

PHƯƠNG PHÁP LOẠI BỎ ẢNH NHIỄU NHẪM TĂNG HIỆU QUẢ NHẬN DẠNG PHONG CẢNH

Dương Anh Kiệt^{1,2}, SeshuRaju Pentakota³, Cao Xuân Nam²

¹Chương trình Cử nhân Tài Năng

²Khoa Công nghệ Thông tin

Trường Đại học Khoa học Tự Nhiên, ĐHQG-HCM

³Dolcera, California, Hoa Kỳ

18120046@student.hcmus.edu.vn, seshurajup@gmail.com, cxnam@fit.hcmus.edu.vn

Tóm tắt

Nhận dạng phong cảnh (landmark recognition) nhằm giúp con người có thể biết thêm thông tin về địa danh có trong ảnh hay hệ thống có thể sắp xếp tự động các ảnh phù hợp vào bộ sưu tập của họ tốt hơn. Bộ dữ liệu phong cảnh Google Landmarks Dataset v2 (GLDv2) [1] có hơn 81,000 đối tượng và dữ liệu mỗi đối tượng không nhiều đang là thử thách lớn cho những nhà nghiên cứu nhằm đạt độ chính xác và ổn định cao. Chúng tôi đề xuất giải pháp sử dụng mô hình VGG16 đã được huấn luyện trên tập dữ liệu Places365 và bổ sung thêm đặc trưng Haar-like nhằm xác định và loại bỏ những ảnh không phải là phong cảnh và tăng hiệu quả cho bước nhận dạng. Chúng tôi đã thực nghiệm trên mô hình Deep Local and Global (DELG) [2]. Nó giúp cải thiện mô hình DELG từ 0,4341 đến 0,4661 điểm Global Average Precision (GAP) trên tập dữ liệu GLDv2. Với kết quả này [3], chúng tôi đã giành giải đồng trong cuộc thi Google Landmark Recognition 2020.

Từ khóa: Nhận dạng phong cảnh, DELG, GAP, GLDv2

Tham khảo

- [1] T. a. A. A. a. C. B. a. S. J. Weyand, "Google Landmarks Dataset v2 - A Large-Scale Benchmark for Instance-Level Recognition and Retrieval," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [2] A. A. J. S. Bingyi Cao, "Unifying Deep Local and Global Features for Image Search," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020.
- [3] Duong, A. K., Nguyen, H. L., & Truong, T. T. (2022, November). Large Margin Cotangent Loss for Deep Similarity Learning. In *2022 International Conference on Advanced Computing and Analytics (ACOMPA)* (pp. 40-47). IEEE..

REMOVE NON-LANDMARK TO IMPROVE LANDMARK RECOGNITION

Duong Anh Kiet^{1,2}, SeshuRaju Pentakota³, Cao Xuan Nam²

¹Honors Program

²Faculty of Information and Technology
University of Science, VNU-HCM

³Dolcera, California, USA

18120046@student.hcmus.edu.vn, seshurajup@gmail.com, cxnam@fit.hcmus.edu.vn

Abstract

Landmark recognition can help people get more information about a place in a photo, or the system can better automatically align photos to their collection. The Google Landmarks Dataset v2 (GLDv2) [1] contains a more than 81,000 classes and the number of training examples per class may not be large is a big challenge for researchers to achieve high accuracy and stability. We propose a solution using the VGG16 model trained on the Places365 dataset and adding Haar-like features to detect and remove non-landmark images and increase efficiency for recognition. We have experimented on the model Deep Local and Global (DELG) [2] model. It helps improve the DELG model from 0.4341 to 0.4661 points Global Average Precision (GAP) on the GLDv2 dataset. With this result [3], we won a bronze prize in the Google Landmark Recognition 2020 competition.

Key words: Landmark Recognition, DELG, GAP, GLDv2

References

- [1] T. a. A. A. a. C. B. a. S. J. Weyand, "Google Landmarks Dataset v2 - A Large-Scale Benchmark for Instance-Level Recognition and Retrieval," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [2] A. A. J. S. Bingyi Cao, "Unifying Deep Local and Global Features for Image Search," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020.
- [3] Duong, A. K., Nguyen, H. L., & Truong, T. T. (2022, November). Large Margin Cotangent Loss for Deep Similarity Learning. In *2022 International Conference on Advanced Computing and Analytics (ACOMPA)* (pp. 40-47). IEEE..