

NHẬN DIỆN VĂN BẢN TRONG HÌNH ẢNH KẾT HỢP TỪ ĐIỂN

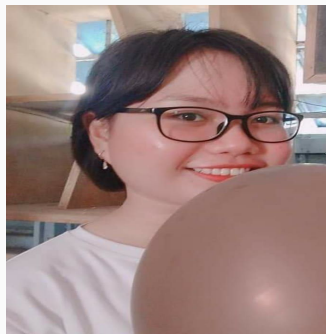
Nguyễn Lê Hoàng Hùng - 18520285

Đinh Hoàng Linh Đan - 19521309

Trần Nguyễn Quỳnh Anh - 19521217

Tóm tắt

- Link Github của nhóm:
https://github.com/18520285/CS519_M11
- Link YouTube video:
- Ảnh + Họ và Tên của các thành viên:



Trần Nguyễn Quỳnh Anh



Đinh Hoàng Linh Đan



Nguyễn Lê Hoàng Hùng

Giới thiệu

- Nhận dạng văn bản trong hình là một bài toán thực dụng.
- Thách thức cho bài toán rất lớn nên khó giải quyết hoặc kết quả không tốt.

=> Ứng dụng từ điển để cải tiến phương pháp giải quyết bài toán.



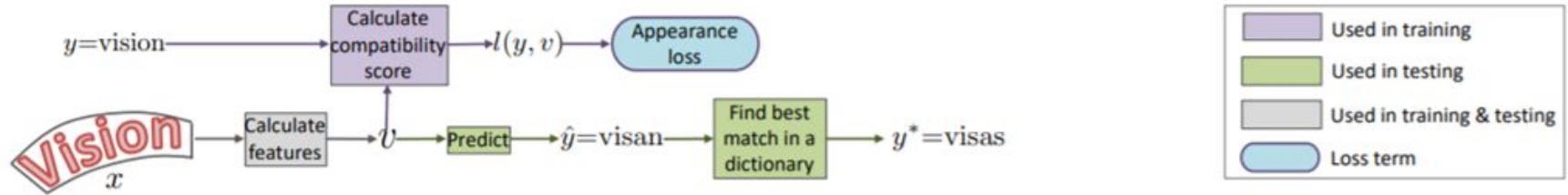
Mục tiêu

- Nghiên cứu phương pháp thông thường hiện có của bài toán Scene Text Recognition và cải thiện độ chính xác và tốc độ xử lý. Cụ thể chúng tôi nghiên cứu cách tiếp cận mới đó là Dictionary-guided Scene Text Recognition.
- Tạo ra 1 bộ dữ liệu tiếng Việt cho bài toán Scene Text Recognition.
- Xây dựng chương trình ứng dụng cho cách tiếp cận này.

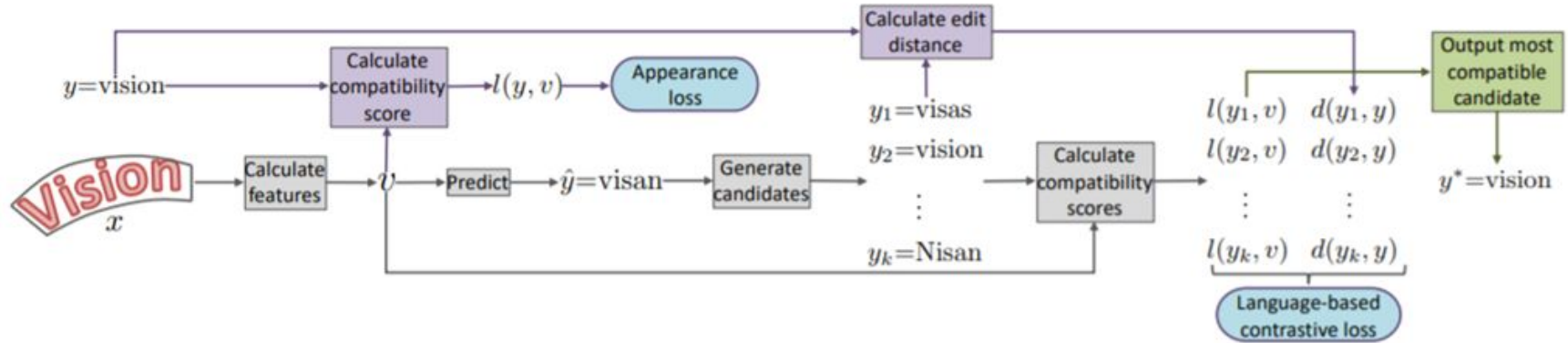
Nội dung

- Nghiên cứu quá trình giải quyết bài toán Scene Text Recognition thông thường.
- Đề xuất phương pháp mới cho bài toán Scene Text Recognition.
- Nghiên cứu và ứng dụng thuật toán ABCNet và MaskTextSpotterV3.
- Xây dựng bộ dữ liệu tiếng Việt - VinText.
- So sánh và đánh giá các mô hình ABCNet và MaskTextSpotterV3 được huấn luyện bởi các bộ dữ liệu như: TotalText, ICDAR2013, ICDAR2015 và VinText.
- Xây dựng chương trình ứng dụng cho cách tiếp cận cải tiến này.

Phương pháp



(a) The normal scene text recognition pipeline



(b) The proposed scene text recognition pipeline

Kết quả dự kiến

- Phương pháp cải tiến Dictionary-guided Scene Text Recognition.
- Bộ dữ liệu VinText.
- Chương trình ứng dụng thực tế.

Tài liệu tham khảo

- [1] Y. Liu, Hao Chen, Chunhua Shen, Tong He, Lian-Wen Jin, and L. Wang. ABCNet: Real-time scene text spotting with adaptive bezier-curve network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [2] Minghui Liao, Guan Pang, J. Huang, Tal Hassner, and X. Bai. Mask textspotter v3: Segmentation proposal network for robust scene text spotting. ArXiv, 2020.
- [3] Chee Kheng Chng and Chee Seng Chan. Total-text: A comprehensive dataset for scene text detection and recognition. In IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017. 2, 6
- [4] Dimosthenis Karatzas, F. Shafait, S. Uchida, M. Iwamura, L. G. I. Bigorda, Sergi Robles Mestre, J. M. Romeu, D. F. Mota, Jon Almazán, and L. D. L. Heras. Icdar 2013 robust reading competition. 2013 12th International Conference on Document Analysis and Recognition, pages 1484–1493, 2013. 2, 6
- [5] Dimosthenis Karatzas, L. G. I. Bigorda, A. Nicolaou, S. Ghosh, Andrew D. Bagdanov, M. Iwamura, Jiri Matas, Lukas Neumann, V. Chandrasekhar, S. Lu, F. Shafait, S. Uchida, and Ernest Valveny. ICDAR2015 competition on robust reading. In International Conference on Document Analysis and Recognition (ICDAR), 2015. 2
- [6] V. I. Levenshtein. Binary codes capable of correcting insertions and reversals. 1966. 1, 2, 3