

Lublin, dnia 10 września 2023 r.

Prof. dr hab. inż. Oleksandra Hotra
Katedra Elektroniki i Technik Informatycznych
Wydział Elektrotechniki i Informatyki
ul. Nadbystrzycka 38 A
20-618 Lublin
e-mail: o.hotra@pollub.pl

Recenzja rozprawy doktorskiej
Pani mgr inż. Kamili Organiściak pt.:
„Automatyczna klasyfikacja wybranych zniekształceń sygnałów
muzycznych z wykorzystaniem konwolucyjno-rekurencyjnych sieci
neuronowych bez porównania do sygnału referencyjnego”

Ogólna charakterystyka rozprawy doktorskiej

Struktura rozprawy jest prawidłowa. Recenzowana rozprawa doktorska obejmuje 94 stron, zawiera 2 nienumerowane rozdziały (wprowadzenie oraz rozdział, zawierający tezę, cel i zakres pracy), 11 numerowanych rozdziałów, liczący 163 pozycji spisu źródeł literaturowych. W rozprawie zamieszczono również streszczenia w języku polskim i angielskim, a także wykaz skrótów anglojęzycznych i ważniejszych oznaczeń. Zamieszczenie **wykazu wybranych skrótów anglojęzycznych** oraz **wykazu wybranych oznaczeń** przed rozpoczęciem właściwej części rozprawy znacznie ułatwiło czytanie pracy.

We **wprowadzeniu** uzasadniono aktualność podjętej tematyki, opisano strukturę rozprawy, co ułatwiło zapoznanie się z tematyką całej rozprawy w sposób szybki.

Po wprowadzeniu określono **cel i zakres pracy** oraz sformułowano **tezę rozprawy**.

W rozdziale: **1. Ogólna charakterystyka oceny jakości sygnału audio** zostały przedstawione główne problemy oceny sygnałów audio. Został przeanalizowany wpływ przetwarzania cyfrowego i transmisji na jakość dźwięku.

W rozdziale: **2. Manualne metody oceny jakości audio** została przeprowadzona analiza najczęściej stosowanych metod manualnych oceny sygnału oraz pokazano charakter subiektywny, wysoki koszt i czasochłonność tych metod.

Dlatego, w rozdziale: **3. Automatyczne metody oceny jakości audio** bardziej szczegółowo przeanalizowano automatyczne metody oceny jakości sygnału audio. W rozdziale tym pokazano, że do analizy sygnałów muzycznych wykorzystywane są metody referencyjne, a metody niereferencyjne stosowane są głównie do analizy mowy. Autorka porównuje zaproponowaną przez siebie metodę z metodą z wykorzystaniem GRU przedstawiając ich komplementarność.

W rozdziale: **4. Autorska prototypowa metoda automatycznej klasyfikacji sygnału z eliminacją porównania do sygnału referencyjnego** Autorka wypunktowała podstawowe

etapy realizacji modeli sieci neuronowych, które opierają się na warstwach konwolucyjnych i rekurencyjnych. Zaprezentowany model wykorzystywany jest do automatycznej klasyfikacji sygnałów audio.

W rozdziale: **5. Autorska baza danych audio** opisano przygotowanie bazy danych sygnałów audio oraz ich niezbędne przetwarzanie w celu następnego wykorzystania tych sygnałów do procesu uczenia oraz do oceny prototypowego modelu sieci neuronowych. Sygnały audio zostały podzielone na 5 kategorii (czyste sygnały, błędy kwantyzacji, zniekształcenia wzmocnienia, dodatkowy dźwięk - szum, brakujący dźwięk), a każdą z nich podzielono na 3 podzestawy (treningowy, walidacyjny i testowy).

W rozdziale: **6. Projekt modelu** Autorka przedstawiła opracowany model klasyfikacji zniekształceń dla sygnałów muzycznych oraz opisała procedurę uczenia i testowania modelu sieci neuronowych klasyfikatora.

W rozdziale: **7. Ewaluacja modelu** została opisana ewaluacja modelu automatycznego klasyfikatora sygnału z wykorzystaniem następujących metryk: accuracy, specificity, precision, recall, F1.

W rozdziałach: **8. Model sieci konwolucyjnych, 9. Model sieci konwolucyjno-rekurencyjnych, 10. Model sieci konwolucyjno-rekurencyjnych z zastosowaniem dodatkowych parametrów wejściowych** zostały przedstawione badania modeli sieci konwolucyjnych, konwolucyjno - rekurencyjnych i konwolucyjno – rekurencyjnych z zastosowaniem dodatkowych parametrów wejściowych. Pokazano, że najwyższą skuteczność klasyfikacji sygnałów uzyskano dla modelu sieci neuronowych z wykorzystaniem warstw konwolucyjnych – rekurencyjno wraz ze spektrogramami skali melowej oraz parametru OBSC jako danych wejściowych. Model ten może być wykorzystywany do automatycznej klasyfikacji sygnałów muzycznych.

W rozdziale: **11. Podsumowanie** zostały podane wnioski z przeprowadzonych badań oraz przedstawione kierunki przyszłych badań, które warto zrealizować w tematyce, poruszonej w rozprawie.

Pracę kończy zawierająca 163 pozycje bibliografia. Dobór źródeł literaturowych jest właściwy. W spisie źródeł literaturowych umieszczona jest pozycja nr 118 współautorstwa Autorki: K. Organiściak, J. Borkowski, „Single-Ended quality measurement of a music content via convolutional recurrent neural networks,” *Metrology and Measurement Systems*, Vol. 27, No. 4, p. 721-733, 2020, w której Autorka jest pierwszym i korespondencyjnym autorem, co oceniam pozytywnie. Czasopismo “Metrology and Measurement Systems” jest indeksowane m.i. w Citation Reports and Web of Science Master Journal List (Clarivate Analytics formerly Thomson Reuters), Scopus, and Google Scholar.

Teza, cel i motywacja podjęcia tematu

Obecnie najczęściej stosowanymi metodami oceny jakości sygnałów audio są metody manualne polegające na przeprowadzeniu testów odsłuchowych, które wykonywane są przez wyszkolonych słuchaczy. Należą one do grupy metod subiektywnych, a w ocenie sygnałów audio może spowodować to występowanie błędów. Dodatkowo metody te charakteryzują się wysokim kosztem i czasochłonnością. W związku z tym istnieje potrzeba opracowania metod obiektywnych, do których można zaliczyć metody automatycznej analizy jakości sygnału

audio. Niniejsza rozprawa doktorska jest jedną z pierwszych dotyczących automatycznej klasyfikacji wybranych zniekształceń sygnałów muzycznych z wykorzystaniem konwolucyjno-rekurencyjnych sieci neuronowych bez porównania do sygnału referencyjnego.

Tytuł rozprawy „Automatyczna klasyfikacja wybranych zniekształceń sygnałów muzycznych z wykorzystaniem konwolucyjno-rekurencyjnych sieci neuronowych bez porównania do sygnału referencyjnego” odpowiada jej treści i określa rozważany w niej problem i zakres badań. Główne zagadnienia naukowe zaprezentowane przez Autorkę w rozprawie doktorskiej dotyczą analizy możliwości zastosowania konwolucyjno-rekurencyjnych sieci neuronowych do automatycznej klasyfikacji zniekształceń sygnałów muzycznych bez porównania do sygnału referencyjnego oraz opracowania prototypowego modelu automatycznej klasyfikacji wybranych zniekształceń sygnałów muzycznych. W mojej opinii podjęcie takiej tematyki badawczej jest bardzo aktualne i przydatne zarówno z punktu widzenia naukowego, jak i praktycznego, szczególnie dzisiaj, kiedy poszukiwane będą rozwiązania (obiektywne automatyczne metody oceny jakości rzeczywistych sygnałów muzycznych), które zastąpią długotrwałe manualne testy odsłuchowe.

Cel pracy został jasno sprecyzowany i brzmi następująco:

Celem pracy jest opracowanie prototypowego modelu neuronowych sieci konwolucyjno-rekurencyjnych oraz zbadanie jego skuteczności w celu klasyfikacji zniekształceń rzeczywistych sygnałów muzycznych. Głównymi założeniami opracowywanego modelu jest możliwość wykrywania wybranych zakłóceń w sposób automatyczny, obiektywny oraz bez konieczności porównania sygnału testowego do referencyjnego. Uzyskany wynik powinien zapewnić podstawę do dalszych badań, w taki sposób, że może on być zastosowany w celu rozszerzenia bazy wykrywanych zniekształceń, lub też do innego pokrewnego problemu (transfer learning), co w przyszłości mogłoby usprawnić długotrwałe manualne testy odsłuchowe przez częściowe zastąpienie ich obiektywną metodą automatyczną.

Teza pracy została sformułowana następująco:

Zastosowanie warstw dwukierunkowych rekurencyjnych w modelu sieci neuronowych wraz z odpowiednim doбором jego architektury, parametrów wejściowych oraz opracowaniem bazy zniekształceń do celów ewaluacji modelu, zwiększa znacząco skuteczność automatycznej klasyfikacji wybranych zniekształceń sygnałów muzycznych bez konieczności porównania do sygnału referencyjnego.

Teza w sposób właściwy postawiona w rozprawie jest tezą naukową, odpowiada ona celowi pracy, a wykazanie jej słuszności ma aspekt poznawczy oraz duże znaczenie praktyczne. Przedstawione w rozprawie metody i przyjęte założenia uważam za uzasadnione. Autorka udowadniając tezę rozwiązała problemy badawcze w niej zawarte. Autorka wykazała się głęboką wiedzą, dotyczącą przede wszystkim metod oceny jakości sygnałów audio oraz modeli sieci neuronowych różnych rodzajów. Autorka udowodniła, że potrafi analizować i oceniać uzyskane rezultaty oraz formułować poprawne wnioski.

Ocena oryginalności rozprawy

Przedstawione przez Autorkę dysertacji podejście zawiera aspekty nowości w zakresie wykorzystania konwolucyjno-rekurencyjnych sieci neuronowych do automatycznej

klasyfikacji zniekształceń sygnałów muzycznych bez konieczności porównania do sygnału referencyjnego. Za najważniejsze aspekty oryginalności pracy uważam:

- utworzenie autorskiej bazy zniekształceń z podziałem na cztery kategorie zniekształceń,
- dobór architektury modeli sieci neuronowych
- implementacja i ewaluacja modeli sieci neuronowych konwolucyjnych, konwolucyjno-rekurencyjnych oraz konwolucyjno-rekurencyjnych z zastosowaniem dodatkowych parametrów wejściowych,
- zbadania wpływu różnych parametrów takich jak ZCR, OBSC, głośności chwilowych i rzeczywistych wartości szczytowych oraz liczby filtrów melowych stosowanych dla danych wejściowych na skuteczność klasyfikacji modeli.

Reasumując, rozprawa doktorska jest aktualna do stanu wiedzy i rozwija dziedzinę automatycznej klasyfikacji zniekształceń sygnałów muzycznych bez konieczności porównania do sygnału referencyjnego.

Uwagi krytyczne, pytania

Rozprawa doktorska nie zawiera istotnych uchybień i błędów. Treść rozprawy jest usystematyzowana, podział rozprawy na 11 rozdziałów uważam za logiczny i właściwy. Rozdziały 8 - 10 rozprawy dotyczące modeli różnych sieci konwolucyjnych kończą się podsumowaniem, co znacznie ułatwia lekturę rozprawie. Rysunki i tabele są starannie przygotowane przez Autorkę. Autorka w odpowiedni sposób odwołuje się do pozycji Bibliografii.

Autorka wykazała się dobrym przygotowaniem w zakresie sieci neuronowych, wybrała właściwe narzędzia i metody, które pozwoliły na automatyczną klasyfikację sygnału z eliminacją porównania do sygnału referencyjnego. Recenzowana rozprawa doktorska generalnie prezentuje dobry poziom naukowy.

Przedstawione poniżej uwagi nie obniżają mojej pozytywnej oceny rozprawy doktorskiej.

Uwagi:

1. W pracy nie zostały przedstawione wartości TN, TP, FN, FP dla każdej z pięciu kategorii. Wskazane byłoby ich przedstawienie.
2. W pracy nie zostały przedstawione jakie wartości parametrów były wykorzystywane przy obliczaniu *accuracy*. Wskazane byłoby przedstawić *confusion metrics* dla każdego rodzaju rozpatrywanych sieci neuronowych.
3. Czym jest uzasadniona liczba sygnałów dla każdej z kategorii w tabeli 5.1?
4. W pracy wystąpiły drobne błędy językowe, które nie utrudniają zrozumienia pracy i nie wpływają na ogólnie dobrą ocenę formalnej strony pracy.

W pracy zawarte są niejednakowe oznaczenia: np. *F1* (str. 73, 74, 76-77, 79-82) zamiast *F1-score*; czcionka niepochyłona zamiast kursywy: *F1* zamiast *F1*.

W nazwach rozdziałów 2 i 3 zostało opuszczone słowo „sygnałów”: 2. Manualne metody oceny jakości audio, 3. Automatyczne metody oceny jakości audio zamiast 2. Manualne metody oceny jakości sygnałów audio, 3. Automatyczne metody oceny jakości sygnałów audio

Wnioski końcowe

Podsumowując stwierdzam, że uzyskane w pracy wyniki są oryginalne i zawierają elementy nowości naukowej. Oryginalne osiągnięcie stanowi opracowany przez Autorkę prototypowy model neuronowych sieci konwolucyjno-rekurencyjnych, a jego weryfikacja doprowadziła do potwierdzenia możliwości zastosowania warstw dwukierunkowych rekurencyjnych w modelu sieci neuronowych. Odpowiedni dobór architektury modelu, parametrów wejściowych oraz opracowanie bazy zniekształceń stosowanej w ewaluacji modelu zrealizowano w celu znaczącego podwyższenia skuteczności automatycznej klasyfikacji wybranych zniekształceń sygnałów muzycznych bez konieczności porównania do sygnału referencyjnego. Przedstawiony w rozprawie cel został osiągnięty, a teza potwierdzona. Zagadnienia poruszane w rozprawie mają aspekt praktyczny.

Autorka wykazała się odpowiednią wiedzą z zakresu znajomości aktualnej i istotnej tematyki podjętej w pracy, umiejętnością samodzielnego prowadzenia badań naukowych oraz sprawnością w zakresie formułowania wniosków.

Uważam, że rozprawa doktorska mieści się w dyscyplinie naukowej „Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne” oraz spełnia wszystkie wymagania stawiane pracom doktorskim przez ustawę o tytule naukowym i stopniach naukowych oraz o stopniach i tytule w zakresie sztuki, zgodnie z Ustawą o stopniach i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 r. (Dz.U. z 2017 r. poz. 1789), oraz zgodnie z Ustawą z 3 lipca 2018 r. - Przepisy wprowadzające ustawę - Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2018 r. poz. 1669 z póź. zm.). Na podstawie powyżej przedstawionej argumentacji, stawiam wniosek do Rady Dyscypliny Naukowej Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Wrocławskiej o przyjęcie ocenianej pracy oraz dopuszczenie Pani mgr inż. Kamili Organiściak do dalszych etapów procesu doktoryzowania.

Oleksandra Kotra

