Hello everyone.

My name is Atsuhiro FUJII from Ritsumeikan University in Japan.

OK, Let’s start.

I will talk about “Cooking Activity Recognition with Convolutional LSTM using Multi-label Loss Function and Majority Vote”.

I will explain about the given dataset.

We were given train dataset and test dataset.

Micro-activities consist of ten labels, and macro-activities consist of three labels.

This is the details of the dataset.

Each segment contains one macro behavior and up to six micro behaviors.

The length of the data ranges from 0 to about 9000.

Zero length means that the data is missing.

The data for Subjects 1, 2, and 3 are for training and the data for Subject 4 are for testing.

Hand crafted feature values are extracted from the raw data.

These features are calculated over a 50 milliseconds-window slid in steps of 3 seconds.

From the preprocessing, 21 dimensions feature are obtained for one sensor.

This is our identification model.

In "Conv1d layer", Mapsize was set to 6, we input features of 21 dimensions with length N’, and get a features of 21 dimensions with 6 map and length N’’.

Then, in the "LSTM layer", that features are inputted to form a 24-dimensional feature.

"Linear layer" compresses 24 dimensional features into 10 dimensional ones.

In training phase, that output data are inputted to "BCEWithLogistsLoss“ in Pytorch library, and train model.

In the "Sigmoid layer", Sigmoid function is applied.

Next, we obtain one-hot vectors in the "activation layer" using the threshold T.

This T is the threshold at which the most accuracy is achieved in the tests within the training data.

In "Final activation layer", a majority vote is taken.

The one-hot vectors of each position obtained by the "activation layer" are added together, and the label with a result of 2 or more is output as the result.

However, if the number of sensors is 2 or less because of missing data, we adopt a label with a result of 1 or more.

On the other hand, for macro activities, in training phase, we apply "CrossEntropyLoss“ after the “Linear layer“, and train model.

Then, the results obtained by the "Sigmoid layer“ are added together and the label with the largest value is adopted.

From these results, the average accuracy of 0.521 and 0.491 were achieved among subjects 1, 2, and 3 in leave-one-subject-out manner for micro and macro activities, respectively.

Considering ten multi-label micro activities, it would be said that 0.521 accuracy is good.

On the other hand, 0.491 accuracy for 3-class macro activity may be improved.

We trained the model on data from subjects 1, 2, and 3, and obtained prediction results for subject 4's data, and the results were then submitted.

Conclusion, our model uses convolution layer and LSTM.

The evaluation results showed that the average accuracy of 0.521 and 0.491 for micro and macro activities.

Now all finished. Thank you for your attention.