

딥러닝 영상인식을 이용한 수화-텍스트 번역 시스템 설계

이종명⁰, 김강규*, 유서연*, 이승건*, 천승현*, 백정윤*, 하옥균*
⁰경운대학교 항공소프트웨어공학과,
*경운대학교 항공소프트웨어공학과
e-mail: {201702018, rkdrb98, yooseo, tmdrjs7369}@ikw.kr⁰,
{cjstmdgusqw, bgy8374}@naver.com*, okha@ikw.ac.kr*

A Design of Sign Language-Text Translation System Using Deep Learning Video Recognition

JongMyeong Lee⁰, Kang-Gyoo Kim*, Seoyeon Yoo*, SeungGeon Lee*,
Seunghyun Chun*, JeongYoon Beak*, Ok-Kyoon Ha*
⁰Dept. of Aeronautical Software Engineering, Kyungwoon University,
*Dept. of Aeronautical Software Engineering, Kyungwoon University

● 요약 ●

본 논문에서는 청각장애인의 사회참여성 증진 및 사회적 차별감소를 목적으로 딥러닝 영상인식 기반으로 MediaPipe 기술을 활용한 수화-텍스트 번역시스템을 설계한다. 제시하는 시스템은 실시간으로 수집된 수화 사용자의 영상정보를 통해 동작과 표정을 인식하여 텍스트로 번역함으로써 장애인과 비장애인의 원활한 의사소통 서비스를 제공하는 것을 주 목적으로한다. 향후 개선된 수화 인식 및 문장 조합을 통해 일상에서 청각장애인과 일반인의 자유로운 커뮤니케이션을 제공하는 서비스로 확장하고자한다.

키워드: 딥러닝(Deep learning), 수화(Sign Language), 번역시스템(Translation system), 미디어파이프(MediaPipe)

I. Introduction

2020년 기준 청각-언어 장애인은 약 42만명으로 매년 그 수가 증가하고 있는 반면 전국수화통역센터는 197개소이며 수화통역사는 976명으로 효율적인 수화통역 서비스 제공이 미흡하다.

본 논문에서는 수화 사용자와의 의사소통을 유도하여 사회 구성원으로써의 사회 참여성 증진과 장애인과 장애인과의 사회적 차별 감소를 목적으로 딥러닝 영상인식 기술을 적용하여 수화를 MediaPipe 를 기반으로 한 ‘딥러닝 영상인식을 이용한 수화-텍스트 번역시스템’ 을 설계한다. 제시하는 방법은 사용자 전방에 있는 수화를 텍스트로 번역하여 화면에 출력한다.

II. Preliminaries

1. Related works

1.1 Sign Language

수화는 가장 작은 단위로 수화소가 있으며 수화소에는 수형(水刑), 수위(手位), 수동(手動), 수향(手向), 얼굴표정과 몸의 움직임인 비수 지신호로 구성되어있다. 눈으로 수화를 인식하고 손을 포함하여 비수 지신호와 함께 표현하는 시각운동체계의 언어이다.

1.2 MediaPipe

Google에서 제공하는 AI프레임워크로써 Video Data를 이용한 비전 AI기능을 파이프 형태로 사용할 수 있도록 제공한다. Face, Pose, Hands인식 등 다양한 비전AI 솔루션 제공되며 본 논문에서는 위에서 언급한 솔루션을 모두 포함한 Holistic을 이용한다.

III. Experimentation

수화를 인식하는 모델을 설계하기 위해 아무동작을 취하지 않는 영상과, 8개의 수화를 표현하고 있는 영상을 사진으로 변환하여 다양한 각도의 수화 사진 6,600장의 자료를 수집하여 학습에 사용하였다. 학습은 Ubuntu OS를 설치한 PC에서 진행하였다. tkinter를 이용하여 수화사용자가 표현한 수화의 순서대로 번역된 텍스트를 출력하여 사용자의 눈앞에 표시하게 구현하였다.

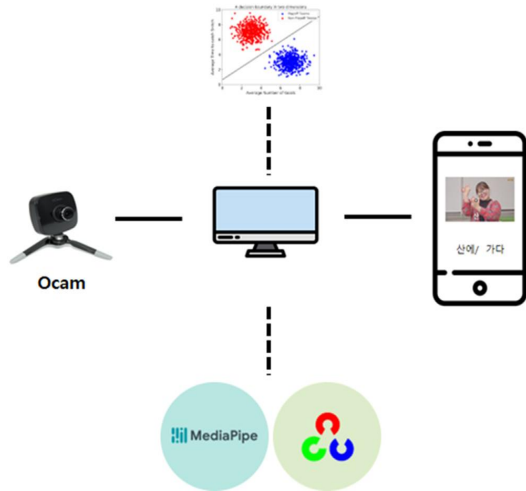


Fig. 1. System Operation Diagram

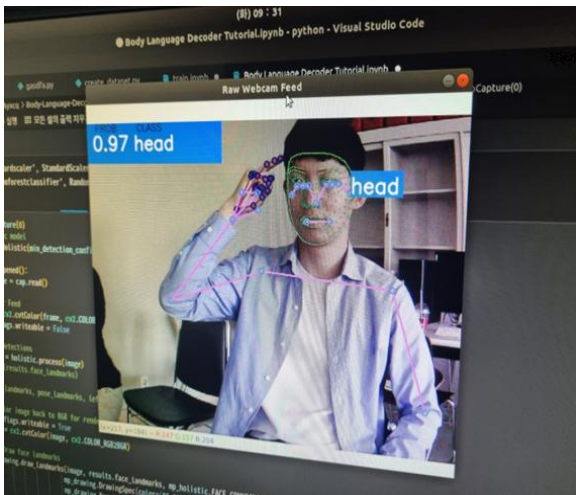


Fig. 2. Sign Language Recognition Test

Fig. 1.은 수화를 tkinter에 표시하기 위한 시스템 운용도이다. 다음으로 수화표현 중 ‘머리’라는 의미를 가진 수화를 정상적으로 검출하고 있다는 것을 Fig. 2.와 같이 확인할 수 있다.

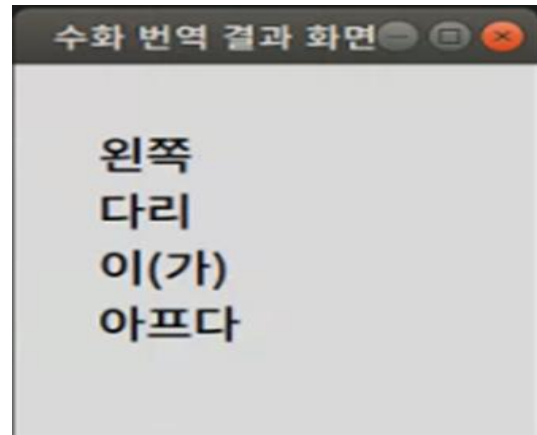


Fig. 3. Sign Language Translation Result

Fig. 3.은 ‘왼쪽’, ‘다리’, ‘아프다’의 의미를 가진 수화를 연속하여 인식 시켰을 때의 텍스트 출력화면이다. 동사를 인식하였을 경우 앞에서 인식한 단어가 명사라면 조사를 출력하도록 하여 수화사용자가 전달하고자 하는 의미를 보다 정확히 전달할 수 있도록 문장으로 출력할 수 있도록 설계해주었다.

IV. Conclusions

본 논문에서는 딥러닝 영상인식을 이용한 수화-텍스트 번역 시스템을 제시함으로써 청각-언어장애인에게 의사소통의 수단을 제공하는 시스템을 설계하고 프로토타입을 제시하였다. 이 시스템은 어플리케이션, 웹페이지 등 장애인이 사용하기 희망하는 다양한 분야에 활용할 수 있을 것으로 기대한다.

REFERENCES

- [1] <https://google.github.io/mediapipe/solutions/holistic.html>
- [2] Yim, HyunJun, "Design and Implementation of The Deep Learning-based Voice to Sign Language Translation System"
- [3] Ryu, Myungwoo, "Design and Implementation of Bidirectional Sign Language System for Emergency Medical Situations"