

prof. dr hab. Piotr Porwik
Instytut Informatyki
Uniwersytet Śląski, Katowice
Wydział Nauk Ścisłych i Technicznych
ul. Będzińska 39
41-200 Sosnowiec

Katowice 21.11.2022

OCENA ROZPRAWY DOKTORSKIEJ

"Analiza sygnału EEG dla potrzeb rozpoznawania emocji"

("Analysis of EEG signals for emotion recognition")

mgr inż. Krzysztofa Kotowskiego

Recenzja została przygotowana na podstawie pisma z dnia 05.10.2022, sporządzonego przez prof. dr hab. inż. Andrzeja Polańskiego, Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej.

Ocena doboru tematu

Rozprawa doktorska napisana została w języku angielskim i opisuje metodę analizy i filtracji elektroencefalogramu (sygnałów EEG) ludzkiego mózgu oraz opisuje urządzenie, wraz z modyfikacjami, do nieinwazyjnej rejestracji tych sygnałów w mózgu człowieka. Aktywność EEG odzwierciedla bioelektryczne pobudzenia neuronów korowych, które są rozmieszczone przestrzennie. Wskazania do wykonania badania EEG obejmują diagnostykę napadów padaczkowych, zaburzenia wzroku, bólu głowy czy ocenę stanu pacjenta po urazach mózgowo-czaszkowych lub urazach ośrodkowego układu nerwowego. Wskazaniem do wykonania badania EEG jest również podejrzenie śmierci mózgu.

Sygnały z elektrod, które są rozmieszczone na powierzchni głowy rejestrują zmiany potencjału elektrycznego, spowodowane aktywnością komórek nerwowych kory mózgu. Są to sygnały, które trudno odseparować od innych sygnałów EEG, (artefaktów), sygnałów tętna, ruchu gałek ocznych, itp.

W praktyce Autor wykorzystuje technikę pomiarową ERP, gdzie zamiast jednokrotnego pobudzenia bodźcem, rejestruje się reakcję na wielokrotne pobudzenie neuronów korowych. Technika ERP używa metody uśredniania. Jej wyniki są bardziej stabilne, a losowe aktywacje są filtrowywane (usuwane). Pozwala to na niezakłóconą analizę rzeczywistej reakcji na bodziec zewnętrzny z eliminacją szumu pomiarowego. Szumy, jako dane odstające pojawiają się również jako efekt niezamierzonych mikroruchów pacjenta. Jest

to oczywiście również technika bezinwazyjna. Autor proponuje w swojej rozprawie dwa swoje rozwiązania związane z techniką ERP. Pierwsze z nich dotyczy budowy nowego urządzenia do rejestracji i oznaczania bodźców pochodzących z emocjonalnych ekspresji twarzy. Drugie rozwiązanie pozwala na rozróżnianie sygnałów neuronów aktywnych w mimice twarzy. Na przykład w stanie złości, radości czy w fazie neutralnej.

Zarówno EEG jak i EPR są powszechnie stosowanymi technikami diagnostycznymi współczesnej medycyny. Propozycje Autora należy zaliczyć do działań, które mają wzbogacić wiedzę o falach mózgowych związanych z ekspresją twarzy obserwowaną w stanach emocjonalnych.

To podejście pozwala na lepsze zrozumienie działania mózgu człowieka. Z naukowego punktu widzenia jest to działanie pożyteczne, pozwalające na planowanie nowych eksperymentów w neuronauce, a więc na pograniczu wiedzy medycznej, biologicznej, biofizycznej, informatycznej lub nawet psychologicznej.

Charakterystyka rozprawy i jej zakres

Praca podzielona została na kilka uzupełniających się części. Składa się z 4 rozdziałów oraz spisu literatury. Na stronie 2 rozprawy doktorskiej Autor przedstawia cele, które zamierza osiągnąć, założenia (w punktach G1-G4) oraz tezę rozprawy doktorskiej. Celem Autora jest opracowanie nowych metod przetwarzania sygnałów, które mogą być stosowane w neuronauce zajmującej się badaniem emocji, w tym komponentów N170 oraz N250. N170 jest komponentem ERP pojawiającym się u badanej osoby na ekspozycję twarzy (np. zdjęcia z określoną mimiką). N170 charakteryzuje się ujemnym odchyleniem amplitudy fali, które pojawia się około 170 ms po wywołania pobudzenia (Bentin, Allison, Puce, Perez i McCarthy, 1996). Podobne znaczenie ma komponent N250, którego opóźnienie od wystąpienia pobudzenia waha się w przedziale 230-320 ms. Fala N250 zakłóca jednak komponent N170.

Tezy rozprawy zostały przedstawione jasno i obejmują dwa aspekty:

- Autor zakłada, że stosując niedroge urządzenia do rejestracji EEG oraz ich odpowiednia adaptacja pozwalają na rozpoznawanie ekspresji emocji badanej osoby, co jest przydane w ocenie jej stanu psychicznego.
- Właściwe wstępne przetwarzanie, w tym filtracja artefaktów ocznych, pozwala na pojedynczą analizę ERP uczenia się twarzy, co znajduje odzwierciedlenie w komponente N250, który jest czynnikiem zakłócającym przetwarzanie emocji twarzy.

W Rozdziale 1.3 Autor wymienia w skrócie najważniejsze osiągnięcia rozprawy. Realizacja tego zadania nie jest jednak dobra. Autor pisze tutaj, że pokaże sekwencję kroków do wykrywania markerów twarzy wyrażającej emocje za pomocą urządzenia Emotiv EPOC+, stosując do tego celu nową filtrację falkową opartą na epokach. W pracy nie zaprezentowano jednak żadnej filtracji opartej na falkach. Nie zrobiono tego ani teoretycznie ani

praktycznie. Trudno więc tę nowość ocenić. Autor odsyła tylko czytelnika do swojej publikacji, która te zagadnienie omawia. Moja uwaga dotyczy także innych punktów Rozdziału 1.3.

W rozdziale kolejnym Autor wymienia swoje publikacje, które ściśle związane są z rozprawą. Dla każdej z nich przedstawiono krótką charakterystykę wraz z opisem osiągnięcia. Zestaw tych samych prac powtarza się później w spisie literatury.

W obszernym Rozdziale 2 Doktorant dokonuje przeglądu teoretycznych podstawy pomiarów aktywności mózgu z zastosowaniem EEG/ERP, w tym przedstawia opisy pól (potyliczny, ciemieniowy, skroniowy, czołowy) i ich funkcje. Przedstawiono także tutaj różne współczesne techniki pomiarów aktywności mózgu. Należy zaznaczyć, że niektóre z tych technik wymagają stosowania bardzo drogich urządzeń medycznych. Autor proponuje zastosowanie taniego urządzenia Emotiv EPOC+, które nie jest urządzeniem medycznym i nie jest w medycynie wykorzystywane. Jest to urządzenie komercyjne, przeznaczone do badań naukowych, stosowane w warunkach laboratoryjnych. Urządzenie jest jednak wystarczająco dokładne do wykonania wiarygodnych pomiarów fal mózgowych. W tym rozdziale pojawiają się też po raz pierwszy nazwy i opisy rejestrowanych w pomiarach ERP fal (komponentów). Niektóre z tych komponentów wykorzystywane są na potrzeby pracy, ale przywołanie tych nazw znalazło się dużo wcześniej bez objaśnień ich pochodzenia i znaczenia. Kompozycja pracy mogła być o wiele lepsza.

Rozdział 2.2.1 zatytułowany "Obszary mózgu związane emocjami" (*Emotion-related brain areas*) został potraktowany bardzo lakonicznie, a przecież jest to obszar bardzo ważnym patrząc na tytuł rozprawy. Pomiary fal mózgowych odpowiedzialnych za rozpoznawanie emocji są przecież głównym obiektem zainteresowań Autora, a rozkład elektrod na głowie pacjenta sugeruje z jakich obszarów pochodzą wzbudzone sygnały elektryczne.

Rozdział 3 jest najobszerniejszy i najważniejszy z naukowego punktu widzenia. W tym rozdziale opisano przygotowanie eksperymentów (w tym pacjentów), zastosowany w eksperymentach sprzęt i oprogramowanie (Python wraz bibliotekami), zasady akwizycji danych, opisano procedury wstępnego przetwarzania surowych danych oraz przedstawiono analizę statystyczną wyników opracowaną w języku R.

Oryginalne urządzenie Emotiv EPOC+ zostało zmodyfikowane i doposażone w moduł znakowania bodźców. Zasada pracy tego układu została przedstawiona poglądowo na Rys. 3.1. Należy zauważyć, że jest to oryginalne, nowatorskie rozwiązanie Autora. Szkoda, że Autor nie pokusił się o głębsze wyjaśnienie zasad działania tego układu. Rejestrowane w technice ERP fale mózgowe danego typu pochodzą z sekwencji pobudzeń bodźcami zewnętrznymi (np. światłem). Zarejestrowane w ten sposób przebiegi są więc wzajemnie przesunięte w czasie. Zapropozowane rozwiązanie znakowania bodźców jest ważne dla wykonania wiarygodnych pomiarów, gdyż pozwala na uzyskanie takiego samego offsetu dla wszystkich powtarzających się rejestracji, co pozwala z kolei na znalezienie uśrednianionego przebiegu sygnału we wszystkich epokach. W tym samym rozdziale Doktorant opisuje

metody zwiększające odporność na dane odstające oraz zwiększające odporność na nieuchronny brak korelacji danych w czasie. Dla realizacji tego zadania Autor zmodyfikował algorytm WACFM (*Weighted averaging based on criterion minimization*) opublikowany w cytowanej przez Doktoranta pracy prof. Jacka Łęskiego z roku 2002. Przedstawiony w pracy nowy algorytm o nazwie corWACFM odpowiada za odporne uśrednianie ważone przebiegów zebranych w epokach ERP. Działanie algorytmu zostało ocenione w badaniach porównawczych przez wyznaczenie wartości RMSE oraz MAX (liczonych bootstrapowo), dla trzech innych metod znanych z literatury. Prównań dokonano dla danych surowych i danych po procesie filtracji. Propozycja Autora wykazuje przewagę nad innymi metodami.

Tutaj pojawia się jednak pytanie. Ponieważ Autor wykorzystuje filtrację sygnałów, to RMSE jest dla takiego przypadku dużo mniej korzystne - jest przeszło 4 razy większe niż dla danych surowych (bez filtracji). Proszę o wyjaśnienie tego postępowania.

W pracy Autor zbadał też oddziaływanie komponentów N250 i EPN na pomiary związane z rozpoznananiem emocji oraz zapamiętywaniem twarzy przez mózg. Przeprowadzono też rozpoznawanie i filtrację artefaktów związanych z ruchem gałek ocznych zakłócających właściwe pomiary. Wymienione powyżej komponenty mają charakter behawioralny i nie da się ich wyeliminować - są niezależne od woli człowieka.

W ostatnim rozdziale zawarto podsumowanie całej rozprawy doktorskiej i podkreślono jej nowości.

Ocena przyjętych metod rozwiązania zagadnienia

Metodologia przyjętych rozwiązań jest poprawna, a eksperymenty wiarygodne. Wykonane eksperymenty udostępniono w postaci kodów programów w języku Python, co pozwala kolejnym eksperymentatorom na badania porównawcze.

Teza i cel pracy zostały przedstawione precyzyjnie, chociaż brak im jasności ze względu na stosowane w tekście symbole. Rozwiązania autorskie wykazują istotną przewagę nad innymi metodami opisanymi w samej rozprawie. Eksperymenty zostały zaprojektowane wielowariantowo, co świadczy o dobrym przygotowaniu Autora do prowadzenia badań naukowych, a także o bardzo dobrej znajomości podejmowanej w pracy tematyki badawczej i metod statystycznych. Pomysł układu do znakowania bodźców w celu uzyskania takiego samego offsetu dla wszystkich pomiarów fal mózgowych w cyklu pomiarów należy uznać za ciekawy i nowatorski. Autor wykazał, że skomplikowane i drogie urządzenia do pomiarów fal mózgowych o niewielkiej energii można zastąpić tanim urządzeniem pomiarowym Emotiv EPOC+, co pozwala na szersze i tańsze stosowanie metody w praktyce związanej z neuronauką.

Zaproponowane rozwiązanie znakowania bodźców jest ważne dla wykonania wiarygodnych pomiarów, gdyż pozwala na uzyskanie takiego samego offsetu dla wszystkich powtarzających się rejestracji.

Lektura rozprawy, a w szczególności Rozdział 3, uzasadnia stwierdzenie, że tezy zo-

stały poprawnie zweryfikowane, a wytyczone cele zostały osiągnięte.

Analiza źródeł literaturowych

W rozprawie znajdują się odwołania do 214 prac, w tym 12 prac, gdzie Doktorant jest współautorem artykułu. W większości tych prac Autor występuje jako główny autor. Większość (11 prac) jest indeksowana w systemie WoS (*Web of Science*). Są to prace o dobrym współczynniku oddziaływania (*IF*) i wysokiej punktacji (140 pkt. wg. MEiN). Baza WoS odnotowuje dla Autora indeks Hirscha $h = 4$ z liczbą cytowań 51, co jest wartością dobrą. Taką aktywność naukową doktoranta należy z uznaniem podkreślić. Źródła literaturowe dotyczą w większości publikacji w czasopismach naukowych oraz materiałach konferencyjnych związanych z zagadnieniami elektroencefalografii, czynności prądów mózgu i fal mózgowych. Niektóre typy tych fal mają użyteczną przydatność w badaniach Autora. Zestaw pozycji literaturowych jest reprezentatywny dla dziedzin związanych nauronaukami i uwzględnia również doniesienia najnowsze, z roku 2022. Wnioski i polemiki wypływające z bardzo szerokiego i reprezentatywnego przeglądu literatury zostały sformułowane przekonywująco i należy je uznać za poprawne. Autor swobodnie i kompetentnie dokonał przeglądu literaturowego.

Ocena wyników przedstawionych w rozprawie

Uzyskane wyniki badań eksperymentalnych zostały zaprezentowane w analizie porównawczej, gdzie Autor konfrontował własne metody z metodami znanymi z literatury. Analiza otrzymanych rezultatów wskazuje, że opisane w rozprawie modyfikacje komercyjnego i taniego urządzenia Emotiv EPOC+ są wystarczające do wykrywania ludzkich emocji na podstawie mimiki twarzy. Można tego dokonać za pomocą analizy komponentu EPN. Autor opisał także komponenty fal mózgowych N170 oraz N250, ich wzajemne relacje oraz przydatność w rozpoznawaniu i zapamiętywaniu twarzy. Autor wykonał i udostępnił oprogramowanie pozwalające na powtórzenie eksperymentów przez innych badaczy. Na tle przedstawionej analizy uzyskane rezultaty są nowe i pogłębiają wiedzę związaną z neuronauką czyli interdyscypliną z pogranicza wiedzy medycznej, biologicznej, biochemicznej, biofizycznej, informatycznej i psychologicznej.

Podsumowując tę część recenzji uważam, że praca jest wartościowa i wykazuje dużą wiedzę merytoryczną doktoranta.

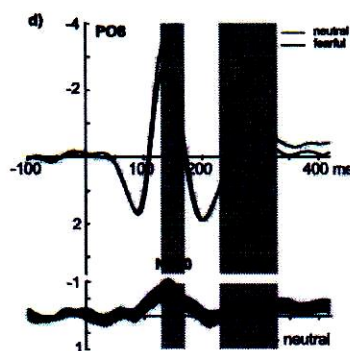
Uwagi

W mojej opinii największym mankamentem pracy jest jej kompozycja i sposób prowadzenia wywodów.

Po pierwsze, bardzo brakuje wprowadzenia czytelnika w bardziej szczegółową problematykę rozprawy. Autor już na wstępie, w tym w prezentacji tezy i celów pracy, posługuje się nieobjaśnionymi terminami takimi jak EPN, ERP, N250 czy N170. Takie wyjaśnienia

znajdują się dopiero na 36 stronie pracy. Brakuje również wstępnego, chociażby nieformalnego, wyjaśnienia znaczenia wymienionych komponentów w kontekście prowadzonych badań i eksperymentów.

Przykładem, który dobrze charakteryzuje powiązanie dwóch ważnych, analizowanych w pracy, komponentów EPN oraz N170 jest poniższy rysunek, zaczerpnięty z pracy Schindler S., Bruchmann M., Bublatzky F., Straube T. (2019). *Modulation of face-and emotion-selective ERPs by the three most common types of face image manipulations*. Social Cognitive and Affective Neuroscience. 14. 1-11. Takie przykłady, wraz z objaśnieniami, bardzo dobrze wprowadzałyby czytelnika w problematykę rozprawy doktorskiej.



Drugą uwagę kieruję w sprawie zasad podawania przez Autora opisów eksperymentów, prezentowania podstaw teoretycznych oraz różnych objaśnień kontekstowych. Autor bardzo często przerywa nagle wyjaśnianie czytelnikowi problemu, odsyłając go do literatury (także własnych prac) podając numery wzorów z przywoływanej pracy. W praktyce, zmusza to czytelnika do odszukania takiej pracy, którą i tak trzeba w całości przeczytać aby zrozumieć cały kontekst zagadnienia.

Na przykład: "I proposed a solution significantly minimizing this problem by limiting the maximum weight by adding a small positive constant to the dissimilarity metric ρ , so the value of weight w_i in equation (12) in [176] never reaches infinity." lub w opisie algorytmu podanego w Tabeli 3-2 "Update v using equation 14 from [176]". Unikanie takich wtrąceń zwiększyłoby precyzję pracy. Takich miejsc w pracy jest dużo więcej.

Konieczność sięgania do różnych prac pokazuje, że przedstawiony do oceny doktorat jest tak naprawdę wybiórczym przeglądem opublikowanych już prac Autora. Taka metoda pokazywania osiągnięć nie jest najlepsza i jest ze szkodą dla całej pracy. Lektura doktoratu wskazuje bowiem na brak ciągłości wywodów i konieczność ich uzupełniania w zewnętrznych pracach Autora.

Wnioski końcowe

Podsumowując przedstawioną charakterystykę rozprawy stwierdzam, że doktorant, Pan mgr inż. Krzysztof Kotowski, zamieścił w niej samodzielnie uzyskane wyniki badań teoretycznych i eksperymentalnych, popartych analizą statystyczną. Doktorant wykazał się

umiejętnością budowy algorytmów oraz programowania. Do dobrych stron pracy zaliczam wartościową chociaż wybiórczą, dyskusję proponowanych rozwiązań oraz interpretację uzyskanych wyników. Autor wykazał się dobrym opanowaniem warsztatu badawczego. Do słabych stron pracy zaliczam nieuporządkowaną dyskusję podejmowanych problemów badawczych i nie najlepszą kompozycję pracy. Moje uwagi mają charakter polemiczny i mam nadzieję, że Doktorant odniesie się do nich w trakcie obrony.

Pod względem merytorycznym praca spełnia wymagania stawiane rozprawom doktorskim i jest zgodna z wytycznymi ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki.

Wnoszę o dopuszczenie opiniowanej dysertacji do dalszych etapów przewodu doktorskiego, w tym do jej publicznej obrony.

