



Báo cáo kết quả nghiên cứu từ 28.10 – 4.11

**22C15033- Hồ Anh Khoa**

# Nội dung

1. Tìm độ đo phù hợp.
2. Đánh giá hiệu quả của thuật toán xóa nền.



Tìm độ đo phù hợp

# Phát hiện văn bản (Text detection)

1. IoU
2. DetEval
3. Tightness-aware IoU
4. TedEval

# Nhận diện văn bản Text recognition

1. CRW (Correctly Recognized Words): Cụm từ nhận được phải khớp hoàn toàn.
2. Edit Distance: Định lượng được sự tương đồng giữa 2 chuỗi thông qua 3 thao tác: Thêm, xóa, sửa.

# Đánh giá End-to-end

1. Iou+ CRW

2. PopEval: Sử dụng thông tin văn bản cho nhận diện. Không có phát hiện văn bản.

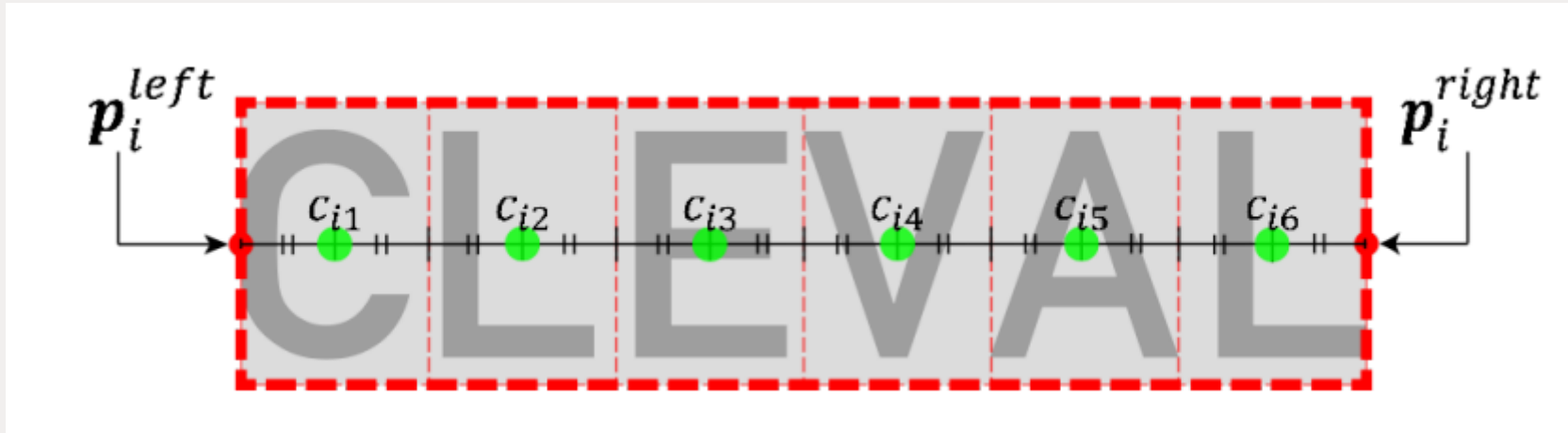
# CLEVal là gì?

- ❑ Đánh giá end-to-end hệ thống phát hiện và nhận diện văn bản.
- ❑ Khắc phục những hạn chế của các phương pháp đánh giá trước đó.
- ❑ **Code:** <https://github.com/clovaai/CLEval>
- ❑ **Paper:** <https://arxiv.org/pdf/2006.06244.pdf>

CLEVal tính như thế nào?

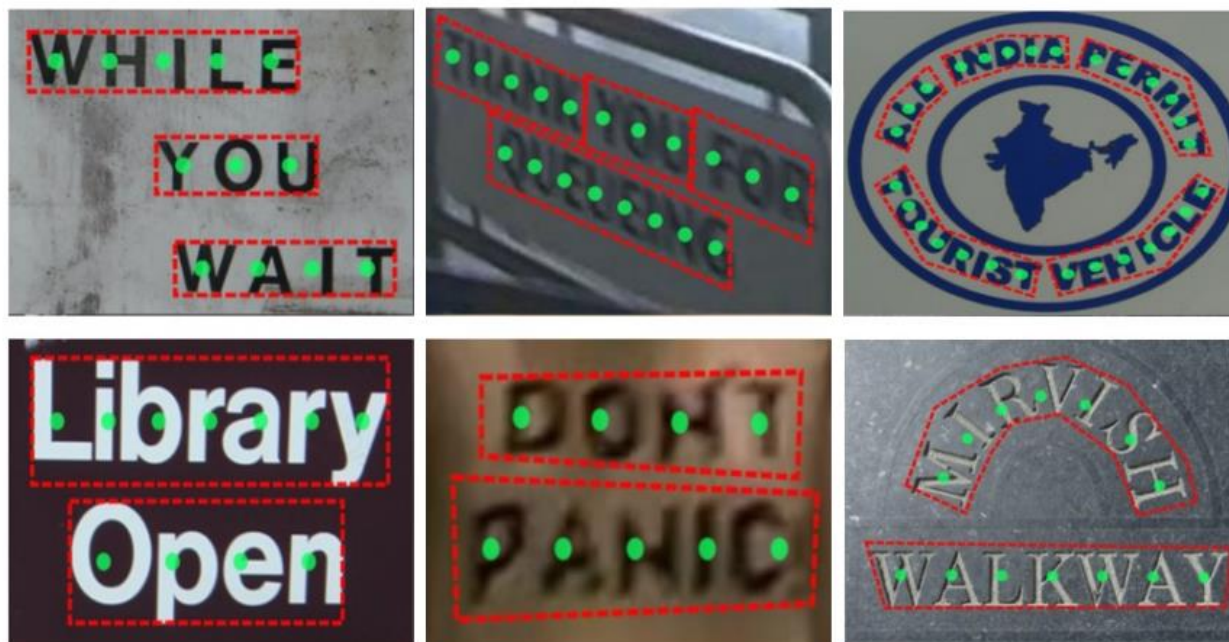


# Pseudo~Character Center (PCC)



$$c_i^k = \left( \frac{2k - 1}{2l_i^G} \right) p_i^{left} + \left( 1 - \frac{2k - 1}{2l_i^G} \right) p_i^{right}$$

# PPC trong dữ liệu ngoại cảnh



(a) ICDAR2013[8] (b) ICDAR2015[7] (c) TotalText[4]

# Nhãn (Ground Truth)

$G_1$  (RIVERSIDE)



$G_2$  (WALK)

Ground Truths

$G_1$

$G_2$

Correct

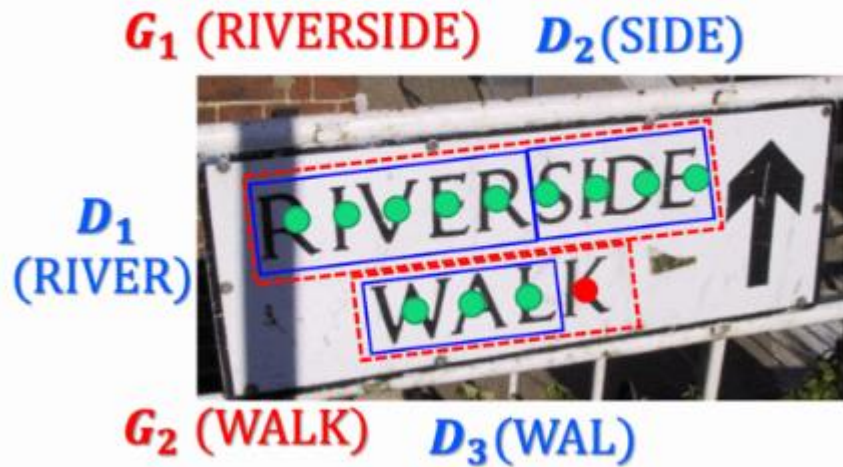
G. Penalty


# Dự đoán (prediction)



Predictions				
	$D_1$	$D_2$	$D_3$	Correct G. Penalty
$G_1$				
$G_2$				
Correct				
G. Penalty				

# Khớp giữa dự đoán và nhãn



Matching Policy (PCC & Area Precision)

	$D_1$	$D_2$	$D_3$	Correct	G. Penalty
$G_1$	5	4	0		
$G_2$	0	0	3		
Correct					
G. Penalty					

# Tính recall



Scoring Policy (Recall)

	$D_1$	$D_2$	$D_3$	Correct	G. Penalty
$G_1$	5	4	0	9/9	1
$G_2$	0	0	3	3/4	0
Correct					
G. Penalty					

G. Penalty viết tắt của : Granularity Penalty (Phạt độ chi tiết)



# Tính recall (tt)



Scoring Policy (Recall)

	<b><math>D_1</math></b>	<b><math>D_2</math></b>	<b><math>D_3</math></b>	<i>Correct</i>	<i>G. Penalty</i>
<b><math>G_1</math></b>	5	4	0	9/9	1
<b><math>G_2</math></b>	0	0	3	3/4	0
<i>Correct</i>				$\frac{9-1}{9} \oplus \frac{3}{4} = \frac{11}{13}$	
<i>G. Penalty</i>					

# Tính precision



Scoring Policy (Precision)

	<b><math>D_1</math></b>	<b><math>D_2</math></b>	<b><math>D_3</math></b>	<i>Correct</i>	<i>G. Penalty</i>
<b><math>G_1</math></b>	5	4	0	<b><math>9/9</math></b>	<b>1</b>
<b><math>G_2</math></b>	0	0	3	<b><math>3/4</math></b>	<b>0</b>
<i>Correct</i>	<b><math>5/5</math></b>	<b><math>4/4</math></b>	<b><math>3/3</math></b>	<b><math>\frac{9-1}{9} \oplus \frac{3}{4} = \frac{11}{13}</math></b>	
<i>G. Penalty</i>	0	0	0		



# Tính precision (tt)



Scoring Policy (Precision)

	$D_1$	$D_2$	$D_3$	Correct	G. Penalty
$G_1$	5	4	0	$9/9$	1
$G_2$	0	0	3	$3/4$	0
Correct	$5/5$	$4/4$	$3/3$	$\frac{9-1}{9} \oplus \frac{3}{4} = \frac{11}{13}$	
G. Penalty	0	0	0		
	$\frac{5}{5} \oplus \frac{4}{4} \oplus \frac{3}{3} = \frac{12}{12}$				

# H~mean = F1~score

$$H\text{-Mean} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$



Scoring Policy (Precision)

	$D_1$	$D_2$	$D_3$	Correct	G. Penalty
$G_1$	5	4	0	$\frac{9}{9}$	1
$G_2$	0	0	3	$\frac{3}{4}$	0
Correct	$\frac{5}{5}$	$\frac{4}{4}$	$\frac{3}{3}$	$\frac{9-1}{9} \oplus \frac{3}{4} = \frac{11}{13}$	
G. Penalty	0	0	0		

$\frac{5}{5} \oplus \frac{4}{4} \oplus \frac{3}{3} = \frac{12}{12}$

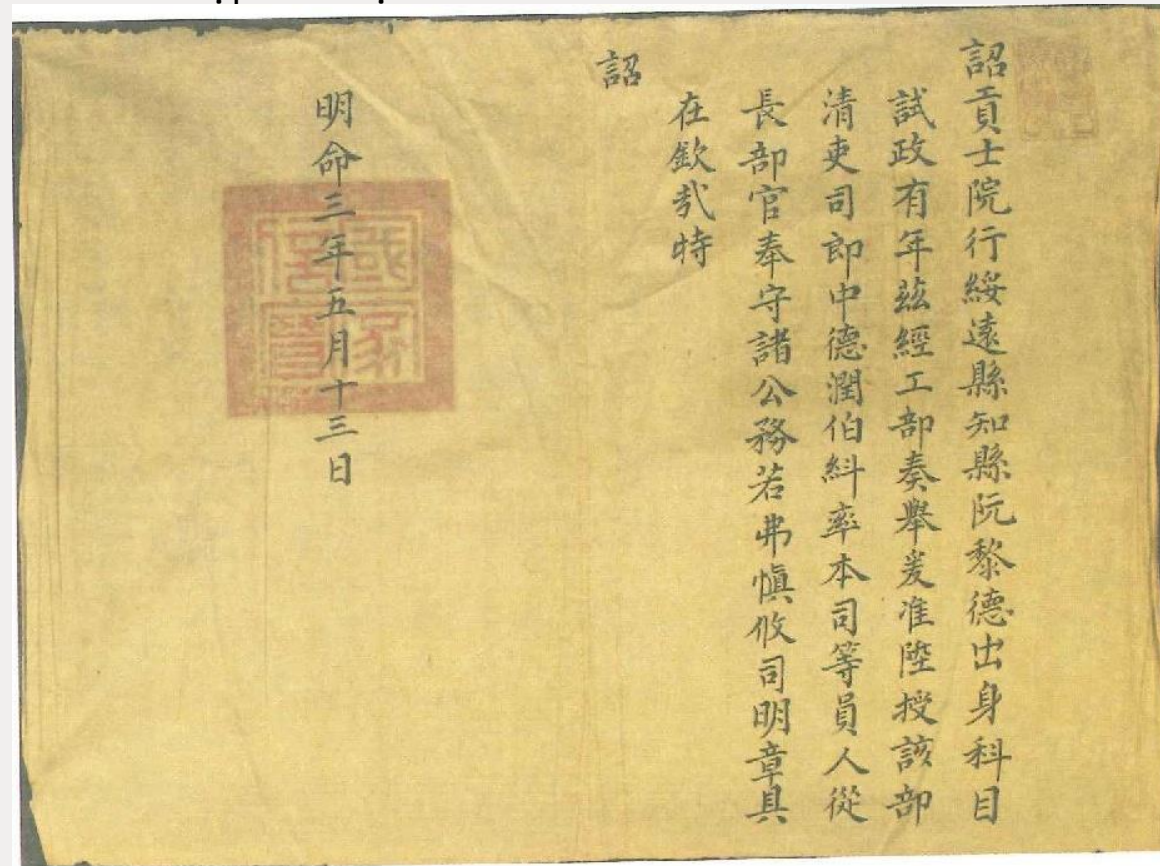
H-mean = 2

$$H\text{-mean} = 2 \times \frac{\frac{11}{13} \times \frac{12}{12}}{\frac{11}{13} + \frac{12}{12}} = \frac{11}{12}$$

# Đánh giá hiệu quả của thuật toán xóa nền

# Baseline

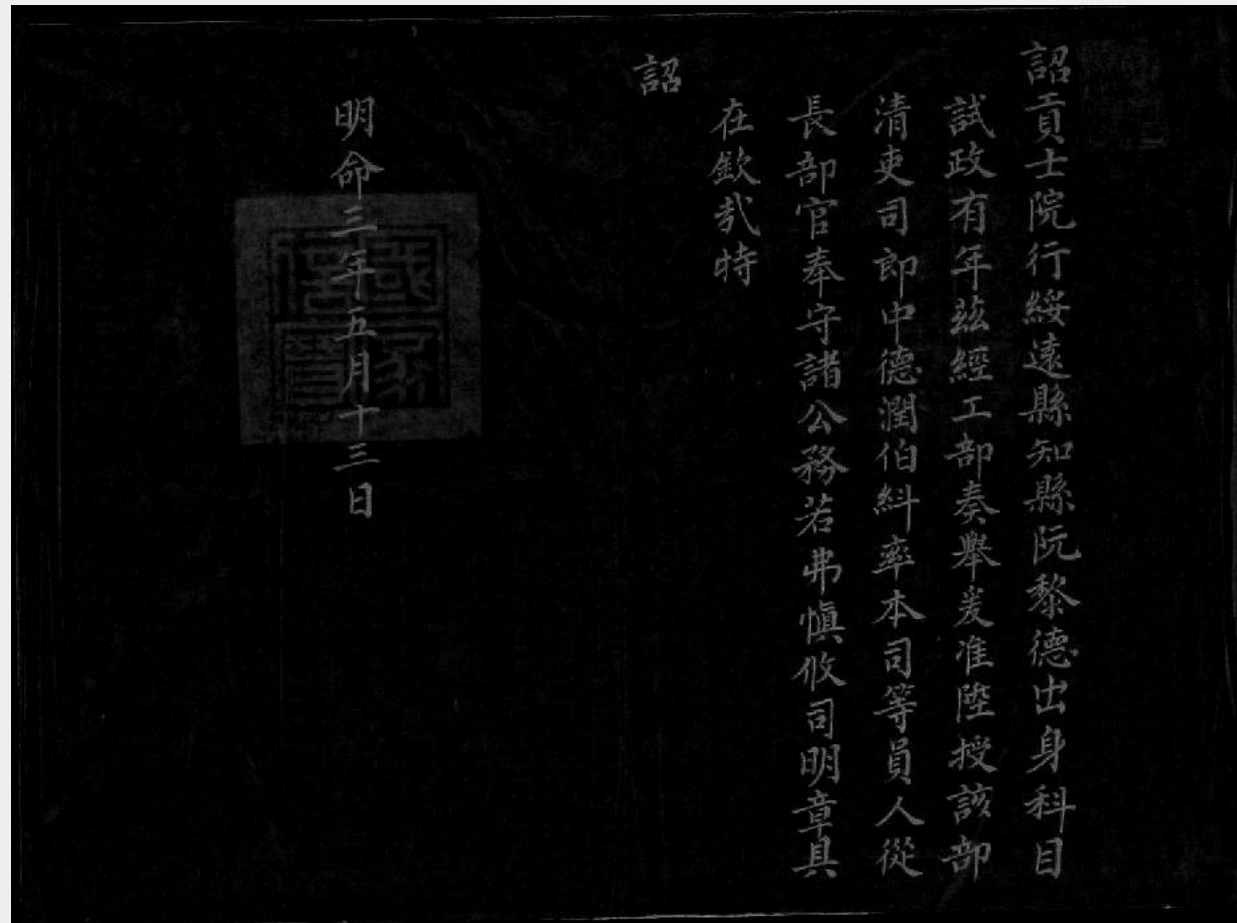
Xét tập dữ liệu chiếu có 19 hình ảnh.



# Kết quả baseline

PP	Số lượng ký tự GT	Số lượng ký tự nhận diện	Số lượng ký tự nhận sai (FP)	Recall	Precision	F1-score
Baseline	1411	1218	316	0.638	0.741	0.686

# XÓA NỀN LỖN

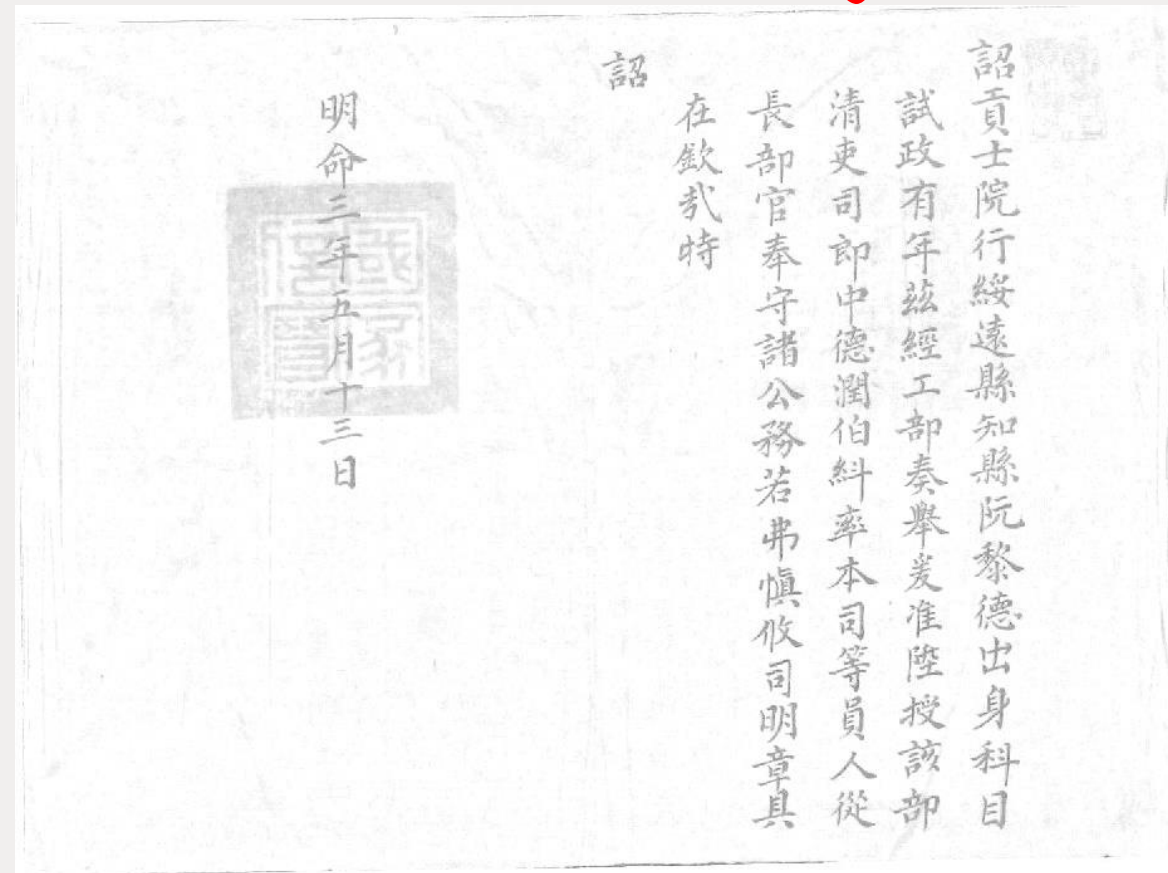


# XÓA NỀN LQN

PP	Số lượng ký tự GT	Số lượng ký tự nhận diện	Số lượng ký tự nhận sai (FP)	Recall	Precision	F1-score
Baseline	1411	1218	316	0.638	0.741	0.686
BG_RM_LQN	1411	1203	290	0.641	0.759	0.695



# XÓA NỀN LỖN





# XÓA NỀN LQN

PP	Số lượng ký tự GT	Số lượng ký tự nhận diện	Số lượng ký tự nhận sai (FP)	Recall	Precision	F1-score
Baseline	1411	1218	316	0.638	0.741	0.686
<b>BG_RM_LQN black</b>	<b>1411</b>	<b>1203</b>	<b>290</b>	<b>0.641</b>	<b>0.759</b>	<b>0.695</b>
BG_RM_LQN white	1411	1183	274	0.634	<b>0.768</b>	<b>0.695</b>

Hình như có tăng nè!!!

# Kết luận

- ❑ Xác định được độ đo phù hợp cho bài toán E2E OCR và STR.
- ❑ Thiết lập được code đánh giá.
- ❑ Trên dữ liệu chiếu chỉ (19 ảnh) thì thấy có hiệu quả cải thiện F1-score cho E2E OCR và STR.

## Kế hoạch tuần này

- Thêm ý tưởng cho phần blend text.
  - Chọn phong chữ phù hợp
  - Neural Style Transfer cho ký tự
- Hoàn thiện đánh giá hiệu quả của giải thuật xóa nền và các thử nghiệm khác.

Cám ơn thầy và các bạn đã theo dõi