

# 농아인을 위한 수어 지문자 인식 시스템 개선에 관한 연구

강나연, 박영신, 장수현, 강지현

덕성여자대학교 소프트웨어전공

skdusb0007@duksung.ac.kr, p0tlsi@duksung.ac.kr, tngus7238@duksung.ac.kr, jhkang@duksung.ac.kr

## A Study on the Improvement of Fingerprint Recognition Systems for Sign Language Users

NaYeon Kang, YoungShin Park, SuHyeon Jang, Jiheon Kang  
School of Software, Duksung Women's Univ.

### 요약

본 논문은 농아인을 위한 수어 지문자 인식 시스템의 개선에 관한 연구를 다룬다. 해당 시스템은 카메라를 통해 수어 지문자를 실시간으로 입력 받아 텍스트로 번역한다. 연구에서는 OpenCV를 사용하여 비디오 데이터를 처리하고, Google의 MediaPipe를 이용해 손 영역을 검출하였다. 또한, 키 포인트의 위치 값을 정확하게 파악하기 위해 키 포인트 벡터와 각도 계산을 수행하였다. 데이터셋은 한글 지문자 31개를 영상으로 수집하여 모델을 학습시켰고, 데이터셋의 순차적 특성을 고려한 딥러닝 모델을 설계하였다. AutoML을 통해 선택된 모델과의 성능 비교를 수행한 결과, 최종적으로 테스트된 모델은 99%의 높은 정확도를 달성하였다. 향후 연구 방향으로는 텍스트를 수어 지문자 영상으로 변환하는 기능을 추가할 예정이다.

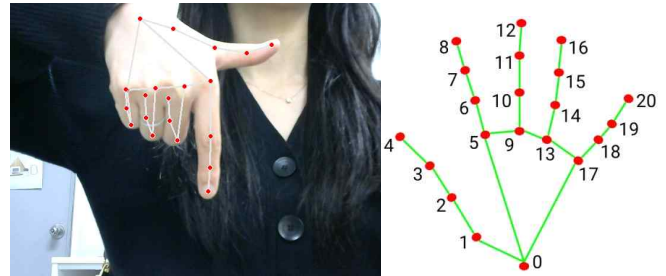
### I. 서론

인간은 언어뿐만 아니라 다양한 방식으로 소통한다. 이 중 수어는 농아인이 일상생활에서 주변 사람들과 의사소통을 하는 데 필수적인 중요한 커뮤니케이션 수단이다. 그러나 수어는 시각적 언어이기에 비장애인이 이해하고 해석하는 데는 어려움이 있다.[1] 이러한 이유로, 농아인에게 편리하고 필수적인 수어를 자동으로 인식하고 해석하는 시스템의 필요성이 커지고 있다. 수어 인식 시스템의 발달은 농아인이 의사소통에 더 자유롭게 독립적으로 참여할 수 있게 도움을 주며, 사회적, 경제적 참여를 촉진한다. 이 시스템은 농아인들이 더 쉽게 정보를 이해할 수 있도록 도움을 줄 뿐만 아니라 사회적 통찰력과 자립성을 증진시킬 수 있다. 또한, 이는 다양한 분야에서 농아인의 삶의 질을 향상시키는 중요한 역할을 하며, 사회적 포용과 인권 증진에 기여한다.[2]

연구의 목적은 수어를 실시간으로 정확하게 인식하고 번역할 수 있는 향상된 시스템을 개발하는 것이다. 이를 위해 최신 딥러닝 기술을 활용하여 LSTM 네트워크를 포함하고 다양한 정규화 및 드롭아웃 계층을 적용한 모델을 설계하였다.

### II. 본론

본 연구에서는 그림 1과 같이 한글 지문자에 해당하는 수어 영상을 데이터로 수집하여, 이를 기반으로 LSTM 기반의 시계열 모델을 학습시켰다. 데이터 수집 과정에서는 OpenCV를 활용하여 입력된 비디오 데이터에서 수어 동작을 추출했다. 이후 Google의 MediaPipe Holistic 모델을 이용해 손 영역을 식별하고, 손의 각 키 포인트의 벡터와 각도를 계산함으로써 동작의 정밀한 분석을 가능하게 했다. 이렇게 수집된 데이터를 바탕으로 모델을 구성하고 학습을 진행하였다.



[그림 1] 지문자 인식 과정

표1은 핸드 랜드마크의 키포인트 벡터 특성을 고려하여 설계한 모델을 나타낸다. 모델 구조는 단순한 순차적 구조로, 하나의 LSTM계층과 두개의 Dense계층으로 구성하였다. LSTM계층은 64개의 뉴런을 가지며, 입력 데이터의 형태에 맞추어 설정되어 있고, ReLU활성화 함수를 사용한다. 또한 과적합을 방지하기 위해 LSTM계층 뒤와 첫 번째 Dense계층 뒤에 각각 0.3의 비율로 Dropout을 적용하였다. 첫번째 Dense계층은 32개의 뉴런을 가지며, 마지막 Dense계층은 다중 분류를 위해 Softmax활성화 함수를 사용하여 최종 출력을 생성한다. 각 계층에는 L2정규화를 적용하여 가중치의 제공에 비례하는 패널티를 추가함으로써 모델의 복잡도를 제어한다. 웹캠에서 실시간으로 이미지를 받아와 처리하며, 감지된 손동작에서 특징을 추출하여 시퀀스를 구성한다. 충분한 길이의 시퀀스가 축적되면, TensorFlow Lite모델에 입력하여 동작을 예측한다. 실험 결과 모델은 84%의 정확도를 달성하였다.

모델의 성능을 개선하기 위해 먼저 각각의 한글 지문자 31개에 해당하는 수어 데이터의 부족한 데이터를 추가로 수집하였다. 이는 데이터 세트의 균일성을 보장하여 모든 지문자가 모델 학습 과정에서 동등하게 대표될 수 있도록 하는 중요한 과정이었다.

표2는 최적의 모델 구조와 하이퍼파라미터를 도출하기 위해 AutoML(Auto-Keras)적용한 모델 구조를 나타낸다. AutoML은 여러 모

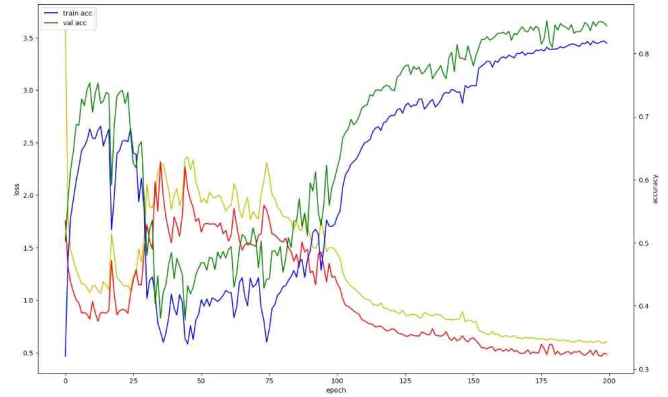
델 구조를 자동으로 생성하고 평가하여 가장 높은 성능을 보이는 모델을 선택하는 프로세스로, 이 과정에서 다양한 네트워크 아키텍처와 하이퍼파라미터가 실험되었다. 이렇게 선택된 모델은 더 복잡한 다층 양방향 LSTM구조를 사용한다. 이 모델은 특히 시퀀스 데이터에서 더 복잡한 특성을 추출하고, 더욱 정교한 학습과 일반화 능력을 발휘할 수 있다. 양방향 LSTM계층은 데이터의 시간적 특성을 양방향에서 포착하여 더 풍부한 정보를 추출하고, Dense계층을 통해 최종적으로 31개의 출력을 생성한다. 이러한 구조적 변경과 추가적인 실험을 통해 이 모델은 99%의 높은 정확도를 달성하였다. 또한 시퀀스 데이터에서 더 복잡한 특성을 추출하고, 더욱 정교한 학습과 일반화 능력을 발휘할 수 있기 때문에 실제 응용에서도 농아인 사용자들에게 더욱 정확하고 신뢰성 높은 번역 결과를 제공할 것으로 기대된다.

Layer (Type)	Output Shape	Parameters
lstm_2 (LSTM)	(None, 64)	30,720
dropout_4 (Dropout)	(None, 64)	0
dense_4 (Dense)	(None, 32)	2,080
dropout_5 (Dropout)	(None, 32)	0
dense_5 (Dense)	(None, 31)	1,023
Trainable Parameters		33,823

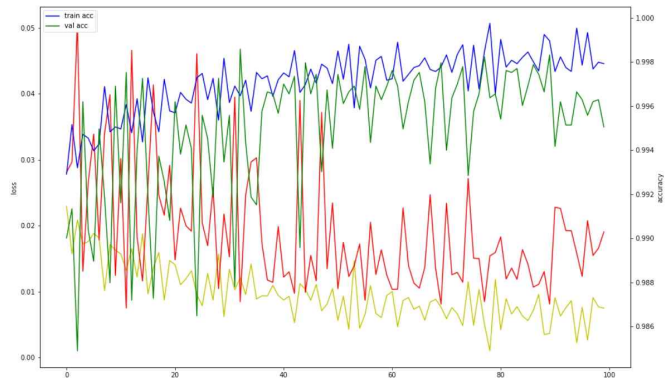
[표 1] 사용자 정의 모델 구조

Layer (Type)	Output Shape	Parameters
input_1 (InputLayer)	[(None, 10, 55)]	0
cast_to_float32 (CastToFloat32)	(None, 10, 55)	0
bidirectional (Bidirectional)	(None, 10, 110)	48,840
bidirectional_1 (Bidirectional)	(None, 10, 110)	73,040
bidirectional_2 (Bidirectional)	(None, 110)	73,040
dense (Dense)	(None, 31)	3,441
classification_head_1 (Soft max)	(None, 31)	0
Trainable Parameters		198,361

[표 2] AutoML 기반 최적화 모델 구조



[그림 2] 사용자 정의 모델의 정확도 그래프



[그림 3] AutoML 기반 최적화 모델의 정확도 그래프

### III. 결론

본 연구에서는 농아인의 수어를 실시간으로 인식하고 텍스트로 번역하는 향상된 시스템을 성공적으로 개발하였다. 개발된 수어 인식 시스템은 기술적 진보를 바탕으로 농아인의 의사소통을 효과적으로 지원할 수 있음을 보여준다. 연구 결과는 기존 시스템에 비해 높은 정확도와 향상된 사용성을 제공하며, 또한 텍스트를 수어 영상으로 변환하는 기술을 추가함으로써 양방향 커뮤니케이션 도구로서의 발전 가능성을 제시한다. 이 시스템의 구현은 농아인의 일상생활에서의 커뮤니케이션을 원활하게 함으로써, 그들의 사회적 소통을 크게 향상시킬 것으로 기대된다.

### 참고 문헌

- [1] 김경민, 송미화 (2022), “MediaPipe를 활용한 수어 번역 시스템 개발.” 한국정보처리학회 학술대회논문집, 29(2). 684-686.
- [2] 구민재. (2019). 딥러닝을 이용한 한국수화언어 통역 시스템 설계 및 구현=Design and Implementation of Korean Sign Language Interpretation System Using Deep Learning. [석사학위논문, 한국방송통신대학교 대학원]. 한국교육학술정보원