Research Proposal

Real-time human action recognition is becoming increasingly crucial in numerous vision applications, such as human-robot interaction (HRI) systems, unmanned aerial vehicle (UAV) monitoring, autonomous driving, and urban security systems. Therefore, it is desirable and worthwhile to explore an intelligent action recognition algorithm for aerial robots.

Although deep learning has already been successfully applied in action recognition, related research towards real-time action recognition on UAV is still rare, because there are many challenges to tackle the real-time action recognition problem.

First, deep learning is a data-driven technique, so enormous available data is the primary condition to obtain superior performance. Second, aerial robots like UAV have limited on-board computational power and low capacity, it is infeasible to process a high-resolution image in such an embedded device, but low-quality images often lead to deterioration in recognition performance. Besides, practical applications such as autonomous systems and CCTV surveillance require not only real-time inference ability but also lightweight deep neural network, which puts too many constraints in network design. More importantly, it is hard to utilize the temporal information in large video file due to the limited storage space in the mobile platform.

In summary, we list the following research questions need to be addressed:

(1) How to design a lightweight deep neural network with real-time inference ability and acceptable recognition accuracy?
(2) How to achieve competitive action recognition accuracy with limited aerial training samples?
(3) How to efficiently integrate temporal information into the network under a resource-constrained environment?
(4) What other data or tasks can be used to benefit aerial action recognition? (i.e., pose estimation, semantic segmentation, optical flow)

Goal and Objectives:

The long term goal of the research is to develop a lightweight action recognition model for aerial robots based on deep learning. Specifically, the study has the following sub-objectives:

1. To develop a lightweight deep neural network that can be applied to embedded systems and mobile devices.
2. To design an efficient module that enables the model to have real-time inference ability, while maintaining competitive recognition accuracy.
3. To improve the recognition performance using a few aerial training samples, this involves the techniques of meta-learning (one-shot learning, few-shot learning), weakly supervised learning and semi-supervised learning.
4. To benefit from other related computer vision tasks. For example, multi-task learning can be utilized to improve the performance and enlarge the applicability.