

Politechnika Wrocławska

Wydział Informatyki i Telekomunikacji

Kierunek: _____ Informatyka techniczna (ITE)
Specjalność: _____ Systemy informatyki w medycynie (IMT)

PRACA DYPLOMOWA

Inżynierska

**Implementacja systemu optycznego rozpoznawania
znaków oraz opracowanie środowiska eksperymentalnego
opartego o samodzielnie skonstruowany zbiór danych**

Krzysztof Zalewa

Opiekun pracy
Dr inż., Paweł Zyblewski

Słowa kluczowe: 3-6 słów kluczowych

WROCŁAW (2025)

Streszczenie

Dodaj streszczenie pracy w języku polskim. Staraj się uwzględnić wymienione na stronie tytułowej słowa kluczowe. Uwaga przedstawiony rekomendowany szablon dotyczy pracy dyplomowej pisanej w języku angielskim. W przeciwnym wypadku, student powinien samodzielnie zmienić nazwy „Chapter” na „Rozdział” itp stosując odpowiednie pakiety systemu L^AT_EX oraz ustawienia w pliku *latex-settings.tex*.

Abstract

Streszczenie w języku angielskim.

Spis treści

1	Wstęp	1
2	Przegląd literatury	3
2.1	Narzędzia OCR	3
2.1.1	Tesseract	3
2.1.2	Easy OCR	3
2.1.3	DocTR OCR	3
2.1.4	Paddle OCR	3
2.2	Zbiory danych	3
2.2.1	IAM	3
2.2.2	FUNSD	3
2.3	Metryki	3
2.3.1	CER	3
2.3.2	WER	4
3	Wyniki eksperymentów	5
4	Podsumowanie	7

1. Wstep

2. Przegląd literatury

2.1. Narzędzia OCR

2.1.1. Tesseract

2.1.2. Easy OCR

2.1.3. DocTR OCR

2.1.4. Paddle OCR

2.2. Zbiory danych

2.2.1. IAM

2.2.2. FUNSD

2.3. Metryki

Do oceny wyników zastosowano dwie główne metryki. [CER](#) i [WER](#) zasadniczą różnicą między nimi jest celność porównań. CER porównuje na poziomie pojedynczych znaków natomiast, z kolei WER porównuje na poziomie poszczególnych słów. Z tego powodu CER jest przydatniejsze w kontekstach w których nawet pojedynczy znak może zmienić znaczenie słowa (np. w medycynie). Natomiast WER jest lepsze do porównywania spójności zdań itp.

2.3.1. CER

CER (Character Error Rate z ang. częstotliwość błędnych znaków) to metryka dzięki której możliwa jest ocena różnic między tekstem wytworzonym poprzez model OCR a tekstem rzeczywistym. W tym przypadku CER obliczane jest poprzez zsumowanie operacji(wstawień, usunięć oraz zamian znaków) potrzebnych do uzyskania tekstu rzeczywistego.

$$CER = \frac{S + D + I}{N_c}$$

Gdzie:

- S - Liczba zamian znaków (ang. Substitutions)
- D - Liczba usunięć znaków (ang. Deletions)
- I - Liczba wstawień znaków (ang. Inserts)

- N_c - Liczba znaków w tekście (ang. Number of characters)

Na przykład

Tekst oryginalny: Życiem wschód, śmiercią południe;

Tekst wygenerowany przez model: Życiem wschod, siercia poudniex;

Aby przekształcić tekst wygenerowany do tekstu oryginalnego należy wykonać 4 zamiany (Brakujące znaki polskie), 1 wstawienie (Brakujące 'm' w tekście wygenerowanym) oraz 1 usunięcie ('x' nie występuje w tekście oryginalnym). Więc $CER = 6/28 = 0.2141 \approx 21.4\%$

2.3.2. WER

WER (Word Error Rate z ang. częstotliwość błędnych słów) podobnie jak CER jest to metryka dzięki której możliwa jest ocena różnic między tekstem wytworzonym poprzez model OCR a tekstem rzeczywistym. Jak sama nazwa wskazuje WER porównuje tekst na poziomie poszczególnych słów.

$$WER = \frac{S + D + I}{N_w}$$

Gdzie:

- S - Liczba zamian słów (ang. Substitutions), czyli słowa które występują w obu tekstach ale te w tekście są różne od tych w tekście oryginalnym.
- D - Liczba usunięć słów (ang. Deletions), czyli słowa które występują w tekście oryginalnym jednakże nie ma ich w tekście wygenerowanym.
- I - Liczba wstawień słów (ang. Inserts), czyli słowa nadmiarowe których nie ma w tekście oryginalnym.
- N_w - Liczba słów w tekście (ang. Number of words)

Na przykład

Tekst oryginalny: Życiem wschód, śmiercią południe;

Tekst wygenerowany przez model: Życiem wschod, siercia poudniex;

Aby przekształcić tekst wygenerowany do tekstu oryginalnego należy wykonać 4 zamiany (Słowa zbliżone do oryginału ale nie takie same). Więc $WER = 4/4 = 1 = 100\%$

3. Wyniki eksperymentów

4. Podsumowanie

Bibliografia

Spis ilustracji

Spis tabel