

Streszczenie

Rozprawa prezentuje wyniki badań nad zastosowaniem sztucznych sieci neuronowych (konwolucyjno-rekurencyjnych) w analizie jakości sygnału audio, tj. automatycznej klasyfikacji wybranych zniekształceń w sygnałach muzycznych. Główną zaletą proponowanego rozwiązania, w porównaniu do innych dostępnych w literaturze dla tego typu sygnałów, jest brak konieczności porównania do tzw. sygnału referencyjnego, który podczas walidacji pozostaje często niedostępny. Znane aktualnie automatyczne metody oceny jakości dźwięku, nie wymagające takiego porównania, dotyczą głównie sygnałów mowy, podczas gdy analiza sygnałów muzycznych, ze względu na ich różnorodność, jest bardziej złożona.

Rozprawa zawiera opis badań przeprowadzonych przez autorkę, będących procesem opracowania prototypowego modelu automatycznej klasyfikacji wybranych zniekształceń sygnałów muzycznych. Na wstępie opisane zostały najważniejsze informacje i problemy związane z manualną i automatyczną analizą jakości dźwięku. Następnie przedstawiono etap projektowania, implementacji oraz ewaluacji modeli sieci neuronowych. Zawarto wyniki badań nad skutecznością zaimplementowanych modeli klasyfikacyjnych w zależności od zastosowanej architektury oraz wybranych danych wejściowych. Ostatnia część pracy zawiera omówienie uzyskanych wyników wraz z podsumowaniem.

Integralnym efektem pracy badawczej wykonanej w ramach niniejszej rozprawy jest także opracowanie własnej bazy danych, umożliwiającej ewaluację zaimplementowanego modelu sieci neuronowych, tj. zbioru sygnałów reprezentujących zniekształcenia kategorii wybranych na podstawie rekomendacji ITU BS.1284-2. Zbiór ten został utworzony na bazie niezniekształconych nagrań sygnałów muzycznych MUSDB18.

Głównym celem opracowywanej metody jest potencjalne usprawnienie kosztownych i czasochłonnych manualnych testów odsłuchowych, które ze względu na wysoką skuteczność, wciąż stanowią najpopularniejszą metodę oceny jakości dźwięku. Celem niniejszej pracy nie było natomiast ich zastąpienie – na ten moment nawet najlepsze metody referencyjne nie odzwierciedlają w pełni wyników subiektywnych testów odsłuchowych. Ograniczenie zakresu testowanego materiału metodą subiektywną poprzez wcześniejszą automatyczną detekcję i klasyfikację wybranych zniekształceń umożliwia przyspieszenie i zmniejszenie kosztów procesu manualnego testowania oraz potencjalnej naprawy błędów badanego toru przetwarzania audio.

Według najlepszej wiedzy autorki, nie została jeszcze opracowana rekomendowana metoda oceny jakości rzeczywistych sygnałów muzycznych w sposób automatyczny, bez porównania do referencji, a proponowany w niniejszej pracy model jest pierwszym opublikowanym zastosowaniem konwolucyjno-rekurencyjnych sieci neuronowych do zadania automatycznej klasyfikacji zniekształceń sygnałów muzycznych bez porównania do sygnału referencyjnego.