WYDZIAŁ PODSTAWOWYCH PROBLEMÓW TECHNIKI

POLITECHNIKI WROCŁAWSKIEJ

**Kontrola i statystyka obiektów**

Projekt zespołowy

Autorzy:

Jakub Duda

Mikołaj Dukiel

Piotr Klepczyk

Mateusz Laskowski

Prowadzący:

dr inż. Łukasz Krzywiecki

WROCŁAW 2018

Spis treści

1. Wstęp do projektu
   1. Technologie i zakres projektu………………………………………………………………….. 2
   2. Cel projektu ……………….…………………………………………………………………... 2
   3. Rozwój projektu …………………..…………………………………………………………... 2
2. Projekt
   1. Komponenty programu .………………………………………………………………………. 3
   2. Funkcjonalności komponentów ………………………………………………………………. 3
      1. Obsługa kamery ……………………………………………………………………… 3
      2. Rozpoznawanie obiektów ……………………………………………………………. 3
      3. Analizator …………………………………………………………………………….. 4
   3. Diagram funkcjonalności ……………………………………………………………………… 4
   4. Diagram aktywności …………………………………………………………………………... 5

Wstęp

**Celem** **projektu** jest utworzenie programu do odczytu wideo, gdzie wideo wejściowe będzie analizowane i przetwarzane na dane statystyczne oraz wideo poglądowe, jako dane wyjściowe. Projekt będzie realizowany w małych krokach. Wpierw do danych wejściowych będziemy udostępniać obrazy, następnie nagrane wideo, a ostatnim, oraz najtrudniejszym celem będzie użytkowanie programu w czasie rzeczywistym. Oczywiście wideo będzie ograniczone jakościowo, jednak samo działanie powinno pozytywnie zaskoczyć. Celem oczywiście będzie również odpowiedni trening sieci neuronowych, które następnie pozwolą na rozpoznawanie cech wielu obiektów. Sieci neuronowe udostępnią nam dane, które my, twórcy będziemy chcieli odpowiednio opisać i wykorzystać do statystyki. Ostatnim celem będzie odtworzenie wideo po nałożonym filtrze, gdzie będzie można rozpoznać przypisane obiekty.

**Technologie i zakres projektu** Naszym wspólnym pomysłem jest stworzenie programu, który będzie wykorzystywał jedną z nowszych technologii w świecie informatyki. Chcemy przedstawić oraz użyć w swojej pracy technologię z wykorzystaniem sieci głębokiego uczenia się przy przetwarzaniu dużych ilości danych. Technika ta opiera się na specjalnych sieciach neuronowych. Jednym z zastosowań tej techniki to możliwość rozpoznawania różnych obiektów w naturalnych scenach naszego życia. Wygodnym rozwiązaniem w tej technologii jest to, że nauczanie sieci rozpoznawania obiektów nie polega na wpajaniu mu szczególnych cech obiektów, lecz przedstawiamy mu przykładowe obiekty oraz jego nazwy, a w efekcie otrzymujemy sieć, która jest w stanie rozpoznawać dość szybko wiele różnych obiektów. Ten proces nazywa się treningiem sieci, gdzie wygląda to jak zwykła nauka dziecka. Przedstawia mu się przykłady, sieć je przetwarza i sama wychwytuje ważne cechy obiektu. Trening jest czasochłonny, ale bardzo opłacalny w porównaniu z manualną nauką poszczególnych cech. Technologia głębokiej sieci, czyta cały obraz jaki jej się podaje do odczytu jako zbiór cech, które później w sieci neuronowej decyduje o istotnych cechach i odnajduje je w nim obiekty. Dzięki czemu można znaleźć wiele obiektów za jedną analizą obrazu. Znając już technologię, na której będziemy opierać cały swój projekt, chciałbym przedstawić jego zakres. Projekt będzie obejmował trzy wyróżniające się segmenty. Pierwszy to obróbka obrazu, a nawet udoskonalenie do możliwości odtwarzania obrazu live. Wideo live lub po produkcyjne wideo będzie danymi wejściowymi do programu. Problemów w tym segmencie jest naprawdę sporo. Jednym z nich to rozdzielenie wideo na poszczególne klatki. Jednak nie celujemy w wideo najlepszej jakości, ponieważ wcześniej wspomniana technologia radzi sobie naprawdę dobrze w średniej jakości wideo. Następny segment to sieć głębokiego uczenia, która umożliwi nam z danych wejściowych odczytać obiekty, które sieć neuronowa rozpoznaje. Ostatnim segmentem będzie analiza i statystyka otrzymanych wyników z segmentu drugiego. Chcemy również, aby danymi wychodzącymi było wideo, na którym będą zaznaczone rozpoznane obiekty przez sieć neuronową.

Rozwój projektu

Program jak najbardziej może wyliczać natężenie ruchu na skrzyżowaniach. Dzięki algorytmowi rozpoznawania obiektów będziemy w stanie określać różne obiekty na pliku wideo, a nawet wideo prosto z transmisji kamery. Projekt może być rozwijany na wiele sposobów. Jednym z nich, które my widzimy to dynamiczne zwiększanie i zmniejszanie przepustowości skrzyżowań w zależności od poziomu obiektów oczekujących na zjazd. Mając ten system na wielu skrzyżowaniach można przewidywać gdzie fala wielu samochodów następnie się pojawi i umożliwić tak zwaną „zieloną falę”, gdzie naprawdę pozwoli to na błyskawiczne usunięcie korków z głównych tras. Kolejnym rozwinięciem to wykrywanie pojazdów MPK, które miałyby priorytet. System odczytuje z kamery autobus MPK, przesyła dane o zwiększeniu priorytetu danego pasa do zaświecenia zielonego światła. Byłby to na pewno bardzo złożone działanie ale w informatyce nie ma rzeczy niemożliwych. Zapewne byłoby jeszcze wiele rozwiązań, których w tym momencie nie widzimy, ale to pokazuje jak dużą przyszłość ma ten projekt.

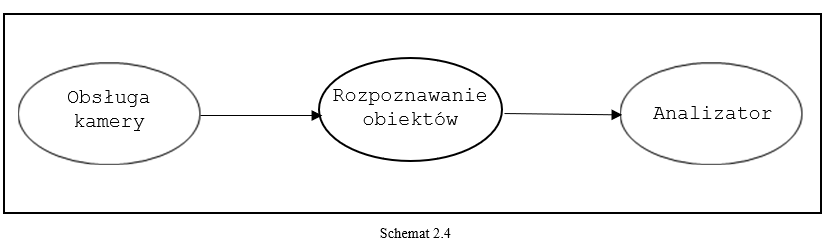
1. Analiza problemu

Opis zagadnień technicznych związanych z projektem:

Machine learning, sieć neuronowa, sieć nauczania, dostępne bibioteki i algorytmy spełniające funkcjonalność projektu, konwersja i obsługa wideo(kodowanie i kontener?, foramt)

* 1. Komponenty programu

Projekt składa się z trzech głównych komponentów. Te komponenty to: Obsługa kamery, Rozpoznawanie obiektów, Analizator. Poniżej przedstawiono graficzny opis zależności pomiędzy komponentami.



* 1. Diagram aktywności
  2. Diagram funkcjonalności
  3. Komponenty systemu

Projekt składa się z trzech głównych komponentów. Te komponenty to: Obsługa kamery, Rozpoznawanie obiektów, Analizator. Poniżej przedstawiono graficzny opis zależności pomiędzy komponentami.

**Komponent** Kamera - Program będzie przechwytywał obraz z kamery w stałych odstępach czasowych. Przechwycony obraz zostanie przetworzony, a następnie zostanie rozbity na pojedyncze klatki. Po tej operacji nastąpi Przekierowanie uzyskanych obrazów do następnej części programu. Ogólną funkcjonalnością tego komponentu jest przygotowanie obrazu w taki sposób by spełniał wymagania kolejnego etapu działania programu. Komponent ten będzie odpowiadał również za wyświetlanie obrazu po wykryciu obiektów w taki sposób że wykryty obiekt zostanie zaznaczony prostokątem z opisem co za typ obiektu został rozpoznany w tym miejscu.

* + 1. Rozpoznawanie obiektów/Sieć neuronowa

Ta część funkcjonalności wykorzystuje algorytm do identyfikacji obiektów z plików wideo lub z plików graficznych. Identyfikacja obiektów odbywa się na obrazie otrzymanym na wejściu. Po wykonaniu pracy zostanie zwrócone, ile obiektów zostało rozpoznanych, jakiego były one typu i z jakim procentem dopasowania zostały one rozpoznane, te dane zostaną przekazane do Analizatora. Będą dane przekazywane również z powrotem do komponentu odpowiedzialnego za obsługę wideo, dane te będą zawierały wielkość i umiejscowienie rozpoznanego obiektu oraz jego typ.

* + 1. Analizator

Program dostanie dane które będą wynikiem pracy algorytmu sieci neuronowych, zliczy ile obiektów zostało rozpoznanych z jednego typu. Na podstawie danych oraz ustawień własnych określi jakie działanie dla sygnalizacji świetlnej będzie najkorzystniejsze. Ta część komponentu odpowiada za generowanie wniosków na podstawie wcześniej opisanych danych. Do ich sformułowania będą potrzebne takie dane jak długość cyklu świateł, ilość pasów w danym kierunku i tym podobne. Ten komponent odpowiada również za prowadzenie statystyk na podstawie gromadzonych danych.

* 1. Diagram czynności

Poniżej przedstawiono diagram obejmujący funkcjonalność programu z podziałem na komponenty oraz z operacjami przeprowadzanymi w obrębie komponentu. (dodać start i koniec)

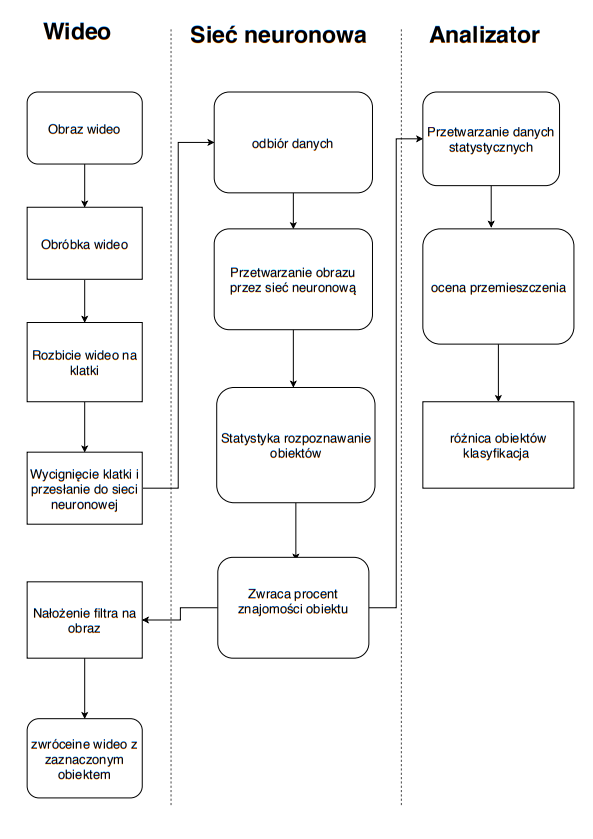


Diagram 2.3

* 1. Diagram aktywności

Poniżej przedstawiono diagram reprezentujący interakcje na linii użytkownik – program.

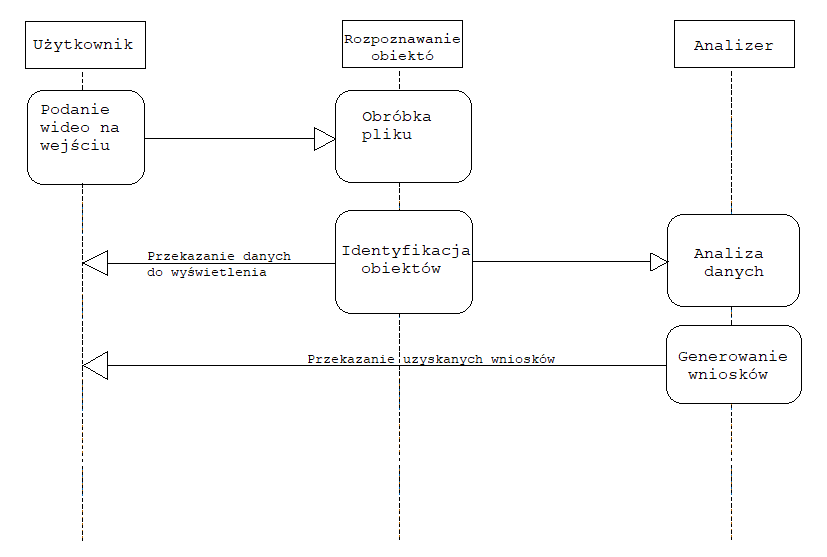


Diagram 2.4

1. Dokumentacja kodu