講演番号	080

文献番号	20225080

Prediction of decreasing arousal level by deep learning for face images

Yuki Mekata [1)]　Miwa Nakanishi [1)]

1) Keio University, Faculty of science and technology
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, 223-8522, Japan (E-mail: yuki.mekata@keio.jp)

KEY WORDS:**human engineering, driver condition, driver condition monitoring,** deep learning [C2]

Sleepiness is a major factor in accidents or errors. Even in these days of promoting system automation, users should monitor the state of the autonomous system and take over in the event of an emergency. Thus, it is important for systems to understand the user's arousal level and manage the user's arousal level appropriately. A method to maintain a driver's arousal level is achieved by establishing a model to detect the driver's sleepiness and an effective means to awaken the driver. We first focused on the latter and proposed an in-vehicle artificial intelligence (AI) agent that enables voice interactions that induce the driver's intrinsic motivation, which was experimentally effective in maintaining the driver's arousal level(1). Additionally, we examined to determine the timing of interaction based on the stage of arousal level (2). If it is realized to apply before the arousal level is decreased, it is hoped that the system will be more effective in preventing decreasing in arousal level. Thus, in this study, we attempt to construct a model that predicts decreasing arousal levels in advance with the goal of achieving such a system. Although sleepiness detection has already been studied, the focus is mainly on predicting whether a driver is in a dangerous state or not, and predicting the stages of arousal level remains a challenge.

In the experiments to acquire data for model construction, participants monitored autonomous driving for one hour. Participants could only change lanes by pressing a button, and all other operations were handled by the autonomous system. During the task, participants responded to the arousal level in five levels by pressing a button. The participants were advised of the timing of answering the arousal level by a synthesized voice every 30 seconds, requesting that they answer the current sleepiness. The experiment included three male and three female participants, and each participant completed the task three times. We measured the participant's face images at 60 Hz.

For model construction, we used features obtained by embedding using deep learning for face images. We used the gated recurrent unit (GRU) and neural network (NN) to construct models. We constructed a model that uses the features extracted from the face images as input and outputs the values of the five-stage arousal level answered during the task. The target time to be predicted was changed by changing the value of arousal level used as output based on the number of seconds ahead of input face image. The target time was varied at 30 second intervals, and examples ranging from 30 to 180 seconds were examined. Furthermore, to reduce the effect of data bias in the arousal levels, we resampled the data for each arousal level to match the training data's lowest number of levels.

From the output of models, we confirmed that the constructed models have been able to learn rough changes in arousal level, although it was not able to adequately capture the five-stage change in arousal level. For model evaluation, the accuracy of the model was calculated by each model using GRU and NN (Fig.1). There was no remarkable difference between using GRU and NN, and the accuracy decreased progressively as the target time increased. In addition, considering that the aim is to apply it to real-time feedback, it is important not only to have high accuracy but also to have a small deviation between the predicted and actual values. Thus, we calculated the percentage of the case that the deviation between predicted and actual values less than 1 for models using GRU and NN (Fig.2). From these results, although the accuracy was not sufficient in the five-stage arousal level prediction, the five-stage the constructed model achieved the prediction deviation of less than 1 for about 60% of the data in the five-stage prediction.

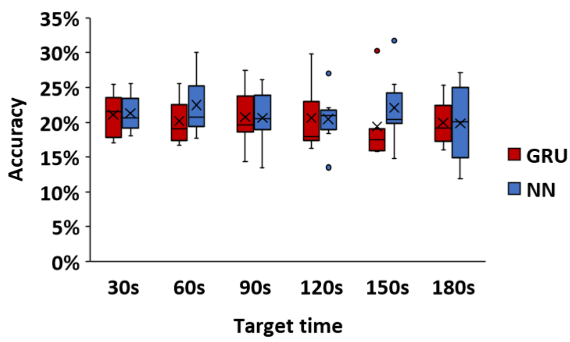

Fig.4 Distribution of prediction accuracy of models using GRU and NN

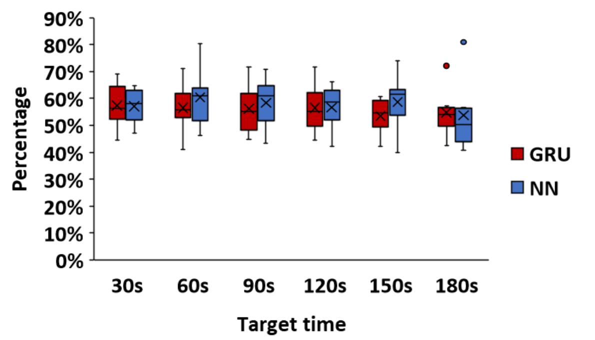

Fig.6 Distribution of percentage of the case that the deviation between predicted and actual values less than 1 for models using GRU and NN