

論文 / 著書情報
Article / Book Information

題目(和文)	複数物体の追跡と姿勢推定の融合による人の追跡と異常行動検知
Title(English)	Human Anomaly Detection and Tracking Using a Joint of Human Pose and Multi-Object Tracking
著者(和文)	NOUSotheany
Author(English)	Sotheany Nou
出典(和文)	学位:博士(工学), 学位授与機関:東京科学大学, 報告番号:甲第287号, 授与年月日:2025年3月26日, 学位の種別:課程博士, 審査員:小尾 高史,山口 雅浩,金子 寛彦,長谷川 晶一,船越 孝太郎
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Institute of Science Tokyo, Report number:甲第287号, Conferred date:2025/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

系・コース：	Information and Communications	系 コース	申請学位 (専攻分野)：	博士
Department of, Graduate major in	Engineering, Information and		Academic Degree Requested	Doctor of
	Communications Engineering			(Engineering)
学生氏名：	Nou Sotheany		審査員主査：	小尾高史
Student's Name			Chief Examiner	

要旨 (英文 800 語程度)

Thesis Summary (approx.800 English Words)

This dissertation entitled “Human Anomaly Detection and Tracking Using a Joint of Human Pose and Multi-Object Tracking” is divided into five chapters, summarized as follows: Chapter 1 “Introduction” presents the background of the study in this dissertation which focuses on improving the performance of human anomaly detection and tracking task. Moreover, this chapter introduces our proposed framework for enhancing the human anomaly detection and tracking model which includes the relationship of each module of the framework, such as Object Detection module, Multi-Object Tracking module, and Human Anomaly Detection module. Furthermore, the problem statements and the limitations from previous studies which are related to each module of this study are also described in this chapter. Additionally, we demonstrate the objective and novelty of our approaches in this chapter. Chapter 2 “Ground Truth Annotation Improvement for Human Detection” introduces the effectiveness of object detection model (detector) which is the core model in supporting the other tasks, such as Multi-Object Tracking, Person Search, Video Analysis, etc. Moreover, we discuss two prominent approaches for improving machine learning models including object detection models. Those are Model-Centric and Data-Centric approaches. In addition, we provide discussions in literature review and limitations of previous works which have been focusing on detector enhancement. The objective of our study in this chapter is to improve the detector through the concept of Data-Centric approach. Due to the limitations of previous works, they proposed enhancing models using a Model-Centric approach, but this resulted in only minor improvements. Our proposed methods focus Data-Centric approach to improve the quality of ground truth annotation in public datasets, such as MS-COCO and Open Image Dataset, in which missing annotation and incorrect size and position of ground truth bounding boxes are addressed. In order to add missing and correct the errors of bounding boxes annotation, we propose to train object detectors with the most notable and prominent dataset relative to current findings. Then, the trained models are used to update the ground truth annotation of target dataset by comparing them to the exiting ground truth annotations based on Intersection over Union (IoU) calculation. Additionally, the experiments and results of our proposed approaches are presented with the improved performance of object detection models which are trained and evaluated using our improved quality datasets. Chapter 3 “The Enhancement of Multi-Object Tracking Performance by Recovering Lost Track Objects” indicates the necessity of recovering lost tracking objects in Multi-Object Tracking (MOT) models. As MOT relies on Detector and Association algorithm for assigning unique identity to each detected object, the ability of detector effects the performance of tacking model. Consequently, lost tracking objects occur which we need to consider for the recovery, especially in the condition that lost tracking objects reappear in the successive frame after a period of loss. Thus, we propose adopting the Kalman Filter to predict the location of objects in the next frame for both losing track and being track objects with our proposed reversed procedure which benefits to recover lost tracking objects that reappear in following frame. Moreover, we propose to generate neighboring object information along with the forward process which is advantageous in the backward (reverse) process to recover the lost tracking objects. Additionally, our experiments and results in the prominent benchmark are presented with improved performance compared to baseline and state-of-the-art methods.

Chapter 4 “Human Anomaly Detection and Tracking Enhancement Using Pose-Based Feature” demonstrates the essential of human anomaly detection, which is the critical task, aimed at identifying abnormal activities of human in surveillance video. Since suspicious behavior is a critical concern which leads to insecurity, it is important to track every individual, including those exhibiting both normal and abnormal behaviors. Thus, we propose to track each human object using MOT, introduced in previous chapter, combining with human pose selection module to leverage the performance of human anomaly detection as well as to ensure tracking every single human object in the video frames. Our pose selection module chooses the most appropriate pose of each tracked human object which enhances the anomaly score calculation, leading to the improvement in anomaly detection model. Moreover, we experiment our approach using ShanghaiTech dataset, the prominent dataset in anomaly detection study. Based on the experiments and results, it indicates the improved performance of our approach against the baseline.

Chapter 5 “Conclusion” wraps up this study by summarizing, highlighting the limitations, and providing the direction of future works. As ShanghaiTech Dataset used in this study is still limited to simple anomalous behaviors, we consider implementing our approach with higher sensitivity anomalous behaviors in the future works. In the situation of high sensitivity anomalous behavior, more context information is necessary for precisely determining human abnormality, in which additional models are needed. Those models include Visual Relationship Detection for identifying suspicious objects, e.g. person holding guns, is considered.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

注意：論文要旨は、東京科学大学リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。

Attention: Thesis Summary will be published on Science Tokyo Research Repository Website (T2R2).