

Magdalena Krampa 193195

Adam Białek 193677

Adam Chabraszewski 193373

SPRAWOZDANIE Z PROJEKTU Z PRZEDMIOTU SZTUCZNA INTELIGENCJA

TEMAT:

**“Detekcja oraz rozpoznawanie tablic
rejestracyjnych pojazdów ze zdjęć.”**

1. WSTĘP

Celem projektu było stworzenie aplikacji wykrywającej ze zdjęć tablice rejestracyjne oraz rozpoznającej zawarte na niej znaki. Chcieliśmy sprawdzić i porównać działanie różnych metod, m.in. modeli sieci neuronowych czy rozwiązań graficznych. Ponadto, z uwagi na zbiór danych treningowych z hiszpańskimi samochodami, chcieliśmy sprawdzić skuteczność modelu, wytrenowanego na hiszpańskich danych, wykorzystanego do rozpoznawanie polskich tablic rejestracyjnych.

2. IMPLEMENTACJA

Projekt został zaimplementowany w Pythonie z wykorzystaniem architektury sieci neuronowych YOLO oraz Tesseract OCR działającego na tradycyjnym przetwarzaniu obrazu (wersja Legacy).

3. DATA SET

Do trenowania modelu wykorzystany został ogólnodostępny zbiór danych z hiszpańskimi tablicami rejestracyjnymi.

Link do repozytorium: <https://github.com/ramajoballester/UC3M-LP>

4. ETAPY I ZASTOSOWANE METODY:

1) Detekcja tablic rejestracyjnych

W celu wykrywania tablic rejestracyjnych na zdjęciach wykorzystaliśmy model YOLOv8, który został poddany dodatkowym treningom. YOLO (You Only Look Once) to jedna z najpopularniejszych metod używanych do detekcji obiektów w obrazach i filmach. Model YOLOv8, dzięki swojej zaawansowanej architekturze, pozwala na szybkie i dokładne wykrywanie tablic rejestracyjnych. Kroki podjęte w ramach tego etapu były następujące:

a. Trenowanie Modelu YOLOv8

Wykorzystaliśmy pretrenowany model YOLOv8, który został dostrojony na specjalnie przygotowanym zestawie danych zawierających obrazy pojazdów i ich tablice rejestracyjnych. Zastosowaliśmy technikę transfer learning, aby dostosować model do naszych specyficznych potrzeb.

b. Wykrywanie Tablic Rejestracyjnych

Model YOLOv8 przetwarzał zdjęcia pojazdów, identyfikując i lokalizując tablice rejestracyjne. Model generował współrzędne prostokątów ograniczających, które wskazywały położenie tablic na obrazach.

c. Wydzielanie Tablic z Obrazów

Na podstawie wygenerowanych współrzędnych wycięliśmy fragmenty obrazu zawierające tablice rejestracyjne. Każdy z tych fragmentów został zapisany jako oddzielny obraz do dalszego przetwarzania.

2) Rozpoznawanie treści tablic rejestracyjnych

a. Pierwsza metoda: YOLOv8 z dodatkowym treningiem znaków

Do rozpoznawania znaków na tablicach rejestracyjnych również zastosowaliśmy model YOLOv8, który został poddany dodatkowym treningom na znakach alfanumerycznych. Proces ten obejmował następujące kroki:

i. Trenowanie Modelu na Znakach:

Wykorzystaliśmy pretrenowany model YOLOv8, który został dostrojony na zestawie danych zawierającym obrazy znaków alfanumerycznych występujących na tablicach rejestracyjnych. Dzięki temu model był w stanie rozpoznawać poszczególne znaki na tablicach.

ii. Rozpoznawanie Znaków:

Model YOLOv8 analizował wycięte fragmenty obrazów tablic rejestracyjnych, identyfikując i klasyfikując poszczególne znaki. Wynikiem tego procesu był tekst reprezentujący treść tablic rejestracyjnych.

b. Druga metoda: Tesseract OCR

Tesseract OCR to open-source'owy system do rozpoznawania tekstu, który jest szeroko stosowany w różnych zastosowaniach. W naszym projekcie wykorzystaliśmy Tesseract jako alternatywną metodę do rozpoznawania tekstu z tablic rejestracyjnych. Kroki zastosowane w ramach tej metody obejmowały:

i. Wstępne Przetwarzanie Obrazów

Przed przystąpieniem do rozpoznawania tekstu, obrazy tablic rejestracyjnych zostały poddane wstępnemu przetwarzaniu. Przeprowadziliśmy operacje takie jak skalowanie, konwersja do odcieni szarości oraz binaryzacja, aby poprawić jakość obrazu i ułatwić rozpoznawanie tekstu.

ii. Rozpoznawanie Tekstu za pomocą Tesseract

Tesseract OCR przeanalizował wstępnie przetworzone obrazy tablic rejestracyjnych, generując tekstowy output, który zawierał rozpoznane znaki alfanumeryczne.

iii. Walidacja i Poprawa Rozpoznanego Tekstu

W celu zwiększenia dokładności, wdrożyliśmy dodatkowe kroki walidacji. Wykorzystaliśmy reguły formatowania tablic rejestracyjnych (np. liczba i rodzaj znaków) oraz porównanie z bazą danych znanych tablic, aby wychwycić i poprawić ewentualne błędy rozpoznawania.

3) Ocena Modelu

Następnie po otrzymaniu rezultatu z przetwarzanego obrazu porównywaliśmy go z faktyczną wartością jaka znajdowała się na tablicy rejestracyjnej. Za wzór naszej oceny heurystycznej przyjęliśmy niżej podaną formułę.

$$Ocena = 1 - \frac{lev(L_f, L_r)}{len(L_r)}$$

Gdzie:

L_f - oznacza odczytaną ze zdjęcia tablice rejestracyjną

L_r - oznacza faktyczną wartość tablicy rejestracyjnej

$lev(x, y)$ - jest funkcją wykorzystującą metrykę odległości Levenshteina

$len(x)$ - jest funkcją zwracającą długość łańcucha znaków zadanej tablicy.

W celu oceny całego modelu, a nie tylko pojedynczych rezultatów, wyciągnęliśmy średnią ze wszystkich otrzymanych wyników.

5. REZULTATY

Rezultat oceny heurystycznej jednoznacznie pokazał, że nasza pierwsza metoda odczytywania tablicy rejestracyjnej (z wykorzystaniem YOLOv8) była lepsza. Ocena metody z wykorzystaniem YOLOv8 wyniosła 0.7999, a dla Tesseracta 0.3517. W pierwszym przypadku odczyt był dokładny dla 18% przypadków, a w drugim dla 0% przypadków.

```
Model evaluation: 0.7999669312169313  
Tesseract evaluation: 0.3517857142857146
```

6. CIEKAWE PRZYPADKI

Podczas tworzenia naszego projektu natrafialiśmy na kilka różnych przeszkód i niedogodności. Kilka z nich okazało się całkiem ciekawe, ale nie wszystkie udało się zdiagnozować. Jednym z nich jest problem z odczytywaniem E. Nasz model nie jest w stanie rozpoznać litery E, co jest spowodowane brakiem tej litery w zbiorze treningowym.



Kolejnym przypadkiem wynikającym z wadliwości zbioru treningowego jest odczytywanie cyfry 4. Jest to spowodowane różnicą w czcionce. Cyfra 4 dla polskich tablic rejestracyjnych jest zamknięta, kiedy na hiszpańskich tablicach jest otwarta.



Tak samo jak w przypadku 4, problem z rozpoznawaniem W wynika z różnicy w czcionkach obydwu państw, gdzie na polskich tablicach rejestracyjnych wyrzuczenie litery W nie jest dociągnięte do końca w przeciwieństwie do hiszpańskich tablic.



Jednak przypadek litery W jest trochę mniej problematyczny, gdyż w przypadku zdjęć bardzo dobrej jakości nasz model jest w stanie poprawnie wykryć literę W.



Ostatnim przypadkiem, którego do końca nie znamy przyczyny, jest zwracanie przez Tesseract znaków spoza kodu ASCII.

7. WNIOSKI

Porównując dwa sposoby odczytania wartości z tablicy rejestracyjnej, nasz model wypada zdecydowanie lepiej od Tesseracta. Jednak nie jest on idealny. W celu zwiększenia skuteczności moglibyśmy poszerzyć zbiór treningowy o polskie tablice rejestracyjne w celu rozwiązania problemów wynikających z różnicy czcionek. Projekt ma potencjał na rozpoznawanie tekstu znajdującego się na tablicach rejestracyjnych, z całej Unii Europejskiej przy odpowiednim doborze zbioru treningowego, w celu minimalizacji liczby błędów wynikających z minimalnych różnic czcionek.