

아동용 스마트 카시트 개발을 위한 상지 이탈 및 표정 감지 경량화 AI 모델 활용에 관한 연구

윤성재*, 구본승*, 윤솔진*, 윤준호*, 김성관*

*공주대학교 기계자동차공학부

e-mail:ysjea011@naver.com

A Study on the Application of Lightweight AI Models for Upper-Limb Ejection and Facial Expression Detection in Smart Child Car Seat Development

SeongJae Yoon*, BonSeung Ku*, Solgeon Yoon*, Junho Yun*, SungGaun Kim**

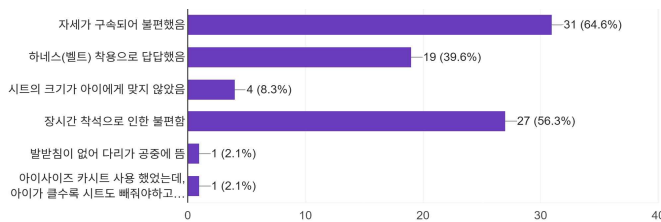
*Dept. of Mechanical & Automotive Engineering, Kongju National University

요약

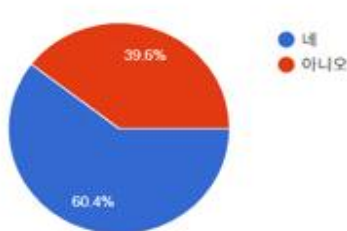
본 연구는 아동 카시트 착용 시 발생하는 하네스 이탈 및 착용 불편 문제를 해결하고자, 아동의 안전성과 착용 편의성 향상을 위한 방안으로 AI 기반 실시간 감지 모델을 개발하였다. 표정 감지 모델은 YuNet⁽²⁾ 얼굴 검출기와 MobileFaceNet⁽³⁾ 분류기를 결합하여 경량화 구조로 설계하였으며, 영유아 표정 데이터셋 Fine-tuning을 통해 주요 감정 클래스에서 85% 이상의 정확도를 확보하였다. 상지 이탈 감지 모델은 YOLOv11n-pose⁽⁴⁾와 MobileNetV2 기반 분류기를 적용하여 좌우 어깨 이탈을 판별하였고, 97% 이상의 재현율(Recall)을 달성하였다. 결론적으로 두 모델은 임베디드 환경을 고려한 AI 경량화 모델을 통해 실시간 동작 성능과 높은 정확도를 동시에 확보하였음을 나타낸다.

1. 서론

만 6세 미만 아동은 자동차에 탑승할 경우, 도로교통법 제 50조에 따라 유아용 카시트를 반드시 사용하도록 의무화되어 있다.⁽¹⁾ 그러나 카시트를 착용한 상태에서도 아동이 하네스에서 팔을 의



(a)



(b)

[그림 1] Survey results from child car seat users: a) Discomfort factors in child car seat usage; b) Experience of upper-limb ejection despite proper harness use

도적으로 이탈시키거나, 보호자가 착용 불편을 이유로 장력을 완화하는 사례가 빈번하게 발생한다. 그림 1에 제시된 설문조사 결과에 따르면, 약 60%의 보호자는 주행 중 아동이 하네스에서 팔을 이탈시킨 경험이 있는 것으로 나타났다. 이러한 상황에서는 주행 중 아동의 상태를 실시간으로 확인하거나 즉각적으로 대응하는 데 구조적 한계가 존재하며, 이는 현행 카시트의 하네스가 아동의 행동 변화에 실시간으로 대응하지 못하고 정적인 고정 방식에 의존하고 있음을 보여준다.

본 연구는 이러한 문제를 해결하기 위한 일환으로, 스마트 카시트 시스템을 직접 구현하기보다는 이를 가능하게 하는 핵심 기반 기술로서 AI 기반 실시간 감지 모델의 구축에 초점을 두었다. 결론적으로, 아동의 표정을 인식하여 불편 신호를 탐지하는 경량화 표정 인식 모델과 하네스에서의 상지 이탈 여부를 판별하는 상지 이탈 감지 모델을 개발하였다. 두 모델은 실시간 추론 성능 및 임베디드 환경 적응성을 고려하여 경량화된 구조로 설계되었으며, 향후 하네스 제어 시스템과 통합되어 아동의 안전성과 착용 편의성을 동시에 향상시킬 수 있는 기반 기술로 기능할 수 있다.

2. AI 경량화 모델 학습

2.1 학습 개요

2.1.1 시스템 개요

본 연구는 유아용 카시트 환경에서 표정 인식과 상지 벨트 이탈감지를 실시간으로 수행하기 위해 합성곱 신경망(CNN, Convolutional neural network) 모델을 사용하여 Two stage 파이프라인을 설계하였다. 1단계에서 탐지 및 크롭을 통해 관심 영역을 분리하고, 2단계에서는 크롭된 관심 영역에 대해 경량 분류기를 적용하여 판정을 수행한다.

2.1.2 표정 감지

표정 감지 모듈은 얼굴 검출과 표정 분류의 두 단계로 구성하였다. 1단계에서는 경량화 얼굴 검출기인 YuNet⁽²⁾을 사용하여 아동 얼굴에 해당하는 바운딩 박스(bbox, bounding box)를 크롭하고, 2단계에서는 MobileFaceNet⁽³⁾ 기반의 분류기로 크롭한 얼굴 이미지에서 표정을 분류한다.

입력 픽셀 크기 320×320 기준 YuNet은 약 7.5만개의 파라미터와 0.15 GFLOPs를 가지며, 42만 개 파라미터와 0.19 GFLOPs 이상의 동급 경량 검출기인 SCRFD-0.5g, RetinaFace-0.25 대비 연산 부담이 작으면서 AP는 유사 수준을 유지한다. MobileFaceNet은 0.24 GFLOPs로 연산량이 작고, 눈, 코, 입 등 지역 특징을 효과적으로 추출하므로 채택하였다.

2.1.3 상지 벨트 이탈감지

상지 벨트 이탈감지 모듈은 상지 어깨부를 크롭하고, 이탈 여부를 이진 분류한다. 먼저 YOLO11n pose⁽⁴⁾로 좌, 우 어깨 위치를 추정한 뒤, 어깨 간 거리 d 에 기반하여 변 길이가 $\frac{2}{3}d$ 인 정사각형을 각 어깨 중심에 배치하여 크롭하고, 224×224 픽셀 크기로 정규화한다. 이후 MobileNetV2 기반의 좌, 우 상지



[그림 2] Structure of emotion-labeled infant dataset with sample images for expression classification

벨트 이탈감지 분류기에서 크롭된 각 상지 이미지에 대해 벨트 이탈 여부를 판단한다.

YOLO11n pose는 7.6 GFLOPs, COCO keypoint 데이터셋에 대해 mAP50 성능 0.81로 연산량 대비 준수한 성능을 보인다. MobileNetV2는 0.3 GFLOPs로 경량화 성능을 보여 채택하였다.

2.2 데이터셋 & 전처리

2.2.1 표정 인식 데이터셋

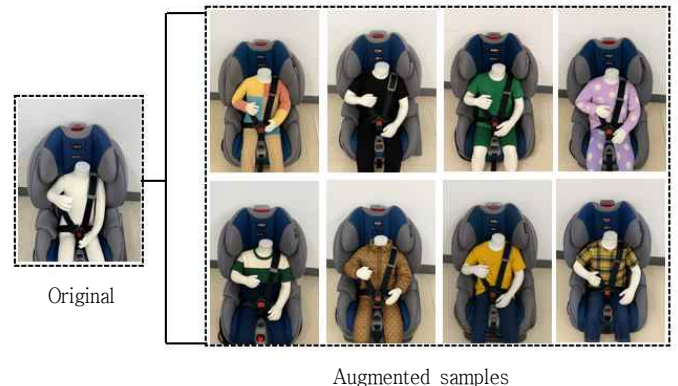
Fine-tuning에는 그림 2와 같이 오픈소스 얼굴 표정 데이터셋인 ExpW에서 유아에 해당되는 약 2,000장을 선별하여 사용하였다. 일반화 성능을 확보하기 위해 YuNet으로 얼굴 영역만 크롭한 뒤, 밝기, 대비, 회전 및 수평 반전의 증강을 5배수로 적용하여 총 10,000장 학습 데이터를 구축하였다. 데이터는 Train, Validation, Test를 각각의 클래스에서 8:1:1로 분할하였다.

2.2.2 상지 이탈감지 데이터셋

데이터는 Britax사의 Advocate ClickTight 카시트에 아동 전신 연결 마네킹을 탑승시켜 벨트가 흉부 위에 보이는 Normal 상태와 어깨 바깥으로 빠진 Fallout 상태를 촬영하여 80장의 이미지를 확보하였다.

상지 이탈감지 모델의 입력 데이터인 어깨부 크롭 이미지가 더미의 색상인 흰색에만 과적합되는 문제를 방지하기 위해, Stable diffusion WebUI와 ControlNet으로 의상을 다양화하였다. ControlNet의 Segmantation으로 마네킹 부위를 추출하여 카시트와 하네스 영역의 왜곡을 방지하고, Stable Diffusion WebUI에서 Image to Image기능으로 마네킹 의상만 변경하여 다양한 복장을 생성하였으며, 그 결과물을 그림 3에 도시하였다.

이후 YOLO11n pose로 양쪽 어깨를 검출하고, 어깨 간 거리를 이용하여 변 길이의 정사각형을 기준으로 224×224 크롭을 일관되게 생성하였다. 밝기와 채도 조정 및 랜덤 크롭 증강을 적용하여 정상/이탈 각각 300장의 데이터셋을 구축하였으며,



[그림 3] Results of mannequin outfit generation using stable diffusion

Train, Validation, Test를 8:1:1로 분할하였다.

2.3 학습 과정

2.3.1 표정 인식 Fine tuning

입력 데이터셋을 112×112로 리사이즈하고, ImageNet 정규화를 적용하였다. 사전학습모델의 가중치를 사용하되, 전 Layer를 Unfreeze하여 Fine-tuning하였다. 옵티마이저는 SGD, Learning rate: 1e-6, Cross entropy loss, Epoch: 200을 사용하였고, Early stopping patience: 5를 적용하였다. 학습률은 Epoch당 0.9배로 감소시키는 스케줄러를 사용하였다.

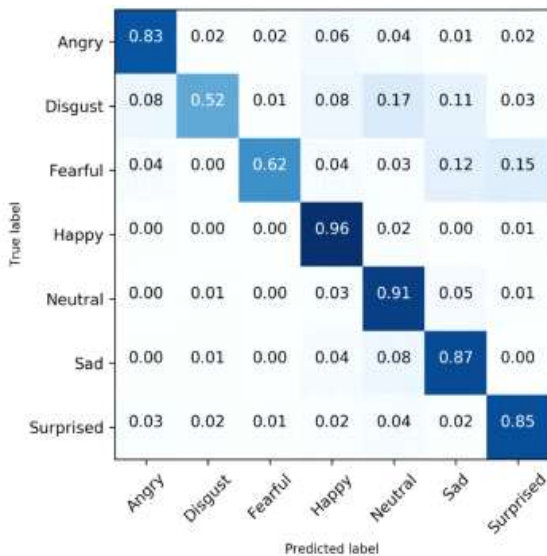
2.3.2 상지 이탈감지 모델 학습

상지 이탈감지 모델 학습은 표정 인식 모델 학습과 비슷하게 진행되었다. ImageNet 사전학습 가중치가 적용된 MobileNetV2에, 하이퍼파라미터는 옵티마이저: SGD, Learning rate: 1e-6, Cross-entropy loss, Epoch: 30, Early stopping patience: 5로 설정하여 학습을 진행했다.

3. 학습 결과

3.1 표정 인식 모델 평가

Fine tuning 이후 테스트셋으로 성능을 검증하였으며, Recall 값을 가지는 혼동행렬을 그림 4에 제시하였다. Disgust와 Fearful 클래스에서는 각각 52 %, 62 %의 Recall 값을 보였는데, 이는 두 클래스의 데이터가 상대적으로 적고, 데이터셋 상에서 Disgust-Angry, Fearful-Surprised 표정 간 경계가 모호하여 발생한 것으로 해석된다. 표정 인식 모델은 카시트 탑승 유



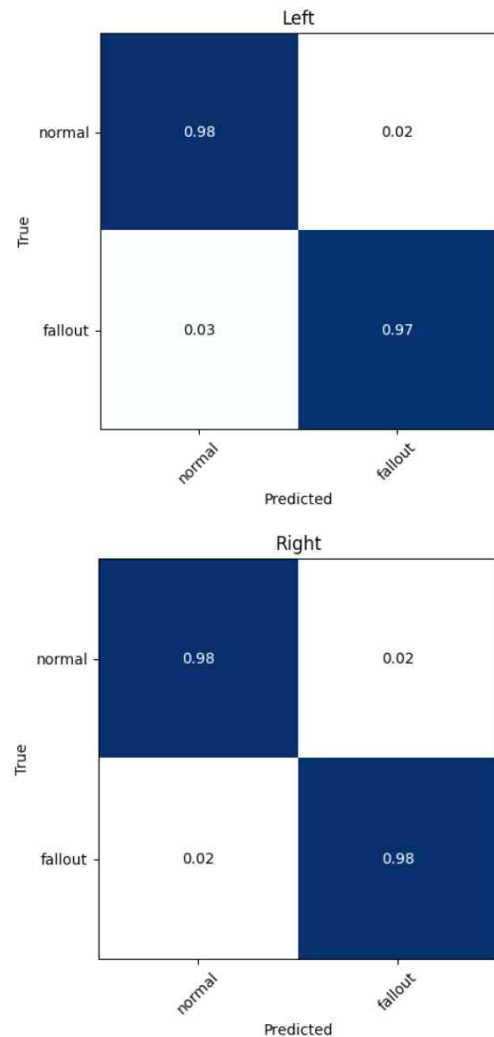
[그림 4] Confusion matrix of the facial expression recognition model

의 불편도를 판단하기 위해 부정, 중립, 긍정의 유사 정서군으로 집계하여 사용하므로 기능적인 영향은 제한적일 것으로 판단했다.

두 클래스를 제외한 나머지에서는 83 % 이상의 값을 보였고, MobileFaceNet에서 얼굴 특징을 효과적으로 추출한 결과로 해석된다.

3.2 상지 벨트 이탈감지 모델 평가

학습된 상지 벨트 이탈감지 모델은 Test 데이터셋이 30장 가량으로 적어, 좌우 Train 데이터셋을 수평 반전시켜 각각 반대편 모델의 평가에 사용하였다. 그림 5는 좌우 모델의 검증 결과를 혼동행렬로 나타낸 것으로 행렬 칸의 숫자는 Recall값을 나타낸다. 상단은 좌상지 벨트 이탈감지 모델, 하단은 우상지 벨트 이탈감지 모델로 두 모델 모두 97% 이상의 Recall을 보인다. 데이터셋에서 Normal, Fallout의 차이점은 벨트의 유무로 해당 특징을 효과적으로 반영한 것으로 판단된다.



[그림 5] Confusion matrix of upper-limb ejection detection model

3.3 경량화 AI 모델 구동 실험

모델은 임베디드 기기인 라즈베리파이5에 Pi OS를 설치한 환경에서 구동하였으며 파이 카메라 V3 모듈을 이미지 센서로 사용하였다. 표정 인식에서는 YuNet과 MobileFaceNet을 사용하여 0.62 GFLOPs의 연산량을, 상지 벨트 이탈감지에서는 YOLOv11n-pose와 MobileNetV2을 사용하여 연산량이 8.2 GFLOPs로, 기기의 초당 부동소수점 연산량에 맞춰 최적화를 진행하였다. 라즈베리파이5는 약 20 GFLOPS의 성능을 가지므로, 영상 스트림을 30 fps로 설정하고 표정 인식과 상지 벨트 이탈감지를 각각 3 fps, 30 fps 주기로 실시하여 총 연산량을 14.4 GFLOPs로 설정했다.

모델 구동 시 큰 프레임 하락 없이 30 fps를 유지하는 모습을 보였으며, 제한한 AI 기반 감지 모델의 실시간 동작 결과를 그림 6에 나타내었다. 그림 6의 (a)와 같은 상황에서는 아동의 표정에 따라 Happy와 함께, 하네스 상태 또한 정상(Normal)으로 분류됨을 확인할 수 있다. 반면, (b)에서는 표정이 Sad로 감지되고, 동시에 좌측 팔의 이탈이 검출되어 하네스 상태가 비정상(Fallout)으로 분류됨을 확인할 수 있다. 이를 통해 제한한 표정 감지 모델과 상지 이탈 감지 모델이 동시에 동작하여 아동의 감정 상태와 하네스 착용 상태를 실시간으로 종합 판별할 수 있음을 나타낸다.

4. 결론



[그림 6] Results of AI-based detection models: a) Facial expression: Happy, Harness state: Normal; b) Facial expression: Sad, Harness state: Fallout (left arm ejected)

본 연구에서는 유아용 카시트 환경에서 아동의 안전성과 편의성 향상을 목표로, AI 기반 표정 및 상지 이탈 감지 모델을 구축하였다. 표정 감지 모델은 YuNet 얼굴 검출기와 MobileFaceNet 분류기를 결합하여 경량화 구조를 설계하였으며, 영유아 표정 데이터셋을 Fine-tuning한 결과 주요 감정 클래스에서 85% 이상의 인식 성능을 확보하였다. 상지 이탈 감지 모델은 YOLOv11n-pose를 통해 추출한 어깨 영역을 MobileNetV2 기반 분류기에 적용함으로써, 좌우 어깨 모두에서 97% 이상의 높은 재현율(Recall)을 달성하였다.

두 모델 모두 임베디드 환경 적용성을 고려한 경량화 설계를 통해 실시간 추론 속도와 높은 정확도를 동시에 만족하였다. 이를 통해 기존 카시트가 지닌 정적인 하네스 고정 방식의 한계를 보완하고, 아동의 행동 및 불편 신호에 능동적으로 대응할 수 있는 기반 기술을 제시하였다.

참고문헌

- [1] 대한민국 법제처, “도로교통법 제50조”, 법률 제17891호, 2021.
- [2] Wu, W., Peng, H. and Yu, S., “YuNet: A Tiny Millisecond-level Face Detector”, *Machine Intelligence Research*, 제 20권, pp. 656–665, 2023.
- [3] Chen, Sheng & Liu, Yang & Gao, Xiang & Han, Zhen. MobileFaceNets: Efficient CNNs for Accurate Real-Time Face Verification on Mobile Devices. 10.48550/arXiv.1804.07573. 2018.
- [4] Ultralytics, “Pose Estimation - YOLOv11 Models (yolo11n-pose.pt, 17 keypoints)”, *Ultralytics Official Documentation*, 2024, [Online]. Available: <https://docs.ultralytics.com/tasks/pose/>