

배제언, 차성현, 유형탁, 최인혁, 최혁진, 김형석
{jeeon98b, cjn112070, air2heo, inhyuck.choi0719, jjin6573, henna6967}@gmail.com

Bae Je Eon, Cha Sung Hyun, Yu Hyeong Tak, Choi In Hyuk, Choi Hyeok Jin, Kim Hyoung Seok
Samsung Software·AI academy for Youth (SSAFY)

본 논문에서는 사용자의 눈 위치와 자세를 실시간으로 인식하여 모니터의 위치(X/Y/Z 축)를 자동으로 조정하는 인체공학 AIoT 디스플레이 제어 시스템(SISEON)을 제안한다. 제안 시스템은 YOLOv11 기반 AI 추론 모듈(Jetson Orin Nano), IoT 제어 모듈(Raspberry Pi 5), 4 관절 로봇 팔, Spring Boot API 서버 및 Flutter 모바일 앱으로 구성된다. 도메인 갭 해소를 위한 3단계 전이 학습을 적용한 시선 추적과 상체 중심의 경량화된 자세 추정 알고리즘을 통해 실시간성과 정확성을 확보했으며, 디스플레이 자동 조정, 프리셋, 자세 통계 등 통합 솔루션을 제공한다. 시제품 구현을 통해 시스템이 높은 정밀도로 사용자 작업 환경을 효과적으로 개선함을 입증했다. (코드 및 시현 영상: <https://github.com/siseon-io/siseon>)

```

graph LR
    User((User)) --> FlutterApp[Flutter App]
    FlutterApp --> PresetFunction[Preset Function]
    FlutterApp --> ManualControl[Manual Control]
    FlutterApp --> LLMChatbot[LLM Chatbot]
    FlutterApp --> PresetDB[Preset DB]
    PresetFunction --> MQTTBroker[MQTT Broker]
    ManualControl --> MQTTBroker
    MQTTBroker --> Process[Process]
    MQTTBroker --> ManualControl2[Manual Control]
    Process <-->|Change Mode| ManualControl2
    ManualControl2 -->|Save Data| SaveData[Save Data]
    PresetFunction -->|UDP Socket| EyeDetection[Eye Detection  
Pose Estimation  
Depth Sensor]
    EyeDetection --> AutomaticControl[Automatic Control]
    AutomaticControl -->|Change Mode| Process
  
```

일반 이미지와 눈 클로즈업 이미지 간의 '도메인 갭' 문제를 해결하고자 점진적 3단계 전이 학습을 적용했다.

(1). 범용 특징 학습: 대규모 데이터셋[6]인 COCO로 YOLOv11 모델을 사전 학습하여 객체 검출에 필요한 기본적인 시각 특징 추출 능력을 확보한다.

(2). 도메인 특화 학습: 얼굴 관련 데이터셋[7] (iSEEDS)을 활용하여 도메인 간극을 1차적으로 해소한다.

(3). 눈4 영역 정밀 파인튜닝: 고해상도 눈 데이터셋[9] (Eye-D)을 사용하여 좌/우안, 및 홍채의 4개 영역을 정밀하게 조정함으로써 검출 정확도를 극대화한다.

3.2 자세 추정 모델: 상체 중심 키포인트 학습

디스플레이 제어에 필요한 상체 정보만 추출하도록, 기존 17~25개 키포인트를 10개(목·양어깨·팔꿈치·손목·가슴·등·머리)로 축소해 학습을 재설계했다. 이 10개의 키포인트는 사용자의 머리 높이와 상체 방향 벡터를 계산하는 데 충분한 정보를 제공해 인체공학적 모니터 위치 제어에 최적화되어 있다. 학습 파이프라인은 COCO 데이터셋으로 사전학습을 진행한 후, 상체 중심 데이터가 풍부한 FLIC[8] 데이터셋으로 미세 조정하여 최적화했다.

3.3 자세 평가 및 피드백 알고리즘

추정된 키포인트 (x, y, z)는 AI 모듈의 2D 좌표(x, y)와 Depth 센서의 거리(z)를 시간동기·외부보정으로 융합(RGB-D) 해 산출하였다. 평활화·좌표계 변환 후 목 굴곡/어깨 전인/상체 기울기 지표를 계산해 [10] 임계치로 판정하고, 지속 편차 시 모니터 $\Delta X/\Delta Y/\Delta Z$ 를 자동 보정하며 결과를 앱에 기록한다.

IV. 성능 평가

4.1 모델 성능

AI 모델의 성능 평가 결과는 [표1] 과 같다. 3단계 전이 학습을 적용한 시선 추적 모델은 2.5 px의 낮은 홍채 중심점 오차(MSE)를 달성했으며, 자세 추정 모델 또한 높은 정확도($AP@0.5=0.65$)를 보였다.

4.2 시스템 주요 성능 지표

- (1). 실시간성: 전체 응답 $\leq 500ms$, Arm 이동 $\leq 1초$
AI 추론 $\leq 50ms/frame$, 자세 피드백 표시 $\leq \sim 1s$
- (2). 보안: 전 구간 HTTPS (TLS 1.3 이상) 적용, JWT 기반 인증과 API Key 병행, 내부망 방화벽, MQTT 브로커 인증·토큰 ACL로 메시지 접근 제한
- (3). 확장성/유지보수: 추상화된 제어 모듈, OTA, 테스트/배포 자동화, 대시보드 통제

모델	주요 역할	학습 데이터셋	성능 지표
시선 추적 (YOLOv11, 3단계 전이학습)	눈 위치/시선 추정	COCO, iSEEDS, Eye-D	홍채 중심점 오차 2.5 px (MSE)
자세 추정 (상체 중심)	10개 상체 키포인트 검출	COCO-POSE, FLIC	$AP@0.5 = 0.65$
자세 이상 탐지 (Rule-based)	잘못된 자세(거북목 등) 판별	의학 논문 및 산업 표준 기반	ISO 9241-5 임계치 기반 물렛·표준 시나리오 정상 검증

[표1: 모델 성능 및 학습 데이터셋]

V. 결론

본 논문에서는 3단계 전이 학습 기반 시선 추적과 경량화된 자세 추정 모델을 적용한 AIoT 인체공학 디스플레이 자동 제어 시스템 'SISEON'을 제안했다. 시제품 구현 및 평가를 통해, 제안 시스템이 높은 정확도와 실시간성으로 사용자의 작업 환경을 개선할 수 있음을 성공적으로 검증했다. 향후 연구에서는 사용자별 피드백을 반영한 개인화 모델 최적화와 다양한 환경 및 디바이스에서의 확장 가능성 검토를 통해, 보다 보편적이고 신뢰도 높은 인체공학 솔루션으로 발전시킬 계획이다.

참 고 문 헌

- [1] ISO. ISO 9241-5:2024 Ergonomics of human-system interaction—Part 5: Workstation layout and postural requirements. International Organization for Standardization, 2024.
- [2] NVIDIA. Jetson Orin Nano Super Developer Kit Datasheet. NVIDIA Corporation, 2024.
- [3] Khanam, R., & Hussain, M. "Yolov11: An overview of the key architectural enhancements." arXiv preprint arXiv:2410.17725, 2024.
- [4] Fingas, J. "Good News! The Faster Raspberry Pi 5 With a Custom Chip Just Dropped." Lifewire, 2023.
- [5] ROBOTIS. "OpenMANIPULATOR-X RM-X52-TNM Overview." ROBOTIS e-Manual, 2025.
- [6] Lin, T.-Y., et al. "Microsoft coco: Common objects in context." European conference on computer vision. Cham: Springer International Publishing, 2014.
- [7] Roboflow Universe. "iSEEDS dataset: Emotion-Types-and-Eye-Location images and annotations in multiple formats for training computer vision models (version 8, May 19, 2023)." Roboflow Universe, 2023.
- [8] Sapp, B. et al. "FLIC: A Dataset for Upper-Body Pose Estimation." 2013.
- [9] Roboflow Universe. "Eye Annotations - YOLO to VOC / Eye Detection KSO3D (Eye-D) dataset (version 4)." Roboflow Universe, 2024
- [10] "Treatment of Chronic Neck Pain in Patients with Forward Head Posture." PMC, 2025.