

## Zawartość folderu i ich funkcje:

1. *present\_data.ipynb* – notatnik Jupyter zawierający przegląd przykładowych dźwięków z poszczególnych klas, przy użyciu różnych typów reprezentacji plików audio. Większość rysunków z pracy, do rozdziału 3, pochodzi z tego pliku.
2. *get\_spectrograms\_for\_data.py* – program w języku Python, w którym dokonywana jest transformacja plików audio do postaci spektrogramów przeskalowanych w skali melowej.
3. *deform\_audio.py* - program w języku Python, w którym odbywa się deformacja plików audio ze zbioru. Tak jak było wspomniane w pracy – autorzy zbioru poddali każdy plik 20 deformacjom. Tutaj również zostało zastosowane takie podejście, a dopiero potem zostały wybrane pliki z poszczególnymi parametrami wymienionymi w pracy. W celu przeprowadzania deformacji potrzebne jest zaaplikowanie adnotacji JAMS plików ze zbioru podstawowego. Adnotacje te można pobrać z GitHuba autora zbioru danych:  
<https://github.com/justinsalamon/UrbanSound8K-JAMS>
4. *get\_spectrograms\_for\_aug\_data.py* – program w języku Python, analogiczny do programu wymienionego w punkcie 2., z tą różnicą, że tutaj transformacjom poddawane są syntetyczne dane z konkretnymi parametrami, wymienionymi w pracy.
5. *training\_first\_model.ipynb* – notatnik Jupyter, w którym przeprowadzane jest trenowanie modelu bazowego i w którym generowane są rysunki z nim powiązane.
6. *training\_second\_model.ipynb* – notatnik Jupyter, w którym przeprowadzane jest trenowanie modelu ulepszanego i w którym generowane są rysunki z nim powiązane.
7. *training\_augmentation.ipynb* - notatnik Jupyter, w którym przeprowadzane jest trenowanie modelu bazowego z syntetycznymi danymi i w którym generowane są rysunki z nim powiązane. Również w tym pliku dokonywane jest porównanie wszystkich trzech modeli pod kątem parametrów wymienionych w pracy.
8. *confusion\_matrix\_standard.npy* – plik w formacie .npy zawierający dane potrzebne do wygenerowania macierzy pomyłek dla modelu bazowego.
9. *confusion\_matrix\_better.npy* - plik w formacie .npy zawierający dane potrzebne do wygenerowania macierzy pomyłek dla modelu ulepszanego.
10. *confusion\_matrix\_standard\_aug.npy* - plik w formacie .npy zawierający dane potrzebne do wygenerowania macierzy pomyłek dla modelu bazowego z syntetycznymi danymi.
11. 4 pliki w formacie .wav, które używane są przy generowaniu syntetycznych danych. Jeden z nich, 150993\_\_sape\_\_street-scene-1.wav, został użyty w pracy.
12. UrbanSound8K.csv – plik CSV zawierający informacje o lokalizacji dźwięków, ich klasach oraz atrybutach. Plik ten jest dołączany przy pobieraniu zbioru danych.

## Uwagi ogólne:

1. Komentarze dotyczące poszczególnych fragmentów kodu zawierają się w programach i notatnikach.
2. Zważając na spory rozmiar samego zbioru danych niemożliwe było dołączenie go z resztą plików. Kolekcję można pobrać na oryginalnej stronie jej autorów:  
<https://urbansounddataset.weebly.com/urbansound8k.html>
3. W kodzie programów/notatników mogą pojawić się odnośniki do poszczególnych miejsc na komputerze/dysku. W przypadku chęci odtworzenia wyników, po pobraniu zbioru danych, należałoby zmienić te ścieżki i/lub umieścić zbiór na swoim Google Drive'ie.
4. Nawiązując do punktu 3., zbiór danych oraz poszczególne pliki w formacie .npy, zawierające zapis przeskalowanych spektrogramów w skali melowej dźwięków zostały umieszczone na moim dysku Google w celu uniknięcia potrzeby ładowania ich za każdym razem.
5. Istnieje szansa, że przy próbie uruchomienia niektórych programów w języku Python pojawi się potrzeba zainstalowania dodatkowych bibliotek. Szczególnie dotyczy to programu z deformacjami.

Piotr Przemielewski