（Form6）

論　文　要　旨
Thesis Abstract

（yyyy/mm/dd）　2023 年　06 月　21 日

※報告番号	第　　　　号	氏　名 (Name)	NGUYEN DUC THO

主論文題名　（Title）

Improving Generalization in Imitation Learning with Adversarial Learning

内容の要旨　（Abstract）

The demand for autonomous agents capable of mimicking human behaviors has grown
significantly in recent years. For example, self-driving vehicles, assistive robots, and
human-computer interaction fields rely on the ability of agents that can not only
make optimal decisions but also behave like humans, which can enable the agents'
actions to be believable and appear natural. In order for autonomous agents to
acquire such human complex behaviors, imitation learning, also known as learning
from demonstration, has been widely used for training autonomous agents in complex
environments. Unfortunately, traditional imitation learning algorithms only focus on
training an agent to perform a single task well in a given domain, which limits the
ability of these agents to generalize on future unseen situations. Thus, they are still far
from being comparable with humans due to the lack of the following abilities:

・Humans can recognize structural differences among the task's environments in
order to adapt their behaviors accordingly. On the other hand, the performance
of the learned agents tends to drop significantly even with a slight amount of
distribution shift between training and testing environments (i.e., domain shift).

・Humans do not learn a new task from scratch. Instead, humans often leverage
the knowledge learned from previous tasks to build up the necessary skills for
performing the new task.

The problem is formalized as domain and task adaptations in imitation learning,
which is the main focus of this dissertation. This thesis advances generalization in
imitation learning by presenting three novel agents.

・First, to overcome the domain shift problem, the DAIL–GAN agent is proposed.
The agent is trained using adversarial learning under a specific objective function

in order to learn both domain-shared and domain-specific features. These features enable the learned agent to perform the task effectively across different domains.

・Second, the TAIL–GAN agent is presented in order to address the task adaptation challenge in imitation learning by utilizing a similar adversarial learning process. Experimental results show that the agent can leverage the knowledge learned from a source task to accelerate the learning process of a new target task.

・Third, to further extend the potential of adversarial learning in tackling the generalization challenge in imitation learning, the DTAIL–GAN agent is introduced as a step toward a more domain- and task- generalizable. The training and adaptation processes leverage the idea of repetition learning in neuroscience in order to help the agent overcome the catastrophic forgetting problem and adapt their previously learned source task's knowledge to a new target task. Moreover, experimental results demonstrate the learned agent can provide a better generalization and consistently perform well on both source and target tasks across different domains.