

ANALIZA OBRAZÓW - DOKUMENTACJA PROJEKTU – „CZYTANIE TEKSTU”

Karolina Kotłowska, Kinga Miszczak, Anna Szumilas

2021/2022

1. Założenia projektu

Celem projektu było stworzenie aplikacji umożliwiającej rozpoznawanie cyfr rzymskich z wczytanego obrazka i ich konwersja na liczby arabskie. Program został napisany przy użyciu programu MATLAB w wersji R2020b.

2. Jak uruchomić projekt?

Należy uruchomić środowisko MATLAB (potrzebny jest zainstalowany pakiet Deep Learning Toolbox) oraz przejść do APPS→Install App, po czym wybrać plik instalacyjny app1.mlappinstall. Nastąpi instalacja aplikacji, którą można następnie uruchomić wchodząc w APPS→MY APPS→app1.

3. Działanie programu

Aby rozpoznać liczbę rzymską ze zdjęcia należy dodać zdjęcie do aplikacji za pomocą przycisku “Dodaj zdjęcie”. Program najpierw binaryzuje obraz od użytkownika a następnie za pomocą wytrenowanej przez nas sieci neuronowej rozpoznaje przedstawione na nim cyfry rzymskie, po czym w oknie “Output” pojawi się odczytana przez algorytm liczba rzymska. Następnie można tą liczbę przekonwertować na arabską za pomocą przycisku “Konwertuj”.

4. Interfejs użytkownika

Interfejs składa się z okna, w którym wyświetlane jest zdjęcie wgrane przez użytkownika, okna output oraz dwóch przycisków na wczytanie zdjęcia oraz konwersję na liczbę arabską.

5. Co nie działa?

Program ma trudności z odczytywaniem tekstu, kiedy tło na zdjęciu nie jest jednolite. Również przy zrobieniu zdjęcia tekstu na białej kartce, w słabo oświetlonym pomieszczeniu, przy wczytaniu zdjęcia do programu, pojawiły się komplikacje w „usuwaniu tła” – z powodu wyraźnego gradientu tła. Dla wyrafinowanych czcionek z ozdobnikami – litery D, X i V są bardzo trudno odczytywalne.

6. Możliwe usprawnienia

- Zastosowanie bardziej zaawansowanej metody „czyszczenia tła”.
- Poprawienie binaryzacji
- Zwiększenie datasetu