国際学会発表報告書

１．氏名
道端　真也（ミチバタ　シンヤ）

２．国際会議名称
The 12th Workshop on Multimedia for Cooking and Eating Activities

３．開催日・開催場所
2020 年 10 月 26 日
アイルランド、ダブリン（オンライン開催）

４．発表内容
タイトル：
Cooking Activity Recognition in Egocentric Videos with a Hand Mask Image Branch in the Multi-stream CNN

概要：
Fine-grained activity recognition, especially cooking activity one, with egocentric videos is a hot topic and a challenging task in computer vision. To tackle this problem, many researchers have tried to leverage the information of cooking tools such as knife, peeler, etc., or that of equipment in the background. Although these are useful to improve the recognition performance on general cooking activity categories, the information does not provide sufficient evidences to recognize fine-grained cooking activities such as slicing, mincing, etc., because these belong same the general category and we often utilize the same type of tools. In addition, since the types of tools and equipment differs for each kitchen, a recognition model can over-fit to some specific environments in training data due to the over-confidence on such information. Therefore, a method having a high discriminating power of object classification and robustness for the environment difference is required. For the first step to realize such a method, in this research, we focus on the characteristics of egocentric video, i.e., capturing hands of camera

wearer without occlusions. Hand shape is useful to recognize the objects manipulated by camera wearer and sequential hand positions are also effective to analyze hand movement. By using these advantages, in this paper, we proposed a new multi-stream CNN, which has a mask image branch to leverage the hand shape and position information, in addition to the RGB and optical flow branches. From the empirical experiments for fine-grained cooking activity recognition in three types of kitchens, our proposed method outperformed the conventional methods and we confirmed that our proposed method has higher robustness for environmental difference compared with conventional methods.

５．会議参加の感想
自分と同じ食メディア分野の場で、英語での口頭発表を通じて日本人だけでなく海外での研究者との意見交換をすることができた。自身の研究に対して、食材画像ごとの行動認識判定の精度などといったこの分野特有の問題に対して、どのようなアプローチをしているか、どのような手法が有用であるかの情報共有をすることができた。また、同じ分野の様々な研究を知ることができ、今後の研究方針について視野を広げることができた。オンライン開催ということもあり、会議用の準備をする期間も長くなったが、そういった点を含めてとてもいい経験をさせていただいた。