

## **Założenia projektowe – „Geoguessr z dashcamów”**

### **Grupa 1, zespół 2**

---

## **1 Skład grupy**

- Aleksandra Śliwska – lider zespołu
- Glib Bersutskyi
- Marcin Kiżewski
- Arkadiusz Korzeniak
- Kamil Krzysztofek
- Patryk Madej
- Adam Niewczas
- Arkadiusz Rudy
- Wiktor Szewczyk

## **2 Założenia projektowe**

Głównym celem projektu jest opracowanie aplikacji, która będzie w stanie na podstawie materiału wideo zarejestrowanego z pokładu samochodu oszacować miejsce jego nagrania. Kluczowym elementem tego procesu jest analiza obrazu oraz wykorzystanie metod przetwarzania tekstu i uczenia maszynowego. System ten będzie przetwarzał obraz wideo klatka po klatce, identyfikując i analizując informacje wizualne, które mogą posłużyć do określenia lokalizacji.

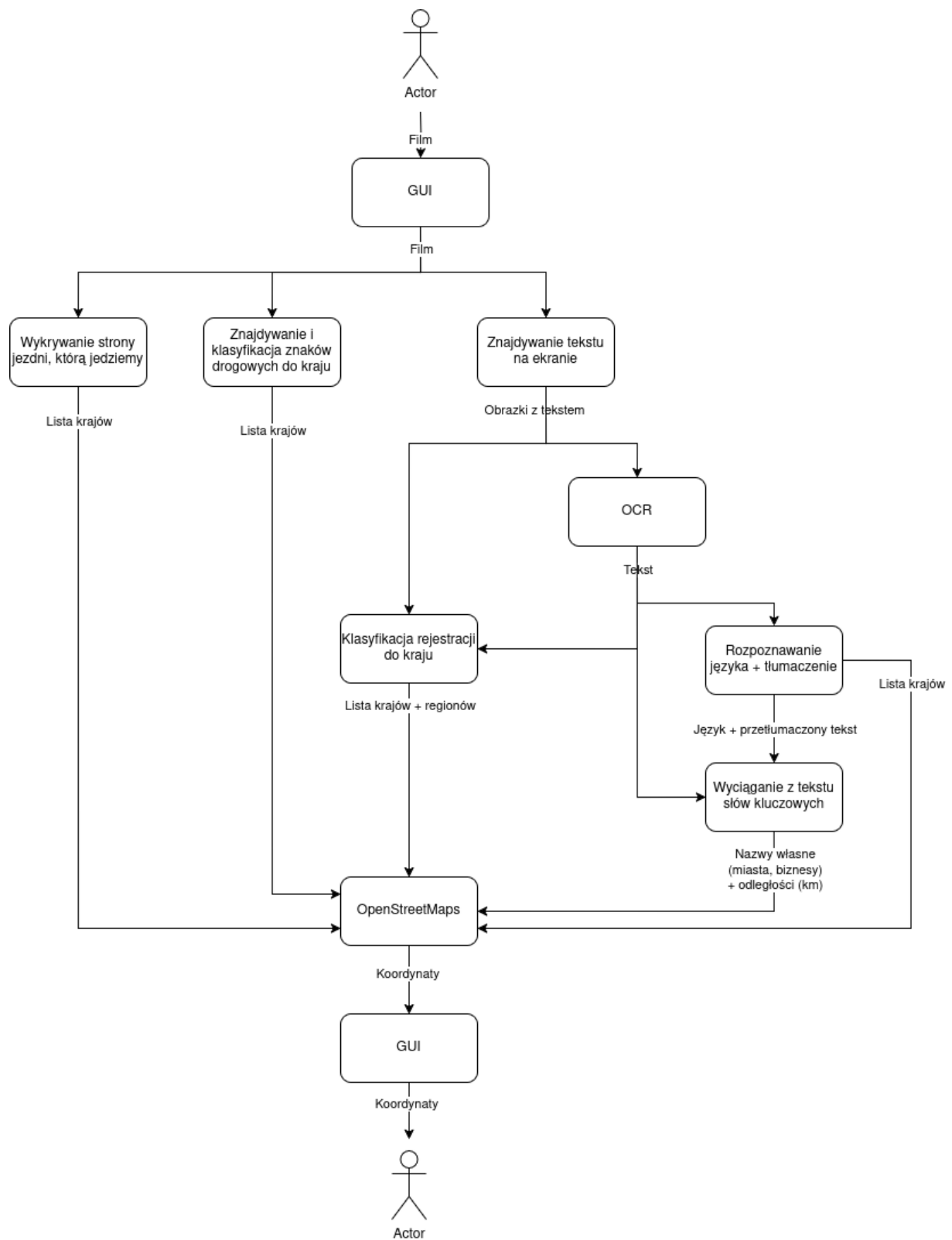
Szczególny nacisk położony został na rozpoznawanie tekstu zarejestrowanego w obrazie — zarówno w językach łacińskich, jak i innych alfabetach — obejmując litery, cyfry oraz znaki specjalne. System ma wykrywać tekst pojawiający się na znakach drogowych, tablicach informacyjnych, szyldach, tablicach rejestracyjnych czy billboardach. Istotnym zadaniem jest tu nie tylko samo wykrycie tekstu, ale również jego klasyfikacja oraz interpretacja pod kątem przydatności lokalizacyjnej.

Po zidentyfikowaniu tekstu następuje jego analiza semantyczna — ekstrakcja słów kluczowych, nazw geograficznych, produktów, walut i innych charakterystycznych informacji, które mogą posłużyć do oszacowania położenia geograficznego nagrania. Przykładowo, wykrycie nazwy miasta, waluty czy nazwy znanego miejsca pozwala zawęzić możliwą lokalizację do określonego kraju lub regionu. Aby zwiększyć dokładność identyfikacji lokalizacji, przewiduje się także analizę języka wykrytego tekstu oraz jego tłumaczenie, co pozwoli na dalsze wnioskowanie niezależnie od języka występującego na nagraniu.

Dodatkową warstwą analizy jest klasyfikacja wykrytych treści i obiektów. System ma rozpoznawać znaki drogowe charakterystyczne dla danego kraju (np. amerykański „Speed Limit” czy turecki „DUR”), a także analizować wygląd i układ tablic rejestracyjnych. Dzięki temu możliwe będzie zawężenie lokalizacji również na podstawie informacji wizualnych nieuwzględniających tekstu. W ramach projektu istotne jest również rozpoznawanie, po której stronie jezdni porusza się pojazd, co może również posłużyć jako wskazówka lokalizacyjna (np. ruch lewostronny charakterystyczny dla Wielkiej Brytanii czy Australii).

Całość danych z różnych zbiorów (istotne będzie zróżnicowanie krajów pochodzenia nagrań) ma być przetwarzana w sposób zintegrowany, a wynik końcowy w postaci przybliżonej lokalizacji ma być prezentowany użytkownikowi w przejrzystym graficznym interfejsie, z wykorzystaniem osadzonej mapy OpenStreetMap. Dane wejściowe w postaci nagrań zbierane będą z ogólnodostępnych zestawów materiałów wideo z kamer samochodowych. Uzupełnieniem systemu będzie mechanizm komunikacji z zewnętrzną bazą geograficzną, umożliwiającą powiązanie wykrytych informacji z rzeczywistymi lokalizacjami na mapie.

### 3 Schemat blokowy działania programu



## 4 Założenia technologiczne i osoby odpowiedzialne

Funkcjonalność	Wykorzystana technologia (biblioteki, źródła)	Odpowiedzialni
Zbiór danych	Źródła: <a href="#">1</a> , <a href="#">2</a> , <a href="#">3</a> ; wybrane zdjęcia z Google Street View	Zadanie zbiorowe
Detekcja tekstu	OpenCV	Marcin Kiżewski
OCR	OpenCV, TesseractOCR	Arkadiusz Rudy
Ekstrakcja słów-kluczy	SpaCy, RAKE, YAKE	Arkadiusz Korzeniak
Klasyfikacja tekstu	SpaCy, SciKit Learn	Patryk Madej Wiktor Szewczyk
Detekcja znaków drogowych	YOLOv5/v8 Roboflow OpenCV	
Klasyfikacja tablic rej.	OpenALPR	
Detekcja kierunku ruchu	YOLOv5/v8 OpenCV Ultra-Fast-Lane-Detection	Glib Bersutskyi
Rozpoznawanie języka	langdetect googletrans (fasttext, deep-translator)	Kamil Krzysztofek
Komunikacja z API OpenStreetMaps	Overpass API / Nominatim API	Aleksandra Śliwska
Interfejs użytkownika	Tkinter (Python) TkinterMapView	Adam Niewczas