

LSTM기반 손동작 구조요청 시스템

이승훈¹, 김현성²

¹경일대학교 컴퓨터공학과 학부생

²(교신저자)경일대학교 컴퓨터공학과 교수

¹ish0542@naver.com, ²kim@kiu.ac.kr

LSTM-based Hand Gesture Emergency Rescue Request System

Seunghun Lee¹, Hyunsung Kim¹

¹Computer Engineering, Smart Engineering, Kyungil University

요 약

현대 사회에서 다양한 상황의 구조 요청은 중요한 문제이다. 기존 시스템들은 음성 통화나 문자 메시지에 의존하여 제약 있는 환경에 적용하기에 문제가 있었다. 본 논문에서는 인공지능을 이용한 손동작 구조요청 시스템을 제안한다. 제안한 시스템은 미리 설정한 손동작을 기반으로 위험 상황을 단계(긴급, 중간, 경미)별로 분류하여 신속한 구조 요청을 가능하다. 특히, 다양한 상황과 사용자를 고려해 접근성을 높이고, 비언어적 의사소통의 효율성을 강화가 제안하는 시스템의 목적이다.

1. 서론

현대 사회에서 개인 안전은 점점 더 중요한 화두로 떠오르며, 특히 긴급 상황에서 신속하고 효과적인 구조 요청은 생명과 직결된 중대한 문제이다. 그러나 기존의 구조 요청 시스템은 주로 음성 통화나 문자 메시지에 의존하므로, 제한된 상황에서는 한계가 명확하다[1-3].

본 연구는 인공지능 기반 손동작 구조요청 시스템을 제안한다. 제안하는 시스템은 위험 상황을 미리 정의한 손동작에 의해 레벨 1(즉각적 구조 필요), 레벨 2(중간 수준 위험), 레벨 3(경미한 위험)으로 분류한다. 또한, 인식된 레벨별 위험 상황에 따른 구조 요청을 제시할 수 있다.

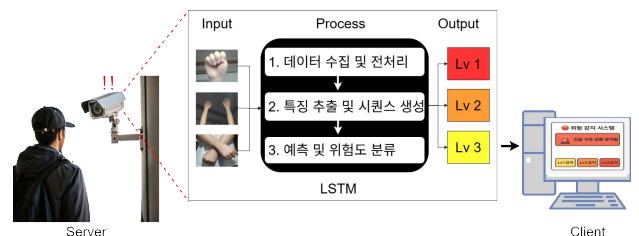
2. LSTM 모델

LSTM(Long Short-Term Memory)은 순환 신경망의 한 종류로, 시간적 의존성을 가진 시퀀스 데이터를 효과적으로 처리하도록 설계된 인공지능 모델이다[4]. 순환 신경망은 짧은 시간 간격의 데이터 처리에는 적합하지만, 기울기 소실 문제로 인해 장기 의존성 학습이 어렵다. LSTM은 이러한 한계를 극복하기 위해 개발되었으며, 연속적인 패턴이나 순차적 이벤트 분석에 강점을 보인다.

셀 상태(Cell State)와 입력, 망각, 출력 게이트로

구성되고 셀 상태는 장기 기억을 유지하며 정보를 다음 단계로 전달한다. 망각 게이트는 과거 정보 중 불필요한 부분을 제거하고 입력 게이트는 새로운 정보를 셀 상태에 추가하며 출력 게이트는 현재 시점의 결과를 생성한다. 이 구조는 복잡한 시간적 관계를 모델링하는 데 유리하다.

동작은 셀 상태를 중심으로 진행되는데 입력 데이터가 셀로 들어오면 망각 게이트가 유지할 정보를 선택하고, 입력 게이트가 새로운 데이터를 업데이트한다. 이후 셀 상태는 과거와 현재 정보를 통합하며, 출력 게이트가 이를 기반으로 은닉 상태를 생성한다. 이러한 과정은 시간적 패턴 학습에 필수적이다.



(그림1) LSTM기반 손동작 구조요청 시스템 개요도.

3. 손동작 구조요청 시스템

본 장에서는 LSTM을 활용한 손동작 구조요청 시스템을 제안한다. 제안하는 시스템은 긴급 상황에서 손동작을 통해 자동으로 구조요청 신호를 보내는 것을 목적으로 한다. LSTM을 통해 손동작을 인식하

고 인식된 결과를 통해 구조 레벨을 결정한 후, 이에 따른 구조요청을 보낸다. 그림 1은 시스템 개요도를 보여준다.

(1) 손동작 캡처

카메라는 사용자의 손동작을 실시간 감지하고, OpenCV의 cv2.VideoCapture(0)을 활용하여 초당 30 프레임의 영상을 획득한다. 보다 정확한 손동작 분석을 위해 cv2.flip(img, 1)을 적용하여 좌우 반전을 수행한다. 감지된 영상은 실시간 처리를 위해 다음 단계로 전달된다.

(2) 데이터 수집 및 전처리

획득한 영상 데이터를 정제하여 인공지능 모델이 학습할 수 있는 형태로 변환한다. MediaPipe Hands를 사용하여 손 관절 좌표를 추출하고 이를 정규화하여 LSTM 모델의 입력값으로 활용한다. 데이터는 30프레임 단위(sequence length=30)로 구성되며, 예측 신뢰도가 90% 이상인 경우 해당 동작을 확정하도록 구현한다.

(3) 인식

손동작의 연속적인 변화를 분석하고 응급 레벨을 예측할 수 있도록 특징을 추출한다. LSTM 층을 128개 및 64개 노드로 구성한다. 기울기 소실 문제를 완화하기 위해 tanh 활성화 함수를 사용하며, Dropout(0.2)을 추가하여 과적합을 방지한다. 학습하는 과정에서는 배치사이즈를 batch_size=64로 설정하고 EarlyStopping를 적용하여 학습 효율성을 최적화한다.

(4) 분석 및 레벨 분류

분석된 손동작 데이터를 기반으로 Lv1, Lv2, Lv3의 3단계 응급 레벨 중 하나로 실시간 분류한다. 레벨 1은 엄지손가락을 제외한 나머지 손가락으로 주먹을 쥐는 동작, 레벨 2는 양손을 들어 올리는 동작, 레벨 3은 양손을 교차하는 동작으로 긴급 구조 요청 손동작을 정의했다.

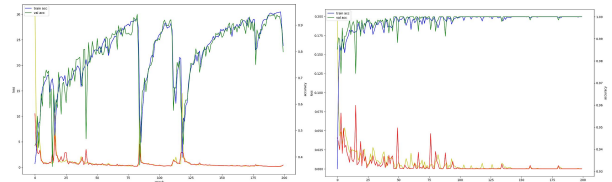
(5) 감지 결과 알림

분류된 응급 레벨을 웹에서 실시간으로 표시하고 위험도가 Lv1 ~ Lv3 중 하나가 감지되면 사용자 인터페이스에 해당 레벨이 출력된다. 사용자는 현재 상황을 빠르게 확인하고 필요한 조치를 즉각적으로 수행할 수 있다.

4. 구현

제안한 시스템은 저전력 임베디드 장치인 라즈베리 파이 4B와 노트북 웹캠을 활용하여 구현하였고,

손동작을 실시간으로 감지한다. MediaPipe Hands를 사용해 21개 손 관절 좌표를 추출하고, 이를 50프레임 시퀀스 데이터로 변환한 후 LSTM 모델로 Level 1, Level2, Level3 동작을 분류하고 97%의 인식 정확도와 평균 30FPS의 처리 속도 달성했다. 서버는 Flask REST API를 통해 연계되고 응답 시간은 약 50ms로 측정됐다. 본 시스템은 실시간 응급 대응에 적합하며 특히 청각 장애인 등 특수 상황에서의 실용성을 입증했다.



(가) 개선 전

(나) 개선 후

(그림 2) 성능 그래프.

5. 결론

본 논문에서는 LSTM 기반 손동작 구조요청 시스템을 제안하였다. 제안한 시스템은 손동작을 통해 응급 상황시 구조의 레벨 1~레벨 3을 판별하고 미리 시스템에 설정된 구조요청 방법을 연계하여 조치하는 손동작 구조요청 시스템이다. 향후 연구에서는 다양한 환경과 사용자 그룹을 대상으로 시스템의 정확도 및 안전성을 확충할 수 있는 추가 검증을 진행하고자 다양한 실제 환경 데이터를 통해 시스템의 효용성을 검증하고자 한다.

참고문헌

- [1] 김태영, 김동우, 김진모, "딥 러닝 기반의 제스처 인식 시스템", 재활복지공학회논문지, 제16권 제1호, pp. 19-261, 2022.
- [2] 이상원, 김태영, "외부 환경에 강인한 딥러닝 기반 손 제스처 인식", The Journal of KING Computing, Vol. 14, No. 5, pp. 31-392, 2018.
- [3] 김계경, 김혜진, 조수현, 이재연, "인간-로봇 상호작용을 위한 제스처 인식 기술", 전자통신동향 분석, Vol. 20, No. 2, pp. 14-20, 2005.
- [4] F. A. Gers, N. N. Schraudolph, J. Schmidhuber, "Learning Precise Timing with LSTM Recurrent Networks", Journal of Machine Learning Research, Vol. 3, pp. 115-143, 2002.