Đánh giá hiệu quả của giải thuật Background Removing-LQN

Mục lục

[I. Tập dữ liệu được sử dụng 1](#_Toc150022393)

[II. Kết quả thực nghiệm 2](#_Toc150022394)

[III. Kết luận 5](#_Toc150022395)

# Tập dữ liệu được sử dụng

Tập dữ liệu bao gồm **19 ảnh chiếu chỉ** và **31 ảnh chế phong** thời Nguyễn. Dữ liệu được gán nhãn bounding box cho bài toán phát hiện văn bản và gán nhãn văn bản cho bài toán nhận diện.

A screenshot of a computer

Description automatically generated

Hình : Gán nhãn phát hiện và nhận diện văn bản

# Kết quả thực nghiệm

Sử dụng độ đo CLEval[[1]](#footnote-1) để đánh giá hiệu suất của mô hình OCR end-to-end. Kết quả được thể hiện trong bảng sau:

## Phát hiện văn bản

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Chế phong (31 ảnh) | | Chiếu phong (19 ảnh) | |
| Mô hình cơ sở(%) | Xóa nền(%) | Mô hình cơ sở(%) | Xóa nền(%) |
| Recall | 63.80 | 65.91 | 96.24 | 96.24 |
| Precision | 99.34 | 98.93 | 100 | 100 |
| F1-score | 77.69 | 79.11 | 98.08 | 98.08 |

## Nhận diện văn bản

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Chế phong (31 ảnh) | | Chiếu phong (19 ảnh) | |
| Mô hình cơ sở (%) | Xóa nền(%) | Mô hình cơ sở(%) | Xóa nền(%) |
| Recall | 32.58 | 33.07 | 63.85 | 64.14 |
| Precision | 58.57 | 61.84 | 74.06 | 75.89 |
| F1-score | 41.87 | 43.09 | 68.58 | 69.52 |

## Ảnh minh họa

A close-up of a paper

Description automatically generated

Hình : Minh hoạ hình ảnh đầu vào của baseline

A black background with writing on it

Description automatically generated

Hình : Minh họa hình ảnh đầu vào của BG\_RM\_LQN

A yellow rectangular object with writing on it

Description automatically generated

Hình : Ảnh chế phong

A black background with white text

Description automatically generated

Hình : Ảnh chế phong sau khi xóa nền

# Kết luận

1. Với tập dữ liệu chiếu chỉ (19 ảnh) và chế phong (31 ảnh) thì có vẻ việc áp dụng thuật toán xóa nền mang lại hiệu quả cho bài toán phát hiện và nhận diện văn bản chữ Hán.
   1. Điểm F1-score cho nhận diện văn bản tăng +0.9% từ 68.6% lên 69.5% trên tập chiếu chỉ và tăng +1.22% từ 41.87% lên 43.09%.
   2. Điểm F1-score cho phát hiện văn bản tăng từ +1,42% từ 77,69% lên 79,11% cho tập chế phong.
2. Mô hình phát hiện văn bản dự đoán khá tốt các bounding box từ ảnh. Cụ thể, F1-score lần lượt là 79,11% và 98,08 % cho ảnh chế phong và sắc phong sau khi đã xóa nền bằng thuật toán.
3. Kết quả mô hình nhận dạng quá thấp, chỉ đạt 43.09% và 69.52% lần lượt trên tập chế phong và sắc phong.
4. Có những ảnh của chế phong hoàn toàn không nhận diện được như ảnh:
   1. Chế phong số 1.
   2. Chế phong số 27.
   3. Chế phong số 28.

1. https://github.com/clovaai/CLEval [↑](#footnote-ref-1)