論 文 の 内 容 の 要 旨

論文題目　Robust Object Detection with Deep Learning
by Utilization of Motion Information
and Generalization to Unknown Environment:
Application to Wide-Area Surveillance of Birds
（頑健な物体検出のための深層学習における
動きの利用と未知環境への汎化-野鳥の広域監視への応用-）

氏　　名　　吉橋　亮太

The advancement in automatic image recognition has paved the ways to diverse real-world applications including wildlife monitoring, automatic driving, robotics, or security. Such applications require further robustness of recognition systems than that of current ones. Currently, the most used approach in image recognition is deep learning, which enabled learning-based acquisition of feature representations. Although it has been shown to be efficient and effective in many generic image recognition tasks, their major successes concentrate in web-based applications, where large-scale and well-annotated data is easily available. Deep learning in image recognition can be characterized as a data-hungry method for acquiring rich visual features. This means that it works best in environments where we can take advantage of the existence of big data. However, this is not the case in most other real-work applications. There are some characteristics of real-work applications that make them different from web-image recognition. Here we refer three of them: 1) Domain specificity, 2) low-resolution targets, and 3) open world. For such situations, there is room for discussion whether and how we can enjoy deep learning's strength. This work studies how deep learning can be robustly applied to such real-world problems. Specifically, we introduce two ideas in deep-learning-based object detection. First, we utilize motion information to differentiate small objects with low visibility. In detecting small-looking birds in wide-area surveillance, such motion cues especially take essential roles. Second, we enhance the detector's generalization to unknown environments by introducing an unknown handling mechanism. Important existing work is the open-set classifier, which can safely reject unknown samples that the classifier did not learn. We extend this to be applicable to detection tasks. The main focus of this study is in wild-bird surveillance, which is a novel and practical application of computer vision. It is also at an opposite extreme of generic image recognition, in term of its nature of domain specificity, low-resolution targets, and open world. Here we introduced three ideas to overcome these difficulties. First, we introduced domain-specific datasets for bird surveillance that offer challenges of robustness due to the "in-the-wild" nature of the task and are

suitable to discuss the robustness with them. Second, we introduced motion-based object detection models that are more robust in detecting visibly small objects than detectors that rely only on appearances. Third, we introduce novelty-tolerant detectors that can handle 'unknown' objects that often appear in the open world.

The highlights of this thesis's contribution are summarized as follows:

- We provide the first practical image dataset for the task of bird recognition at wind farms (Chapter 3).
- We introduce the first deep-learning-based motion feature that is useful in detection tasks and improves detection accuracy by ~10% in pedestrian detection and bird detection (Chapter 4).
- We introduce a novel joint detection and tracking framework, named *Recurrent Correlation Network*, where detection and tracking help each other in terms of motion-feature learning (Chapter 5).
- We introduce *novelty-tolerant detection*, which handles 'unknown' objects by exploiting open-set classifiers.
- We further improve the existing open-set classifiers by developing *Classification-Reconstruction learning for Open-Set Recognition* (CROSR), a novel learning framework for open-set learning (Chapter 6).