(Attachment)

About SignTown's Sign Language Recognition Technology

Many sign language recognition models to date have developed recognition technology focusing only on the shape and movement of the hands, but because sign language is expressed in combination with things like upper body orientation, nodding, and facial expressions, hand shapes and movements alone cannot deliver high-accuracy recognition. Highly accurate movement detection requires a machine learning model that can recognize the upper body, head, face, and mouth.

To recognize these important grammatical features of sign language, Google developed sign language movement detection technology that combines three machine learning models using TensorFlow. The first machine learning model is PoseNet, which recognizes human poses and gestures; the second is Facemesh, which recognizes the mouth and facial expressions; and the third is hand tracking, which detects hand shapes and fingers. To obtain suitable training data, the project also collected sign language video data of deaf people in Japan and Hong Kong who use sign language on a daily basis and used it for TensorFlow training.

This sign language recognition technology was used to develop Sign Town. The application can be used on a standard browser, so there is no need for any kind of download. Furthermore, because everything is run within the browser using JavaScript, it enables use at fast communication speeds, making it safe from a user-privacy perspective.
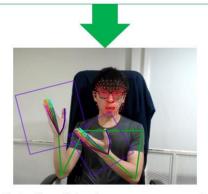

(1)PoseNet: Poses and gestures

(2)Facemesh: Mouth and facial expressions

(3)Hand tracking: Hand shapes and fingers

Detection of sign language movements