

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Learning-Based Pedestrian Detection on Visible and Thermal Images with Multiple Regressors Considering Misalignment
著者(和文)	WANCHAITANAWONGNapat
Author(English)	Napat Wanchaitanawong
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12852号, 授与年月日:2024年9月20日, 学位の種別:課程博士, 審査員:田中 正行,塚越 秀行,中臺 一博,原 精一郎,川上 玲,奥富 正敏
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12852号, Conferred date:2024/9/20, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	Napat WANCHAITANAWONG		
論文審査 審査員		氏名	職名		氏名	職名
	主査	田中 正行	教授		川上 玲	准教授
	審査員	塚越 秀行	教授	審査員	奥富 正敏	特任教授
		中臺 一博	教授			
		原 精一郎	准教授			

論文審査の要旨 (2000 字程度)

本論文は「Learning-Based Pedestrian Detection on Visible and Thermal Images with Multiple Regressors Considering Misalignment」と題して、全6章から構成されている。

第1章「Introduction」では、本研究の背景と目的について述べている。可視カメラを用いた歩行者検出は、コンピュータビジョン分野における重要な問題であることが指摘されている。特に、近年注目されている自動運転等の応用を考えた場合、夜間や逆光などにおいても頑健に歩行者を検出することが重要であると述べられている。一方、可視カメラだけでは夜間や逆光などの状況においては、情報を十分に取得することは困難であることも指摘されている。そのため、可視カメラだけではなく、遠赤外線カメラも同時に活用する歩行者検出の研究に注目が集まっている。このような複数の異なる特性のカメラを用いることは有効性、それらの情報を融合する手法が提案されていると述べている。しかしながら、複数カメラを用いる課題として、視点の異なりや不十分な時間同期により、同一対象物がずれてしまう問題を指摘している。このような観測される歩行者に位置ずれがある場合に対して、正しく評価できる評価指標を提案すること、ならびに、位置ずれに対して頑健であり、異なる特性の画像に基づく、歩行者検出のアルゴリズムを新しく提案することを研究目的とすると述べられている。

第2章「Existing Methods Review」では、ハンドクラフトな特徴を利用した古典的な歩行者検出手法ならびに、近年、急速に発展している深層学習に基づく歩行者検出手法を紹介している。さらに、本論文で活用するFaster RCNN手法およびSingle Shot MultiBox Detector (SSD)手法が調査され、さらにそれらの手法を可視カメラと遠赤外線カメラに拡張する手法に関しても調査されている。

第3章「Proposed Evaluation Metrics」では、まず、既存の評価基準であるIntersection over Union (IoU)と、IoUに基づくlog-averaged Miss Rate (MR)を、位置ずれのある対象に用いる際の問題点を指摘している。その問題とは、画像間で歩行者の観測位置がずれており、複数の観測画像があるにもかかわらず、ひとつの観測位置のみを用いて評価している点である。そこで、複数の観測画像に存在する位置ずれのある対象に対して評価することが可能な新しい評価指標として、Multi-modal IoU (IoU^m)を提案し、その評価指標に基づくMR^mを提案している。また、以降は、この提案された指標を用いて、歩行者検出手法を評価することが述べられている。

第4章「Proposed Multi-Modal Faster R-CNN Considering Misalignment」では、既存のFaster RCNNを拡張し、画像間の対象の位置ずれを考慮する手法を提案している。既存のFaster RCNNは、前景領域抽出ネットワーク(RPN)と抽出された領域を識別する識別器から構成されていると説明されている。提案手法では、RPNと識別器に対して、従来の手法では検出位置を調整する回帰ネットワークが1つだけ利用されていた。一方、提案手法では、可視画像用の回帰ネットワークと遠赤外線画像用の回帰ネットワークを、それぞれ学習することが特徴的であると述べられている。このような提案手法を、既存の手法と、実験的に比較を行い、提案手法の有効性が示されている。

第5章「Proposed Multi-Modal Single Shot MultiBox Detector Considering Misalignment」では、SSD手法に対して、複数の回帰ネットワークを活用する手法を適用している。まず、複数の観測

画像に対する SSD 手法として、Multi-Label Pedestrian Detection (MLPD) 手法が紹介されている。この MLPD 手法では、認識部においてひとつの回帰ネットワークが用いられている。提案手法では、可視画像用の回帰ネットワークと遠赤外線画像用の回帰ネットワークを、それぞれ学習する手法に拡張している。また、MLPD 手法は、可視画像および遠赤外線画像の検出結果に基づき損失関数が設定されているが、提案手法では可視画像と遠赤外線画像の検出ペアに基づく損失関数を新しく提案していると述べられている。さらに、可視画像と遠赤外線画像の間の位置ずれを人工的に生成するデータ拡張を提案し、学習に利用することで検出性能が向上すると述べられている。このような提案手法を、既存の手法と、比較し、実験的に提案手法の有効性を示している。

第 6 章「Discussion and Conclusion」では、新しい指標の提案、複数の回帰ネットワークを利用する有効性などの研究の成果と将来展望について述べられている。

以上のように、本論文は、可視画像と遠赤外線画像に基づく歩行者検出に関して、可視画像と遠赤外線画像の検出結果を統合して評価する新しい評価指標の提案、ならびに複数の回帰ネットワークを利用する手法を提案し、実験によりその有用性を示している。その成果は工学上・工業上貢献するところが大きい。よって、本論文は博士(工学)の学位論文として、十分な価値があると認められる。

注意：「論文審査の要旨及び審査員」は、東工大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。