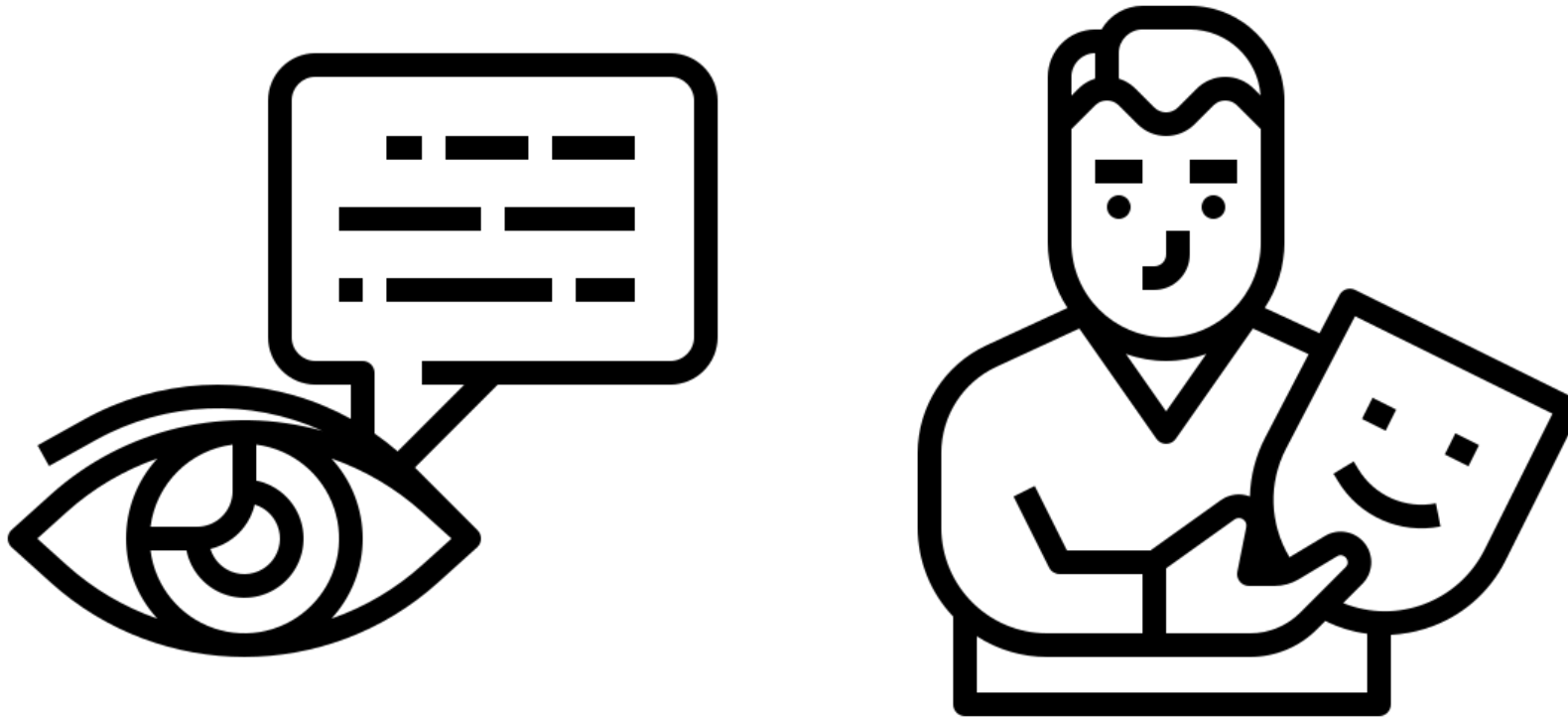


Automatic Generation of Eye Expressions with End-to-End Learning

Ung Park*, Eui Jun Hwang*, JongSuk Choi

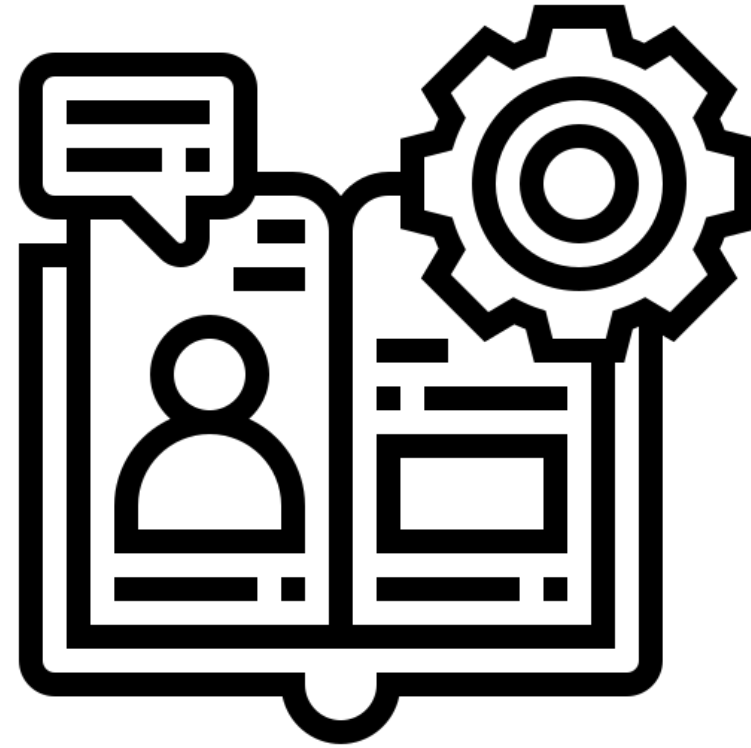
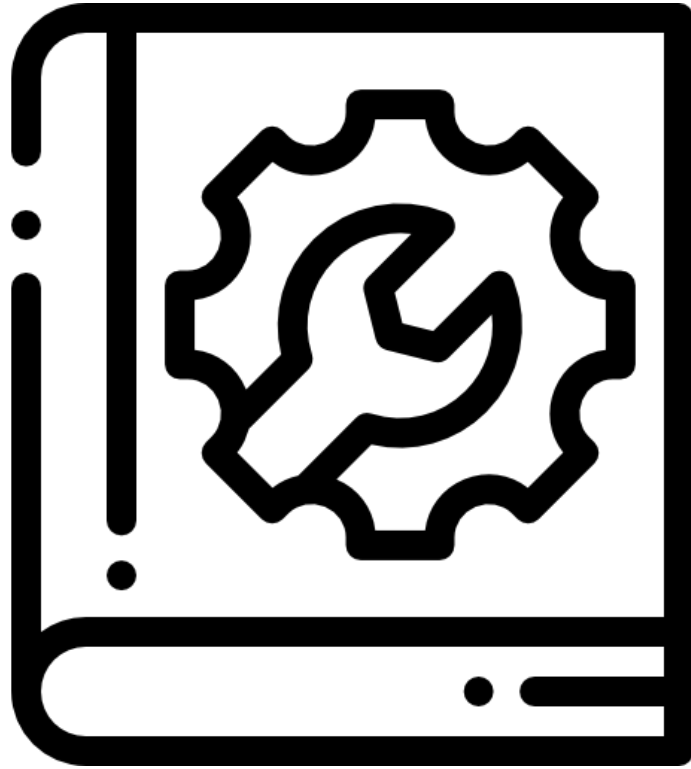
University of Science and Technology
Korea Institute of Science and Technology
{wpark, ktye2220, cjs}@kist.re.kr





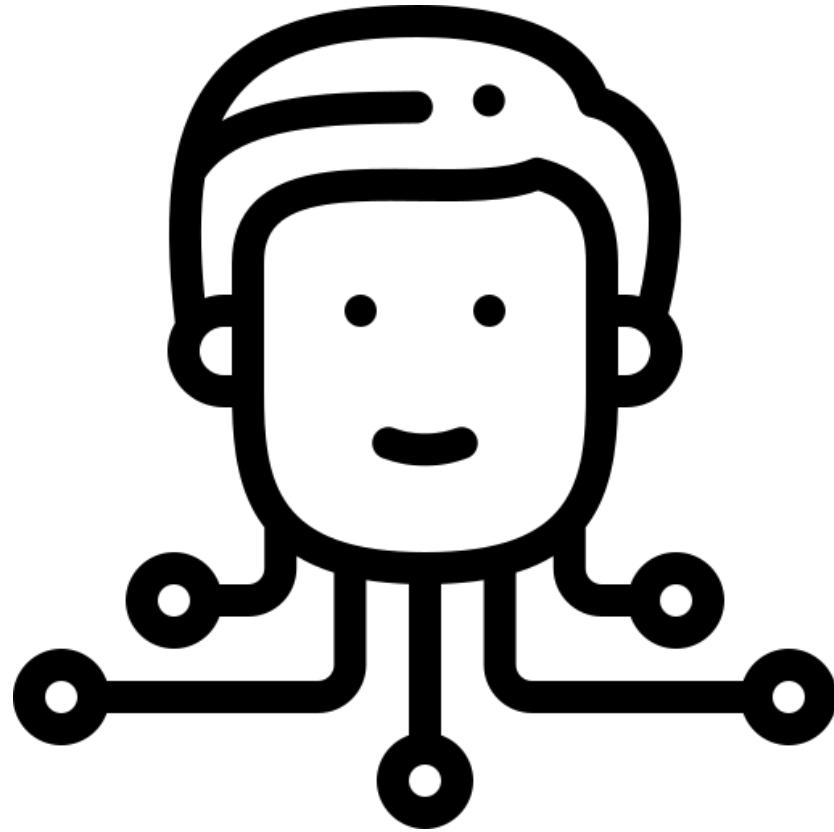
expressing the internal state of itself
affects the development of a positive relationship
between humans and robots





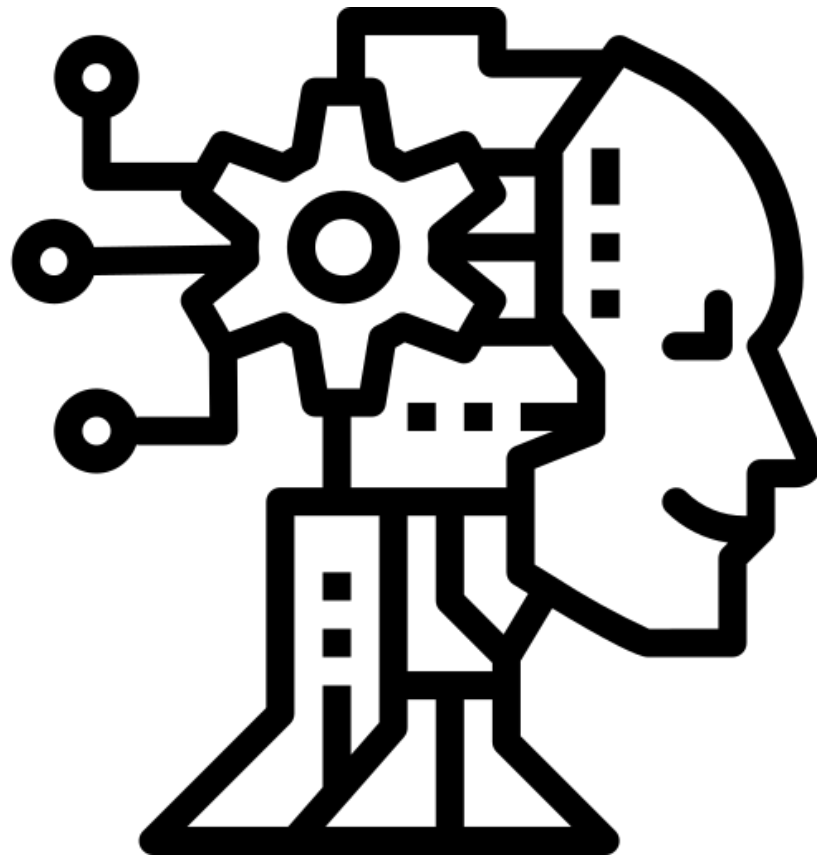
Heuristic expression can be expressed naturally in a similar manner to a human, but there's a limitation as only a predefined expression can be generated





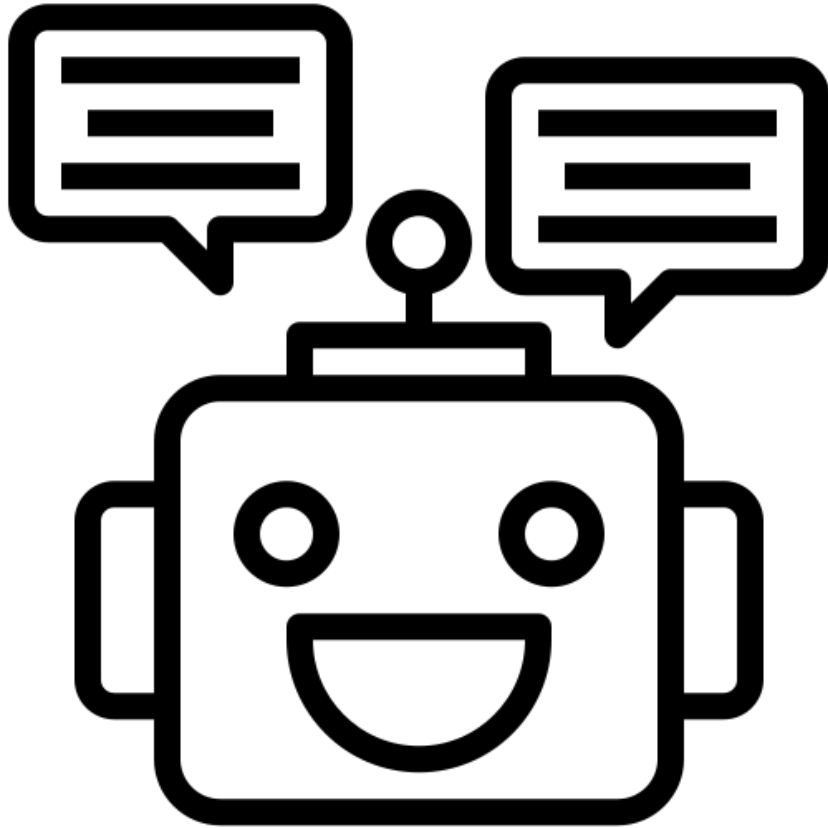
model for generating natural eye expressions of robots





the fundamental task of recognizing or expressing facial expressions is to define the range of each component, with certain facial features involved in facial identification.





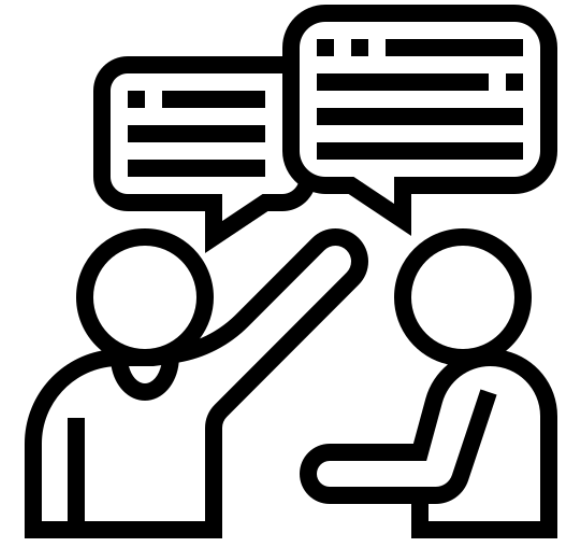
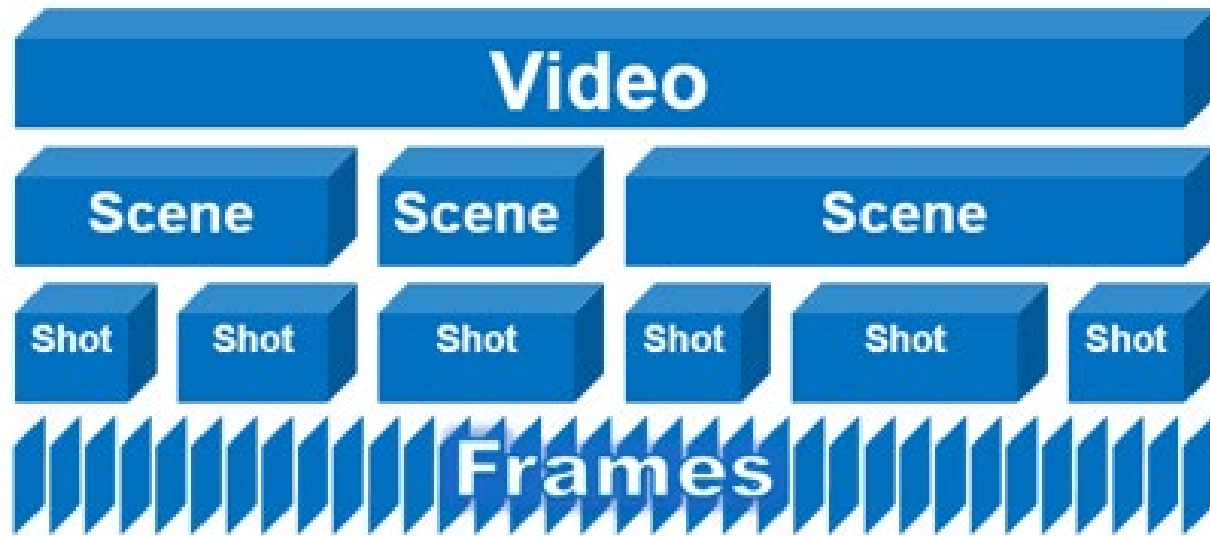
- vertical eye aperture
- distance between the eyebrows
- slope of the eyebrows
- the direction of the gaze





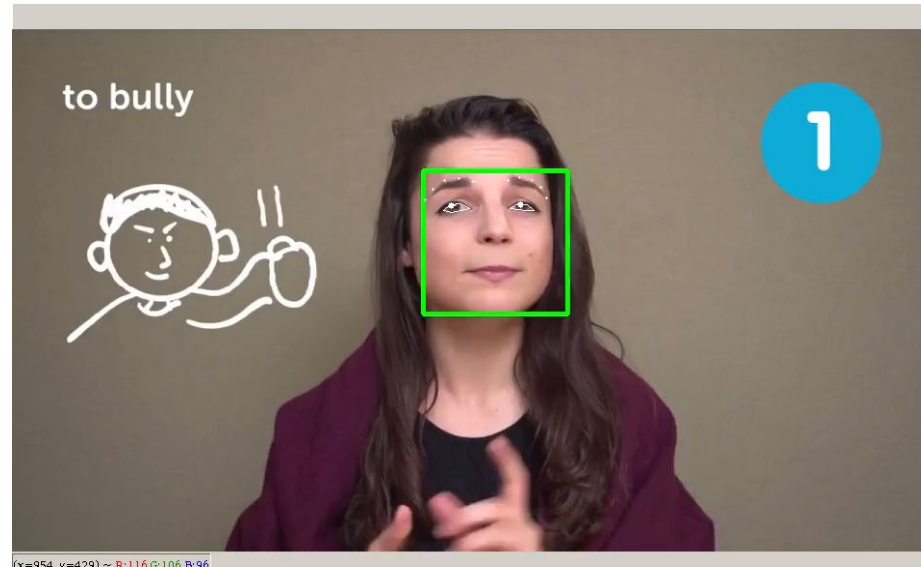
expression in human–human and human–robot interaction is a dynamic region rather than a static region, a method for generating an expression considering a sequence in a video region rather than an image is required.





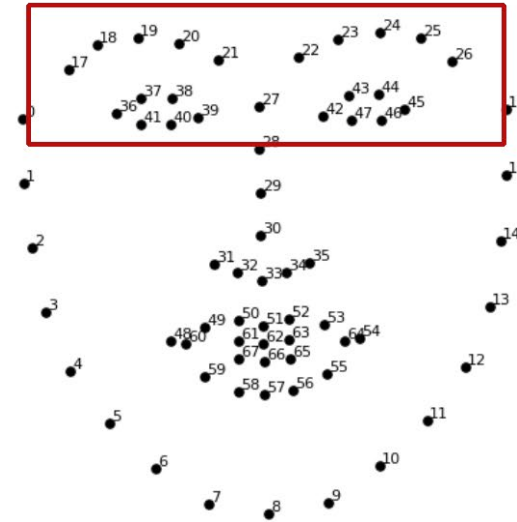
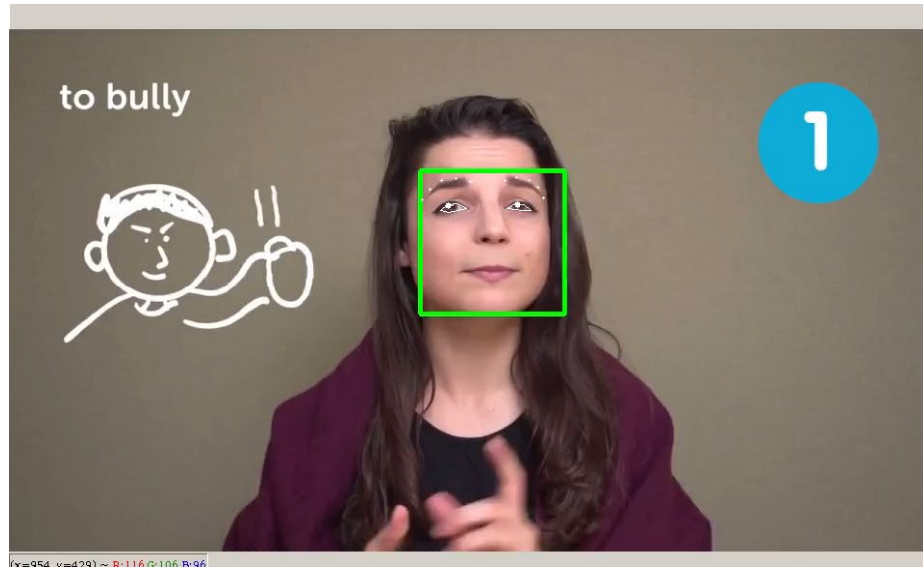
defined facial expressions as a sequential problem





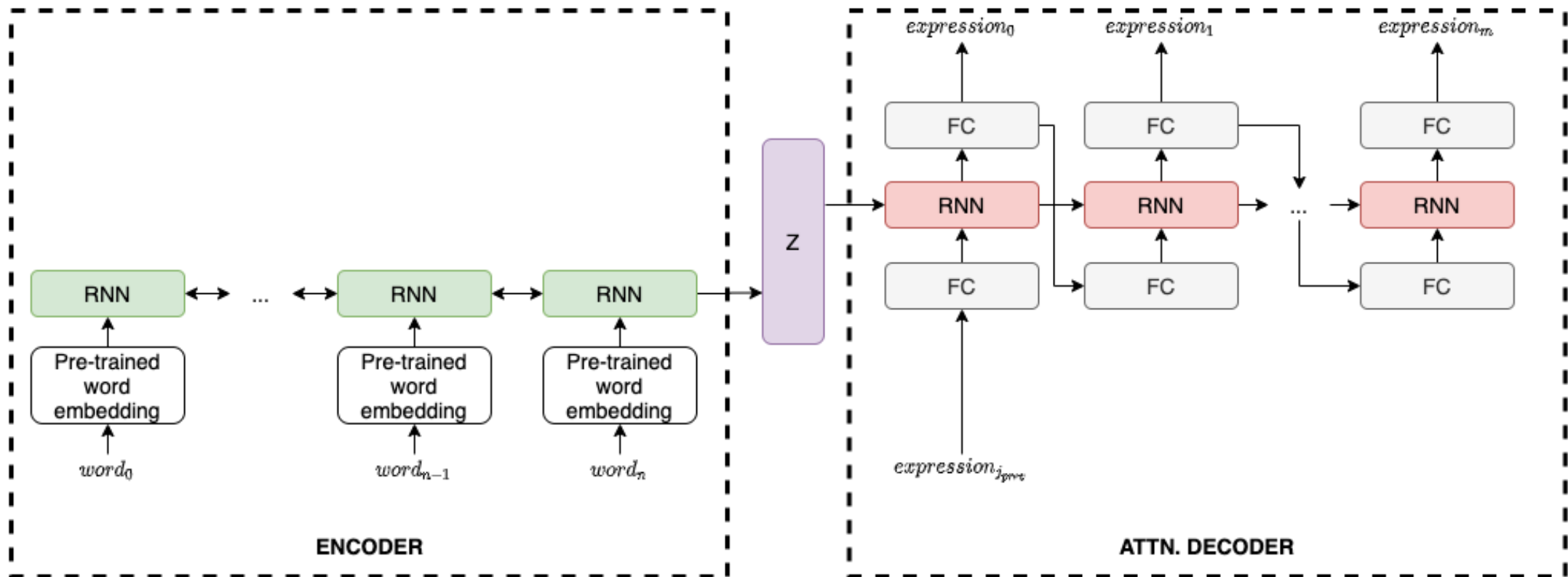
- total of 763 videos & subtitles
- from YouTube's English education channel
- dlib & OpenPose

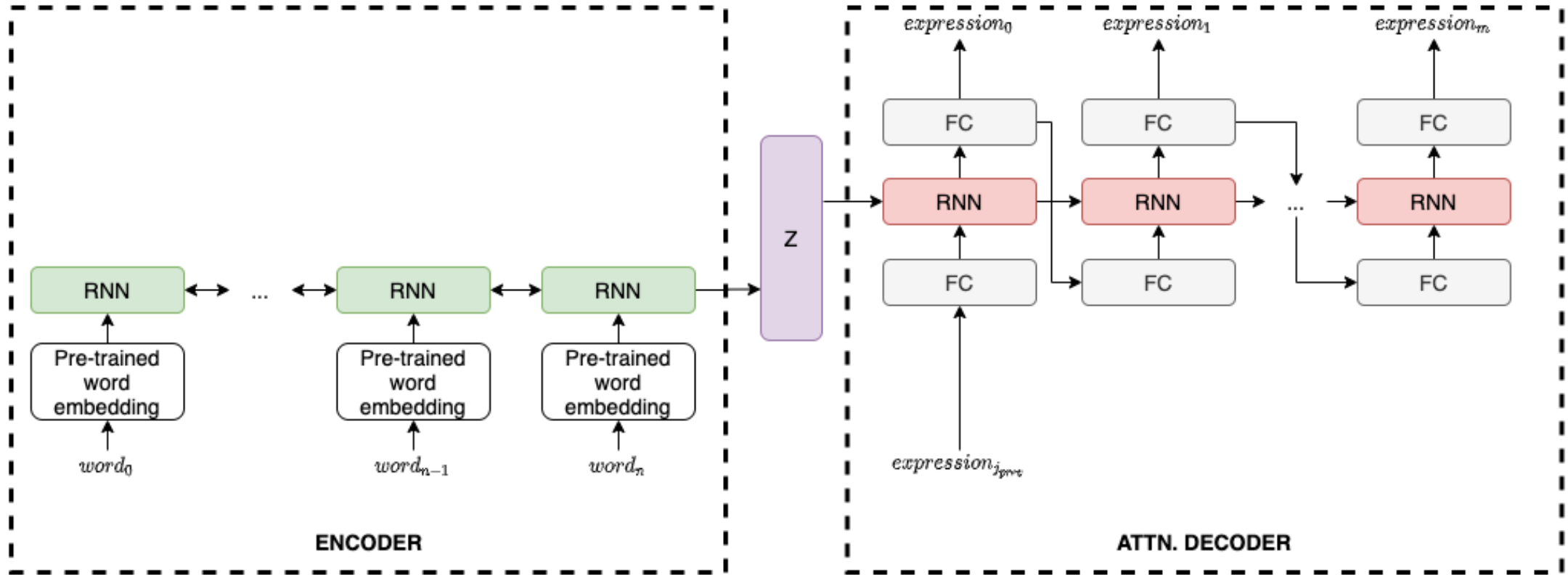




- 1 for eye pupil
- 6 for eye shapes including eyelids
- 5 for eyebrows

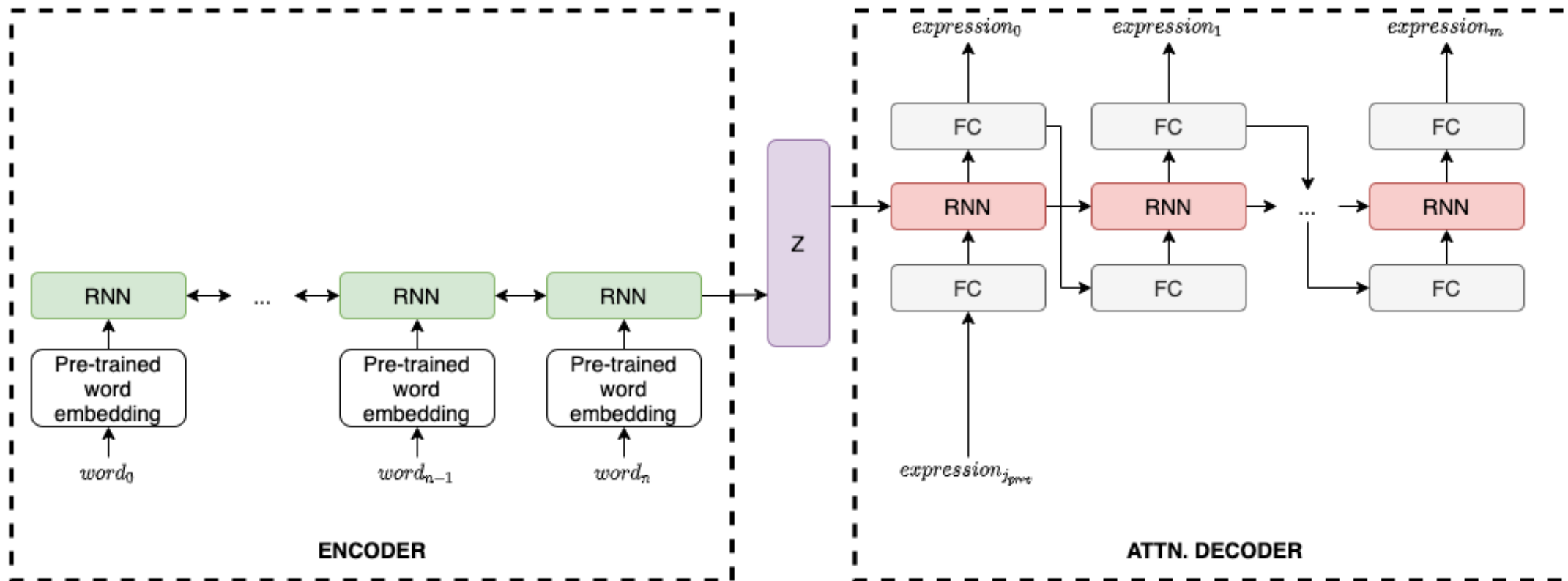






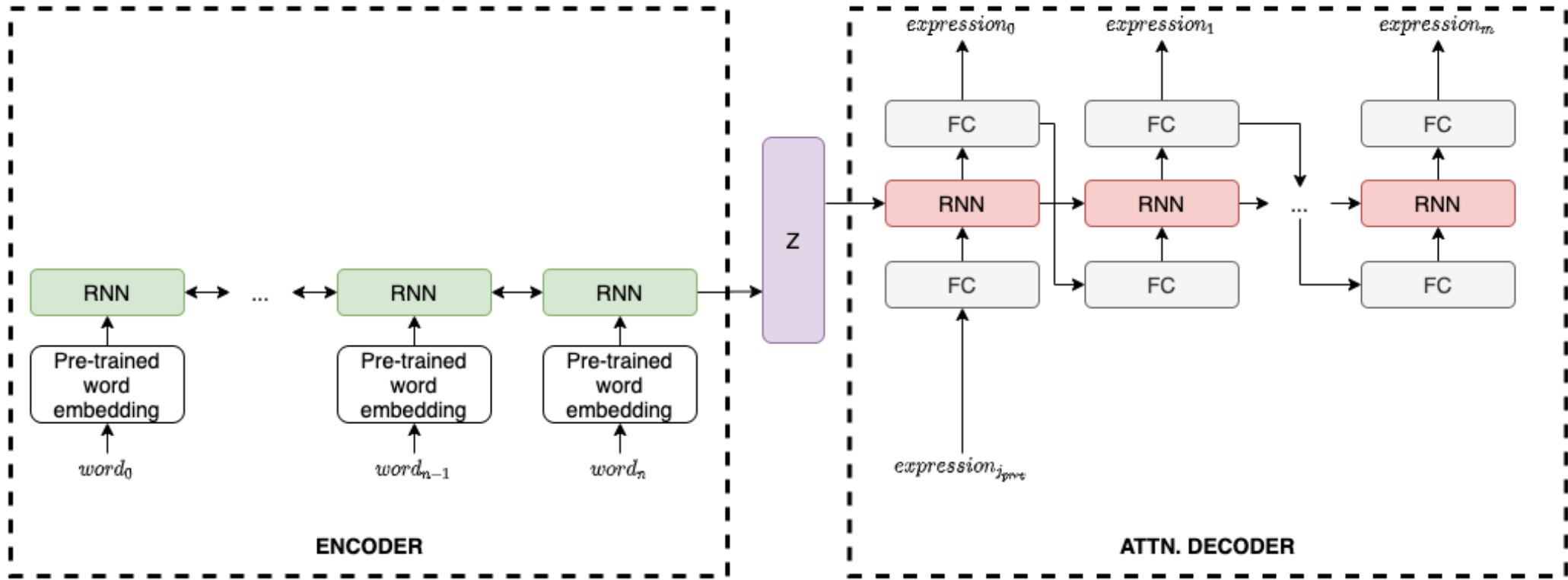
the encoder uses a bidirectional RNN and processes input word by word.





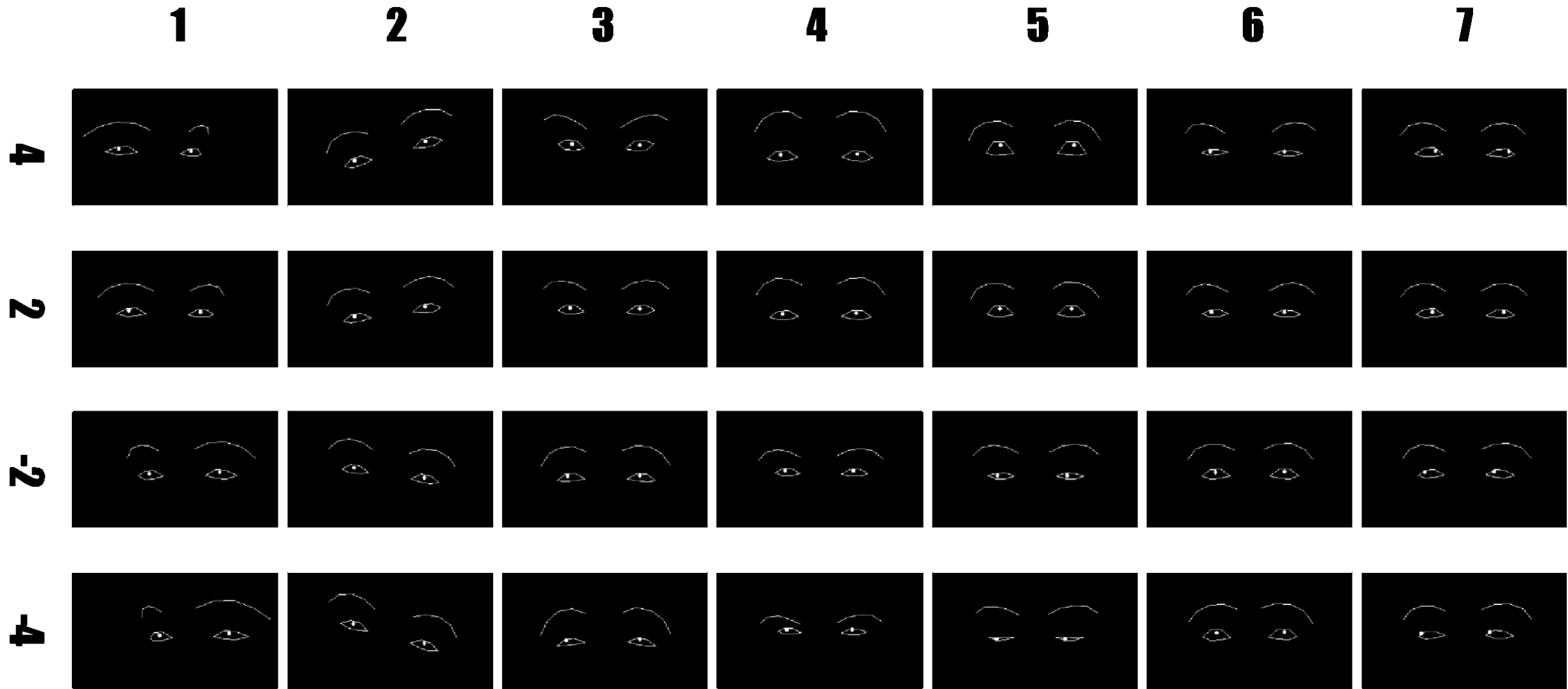
both encoder and decoder use $2 \times$ LSTM with 200 hidden layers





- mean squared error
- 33,195 pairs of sequences of words and eye motion
- Adam optimizer with learning rate of 0.0001
- 512 batch size
- gradients were clipped to 2.0
- 0.1 dropout rate
- train iteration was 1400 epochs
- k-fold cross-validation





the trained network generated different facial expressions for basic emotions
 joy, trust, fear, anticipation, sadness, disgust, anger



- it is a challenge to evaluate how organically the generated expressions connect with spoken text
- trying to conduct a subjective assessment of multiple subjects to measure the correlation between text and facial expressions generated by this model



Thank You

Ung Park*, Eui Jun Hwang*, JongSuk Choi
{wpark, ktye2220, cjs}@kist.re.kr

