# Raport z wykonania prac badawczych

Poziom gotowości technicznej: 3

Opis prac objętych raportem: Stworzenie metodologii oceny dostępnych narzędzi i bibliotek pod kątem przydatności do stworzenia aplikacji przetwarzającej zebrane obrazy na dane ilościowe.

Raport z wykonania:

Metodologia oceny dostępnych narzędzi i bibliotek pod kątem przydatności do stworzenia aplikacji przetwarzającej zebrane obrazy na dane ilościowe.

Na potrzeby projektu przeanalizowano biblioteki do rozpoznawania i analizy obrazu pod kątem przydatności do analizy „kłaczków” będących cechą charakterystyczną poprawności procesu oczyszczania ścieków. Analizie poddano następujące biblioteki:

* PyTorch (fastai),
* Tensorflow (Keras),
* Sonnet,
* MxNet,
* CNTK.

Do dalszych prac wybrano bibliotekę Tensorflow (Keras). Zbudowano model sieci, w którym ostatnia warstwa sieci posiadała wstępnie dwa wyjścia. W ten sposób uzyskano klasyfikację binarną 0/1. Wykorzystano zbiór uczący 24 tysiące obrazów przedstawiających 2 rodzaje obrazów. Po pozytywnych wynikach testów powtórnie zmieniono model i przetrenowano jego ostatnią warstwę dla 8 klas celem uzyskania szerszej skali dla potrzeb oceny kłaczków. W projekcie jest planowana 5 stopniowa skala oceniania jakości kłaczków. Większość prac było prowadzonych w środowisku Python

W kolejnym kroku wykorzystano możliwość „transfer learningu” i wykorzystano modele MobileNet oraz ResNet, które zostały wytrenowane i udostępnione na licencjach możliwych do wykorzystania w projekcie na znacznie większych zbiorach uczących niż jesteśmy w stanie zrobić to w projekcie. Dodatkowo w każdym przypadku dodano ostanie warstwy sieci tak, by odpowiadały klasyfikacji w zależności od ilości różnych klas. Zaimplementowana została również możliwość wyboru, ile warstw wytrenowanego modelu ma możliwość być dotrenowana/modyfikowana podczas uczenia na docelowych danych.

W ramach Sprintu docelowo stworzono dwie aplikacje. Jedna umożliwia przeanalizowanie obrazu z oczyszczalni i zakwalifikowanie go do jednej z klas. Co ma odpowiadać przyszłej ocenie eksperta. Drugi służy do wytrenowania/douczenia modelu na całym secie zdjęć pozyskanych z oczyszczalni. Aplikacje mają na celu uproszczenie procesu uczenia sieci oraz weryfikację wyników jej działania. Obydwie aplikacje zostały stworzone w języku Python.

Ze względu na charakter zdjęć oraz fizyczną metodykę oceny przeprowadzaną przez ekspertów przyjęto, że w projekcie najlepiej sprawdzą się neuronowe sieci konwolucyjne (Convolutional Neural Network - CNN). Ze względu na strukturę budowy oraz sposób działania przypominający ludzkie uczenie się. A co za tym idzie pozyskiwanie wiedzy i nabywanie doświadczenia najbardziej przypominające obecną pracę ekspertów. Na podstawie analizy dostępnych bibliotek wybrano dwie MobileNet oraz Resnet, które zostały wstępnie wytrenowane przez autorów na wielu (kilkudziesięciu) warstwach. Każda z warstw jak ludzki mózg przechowuje informacje o rozpoznawanych kształtach, które finalnie składają się na rozpoznawany obraz. Założono, że kilka górnych warstw (ich ilość będzie podlegać badaniom) zostanie wytrenowane na podstawie pozyskanych w projekcie obrazów pochodzących z działających prototypów urządzeń w oczyszczalniach. Pozyskane zdjęcia podzielono na pięć kategorii w zależności od stopnia oczyszczenia ścieków wynikającego wprost z podanej chemii:

5 = Bardzo dobry

4 = Dobry

3 = Dostateczny

2 = Mierny

1 = Niedostateczny

Dodatkowo w ocenie używamy oceny 0 (zero) w przypadku, kiedy pozyskane zdjęcie nie kwalifikuje się do żadnej z pozostałych klas.

Pozyskane zdjęcia jak już wcześniej podano oceniane są przez eksperta, który nie wie jakie były proporcje podawanej chemii. Jako dane wejściowe do oceny bibliotek mamy więc dwie oceny. Pierwsza wynikająca z wiedzy o podanej chemii w skali od 1-5 oraz ocena eksperta w jaki sposób on ocenia to co widzi na zdjęciu w skali od 1 do 5 oraz jeśli odrzuca zdjęcie to 0.

Tak przygotowane próbki będą używane do trenowania sieci w wyżej wymienionych bibliotekach. Weryfikację przeprowadzamy zarówno kilku-kilkunastu-kilkudziesięciu ostatnich warstw jak również na wszystkich zdefiniowanych w modelu w celu optymalnego dobrania kontekstu czas analizy-uzyskane wyniki-ilość dostępnych zdjęć w poszczególnych klasach jak również wielkość zdjęcia wejściowego co znacząco wpływa na czas analizy i uczenia się sieci. A co za tym idzie dokonania wyboru właściwej biblioteki. W miarę pozyskiwania zdjęć testowych zakres będzie można poszerzać. Z każdej próbki zostaną zostawione przykładowe zdjęcia w celu weryfikacji jakości wytrenowania modelu w porównaniu do oceny eksperta.

Miarą jakości przydatności zastosowania biblioteki będzie wynik zakwalifikowania zdjęć testowych do poszczególnych klas ocen w porównaniu do oceny eksperta. Finalny wybór biblioteki będzie uwzględniał również czas potrzebny na analizę. Może się również okazać, że korzystne będzie ograniczenie skali ocen do 3 lub 4 klas.

# Ryzyka dla kolejnego poziomu gotowości

W toku prac badawczych zidentyfikowano następujące ryzyka na drodze do podniesienia gotowości do poziomu IV:

Brak ryzyk dla omawianego obszaru prac