

DỰ ĐOÁN HƯỚNG CHÍNH XÁC TRONG PHÁT HIỆN ĐỐI TƯỢNG CÓ HƯỚNG SỬ DỤNG PHASE-SHIFTING CODER

Lê Minh Quang - 21522510

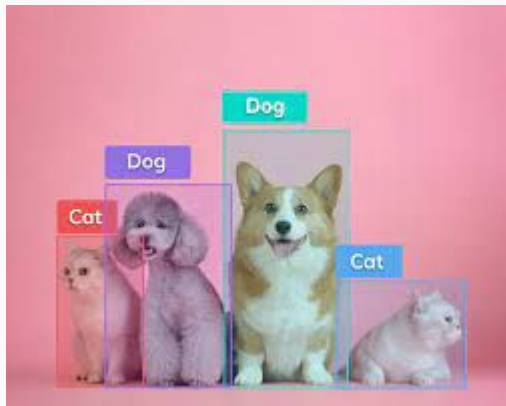
Tóm tắt

- Lớp: CS519.011
- Link Github của nhóm:
<https://github.com/leminhquang21522510/CS519.011>
- Link YouTube video:
<https://youtu.be/B1jJ05stxXs>



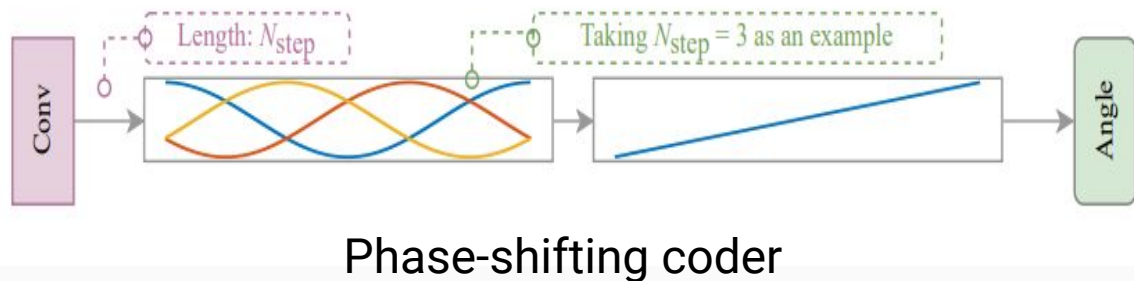
Lê Minh Quang

Giới thiệu



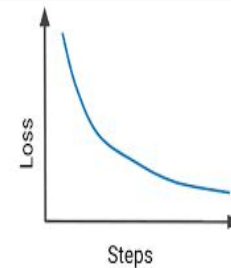
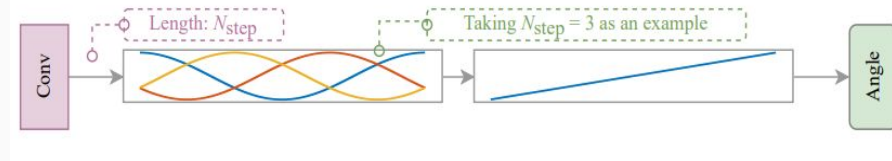
Input: Ảnh chụp thuộc một trong các lĩnh vực: ảnh hàng không, văn bản ngữ cảnh, kiểm tra công nghiệp...

Output: Ảnh input tương ứng kèm theo bounding box thích hợp nhất tương ứng với vật thể trong ảnh input.



Mục tiêu

- Nghiên cứu các phương pháp giải quyết bài toán phát hiện đối tượng có hướng trước đó và khảo sát bộ dữ liệu DOTA.
- Xây dựng bộ mã hóa chuyển pha (Phase-shifting coder) được gắn thêm vào đầu ra của mô hình.
- Đề xuất một hàm mất mát mới (Loss function) bởi vì đầu ra của mô hình đã bị thay đổi và phương pháp đánh giá.



Nội dung và Phương pháp

- **Nội dung và phương pháp 1: Nghiên cứu các phương pháp giải quyết bài toán phát hiện đối tượng có hướng trước đó và khảo sát bộ dữ liệu DOTA.**

Các phương pháp trước đó: Phát hiện đối tượng ngang; Sử dụng bộ quay bất biến; Kỹ thuật giải quyết biên và dạng hình vuông.

Khảo sát bộ dữ liệu DOTA: số lượng
mẫu ảnh, kích thước của mỗi ảnh, số
lớp,...

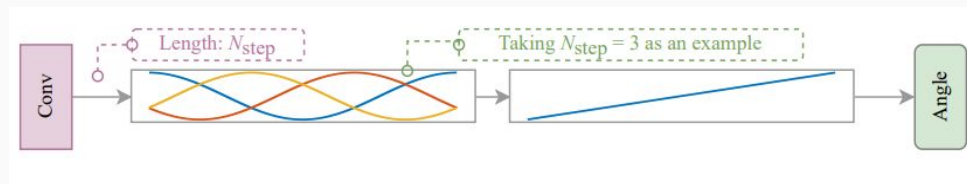


Nội dung và Phương pháp

- Nội dung và phương pháp 2: Xây dựng bộ mã hóa chuyển pha (Phase-shifting coder).

- Mapping
- Encoding
- Decoding

⇒ Xác định giá trị của góc.



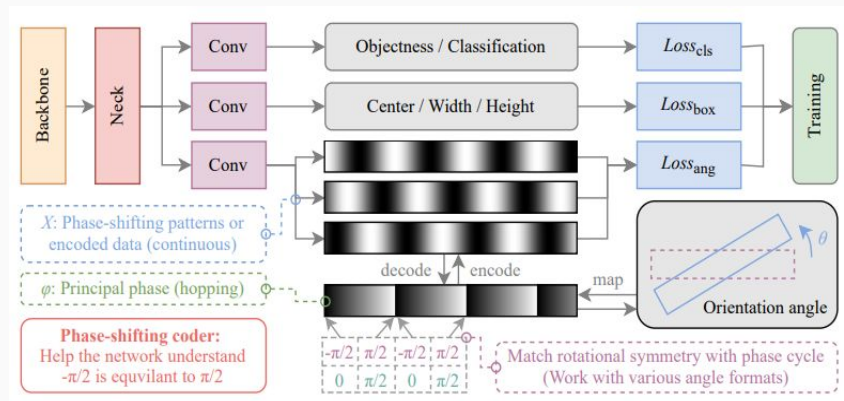
Nội dung và Phương pháp

- **Nội dung và phương pháp 3: Đề xuất một hàm mất mát mới (Loss function) và phương pháp đánh giá.**

- Hàm loss cho bài toán phân loại.
- Hàm loss cho bài toán hồi quy.
- Hàm loss riêng cho việc tính góc.

⇒ Tổng 3 hàm loss.

Đánh giá mô hình dựa vào mAP0.5,
mAP0.75.



Kết quả dự kiến

- Chúng tôi tiến hành khảo sát các thuật toán trên bộ dataset DOTA.
- Tạo bảng kết quả mAP0.5 và mAP0.75 của phương pháp được đề xuất ở trên được sử dụng so với các phương pháp giải quyết bài toán phát hiện đối tượng có hướng khác. Từ những kết quả thực nghiệm đó, chúng tôi tiến hành so sánh và đánh giá giữa các phương pháp. Qua kết quả đánh giá đó chúng tôi hy vọng phương pháp mà chúng tôi đề xuất có thể cải thiện từ 2-5% về độ chính xác.



Tài liệu tham khảo

- [1]. Yi Yu, Feipeng Da. Phase-Shifting Coder: Predicting Accurate Orientation in Oriented Object Detection. CVPR 2023: 13354-13363.
- [2]. Kun Fu, Zhonghan Chang, Yue Zhang, Guangluan Xu, Keshu Zhang, and Xian Sun. Rotation-aware and multi-scale convolutional neural network for object detection in remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing, 161:294–308, 2020.
- [3]. Minghui Liao, Zhen Zhu, Baoguang Shi, Gui-Song Xia, and Xiang Bai. Rotation-sensitive regression for oriented scene text detection. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5909–5918, 2018.
- [4]. Yuekai Liu, Hongli Gao, Liang Guo, Aoping Qin, Canyu Cai, and Zhichao You. A data-flow oriented deep ensemble learning method for real-time surface defect inspection. IEEE Transactions on Instrumentation and Measurement, 69(7):4681–4691, 2020.
- [5]. Xue Yang and Junchi Yan. Arbitrary-oriented object detection with circular smooth label. In Computer Vision – ECCV 2020, pages 677–694, 2020.

Tài liệu tham khảo

- [6]. Xue Yang, Junchi Yan, Ming Qi, Wentao Wang, Xiaopeng Zhang, and Tian Qi. Rethinking rotated object detection with gaussian wasserstein distance loss. In Proceedings of the 38th International Conference on Machine Learning, volume 139, pages 11830–11841, 2021.
- [7]. Chao Zuo, Lei Huang, Minliang Zhang, Qian Chen, and Anand Asundi. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. Optics and Lasers in Engineering, 85:84–103, 2016.
- [8]. Chao Zuo, Shijie Feng, Lei Huang, Tianyang Tao, Wei Yin, and Qian Chen. Phase shifting algorithms for fringe projection profilometry: A review. Optics and Lasers in Engineering, 109:23–59, 2018.
- [9]. Gui-Song Xia, Xiang Bai, Jian Ding, Zhen Zhu, Serge Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, and Liangpei Zhang. Dota: A large-scale dataset for object detection in aerial images. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3974–3983, 2018.