Białystok, 12.06.2023

dr hab. inż. Małgorzata Krętowska, prof. PB
Wydział Informatyki
Politechnika Białostocka
ul. Wiejska 45a, 15-351 Białystok
e-mail: m.kretowska@pb.edu.pl

Recenzja rozprawy doktorskiej mgr. inż. Macieja Gołgowskiego

Wybrane metody uczenia maszynowego w zadaniach wykrywania anomalii procesów

Promotor rozprawy: prof. dr hab. inż. Stanisław Osowski

Podstawą niniejszej recenzji jest uchwała Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego nr 21/RDN ITiT/2023 z dnia 4 kwietnia 2023 roku powołująca mnie na recenzenta rozprawy doktorskiej mgr. inż. Macieja Gołgowskiego nt. „Wybrane metody uczenia maszynowego w zadaniach wykrywania anomalii procesów" w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie informatyka techniczna i telekomunikacja.

1. Cel, zakres i charakter rozprawy

Celem pracy było opracowanie metod wykrywania anomalii procesów, reprezentowanych przez sygnały jedno- i dwuwymiarowe, które pozwolą na uzyskanie polepszonej dokładności wychwycenia anomalii. W ramach zaproponowanych rozwiązań można wyróżnić dwa etapy, na które składa się wstępne przetworzenie danych „surowych" wykorzystujące ciągłą (CWT) oraz dyskretną (DWT) transformację falkową oraz przypisanie przetworzonych danych do klasy normalnej lub anomalnej przy zastosowaniu różnego typu klasyfikatorów (klasycznych lub głębokich) i ich zespołów.

Zakres pracy obejmował przedstawienie pojęcia anomalii procesu oraz przegląd istniejących metod wykrywania anomalii. W ramach proponowanych rozwiązań, przedstawiono dwa sposoby wstępnego przetworzenia danych wejściowych, które bazują na wykorzystaniu ciągłej i dyskretnej transformacji falkowej. Otrzymane dane wyjściowe, obrazy (CWT) oraz wektor cech diagnostycznych (DWT) były następnie przekazywane jako dane wejściowe do zespołów klasyfikatorów klasycznych oraz głębokich sieci CNN. Zakres pracy obejmował również eksperymentalną weryfikację zaproponowanych rozwiązań. Ich działanie zweryfikowano na bazie trzech problemów praktycznych: wykrywania anomalii w sygnałach EKG, wykrywania uszkodzeń łożyska tocznego na podstawie sygnałów z czujników akcelerometrycznych oraz wykrywanie obrazów typu „deep fake" wyekstrahowanych z filmów video. Rozprawa obejmuje swoim zakresem również dyskusję na temat otrzymanych wyników, ich podsumowanie oraz wnioski końcowe.

Tematyka rozprawy stanowi niewątpliwie istotny element badań w dyscyplinie informatyka techniczna i telekomunikacja. Wykrywanie anomalii procesów ma miejsce w wielu różnych dziedzinach życia - począwszy od medycyny, przez zagadnienia typowo inżynierskie, na analizie informacji obrazowej kończąc. Praca ma charakter badawczy. Opracowany został schemat rozwiązania problemu naukowego, jakim jest wykrywanie anomalii procesu, którego działanie zostało eksperymentalnie zweryfikowane na bazie rzeczywistych zbiorów danych.

2. Zawartość rozprawy

Rozprawa doktorska liczy 104 strony, w jej skład wchodzi 6 rozdziałów oraz spis literatury.

Rozdział pierwszy stanowi wstęp do pracy doktorskiej. Zostało w nim omówione pojęcie anomalii procesu, jak również nakreślono główne problemy, które mogą pojawić się w trakcie rozpoznawania sygnałów anomalnych. Omówione zostały, w dosyć ogólny sposób, podstawowe metody wykrywania anomalii procesów, bazujące na uczeniu bez nauczyciela, jak też, na wykorzystywanych w pracy, metodach klasyfikacji, w których znana jest przynależność sygnału do odpowiedniej klasy, normalnej lub anomalnej. W podrozdziale trzecim przedstawiono cel pracy oraz zawartość poszczególnych jej rozdziałów.

W rozdziale drugim omówiona została zaproponowana przez Doktoranta metoda wykrywania anomalii procesów. Można tu wyróżnić dwa etapy. W pierwszym następuje wstępne przetworzenie danych wejściowych. Jego celem jest wygenerowanie (ewentualnie selekcja) odpowiednich cech diagnostycznych procesu, które będą stanowiły dane wejściowe do kolejnego etapu klasyfikującego sygnał do klasy normalnej lub anomalnej. Do wstępnego przetworzenia sygnału wejściowego wykorzystana została transformacja falkowa, której podstawy teoretyczne zostały omówione w podrozdziale drugim. Ciągła transformacja falkowa jest narzędziem pozwalającym na generowanie obrazów w funkcji współczynnika skali i przesunięcia, natomiast w przypadku dyskretnej transformacji falkowej otrzymujemy wektory wartości dla odpowiednio dobranej skali. W tym przypadku, dane wejściowe do etapu drugiego, będą statystykami opisującymi rozkład wartości uzyskanych przy użyciu dyskretnej transformacji falkowej na różnych poziomach dekompozycji. Podrozdział trzeci zawiera ogólne opisy modeli klasyfikatorów wykorzystywanych w pracy. Są to: las losowy, metoda „gradient boosting", sieć SVM, metoda k-NN, klasyfikator bazujący na procesie gaussowskim, perceptron wielowarstwowy, naiwny klasyfikator Bayesa oraz sieci głębokie CNN. W większości z omówionych modeli, Doktorant zwraca uwagę na występowanie dodatkowych parametrów uczenia, których wartość powinna być odpowiednio dobrana do analizowanego problemu.

Niewątpliwie głównych elementem pracy są przykłady wykorzystania zaproponowanej metody w analizie sygnałów rzeczywistych. W rozdziale trzecim zawarty został opis wykrywania anomalii w sygnałach EKG. Dane pochodziły z baz ogólnodostępnych i składały się z trzech rodzajów sygnałów reprezentujących arytmię, niewydolność serca oraz normalną pracę serca. W wyniku zastosowania dyskretnej transformacji falkowej, otrzymano 6 zbiorów sygnałów na różnych pasmach częstotliwości, z których każdy został opisany przy pomocy 13 statystyk. W sumie otrzymano 78 cech. W etapie klasyfikacji sygnałów, pod uwagę brany był

cały zbiór deskryptorów oraz jego podzbiory (12, 24 i 48 cech), wybierane na podstawie wartości statystyki chi^2. Porównano działanie dziewięciu różnych modeli klasyfikatorów, a następnie wykonano eksperymenty, w których sześć najlepszych połączono w zespół. Tak otrzymany model złożony, w którym przynależność do klasy ustalano na podstawie głosowania większościowego, uzyskał lepszą jakość klasyfikacji niż modele klasyczne. Dodatkowym elementem było wykorzystanie ciągłej transformacji falkowej z trzema różnymi typami falek oraz czterema wartościami skali. Uzyskane obrazy zostały następnie przekazane na wejście utworzonej samodzielnie sieci CNN oraz zespołu ośmiu wstępnie wytrenowanych sieci CNN, wykorzystując tzw. technikę „transfer learning". Wyniki klasyfikacji uzyskane przez zespół okazały się lepsze niż wyniki budowanej od podstaw sieci CNN.

W rozdziale czwartym opisany został przykład wykorzystania zaproponowanego algorytmu do wykrywania uszkodzeń łożysk. Dane pochodziły z ogólnodostępnej bazy zawierającej sygnały z akcelerometru. Uzyskane sygnały reprezentują 5 klas, z czego cztery dotyczą anomalii, jedna - działania łożyska bez stwierdzonych defektów. W części eksperymentów wszystkie sygnały anomalne, połączone zostały w jedną klasę. Sposób analizy sygnałów jest podobny do analizy sygnałów EKG. Wyniki eksperymentów pokazują dobrą jakość klasyfikacji przy użyciu zespołów klasyfikatorów. Uzyskane wyniki wydają się być w grupie najlepszych wyników uzyskiwanych dla tego typu problemów przez innych naukowców, jednak, jak zauważono w pracy, należy zachować ostrożność przy porównywaniu metod, ze względu na różne analizowane zestawy danych.

Rozdział piąty zawiera opis wykrywania anomalii typu „deep fake" w obrazach. Na wstępie została przedstawiona definicja problemu oraz metody generowania obrazów „deep fake": *FaceSwap*, *FakeApp* oraz *Face2Face*. Do analizy wykorzystano dane pochodzące ze zbioru *Faceforensics++*, z którego losowo wybrano 4000 obrazów (klatek) oryginalnych oraz wygenerowanych przy użyciu trzech, wymienionych wyżej, metod. Przygotowanie danych obejmowało również detekcję obrazu twarzy z klatek video metodą HOG. Obrazy były wstępnie skompresowane (c23). Proponowana procedura składała się z trzech kroków: wykrywania punktów orientacyjnych twarzy, zastosowania ciągłej transformacji falkowej przy różnych wartościach skali dekompozycji oraz wykorzystania wygenerowanych w ten sposób obrazów jako atrybutów wejściowych do zespołu głębokich sieci CNN. Eksperymenty przeprowadzono oddzielnie dla trzech algorytmów generowania fałszywych obrazów, przy użyciu trzech tensorów wejściowych będących złożeniem trzech obrazów wynikowych CWT (amplitudowych lub kątowych) dla kolejnych wartości skali, uzyskanych z zastosowaniem falki Morleta. Najlepsze wyniki rozpoznawania anomalii uzyskano dla metody *FakeApp* (ACC=97,33%), najgorsze dla *Face2Face* (ACC=83,5%). Są one porównywalne z wynikami prezentowanymi w literaturze.

W rozdziale szóstym Doktorant przedstawia końcowe wnioski oraz nakreśla dalsze kierunki badań.

Bibliografia składa się z 96 pozycji. Doktorant powołuje się na trzy publikacje, których jest współautorem. Są to artykuły opublikowane w czasopiśmie *Computational Problems of Electrical Engineering* (2020 rok), *Przeglądzie Elektrotechnicznym* (2021 rok, 70 punktów)

oraz na konferencji *IEEE International Joint Conference on Neural Networks* (2022 rok, 140 punktów).

3. Główne osiągnięcia

Podsumowując merytoryczną ocenę rozprawy stwierdzam, że Doktorant osiągnął założone cele badawcze dzięki opracowaniu nowych metod wykrywania anomalii procesu oraz weryfikacji ich skuteczności na przykładzie trzech problemów praktycznych. Należy podkreślić, że osiągnięcie celów pracy wymagało posiadania przez mgr. inż. Macieja Gołgowskiego szerokiej wiedzy z dyscypliny informatyka techniczna i telekomunikacja.
Do głównych osiągnięć Autora zaliczyć można:

- opracowanie metody wykrywania anomalii procesów reprezentowanych przez ciągi czasowe, bazującej na wykorzystaniu dyskretnej transformacji falkowej generującej wyniki, których charakterystyki liczbowe stanowią zbiór atrybutów wejściowych dla zespołu klasyfikatorów płytkich,
- opracowanie metody wykrywania anomalii procesów bazującej na ciągłej transformacji falkowej, wykorzystywanej zarówno w przypadku ciągów czasowych, jak i obrazów, której rezultaty są danymi wejściowymi do zespołu wstępnie wytrenowanych głębokich sieci CNN,
- sprawdzenie skuteczności zaproponowanych metod w oparciu o trzy problemy praktyczne: wykrywanie anomalii w sygnałach EKG, wykrywanie uszkodzeń łożyska tocznego na podstawie zarejestrowanych sygnałów czujników akcelerometrycznych oraz wykrywanie sfałszowanych obrazów twarzy wyekstrahowanych z filmów wideo. Każdy z tych przykładów wymagał indywidualnego rozwiązania oraz głębszej analizy problemu.

4. Uwagi do pracy

Poniżej przedstawione zostały uwagi do rozprawy doktorskiej:

1. Istotnym elementem pracy naukowej jest przegląd literatury, w tym przypadku dotyczącej wykrywania anomalii procesów. W niniejszej rozprawie zabrakło rozdziału, który w wyczerpujący sposób pokazałby stosowane w literaturze metody, zarówno przygotowania danych, jak też już właściwej ich analizy. Pewne elementy przeglądu zostały zawarte we wstępie całej pracy (podrozdział 1.2) oraz we wstępach rozdziałów eksperymentalnych.
2. Opis miar wykorzystywanych do oceny jakości działania klasyfikatorów powinien znaleźć się w rozdziale 2. Miary te zostały użyte do oceny jakości klasyfikacji w każdym z trzech analizowanych problemów, a nie tylko w rozdziale 3 „Wykrywanie anomalii w sygnałach EKG", w którym zostały opisane. Definicja dokładności (ACC) podana jest błędnie. W liczniku powinna być suma TP i TN.
3. W opisie modeli klasyfikatorów wykorzystywanych w pracy, wielokrotnie podnoszony był temat parametrów poszczególnych modeli, które powinny być dobierane w takcie procesu uczenia. W rozdziałach 3 i 4 prezentowane są wyniki działania klasyfikatorów klasycznych oraz ich zespołów, natomiast nie ma informacji na temat wartości i sposobu doboru parametrów tych modeli.
4. Jakość klasyfikacji poszczególnych klasyfikatorów weryfikowana była przy użyciu takich miar, jak dokładność, precyzja, czułość oraz miara F1. W rozdziale 3 i 4

prezentowane były jedynie wartości średnie tych miar wyliczone na bazie 10 powtórzeń eksperymentów. Być może warto byłoby przedstawić również odchylenia standardowe, jak to zostało zrobione w rozdziale 5, w celu zobrazowania zróżnicowania otrzymanych rezultatów. W pracy zabrakło również weryfikacji, czy otrzymane różnice jakości działania klasyfikatorów klasycznych i ich zespołów są istotne statystycznie.

5. W rozdziale 3.3.3 pt. „Selekcja cech diagnostycznych procesu" zawarta jest informacja o wykorzystaniu statystyki chi^2 do selekcji cech. W opisie Doktorant używa nazwy test zgodności, który wykorzystywany jest do badania zgodności rozkładu danej cechy z pewnym, z góry zadanym rozkładem. Innym testem wykorzystującym statystykę chi^2 jest test niezależności, który bada niezależność zmiennych jakościowych. Bardzo proszę o doprecyzowanie opisu wykorzystywanego testu, z uwzględnieniem wzoru wyliczającego statystykę chi^2.
6. W pracy brakuje informacji o czasach uczenia poszczególnych klasyfikatorów, czy też przygotowania danych uczących. W podsumowaniu jest informacja o potrzebie przyspieszenia procesu przetwarzania danych, jednak Doktorant nie podał wartości tych czasów w pracy. Mogłoby być to szczególne interesujące w kontekście uczenia sieci CNN tworzonej od podstaw oraz metody „transfer learning".
7. W problemie wykrywania anomalii sygnału EKG wykorzystano między innymi budowaną od podstaw sieć CNN. Liczba elementów dla poszczególnych klas wynosiła odpowiednio 96, 30 i 36 przypadków, przy czym jako zbiór uczący branych było 70% przypadków. Jaka była liczba parametrów (wag) zaproponowanej sieci? Jakie były argumenty za wyborem budowanej od podstaw sieci CNN jako klasyfikatora przy stosunkowo niewielkiej ilości danych uczących?

Uwagi o charakterze redakcyjnym:

1. W pracy występuje spora liczba błędów interpunkcyjnych, stylistycznych, literówek oraz kilka błędów ortograficznych (str. 25, 37, 47, 57).
2. Brak rysunku 3.5, do którego jest odwołanie w tekście pracy.
3. Liczba klasyfikatorów w zespole: w tekście na str. 48 - 6, w tytule tabeli 3.2 – 7.
4. 26. pozycja w bibliografii – nazwa czasopisma wyróżniona jest czerwonym kolorem czcionki.
5. Na stronie 61 jest *„otrzymano 5 szeregów czasowych szczegółowych"*, powinno być *„otrzymano 4 szeregi czasowe szczegółowe"*.
6. Wykorzystanie słowa „ilość" zamiast „liczba" w odniesieniu do rzeczowników policzalnych.
7. W języku polskim separatorem dziesiętnym jest przecinek, w pracy występuje kropka.

5. Wnioski końcowe

Podsumowując stwierdzam, że, mimo przedstawionych powyżej uwag, praca pt. „Wybrane metody uczenia maszynowego w zadaniach wykrywania anomalii procesów" spełnia wszystkie wymogi stawiane rozprawom doktorskim przez obowiązujące przepisy. W związku z powyższym wnioskuję o przyjęcie tej rozprawy i dopuszczenie mgr. inż. Macieja Gołgowskiego do publicznej obrony.

Małgorzata Krętowska