 제작



최종적으로 모든 인풋이 각 유저ID 기준으로 결합  
Binary Prediction Model 제작

## 방법 | Methods – 분석절차 | Analytical Procedure

### 사용 변수



### 모델 하이퍼파라미터 최적화

CV=10 기준 그리드 서치 활용

(AUROC, Area under the precision-recall curve, F score, Precision, Recall, Specificity 점수 지표 활용)

### 모델 개발 외 추가적인 연구

Grad-CAM 방법 활용하여 예측에 효과적이었던 미세 운동조절 움직임을 확인함

# 연구 결과



결과 | Results – 기저 특성 | Baseline Characteristics

인구통계학적 변수

Table 1. Demographic characteristics in eligible users.

Variables	Children with typical development (n=223)	Children with developmental disabilities (n=147)	Total (n=370)	P value	Effect size
Age (months), median (IQR)	40.0 (12.0)	72.0 (32.5)	45.0 (26.5)	<.001	0.839 <sup>a</sup>
<b>Diagnosis, n (%)</b>					
Intellectual disability	0 (0.0)	44 (0.33)	44 (0.11)	N/A <sup>b</sup>	N/A
Autism spectrum disorder	0 (0.0)	41 (0.27)	41 (0.11)	N/A	N/A
Developmental disorder	0 (0.0)	33 (0.22)	33 (0.09)	N/A	N/A
Brain lesions	0 (0.0)	25 (0.17)	25 (0.09)	N/A	N/A
Monogenic disorder	0 (0.0)	4 (0.02)	4 (0.01)	N/A	N/A
<b>Children playing subgame, n (%)</b>					
Second subgame	222 (99.55)	144 (97.96)	366 (98.92)	.03	<0.001
Third subgame	216 (97.31)	145 (98.64)	361 (97.84)	.24	<0.001
Sixth subgame	198 (88.79)	139 (95.56)	337 (91.98)	.04	0.002
Device size (inches), median (IQR)	6.1 (4.2)	6.1 (4.2)	6.1 (4.2)	<.001	0.469 <sup>a</sup>
Game playtime, (s/per game), median (IQR)	11.8 (8.5)	8.1 (6.5)	10.3 (8.2)	<.001	0.687 <sup>a</sup>
Games played, median (IQR)	9.0 (9.0)	1.0 (14.2)	7.0 (12.5)	<.001	0.682 <sup>a</sup>

두 집단간 연령 차이 존재

발달장애 집단에만 지적장애, 자폐증, 발달장애, 뇌병변, 단일 유전자 인자 장애 등을 진단받은 경험이 존재

평균 플레이시간 차이 존재

<sup>a</sup>Common language effect size of continuous variables;  $\eta^2$  for effect size of categorical variables.

<sup>b</sup>N/A: not applicable.

결과 | Results – 드래그앤드롭게임 특성 | Characteristics of Drag-and-Drop Game Play

Table 2. Comparison of movement features between children with typical development and children with developmental disabilities in each game.

Characteristic	Children with typical development	Children with developmental disabilities	P value	CLES <sup>a</sup>
Second game				
Distribution, n	222	144	N/A <sup>b</sup>	N/A
Play information, median (IQR)				
Playtime (second/game)	2.03 (6.39)	1.98 (5.11)	.33	0.513
Line number (n/game)	1.0 (2.0)	2.0 (2.0)	.24	0.376
Line length (n/game)	3.04 (3.17)	3.1 (4.24)	.28	0.518
Release to touch time (sec)	0.0 (3.31)	0.52 (4.46)	.15	0.409
Touch region, median (IQR)				
Height mean, ratio (%)	0.38 (0.1)	0.38 (0.08)	.26	0.540
Height max, ratio (%)	0.58 (0.08)	0.57 (0.06)	.10	0.520
Width max, ratio (%)	0.74 (0.08)	0.72 (0.08)	.13	0.534
Height mean, ratio (%)	0.38 (0.1)	0.38 (0.08)	.006	0.540
Third game				
Distribution, n	216	145	N/A	N/A
Play information, median (IQR)				
Playtime (second/game)	2.75 (6.63)	2.44 (5.99)	.22	0.524
Line number (n/game)	1.0 (2.0)	1.0 (2.0)	.46	0.344
Line length (n/game)	3.12 (4.23)	3.62 (6.28)	.04	0.553
Release to touch time (sec)	0.0 (2.84)	0.0 (3.11)	.45	0.372
Touch region, median (IQR)				
Height mean, ratio (%)	0.29 (0.09)	0.27 (0.06)	.02	0.520
Height max, ratio (%)	0.35 (0.23)	0.35 (0.17)	.26	0.565
Width max, ratio (%)	0.72 (0.1)	0.71 (0.08)	.18	0.528
Height mean, ratio (%)	0.29 (0.09)	0.27 (0.06)	.04	0.520

두번째, 세번째 게임 데이터 결과  
플레이시간 통계적 유의미한 차이 없음  
터치 영역과 관련된 변수 유의미한 차이 없음

# 결과 | Results – 드래그앤드롭게임 특성 | Characteristics of Drag-and-Drop Game Play

## Sixth game

Distribution, n	198	139	N/A	N/A
Play information, median (IQR)				
Playtime (second/game)	7.85 (15.25)	4.33 (10.03)	<.001	0.616
Line number (n/game)	3.0 (3.0)	2.0 (3.0)	.001	0.514
Line length (n/game)	4.64 (8.5)	4.5 (7.66)	.41	0.507
Release to touch time (sec)	2.9 (8.33)	1.4 (6.27)	.001	0.549
Touch region, median (IQR)				
Height mean, ratio (%)	0.43 (0.06)	0.34 (0.16)	<.001	0.695
Height max, ratio (%)	0.53 (0.09)	0.49 (0.2)	<.001	0.766
Width max, ratio (%)	0.78 (0.1)	0.73 (0.09)	<.001	0.658
Height mean, ratio (%)	0.43 (0.06)	0.34 (0.16)	<.001	0.695

여섯번째 게임 데이터 결과  
플레이시간 통계적 유의미한 차이 있음  
터치 영역과 관련된 변수 유의미한 차이 있음

기타  
X축 기준의 터치 속도 IQR(사분위수 범위) 차이 있음  
4개의 가속도 변수에서 정상 범주 아동들에 비해 수치가 큼

## 결과 | Results – 모델 성능 | Model Performance

**Table 3.** Average model performance result by 10-fold cross-validation in each drag-and-drop subgame.

Deep learning model for individual subgame	Performance, mean (SD)					
	AUROC <sup>a</sup>	Accuracy	<i>F</i> score	Precision	Recall	Specificity
Second subgame	0.746 (0.116)	0.719 (0.094)	0.627 (0.165)	0.616 (0.136)	0.683 (0.151)	0.755 (0.089)
Third subgame	0.793 (0.117)	0.747 (0.082)	0.675 (0.126)	0.686 (0.142)	0.688 (0.167)	0.783 (0.069)
Sixth subgame	0.817 (0.070)	0.769 (0.078)	0.708 (0.153)	0.675 (0.183)	0.757 (0.123)	0.740 (0.099)

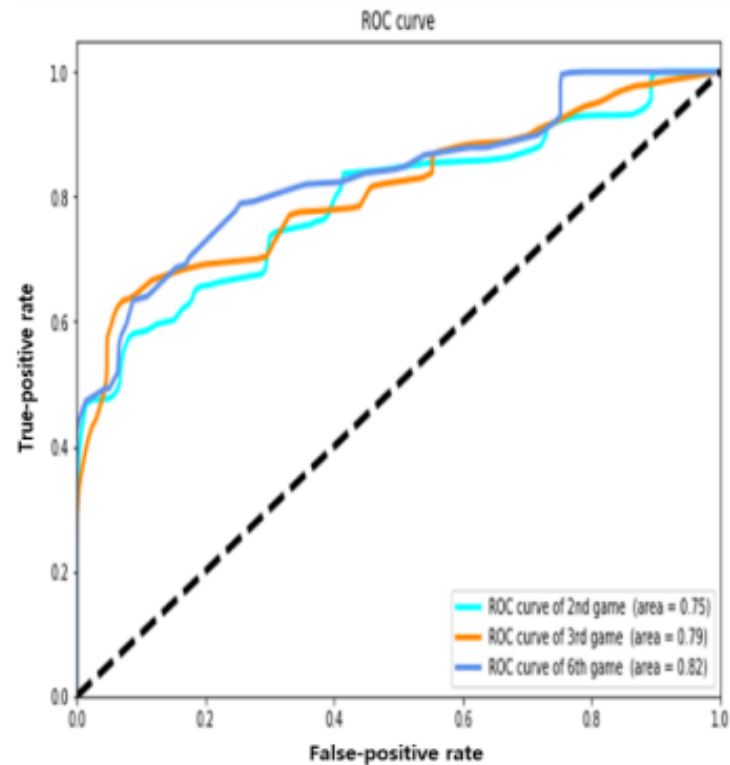
<sup>a</sup>AUROC: area under the receiver operating characteristics curve.

### 모델 성능 비교 결과

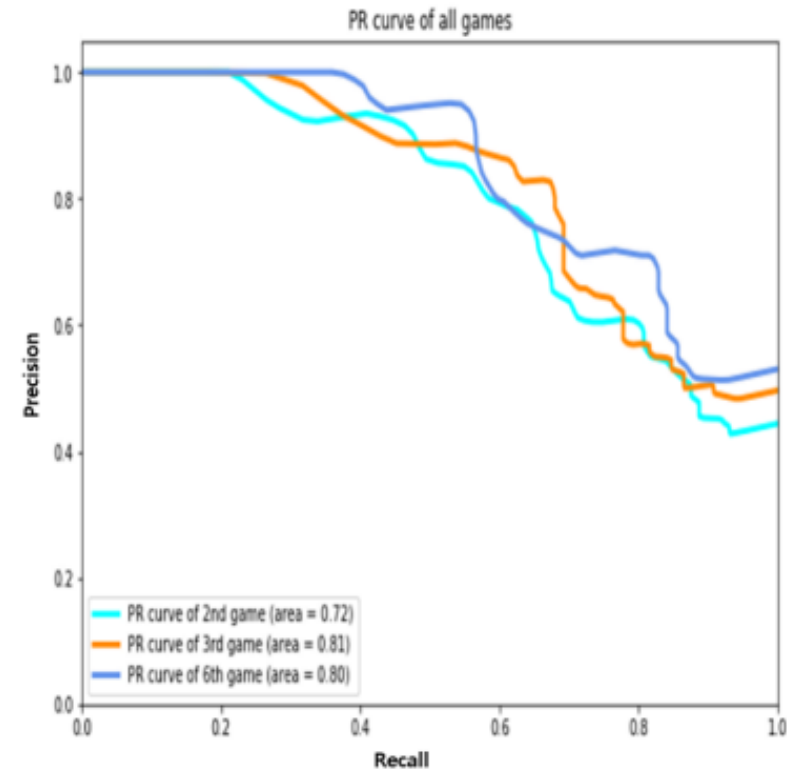
- 여섯번째 게임이 발달장애 예측 모델에 있어 가장 좋은 성능을 보임
- 정상을 찾아내는 지표 특이도의 경우, 세번째 모델에서 비교적 높은 성능을 보였음

## 결과 | Results – 모델 성능 | Model Performance

**Figure 3.** ROC curves and PR curves of the deep learning model. PR: precision-recall. ROC: receiver operating characteristic.



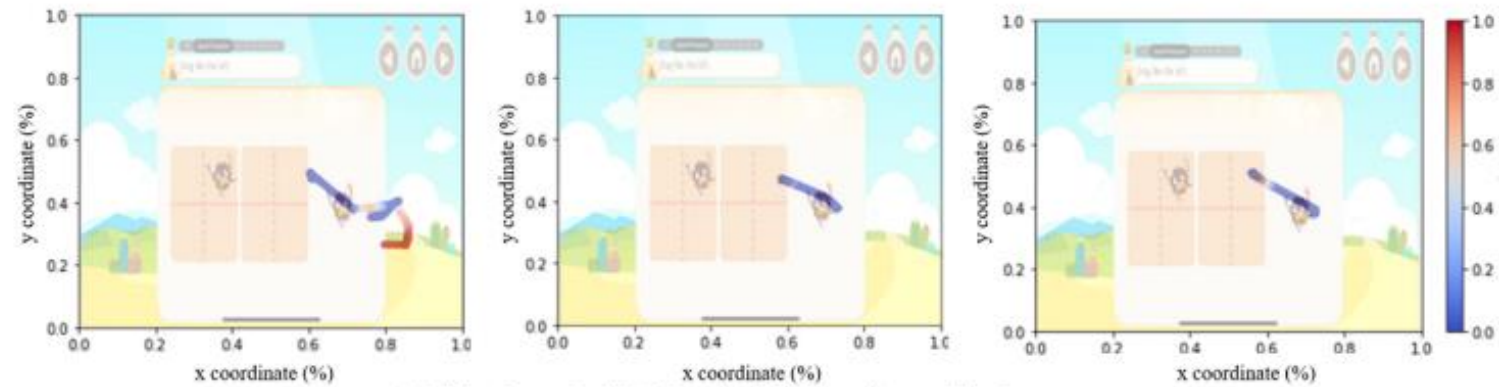
(A) ROC curve of model based on each subgame



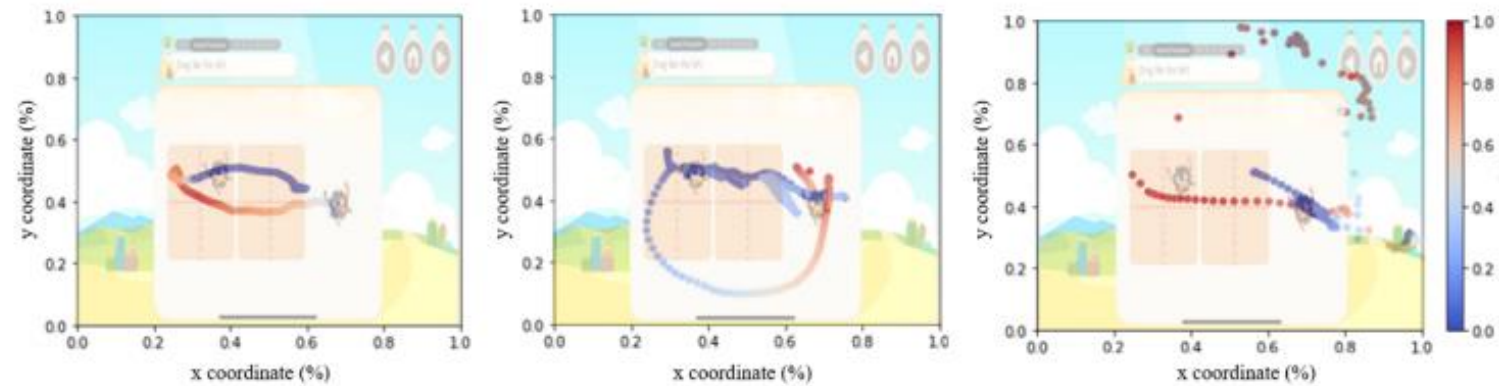
(B) PR curve of model based on each subgame

## 결과 | Results – 손가락 꺾적 예측 결과 | Prediction of Developmental Disabilities From Finger Strokes

**Figure 4.** Sixth subgame with Grad-CAM results of children with normal and abnormal development disability. Grad-CAM: gradient-weighted class activation mapping.

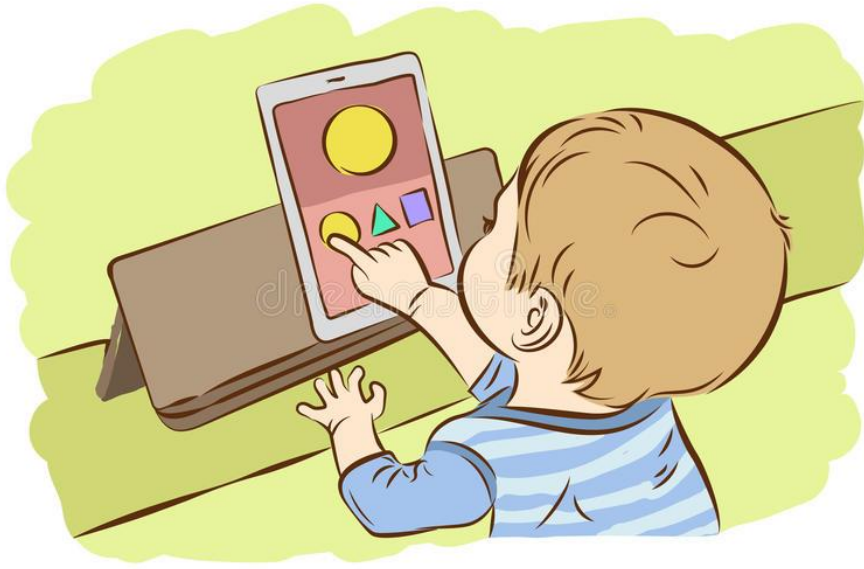


(A) Sixth subgame Grad-CAM results of children with normal development



(B) Sixth subgame Grad-CAM results of children with developmental disability

# Discussion | 고찰



### 1. 새로운 디지털 생물학적 지표

모바일 터치 기기에서의 미세 운동 조절 움직임

순간적인 드래그앤드롭 데이터를 미세 운동 조절 데이터로 활용 가능

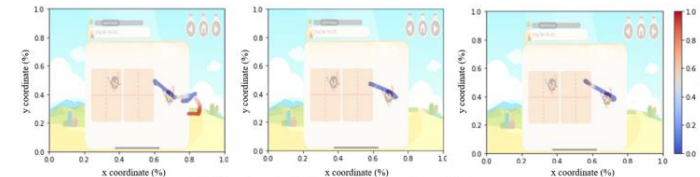
### 2. 지능개발 게임의 활용성

발달장애를 탐지하는 진단보조 툴 또는 검사 툴로서 활용 가능

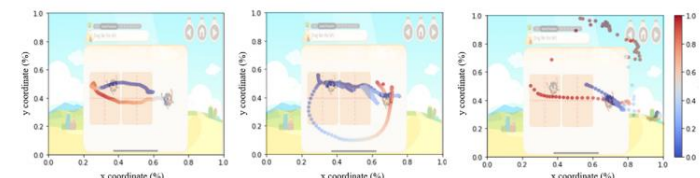
### 3. 의료진에 대한 적응성

분류에 각 좌표들이 얼마나 진단(분류)에 영향을 주었는지 시각적으로 직접 확인할 수 있음

Figure 4. Sixth subgame with Grad-CAM results of children with normal and abnormal development disability. Grad-CAM: gradient-weighted class activation mapping.



(A) Sixth subgame Grad-CAM results of children with normal development



(B) Sixth subgame Grad-CAM results of children with developmental disability



### 기존 운동능력을 측정하고 검사하는 방법

Behavior Rating Inventory of Executive Function II (BRIEF-II)

Denver Developmental Screening Tests II (DDST-II)



### 정보의 기록 출처

이러한 정보는 모두 부모나 교사로부터 점수를 직접 계산하여 반영하는 방식으로 한계점 명확

새로운 툴의 신뢰도 높음

### 기존 의사 진단과정에서의 어려움

병원에서의 대기가 필수, 아이들의 지속적인 관찰 및 검사가 어려움,  
효과적인 툴의 부재로 인한 진단결과에 대한 확신 부족



### 이번에 제안된 방법

위의 난제를 극복함과 동시에, 비용이 저렴하고 병원을 방문하지  
않고도 손쉽게 검사환경을 갖출 수 있음

새로운 툴의 저임금 국가로의  
보급 기대



### 1.의사의 진단을 통한 결과 데이터가 아님

환자 개개인의 정보를 의사의 진단이 환자 및 보호자가 작성한 리포트 데이터를 바탕으로 하기에, 데이터의 신뢰도 비교적 낮음

### 2.모바일 기기 사용불가 아동 적용 불가

모바일 기기를 사용할 수 있다는 전제하에 데이터를 얻을 수 있는 구조로 모바일 사용이 불가능한 아동의 경우 이 툴을 사용하기 어려움

# Conclusion | 결론



// 디지털 생물학적 지표로 딥러닝 기반의 검사 툴을 사용할 시  
발달 장애 아동을 진단하는데 활용 가능하다 //

감사합니다

