Road Object Detection in Fish-Eye Cameras

RE6124019
Ming-Hsuan Wu
Institute of Data Science
RE6124019@gs.ncku.edu.tw

RE6124027
Wen-Hai Tseng
Institute of Data Science
RE6124027@gs.ncku.edu.tw

RE6121011
Jen-Lung Hsu
Institute of Data Science
RE6121011@gs.ncku.edu.tw

N96124153
Hsu-Chun-Che
Institute of Engineering Science
N96124153@gs.ncku.edu.tw

Abstract

This study aims to address the object recognition challenge in circular fisheye images, a task that presents significant difficulties in the field of computer vision. We participated in the AI City Workshop competition held at CVPR, specifically focusing on road target detection using fisheye cameras. We propose a multiscale and multiresolution method that combines geometric calibration, transformations, and deep learning models (such as the YOLO series) to tackle this problem. Our research evaluates the performance on the FishEye8K and FishEye1Keval datasets, demonstrating competitive object recognition capabilities in circular fisheye imagery.

1. Introduction

The surveillance of vehicular movement is a crucial aspect of modern urban transportation systems, providing essential data on traffic dynamics, incident identification, and facilitating effective traffic management. Traditional cameras, with their restricted fields of view (FoV), often fail to offer a comprehensive perspective of roadways and intersections, thereby necessitating the deployment of multiple cameras for adequate coverage. Fisheye lenses, capable of capturing wide-ranging, omnidirectional vistas with a single camera, have recently gained traction, offering an attractive alternative for traffic surveillance tasks [11].

Nonetheless, fisheye cameras pose their unique set of challenges: they generate distorted, curved imagery that requires specialized image processing techniques for rectification and dewarping [4,6]. This intricate task has hindered the widespread incorporation of fisheye cameras in traffic surveillance systems. Moreover, there is a dearth of publicly available datasets specifically tailored for fisheye road object detection, which hampers the development and evaluation of state-of-the-art computer vision algorithms in this domain.

To surmount these hurdles, we propose a comprehensive research project aimed at developing an efficient traffic surveillance system employing fisheye cameras and cutting-edge deep learning object detection methodologies. Our project will leverage the recently launched FishEye8K [5] and FishEye1Keval datasets, encompassing a diverse range of traffic scenarios, illumination conditions, and viewing angles of five road object categories at varying scales. These datasets will serve as a priceless resource for training and evaluating our proposed object detection models.

2. Related Work

2.1. Feature-Based Approaches

These methods typically utilize local feature descriptors such as SIFT [8], SURF [1], and ORB [10] for object detection and matching in images. However, in the case of fisheye imagery, traditional feature detection and descriptor matching may be affected due to distortion and deformation [3].

2.2. Deep Learning Methods

In recent years, deep learning has achieved significant success in object recognition tasks [9]. For circular fisheye images, convolutional neural networks (CNNs) can be employed for feature extraction and object recognition. Leveraging the powerful capabilities of CNNs, discriminative features can be learned directly from raw fisheye images.

2.3. Geometric Calibration and Transformations

The unique projection characteristics of fisheye images allow for geometric calibration and transformations to convert circular fisheye images into regular perspective projection images [4,6]. Following this transformation, traditional object recognition methods or deep learning approaches can be applied for object detection.

2.4. Multiscale and Multiresolution Methods

Features extracted from images at different scales and resolutions contribute to effective learning in deep neural network models [2, 7]. Therefore, adopting multi-scale and multi-resolution techniques can enhance target recognition accuracy and improve efficiency in fisheye image processing.

3. System framework

We anticipate propose a multiscale and multiresolution approach that combines geometric calibration and transformations with deep learning models (such as the YOLO series) to address the object recognition challenge in circular fisheye images. Our strategy involves training separate models for daytime and nighttime scenarios, followed by evaluation in experiments.

Firstly, we perform geometric calibration and transformations on fisheye images1, converting them into regular images, thereby reducing the distortion caused by fisheye pictures. Secondly, we employ models from the YOLO series for multiscale object recognition. By adjusting the model architecture and enhancing the images, we improve the model's detection capabilities across different scales and resolutions. Finally, we conduct specialized model training for different objects and scenarios. During the model embedding process, we integrate these specialized models to enhance the model's accuracy in object recognition.

Figure 1. Fisheye Images

4. Expected results

In this study, we participated in the CVPR competition and observed the dataset. We found a significant imbalance in the dataset, And anticipate that the method we proposed could effectively address this issue and accurately identify different road objects. Our recognition model tries to overcome the lighting differences between day and night. Even with fewer road images at night, our model may correctly identify objects. This achievement will help improve our ability to recognize road objects to the field of traffic monitoring systems. We hope that this research can contribute to the field of traffic monitoring systems, enhancing the ability to recognize road objects.

References

[1] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pages 404–417. Springer, 2006. 1

[2] Ping-Yang Chen, Ming-Ching Chang, Jun-Wei Hsieh, and Yong-Sheng Chen. Parallel residual bi-fusion feature pyramid network for accurate single-shot object detection. *IEEE transactions on Image Processing*, 30:9099–9111, 2021. 2

[3] Liuyuan Deng, Ming Yang, Yeqiang Qian, Chunxiang Wang, and Bing Wang. Cnn based semantic segmentation for urban traffic scenes using fisheye camera. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 231–236. IEEE, 2017. 1

[4] Gerardo Garcia-Gil and Juan M Ramirez. Fish-eye camera and image processing for commanding a solar tracker. *Heliyon*, 5(3), 2019. 1, 2

[5] Myoung Gochoo, Zhenxue Wang, and Seoung Wug Cho. Fisheye8k and fisheye1keval: Large-scale datasets for road object detection in fisheye cameras. *arXiv preprint arXiv:2303.12745*, 2023. 1

[6] Hyungtae Kim, Jaehoon Jung, and Joonki Paik. Fisheye lens camera based surveillance system for wide field of view monitoring. *Optik*, 127(14):5636–5646, 2016. 1, 2

[7] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017. 2

[8] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999. 1

[9] Ajeet Ram Pathak, Manjusha Pandey, and Siddharth Rautaray. Application of deep learning for object detection. *Procedia computer science*, 132:1706–1717, 2018. 1

[10] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee, 2011. 1

[11] Senthil Yogamani, Ciarán Hughes, Jonathan Horgan, Ganesh Sistu, Padraig Varley, Derek O'Dea, Michal Uricár, Stefan Milz, Martin Simon, Karl Amende, et al. Woodscape: A multi-task, multi-camera fisheye dataset for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9308–9318, 2019. 1