

Multistep classification of static and dynamic finger gestures using a soft sensor embedded glove

전수진(석사과정), 박소연(석박사 통합과정), 조혜원(박사과정), 임성훈(교수)

울산과학기술원 산업공학과

{jeonsujin37, syparksy98, hyewon, sunghoonlim}@unist.ac.kr

Contents

1. Introduction
2. Related works
3. Method
4. Results
5. Conclusions and future works

Introduction

Gesture Recognition

- Gesture Recognition

- Gesture recognition indicates the interpretation of gestures by the computing device.
- It is useful for the application like sign language translation, human-computer interaction, and virtual reality.
- Especially, **hand gesture recognition** has been discussed actively because of its applicability and variety of gestures that can be expressed.
- Data are collected from sensors to be used for the input data for gesture recognition of the computer.

Introduction

Gesture Recognition

- Gesture Recognition

- Gesture recognition research field can be classified, based on the sensor, into vision-based and non-vision-based gesture recognition.
- Non-vision-based research usually considers wearable sensors for the data collection, for example, **wearable gloves with soft sensors**.



Soft sensor embedded glove
<출처 : <https://kr.ftsame.com>>

Introduction

Hand Gesture Recognition

- Hand Gesture Recognition

Sensor	Advantage	Disadvantage
Vision-based	More useful when other vision-based tasks are considered together	More affected by the environment
Glove-based	Faster gesture recognition performance	Expensive for real-world usage Inconvenience that comes from wearing gloves

Related works

Hand Gesture Recognition

Reference	Sensor	Description
Kopuklu, Okan, Ahmet Gunduz, Neslihan Kose, and Gerhard Rigoll. "Real-Time Hand Gesture Detection and Classification Using Convolutional Neural Networks." 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), 2019.	Vision-based	<ul style="list-style-type: none">- Highlight single time activation per gesture as an important index for the real-time hand gesture recognition and propose Levenshtein distance as an evaluation criterion.- Use a hierarchical structure of detectors and classifiers.
Zhang, Xiaoliang, Ziqi Yang, Taiyu Chen, Diliang Chen, and Ming-Chun Huang. "Cooperative Sensing and Wearable Computing for Sequential Hand Gesture Recognition." IEEE Sensors Journal 19, no. 14 (2019): 5775-83.	Glove-based	<ul style="list-style-type: none">- Use both a wearable armband and a glove for hand gesture data collection- A long short-term memory (LSTM)-based algorithm is used.
Dong, Yongfeng, Jielong Liu, and Wenjie Yan. "Dynamic Hand Gesture Recognition Based on Signals from Specialized Data Glove and Deep Learning Algorithms." IEEE Transactions on Instrumentation and Measurement 70 (2021): 1-14.	Glove-based	<ul style="list-style-type: none">- Propose low-cost data gloves- Use the TCN-based algorithm and outperform previously proposed methods for glove-based dynamic hand gesture recognition

In real-world application, **static and dynamic hand gestures** should be considered simultaneously.

Method

Gestures

American Sign Language (ASL)

Dynamic gestures

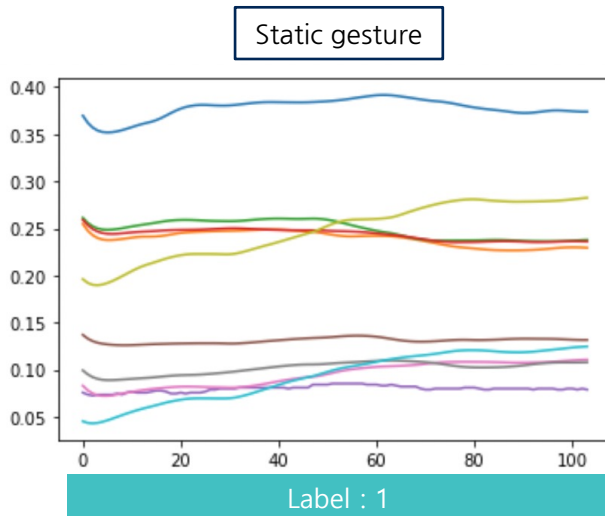
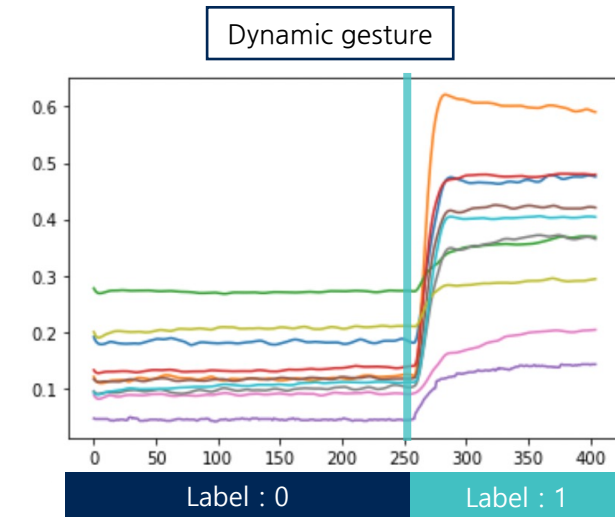


Static gestures



Method

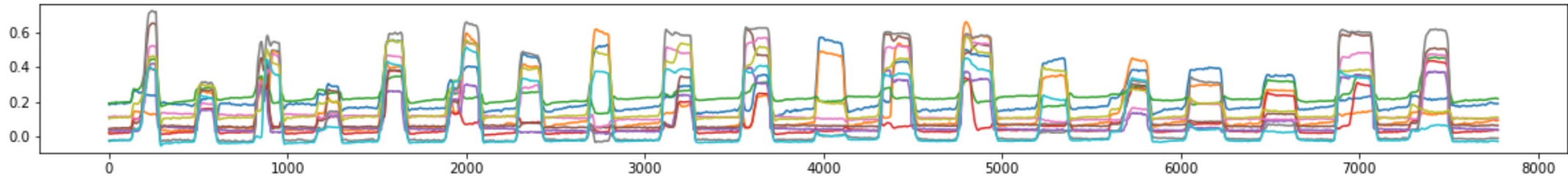
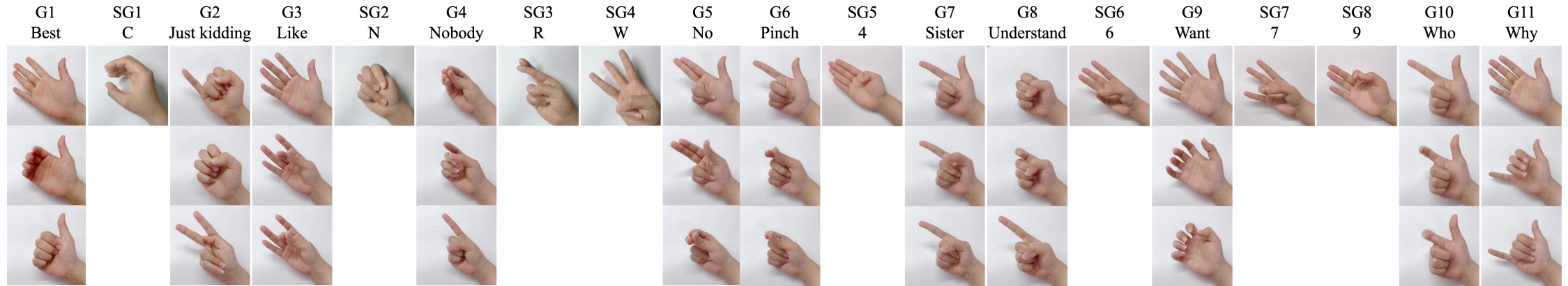
Data - Training set



- In training set, all static and dynamic gestures are separately collected with 10 soft sensors embedded in data glove by 4 subjects.
- After discarding some noisy data, 1,210 dynamic gesture samples and 452 static gesture samples are collected.
- Stratified split of training and validation sets with respect to both subjects and gestures by 0.2:0.8 ratio.
- Min-max normalization for each 10 sensor values is done.
- In dynamic gesture samples, labeling is done as 0 (non-gesture) before the actual gesture start point.

Method

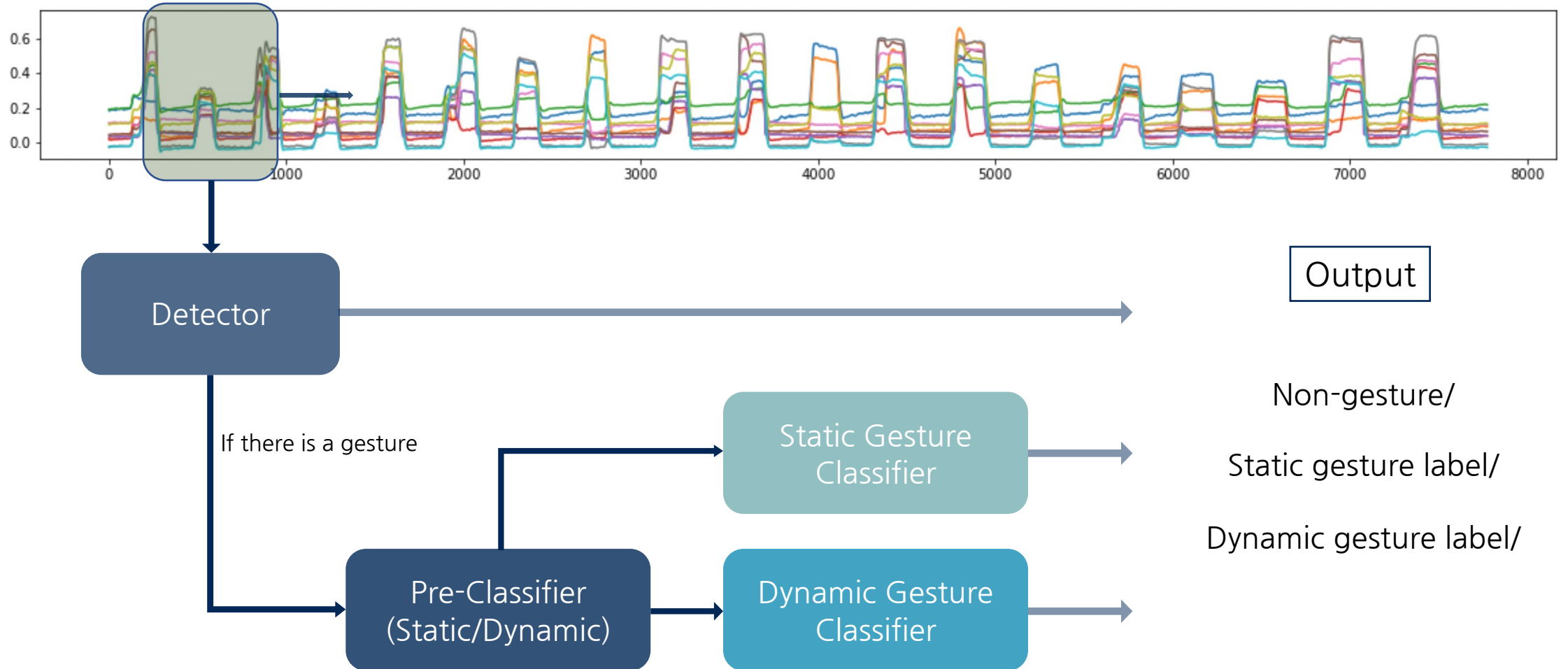
Data - Test set



- In test set, all static and dynamic gestures are collected continuously in a sample to consider the real-world setting.
- 20 samples will be collected by one subject and subjects not in the training set will be participated.

Method

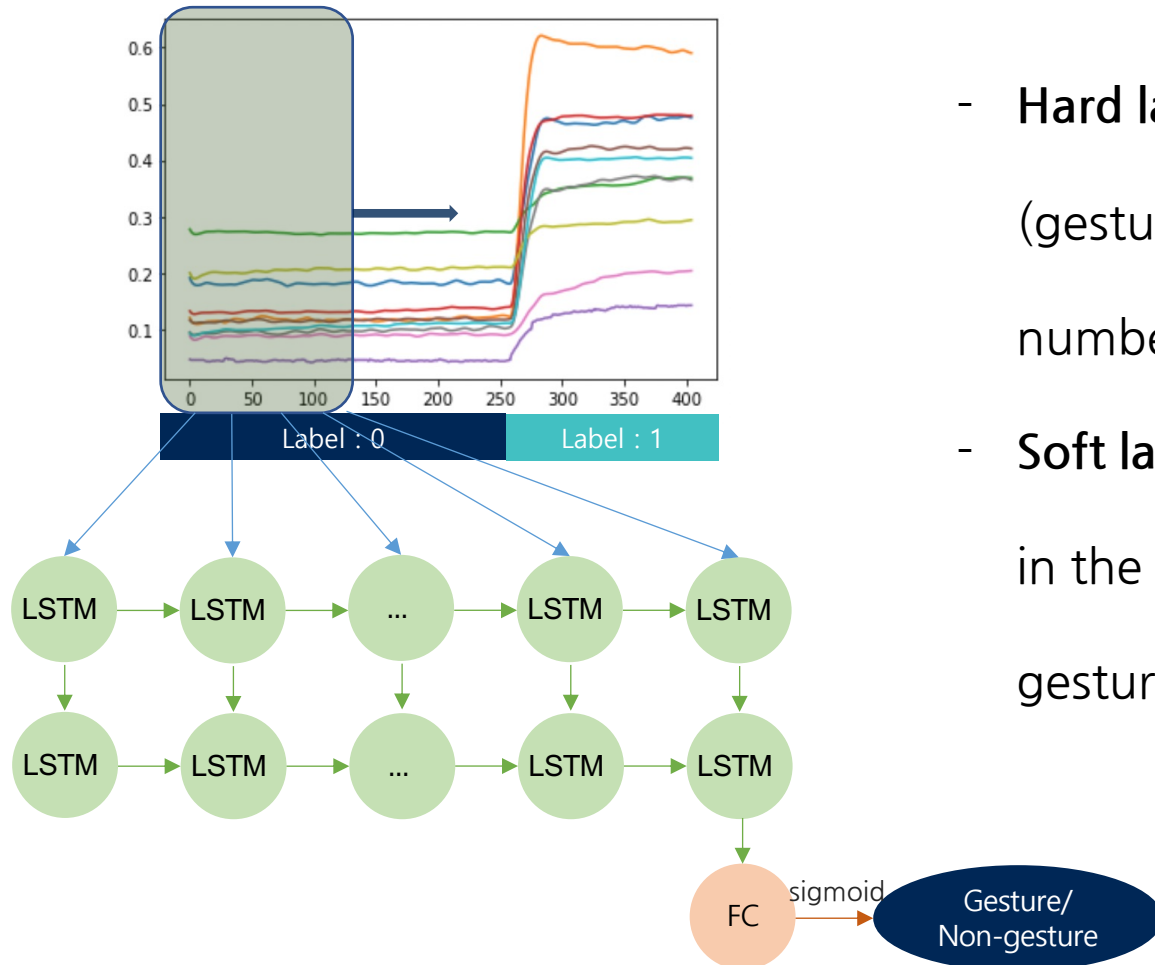
Overall architecture



Method

Detector

Detector : binary classification (classify whether there is a gesture in the input data)

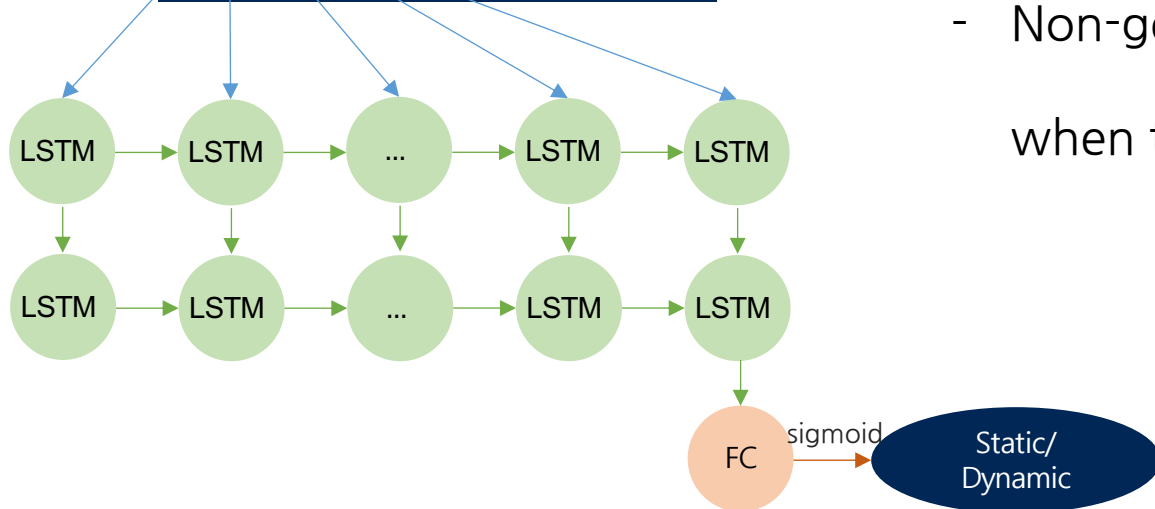
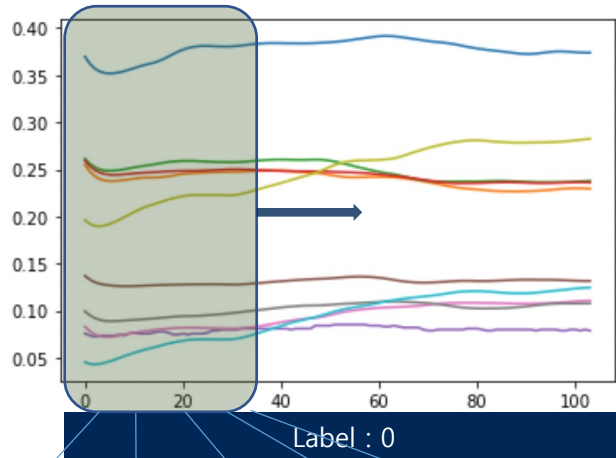


- **Hard labeling** : label as 0 (non-gesture)/1 (gesture) with a pre-defined threshold of number of 1 in the window data label
- **Soft labeling** : indicate normalized number of 1 in the window data label as probability of the gesture's existence in the window data

Method

Pre-Classifier

Pre-Classifier : binary classification (classify whether a gesture is static or dynamic)



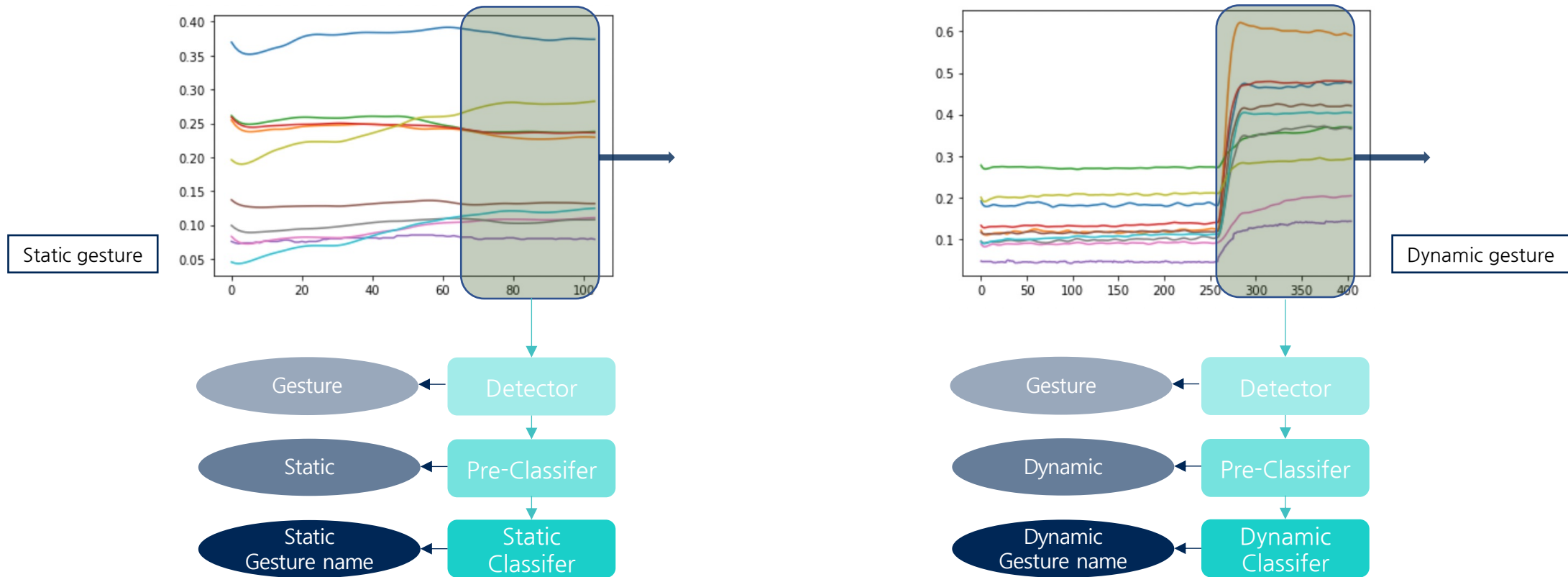
of features in the hidden state = 20

- When the detector classifies input data as 'gesture', Pre-Classifier classifies input data as static or dynamic.
- Non-gesture part in training set is dropped when training Pre-Classifier.

Method

Static Classifier and Dynamic Classifier

Classifier : multi-class classification (classify each static or dynamic gesture class)



- Depending on the Pre-Classifier's prediction, input data move to Static Classifier or Dynamic Classifier.

Results

Detector

- Hard labeling

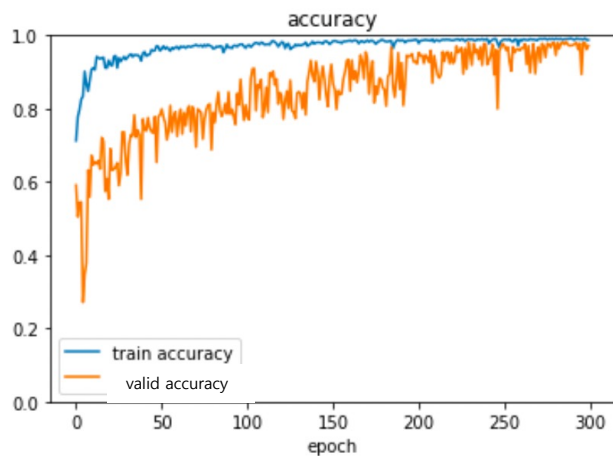
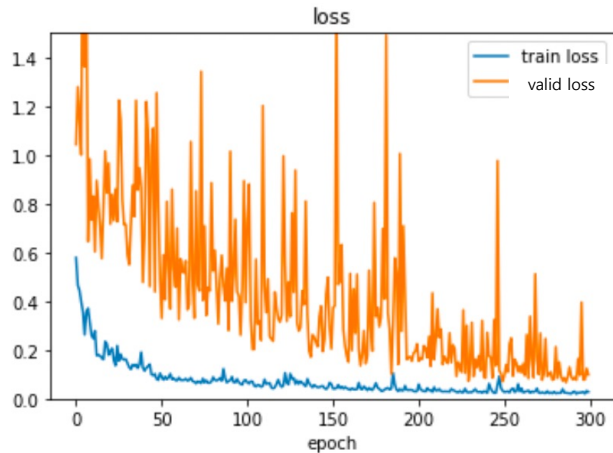
Window size = 20

Stride = 1

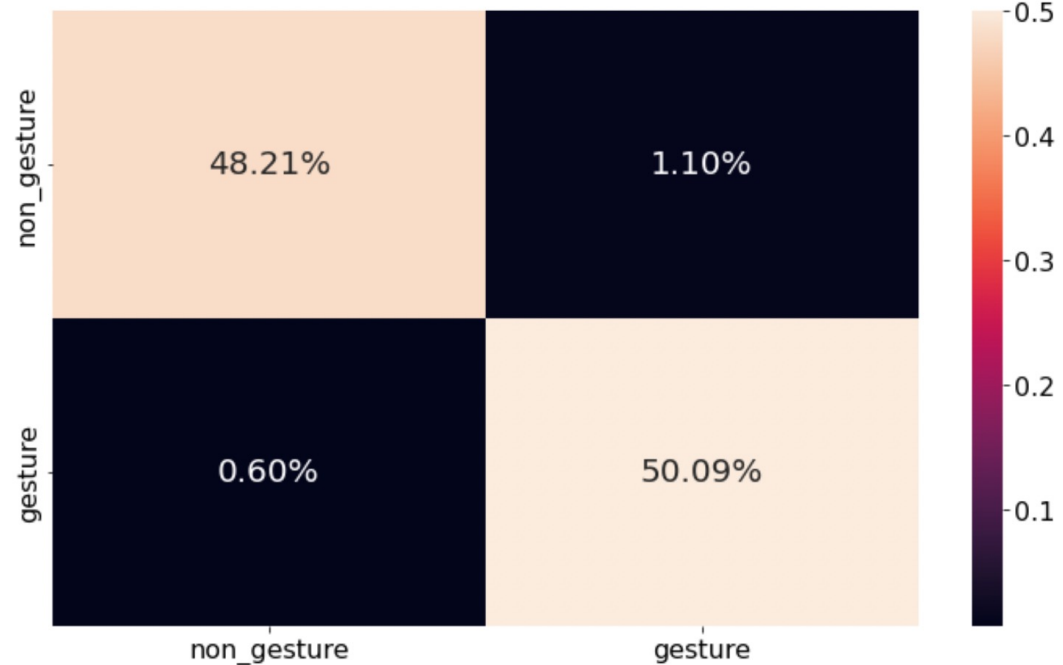
Epoch = 287

Binary Cross Entropy Loss

Adam Optimizer



Confusion matrix on the validation set



F1 score : 0.9833

Results

Detector

Window size = 20

Stride = 1

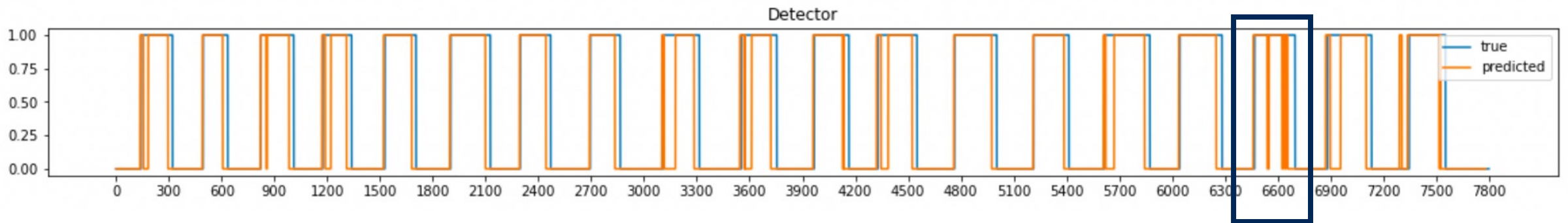
Epoch = 287

Binary Cross Entropy Loss

Adam Optimizer

- Hard labeling

Visualization of prediction on a sample of continuous gesture data (1: gesture, 0: non-gesture)



Post processing needed

Results

Detector

- Soft labeling

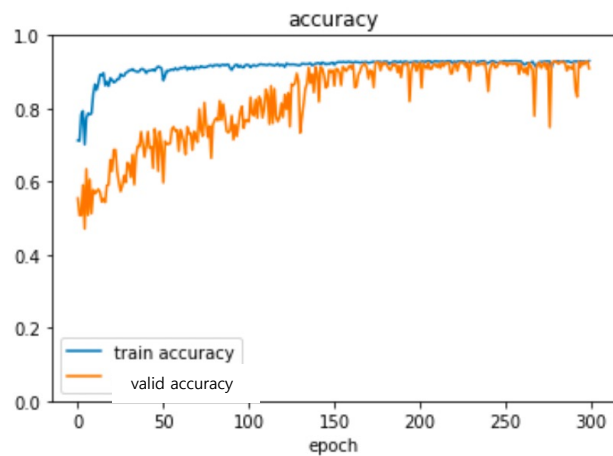
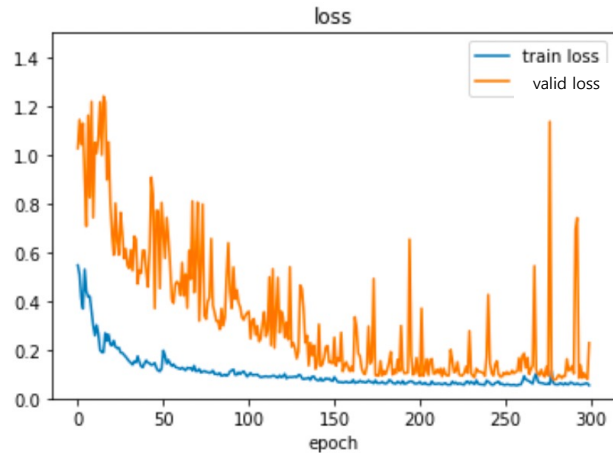
Window size = 20

Stride = 1

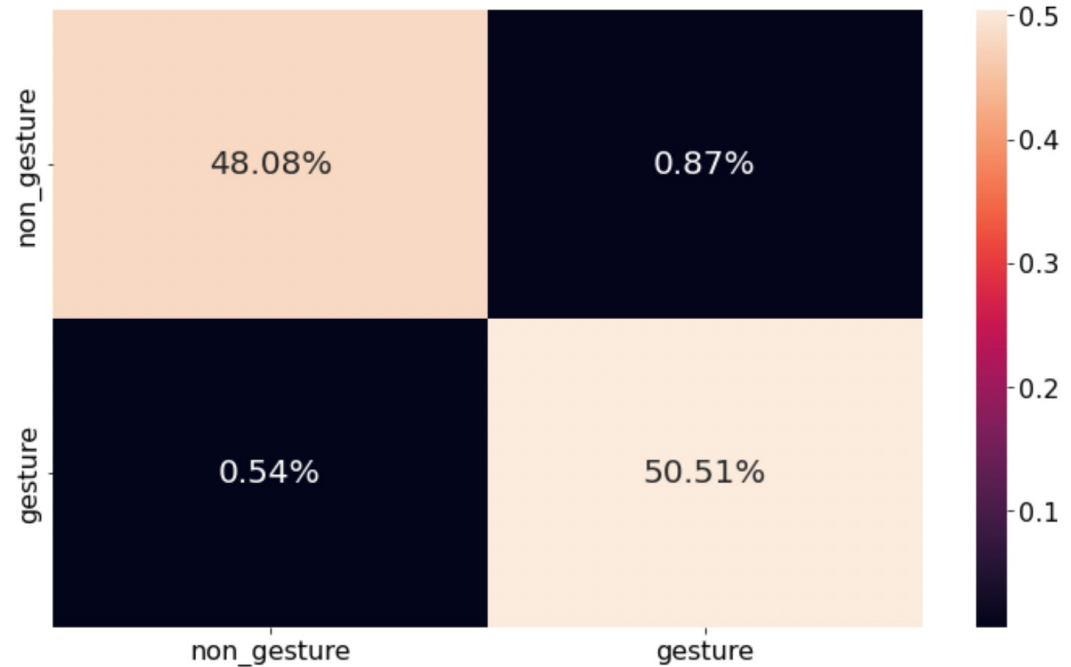
Epoch = 280

Binary Cross Entropy Loss

Adam Optimizer



Confusion matrix on the validation set



F1 score : 0.9862

Results

Detector

Window size = 20

Stride = 1

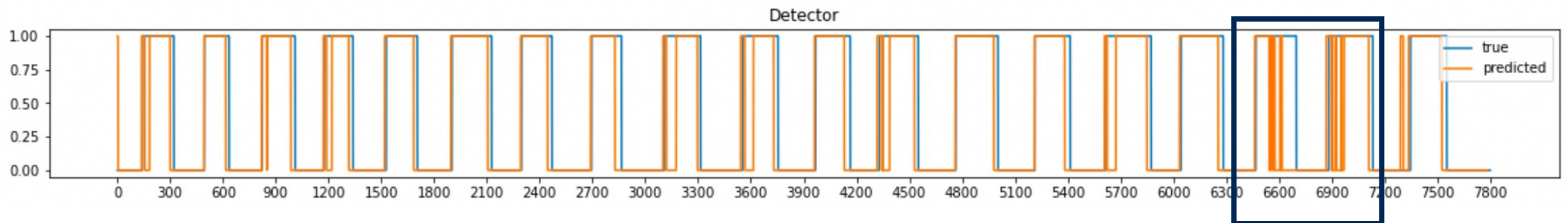
Epoch = 280

Binary Cross Entropy Loss

Adam Optimizer

- Soft labeling

Visualization of prediction on a sample of continuous gesture data (1: gestures, 0: non-gestures)



Post processing needed

Results

Pre-Classifier

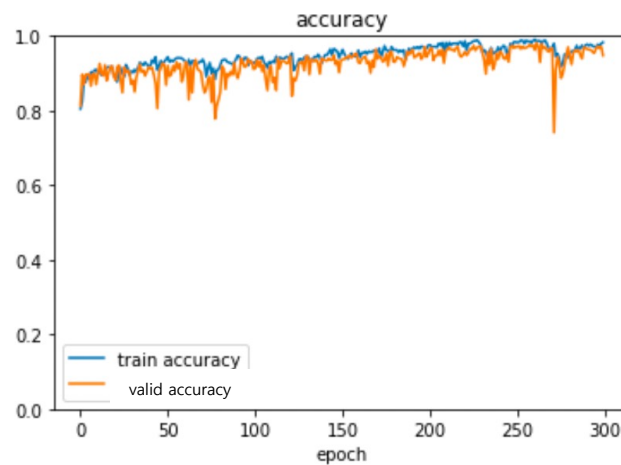
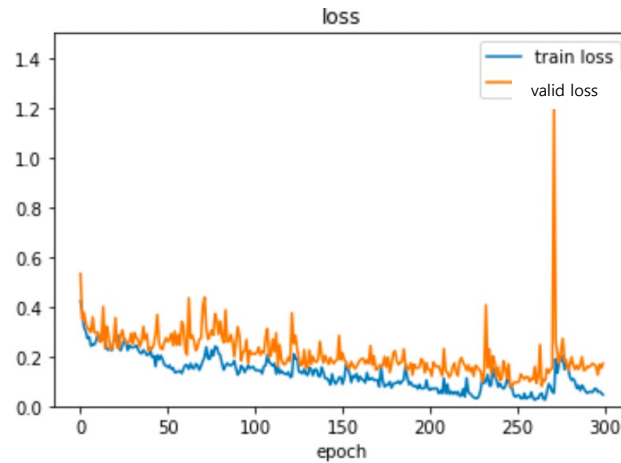
Window size = 20

Stride = 1

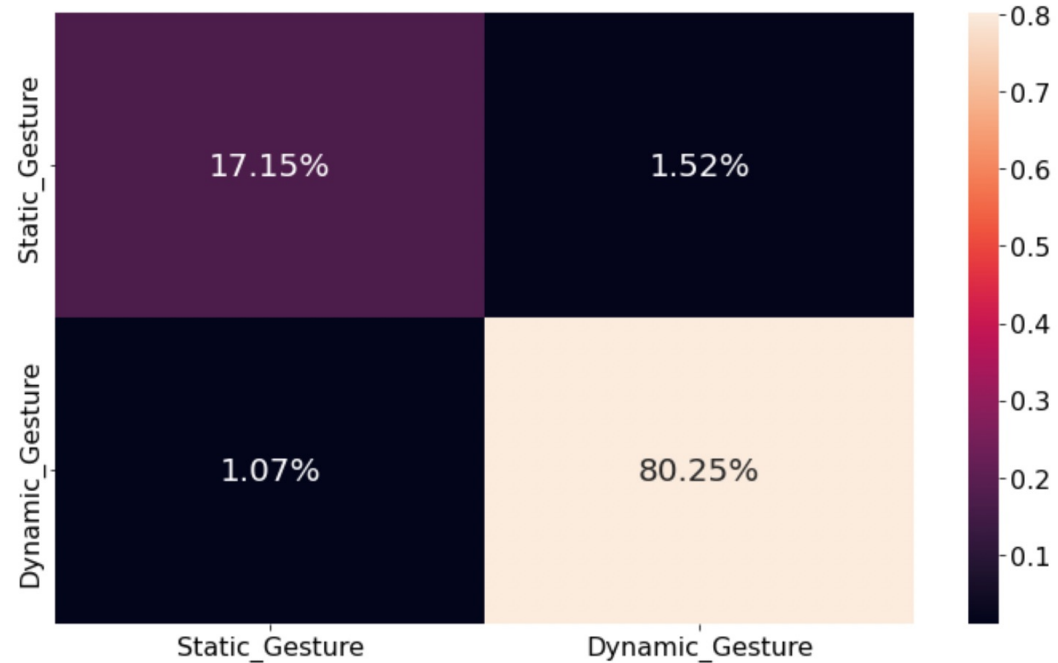
Epoch = 248

Binary Cross Entropy Loss

Adam Optimizer



Confusion matrix on the validation set



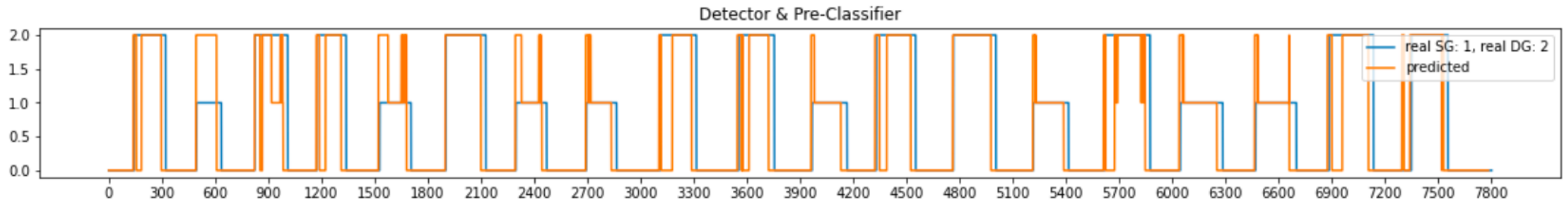
F1 score : 0.9841

Results

Detector & Pre-Classifier

Visualization of prediction on a sample of continuous gesture data

(0: non-gestures, 1: static gestures, 2: dynamic gestures)



Conclusions and future works

- In this research, a multi-step classification architecture is proposed to recognize static and dynamic finger gestures robustly and simultaneously in real time using a soft sensor embedded glove.
- Trained detector has shown that it can distinguish whether there is gesture or not with the LSTM-based model, when both static and dynamic gestures are labeled same as a gesture for training.
- Attention-based algorithm can be applied for the classifier component for the future works.
- Metrics for natural language processing for evaluation with continuous datasets will be considered.
- Two data-glove can be used for developing a two-hand gesture recognition model as future works.

Reference

- [1] Lee, Minhyuk, and Joonbum Bae. "Deep Learning Based Real-Time Recognition of Dynamic Finger Gestures Using a Data Glove." IEEE Access 8 (2020): 219923-33.
- [2] Liu, Hongyi, and Lihui Wang. "Gesture Recognition for Human-Robot Collaboration: A Review." International Journal of Industrial Ergonomics 68 (2018): 355-67.
- [3] Kopuklu, Okan, Ahmet Gunduz, Neslihan Kose, and Gerhard Rigoll. "Real-Time Hand Gesture Detection and Classification Using Convolutional Neural Networks." 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), 2019.
- [4] Zhang, Xiaoliang, Ziqi Yang, Taiyu Chen, Diliang Chen, and Ming-Chun Huang. "Cooperative Sensing and Wearable Computing for Sequential Hand Gesture Recognition." IEEE Sensors Journal 19, no. 14 (2019): 5775-83.
- [5] Dong, Yongfeng, Jielong Liu, and Wenjie Yan. "Dynamic Hand Gesture Recognition Based on Signals from Specialized Data Glove and Deep Learning Algorithms." IEEE Transactions on Instrumentation and Measurement 70 (2021): 1-14.

THANK YOU

This work was supported by the Advanced Technology Center Plus (ATC+) Program (20017932, 50% Accident Prevention Focus to reduce accident rate Development of Risk Detection System for Road Facilities Based on Artificial Intelligence) funded by the Ministry of Trade, Industry and Energy (MOTIE) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1F1A1046416).