

# Analiza obrazów - projekt

## Rozpoznawanie tekstu

Marek Marchewka, Maciej Pieczonka, Michał Żoczek

25 stycznia 2021

## 1 Opis problemu

Celem projektu jest stworzenie programu służącego do rozpoznawania tekstu. Na wejście programu podawany będzie obraz zawierający tekst (w kolorze bądź nie), a na wyjściu zostanie wypisany tekst w postaci znaków. W celu rozpoznawania znaków sprawdzane będzie ich podobieństwo do znaków testowych.

## 2 Opis algorytmu

### 2.1 Binaryzacja

W programie zaimplementowano trzy rodzaje binaryzacji:

1. Ręczna: Binaryzacja jest sumą binaryzacji na wybranych warstwach z zadaną wartością graniczną.
2. Automatyczna: Binaryzacja jest wykonywana przez funkcję `iminarize` z domyślnymi parametrami.
3. Dynamiczna: Binaryzacja wykonywana przez funkcję `imbinarize` z zadaną wartością graniczną.

Dodatkowo, możliwe jest odwrócenie kolorów na obrazie. W przypadku pierwszych dwóch opcji następuje ono po binaryzacji, dla trzeciej - przed.

### 2.2 Segmentacja w projekcie

Isotnym elementem aplikacji jest segmentacja. Napisany skrypt uzyskuje odczytany tekst z obrazka z podziałem na linie, słowa i litery. Docelowo skrypt zwraca tablice, która zawiera pojedyncze litery, spacje lub znak nowej linii. Podział na linie zrealizowany jest przez sumowanie wartości danego wiersza macierzy obrazu po binaryzacji. Jeśli wartość wiersza jest większa od zera to znaczy, że jest to część linii. Następnie każda linia tekstu podlega analizie osobno. Podział na słowa, realizuje użycie dylatacji tak by znaki się połączyły i wyznaczyły słowo. Wartość dylatacji została określona doświadczalnie dla testowanych przykładów. W trakcie realizacji segmentacji pojawiły się problemy, np. kropka nad i (jak ją zachować), czy też problem kerningu.

## 3 Obsługa programu

W celu uzyskania wyniku w programie należy wykonać następujące kroki:

1. Wybór obrazu dla którego nastąpi rozpoznanie tekstu przy pomocy przycisku **Select Image**. Obraz powinien zostać wyświetlony w lewej górnej części aplikacji.
2. Wybór czcionki odpowiadającej tej widocznej na obrazie lub, w przypadku nieznanego czcionki wartości **Other**.
3. Dobór odpowiednich parametrów dla binaryzacji obrazu. Dla trybu **manual** wiąże się to z zaznaczeniem odpowiednich warstw koloru (R, G, B) i wyborem wartości granicznej przy pomocy suwaka **Threshold**. Dla trybu **dynamic** należy wybrać tylko dla wartości graniczną. Dla trybu **auto** nie trzeba nic wybierać. Zbinaryzowany obraz jest wyświetlony poniżej oryginalnego. Wynikowy obraz powinien zawierać białe litery na czarnym tle - jeżeli jest odwrotnie należy zaznaczyć opcję **Color Inversion**.

4. Przeprowadzenie operacji rozpoznania tekstu przy użyciu przycisku **Run OCR**. Wynikowy tekst zostanie wypisany w prawym panelu aplikacji.

## 4 Podział pracy

### Marek Marchewka

- Segmentacja tekstu
- Dokumentacja

### Maciej Pieczonka

- Realizacja skryptu odpowiedzialnego za rozpoznawanie znaków
- Przygotowanie zbioru danych
- Dokumentacja

### Michał Żoczek

- Stworzenie GUI
- Implementacja funkcjonalności aplikacji do GUI
- Binaryzacja
- Dokumentacja