

Dokumentacja projektu “Detekcja palenia”

Jakub Kiljański
3 rok informatyka

Wprowadzenie

Cel projektu

Cel projektu to stworzenie modelu zdolnego do rozpoznawania czynności palenia wyrobów tytoniowych na zdjęciach.

Temat projektu został zainspirowany prawem w Polsce, które zabrania palenia wyrobów tytoniowych w miejscach publicznych, a do tego nieprzeznaczonych. Palenie wyrobów tytoniowych szkodzi nie tylko palaczowi ale również przypadkowym osobom będącym w pobliżu dymu tytoniowego. Pomimo zakazu palenia w miejscach publicznych dochodzi do sytuacji w których ludzie nie przestrzegają prawa, często łamiąc je w miejscach o dużym zagęszczeniu ludzi, jak np. przystanki komunikacji miejskiej.

Współczesna technologie wykorzystujące machine learning są w stanie zaradzić na wyżej wspomniany problem. Publiczna przestrzeń może być obserwowana kamerami, które następnie mogą rejestrować niepożądane czynności. Do rejestracji takich właśnie niepożądanych czynności wytrenowałem model detekcji czynności palenia. Zastanawiałem się nad zastosowaniem klasycznej sieci neuronowej klasyfikującej zdjęcia np. na kategorie "palenie", "brak palenia" lub nad zastosowaniem metody detekcji obiektów.

W tym projekcie zastosowałem metodę detekcji obiektów, ze względu na to, że powinna dać lepszy efekt niż tradycyjna sieć neuronowa. Uważam tak, ponieważ miejsca publiczne znacząco się od siebie różnią. Nas natomiast w tym problemie interesuje jedynie czynność, która może być małym fragmentem zdjęcia. Dlatego zamiast brać pod uwagę całe zdjęcia jak w przypadku klasycznych sieci neuronowych model detekcji będzie szukał określonych fragmentów.

Wybrana technologia

Projekt został zrealizowany na platformie Google Colab w Pythonie w pliku typu Jupiter Notebook.

Wykorzystane zostały biblioteki:

- torch - otwartoźródłowa biblioteka programistyczna do maszynowego uczenia się
- detecto - architektura pretrenowana wykorzystana do zbudowania modelu
- numpy - biblioteka do obsługi dużych, wielowymiarowych tabel i macierzy
- matplotlib - biblioteka do tworzenia wykresów
- os - biblioteka do operacji wejścia wyjścia

Do tagowania fragmentów zdjęć skorzystałem z programu labeling.

Metoda

Zbiór danych

Do wytrenowania modelu skorzystałem z znalezionej na platformie Kaggle datasetu (<https://www.kaggle.com/datasets/vitamins/cigarette-smoker-detection>) zawierającego prawie 2 tysiące zdjęć ludzi w sytuacji palenia wyrobów tytoniowych. Spośród tych zdjęć

wybrałem około 100 zdjęć przedstawiających ludzi trzymających w dłoni wyrób tytoniowy blisko ust, pośród dymu tytoniowego. Następnie, wykorzystałem program labeling do oznaczenia granic interesującego fragmentu obrazków.

Parametry modeli ML

Przygotowane dane rozdzieliłem w proporcji 70/30 na zbiór treningowy i zbiór testowy. 70 zdjęć w zbiorze treningowym i 30 w zbiorze testowym.

Dane poddałem augmentacji:

- `resize = 900`
- `random horizontal flip = 0.5`
- `color jitter = 0.2`

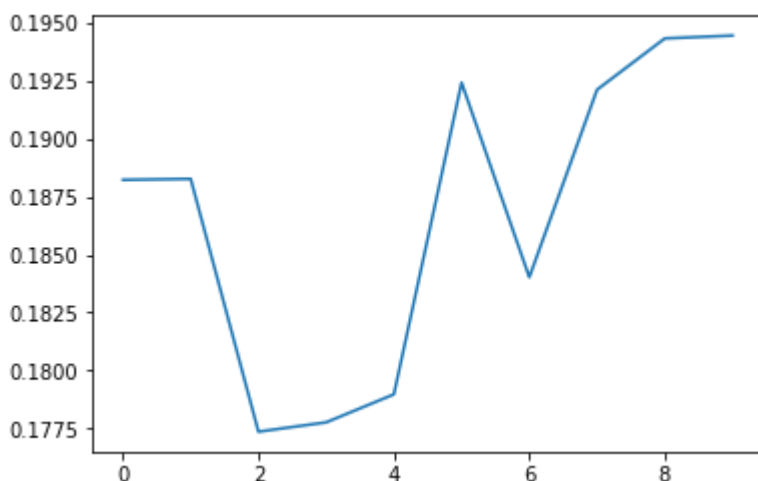
Do zbudowania modelu skorzystałem z pretrenowanej architektury Detecto.

Parametry uczenia modelu które udało mi się ustalić na drodze eksperymentów, to:

- `epochs = 30`
- `batch_size = 1`
- `lr_step_size = 5`
- `learning_rate = 0.001`
- `verbose = True`

Zwiększenie ilości epok zwiększa straty. Co pokazuje wykres poniżej.

Straty



Zwiększenie `batch_size` powodowało mniejszą precyzję wykrywania aktywności palenia na zdjęciach które testowałem.

Optymalny thresh to 0.6. Przy takiej wartości model najlepiej wykrywa aktywności palenia na testowanych przeze mnie zdjęciach.

Metryka zastosowana w modelu, `validation loss`, nie jest wiarygodną metryką oceny jakości modelu, jako że model wytrenowany na 3 epokach gorzej rozpoznawał czynności palenia na

testowanych przeze mnie zdjęciach niż model trenowany na 10 epokach o większej wartości straty.

Opis funkcjonalności

Podstawowa funkcjonalności modelu to wykrywanie na zdjęciu czynności palenia.

Zdjęcia do modelu zostały dobrane w taki sposób, aby ukazywały dłoń trzymającą papierosa, blisko ust, usta i obłok dymu. Model potrafi rozpoznawać fragmenty zdjęć, gdzie ludzie palą właśnie w takiej pozycji. Model gorzej sobie radzi w rozpoznawaniu zdjęć gdzie nie widać papierosa, nie ma dymu, papieros jest z dala od ust.

Zaletą modelu jest to, że powinien rozpoznawać ludzi jedynie w sytuacji palenia, a nie w sytuacji samego trzymania papierosa, czy w sytuacji gdy bucha im para z ust gdy jest zimno.

Wadą tego modelu jest to, że czasami rozpoznaje czynność palenia gdy na zdjęciu ręka znajdzie się blisko twarzy i to jeszcze w jakimś mglistym świetle.

Ten model mógłby znaleźć zastosowanie we wcześniej wspomnianej sytuacji potrzeby wykrywania czynności palenia w miejscach publicznych. Mógłby również zostać wykorzystany w celu monitorowania klientów na stacjach benzynowych czy parkingach.

Załączniki

kod - <https://github.com/KubKill/Smoking-Detection-ML-Model>

model - <https://drive.google.com/file/d/1--31GLuDR-xtec9i5TACfCRRg9kyP-3e/view?usp=sharing>