Sztuczna Inteligencja

*Projekt*

**Skład zespołu:**

Łukasz Dudziak, Marta Żuławska, Anna Żurawik, Aleksander Smiatacz

**Wybrany temat:**

Wykrywanie i zliczanie wystąpień samochodów na zdjęciach ze skrzyżowań

kamer miejskich za pomocą sieci neuronowej.

***Sformułowanie problemów:***

Wybrany przez nas temat wiąże się z dużą liczbą problemów, które będzie trzeba rozwiązać w trakcie tworzenia projektu. Pierwszym z nich jest odpowiednie przetworzenie danych wejściowych – tutaj będzie trzeba zająć się oddzieleniem samochodów od otoczenia, żeby skupić się na obiektach, które nas interesują. Ponadto istotnym kłopotem jest nakładanie się pojazdów w związku z ich licznymi wystąpieniami na zdjęciach, jak i ich różne typy samochodów (osobowe, ciężarowe, z przyczepami itd.), co może negatywnie wpływać na podanie poprawnego wyniku. Problemem może być również kwestia perspektywy (przy metodzie sliding window), jak i oświetlenia szukanych przez nas środków transportu. Wiązać się to będzie z licznymi uproszczeniami w modelu wzorcowym tego projektu. Uproszczenia najprawdopodobniej będą się wiązać nie tylko z grafiką, ale i budową samej sieci neuronowej. Tutaj istotnym problemem będzie również dobranie odpowiedniej funkcji oceniającej (kryterium), dla wcześniej spreparowanych danych wejściowych, jak i przygotowanie odpowiednio dużego zbioru uczącego, aby można było zwrócić wiarygodne wyniki.

***Analiza problemów:***

Dokładna realizacja zadania zależy od przyjętej metody i ewentualnych uproszczeń związanych z ww. problemami i nie może być jednoznacznie określona na tak wczesnym etapie projektu. Kwestie związane ze wstępnym przetwarzaniem obrazu najprawdopodobniej wiązać się będą z wykorzystaniem biblioteki OpenCV. Dane wejściowe przypuszczalnie będą również związane z dużymi uproszczeniami, jak np. mniejsze zagęszczenie drogi przy odpowiednim oświetleniu. Jeśli chodzi o samą budowę sieci neuronowej, istnieje biblioteka FANN (ang. Fast Artificial Neural Network), która również może nam pomóc przy pracy nad rozwiązaniem postawionych sobie problemów. Ponadto istotną pomocą będą istniejące już rozwiązania, które są opisane poniżej.

***Analiza istniejących rozwiązań:***

# Multi-Net for Object Detection (MNOD):

Metoda została opisana tutaj: <http://www.academia.edu/1272586/Learning_Object_Detection_using_Multiple_Neural_Networks>

Metoda odnosi się do różnych obiektów, nie tylko samochodów. Multi-net to multisieć - sieć sieci neuronowych, z których każda jest konfigurowana i uczona osobno. Analiza tej metody pozwoliła stwierdzić, że jest ona lepsza w przypadkach zmian perspektywy oraz, gdy jakiś obiekt pojawia się na zdjęciach w różnym położeniu (od tyłu, od boku itp.). Informacje podawane na wejście każdej sieci pochodzą bezpośrednio z obrazu lub z wyjść inny sieci. Oczywiście, aby wytrenować sieć, musimy najpierw zrobić to z sieciami-dziećmi. W tej metodzie, gdy chcemy przeanalizować jeden obraz, poddajemy go różnym operacjom podkreślającym jego pewne istotne cechy, na przykład poprzez zmianę nasycenia, odcieni, jasności czy zastosowaniu różnych filtrów, przez co powstaje wiele różnych obrazów, tworzących później wejścia różnych (często kilku) węzłów multisieci. Rozmiar wejściowego obrazka jest ustalany taki sam dla każdej sieci. Obraz jest odpowiednio skalowany, wykorzystuje się też tzw. sliding window - okienko przesuwające się po obrazie. Zaletą MNOD jest to, że pozwala na podział problemu na podproblemy. Niestety ma też wadę - w przypadku dużej różnicy w skali obiektu nie działa zbyt dobrze.

# Vehicle Detection, Tracking and Counting:

Metoda została opisana tutaj:

<https://www.behance.net/gallery/Vehicle-Detection-Tracking-and-Counting/4057777>

[Andrews Sobral](https://www.behance.net/andrewssobral), autor tej metody zajął się również śledzeniem samochodów na filmach z kamerek, nas interesuje tylko wykrywanie i liczenie. W swoim projekcie wspomina on o dwóch metodach wykrywania obiektów. Pierwszą jest Haar Cascades a dokładniej Haar-like features z wykorzystaniem Cascade Classifier. Polega ona na znajdywaniu na obrazie kilku sąsiadujących, prostokątnych obszarów, których różnice sum jasności wszystkich pikseli spełniają pewne warunki. Metoda ta wymaga uczenia klasyfikatora. Andrews Sobel korzysta z gotowej implementacji klasyfikatora OpenCV. Drugim sposobem wykrywania obiektów jest Background Substraction. W tej metodzie potrzebne są sekwencje ramek filmu lub zdjęć. Są na nich wykonywane pewne operacje a później od każdej ramki jest odejmowane tło, wyznaczane na podstawie wszystkich ramek. Autor metody sam korzysta tu z biblioteki stworzonej przez siebie, ale w OpenCV również jest dostępne narzędzie odejmowania tła. Ten sposób nieco lepiej się sprawdza niż Haar Cascades, ale przy obu z nich pojawiają się problemy w przypadku dużej liczby samochodów obok siebie, np. gdy jest korek.  
  
Inne prace:

- <https://cseweb.ucsd.edu/classes/fa05/cse252c/dlagnekov_belongie.pdf> (jednak tu autor skupia się bardziej na rozpoznawaniu konkretnych marek aut)