

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

Katowice 01.09.2025

OCENA ROZPRAWY DOKTORSKIEJ

"Rozpoznawanie człowieka na podstawie chodu przy wykorzystaniu sensorów IMU oraz metod tworzenia sztucznych próbek w systemach biometrycznych"

Pana mgr inż. Aleksandra Sawickiego

W piśmie prof. dr hab. inż. Marka Krętowskiego, Przewodniczącego Rady Dyscypliny Naukowej Wydziału Informatyki Politechniki Białostockiej z dnia 03.07.2025 zostałem poinformowany, że zgodnie z decyzją Rady, zostałem powołany na recenzenta rozprawy doktorskiej Pana Aleksandra Sawickiego.

Ocena tematyki rozprawy

Rozprawa podejmuje aktualny i ważny problem podniesienia skuteczności systemów biometrii chodu opartych na czujnikach IMU. Autor bada dwa komplementarne nurty: (i) algebraiczną augmentację sygnałów oraz (ii) generowanie próbek syntetycznych metodą LSTM-MDN.

W wariancie (i) najpierw perturbuje się przestrzeń kwaternionów przez zaburzenie trzech składowych q_{Offset} , q_{Noise} oraz q_{Drift} kwaternionu orientacji, za co odpowiada algorytm losowania kątów i składania ich do wypadkowego q_{per} . Następnie z perturbowanej orientacji q_{per} syntezyzuje się sygnał żyroskopu z różniczki kwaternionu. Potem estymuje się składową grawitacyjną akcelerometru przez zmianę układu odniesienia i usuwa się składową grawitacji. Jest to autorska metoda postępowania i została nazwana przez Doktoranta algebraiczną augmentacją sygnałów IMU.

Aby "dogaścić" rozkład danych uczących i nauczyć model tego, czego w realnych nagraniach brakuje lub rzadko występuje, Autor stosuje syntetyczne generowanie próbek, co pozwala na lepszą generalizację danych modelu klasyfikatora CNN w tzw. scenariuszach cross-day. Scenariusz cross-day to sposób oceny, w którym model jest trenowany na danych z jednego dnia, a walidowany jest na nagraniach tych samych osób, ale z innych dni. Celem jest sprawdzenie, czy system radzi sobie z realną zmiennością różnych czynników „z dnia na dzień”, przy czym takimi czynnikami mogą być np. pozycja sensora, rodzaj obuwia, zmęczenie, pogoda, trasa, itp. Metoda cross-day jest znacznie trudniejsza i bardziej realistyczna niż strategia same-day, gdyż pozwala na wykrywanie zmian rozkładu cech X (covariate shift) między dniami,

czyli wykrywane są zmiany rozkładu cech między zbiorem uczącym a testowym, gdy zależność etykiety od cech $p(y|X)$ pozostaje w przybliżeniu taka sama. Strategia same-day często ukrywa ten efekt, gdyż warunki akwizycji danych są prawie te same.

Są to dojrzałe podejścia świadczące o dobrym przygotowaniu doktoranta do podejmowania tematów wymagających znajomości statystyki, wyższej matematyki i metod uczenia maszynowego. Tematyka i hipoteza są jasno postawione i dobrze uzasadnione bieżącym stanem badań znanym z literatury przedmiotu.

Narracja Autora jest logiczna: od problemu i wymagań aplikacyjnych, przez przemyślane przetwarzanie, aż do dwóch dróg zwiększania skuteczności (augmentacja i generacja) oraz ich świadomą weryfikację na zróżnicowanych zbiorach. Doktorant proponuje wielowariantowe rozwiązania, a tam gdzie są ograniczenia — mówi o nich wprost. To buduje zaufanie do wyników.

Autor stawia tezę i trzy cele badawcze:

Zastosowanie próbek tworzonych w augmentacji lub przez modele generatywne zwiększa skuteczność systemu biometrycznego chodu.

Z punktu widzenia poprawności językowej, treść tej tezy jest niefortunna, bo co oznacza "system biometrycznego chodu", albo "biometryczny chód"? Moim zdaniem lepiej brzmiącym tekstem byłby:

Zastosowanie próbek syntetycznych, wygenerowanych w ramach augmentacji danych lub przy użyciu modeli generatywnych, zwiększa skuteczność systemu biometrii chodu.

Autor stawia sobie również następujące cele:

- zbudowanie systemu IMU (acc+gyro),
 - przeprowadzenie algebraicznej augmentacji bazującej na orientacji,
 - wykorzystanie modelu LSTM-MDN z modelowaniem wielu dystrybucji.
- To model, który łączy pamięć o ciągu zdarzeń z umiejętnością podania kilku możliwych dalszych wersji wydarzeń naraz.

Charakterystyka rozprawy i jej zakres

Rozprawa dotyczy biometrii chodu, która ustalana jest na podstawie sygnałów IMU (akcelerometr + żyroskop) i bada dwa komplementarne sposoby powiększania zbiorów uczących: algebraiczną augmentację (z perturbacją orientacji i fizycznie spójną rekonstrukcją wskazań akceleratora i żyroskopu oraz generację próbek syntetycznych modelami LSTM-MDN.

Algebraiczna augmentacja – najpierw lekko zmienia orientację sensora (Offset/Noise/Drift), a potem przelicza z tego nowe sygnały żyroskopu i modyfikuje dane akcelometru. Dzięki temu można zasymulować np. inny montaż telefonu, drobne drgania czy powolny przechył ciała podczas chodu.

Strategia LSTM-MDN generuje syntetyczne okna IMU, modelując rozkład możliwych sygnałów. Tak utworzone próbki wzbogacają dane uczące klasyfikatora.

Celem obu podejść jest podniesienie skuteczności systemu w scenariuszu cross-day, w którym trening i test pochodzą z różnych dni.

Praca doktorska podzielona jest na rozdziały.

Rozdział 1: wprowadzenie, wyzwania biometrii chodu, przegląd literatury, teza i cele, opis korpusów precyzyjnie charakteryzujących zbiory danych użyte w eksperymentach. Czytelnik uzyskuje wiedzę co to za dane, skąd pochodzą, jak są zbudowane oraz w jakich warunkach zostały zebrane i jak zostały użyte (trening/walidacja/test). Przedstawiono także architektury CNN oraz opis sprzętu (od stacji roboczej po klastrer GPU NVIDIA A100/H100). Jasno podano, że cross-day to celowy, trudniejszy scenariusz oceny.

Rozdział 2: obejmuje szczegółowy opis autorskiej augmentacji algebraicznej, ze schematami blokowymi oraz przykładami wpływu Offset/Noise/Drift na przebiegi IMU, wraz opisem baz danych użytych w testach.

Rozdział 3: przedstawia modele generatywne, ze szczególnym naciskiem na LSTM-MDN. Autor omawia przetwarzanie wstępne danych, dobór funkcji kosztu oraz strategię próbkowania składowych rozkładu, aby uzyskać stabilne i wiarygodne sekwencje.

W *Rozdziale 4* Autor odpowiada na pytanie ile sesji treningowych faktycznie pomaga w podniesieniu skuteczności identyfikacji osób w warunkach zmienności między dniami. Autor ustalił to na podstawie porównań scenariuszy z jedną sesją treningową oraz z dwiema sesjami. Np. uczenie na danych z sesji 1(2), a test na danych z sesji 3. Doktorant wykazał, że dołożenie drugiej sesji wyraźnie poprawia wynik, bo model widzi większą różnorodność międzydniową już na etapie uczenia.

W *Rozdziale 5* Autor bada wpływ podłoża (nawierzchni) na wyniki: różne powierzchnie wprowadzają przesunięcie rozkładu cech, co naturalnie obniża jakość rozpoznawania chodu. Doktorant pokazuje, że transformacja do wspólnego układu odniesienia oraz dopełnienie danych (augmentacja algebraiczna i/lub próbki syntetyczne) łagodzą spadki jakości rozpoznawania osoby.

W *Rozdziale 6* Doktorant analizuje lokalizację sensora (np. mocowanie na ciele vs noszenie w kieszeni) i wynikających z tego konsekwencji dla stabilności sygnału. Wszystkie eksperymenty zostały zaprojektowane tak, aby unikać nakładania okien między zbiorami danych i generować raporty z wielu powtórzeń eksperymentów.

Rozdział 7 przenosi rozwiązania do wirtualnej rzeczywistości. Autor wykorzystuje dane o położeniu i ustawieniu ręcznych kontrolerów (trzymanych w dłoniach urządzeń sterujących) oraz hełmu/gogli zakładanych na głowę. Na tej podstawie oblicza sygnały zastępcze akcelerometru i żyroskopu — tak jak gdyby urządzenia miały wbudowane czujniki przyspieszeń i prędkości obrotowych. Następnie sprawdza, czy praca na tak przeliczonych sygnałach poprawia rozpoznawanie osób, gdy zbiory uczące i testowe pochodzą z róż-

nych dni.

Całość pracy została dobrze osadzona w reprezentatywnej literaturze przedmiotu. Bibliografia jest bogata i zawiera najważniejsze źródła literaturowe dotyczące technik biometrycznych. Są to zarówno prace konferencyjne jak i prace publikowane w czasopismach. W dziedzinie biometrii Autor opublikował siedem prac w czasopismach, gdzie w sześciu był pierwszym autorem publikacji. Doktorant brał także udział w dwudziestu pięciu krajowych i zagranicznych konferencjach naukowych, gdzie prezentował wyniki swoich badań. Jest także współautorem pięciu rozdziałów monograficznych.

Praca stanowi ważny wkład w rozwijanie nowych technik biometrii chodu, co przekłada się na skuteczne i wiarygodne rozpoznawanie osób na podstawie ich sposobu poruszania się. Proponowane podejścia jest skalowalne i potencjalnie może również znaleźć zastosowania w ortopedii i rehabilitacji przy ocenie postępów leczenia pourazowego.

Problem naukowy rozprawy

Celem pracy jest wiarygodne rozpoznawanie osób na podstawie chodu, korzystając z danych z czujników ruchu IMU (akcelerometr + żyroskop) umieszczonych np. w telefonie lub opasce. To nie jest tylko kolejny sposób klasyfikacji. W praktyce człowiek nie chodzi identycznie każdego dnia. Zmienia się buty, tempo, nawierzchnia, sposób włożenia telefonu do kieszeni, a nawet samopoczucie. Te różnice sprawiają, że system, który świetnie działał wczoraj, może zawieść jutro.

Problem naukowy brzmi więc: *jak zbudować i badać taki system biometryczny, który będzie skuteczny i odporny na zmienność między dniami (tzw. scenariusz cross-day), mimo ograniczonej ilości danych i wielu źródeł zakłóceń*. Jest to zadanie nietrywialne, gdyż należy uwzględnić wiele czynników mających wpływ na wynik pomiarów:

1. Rozkład cech sygnału zmienia się z dnia na dzień, choć osoba pozostaje ta sama. Klasyczny model uczony na danych w Dniu nr 1 może tracić na jakości dla danych w Dniu nr 2 lub kolejnych.
2. Zbiory danych z IMU są stosunkowo niewielkie i są nie zrównoważone, trudno więc uchwycić różnorodność chodu.
3. Podłoże (asfalt, trawa), prędkość chodu, zmęczenie, a przede wszystkim pozycja i orientacja sensora (np. luźno w kieszeni vs ciasno przy udzie) – to wszystko istotnie zmienia sygnał.
4. Proste dodawanie szumu do danych bywa niefizyczne: akcelerometr i żyroskop przestają generować spójne sygnały, co ma duży wpływ na wskazania klasyfikatora.
5. Należy tak zaplanować, i zrealizować eksperymenty, aby odzwierciedlały rzeczywistą odporność systemu na perturbacje danych.

W rozprawie doktorskiej postawione są więc następujące pytania badawcze:

1. Czy algebraiczna (a nie losowa) augmentacja sygnałów IMU poprawia skuteczność w scenariuszu cross-day?
2. Czy próbki syntetyczne generowane przez LSTM-MDN poprawiają wyniki; w jakiej liczbie i konfiguracji przynoszą największą korzyść?
3. Jak na skuteczność wpływają: liczba sesji treningowych, rodzaj nawierzchni i lokalizacja sensora?
4. Czy przekształcenie danych pozycji i orientacji kontrolera VR na sygnały przyspieszeń i prędkości kątowych (jak z akcelerometru i żyroskopu) prowadzi do bardziej skutecznego i stabilnego rozpoznawania osób w VR (między różnymi dniami)?

Wyżej wymienione pytania mają znaczenie praktyczne, gdyż w realnym użyciu (telefon, zegarek, opaska, kontrolery VR) warunki pomiarów nigdy nie są stałe. Jeśli system ma działać wiarygodnie, musi być odporny na zmiany w czasie. Rozprawa pokazuje, że:

Fizycznie spójne dopełnianie danych (augmentacja) i świadome generowanie danych syntetycznych (LSTM-MDN) zmniejszają wrażliwość systemu biometrycznego na różnice między dniami. Autor wykazał ponadto, że nawet w innej domenie (VR) te same idee uzyskują dobrą skuteczność, gdy zamiast pozycji i kątów kontrolera VR użyjemy, po przeliczeniu, sygnałów jak z IMU.

Rozprawa identyfikuje lukę związaną z brakiem skutecznych, odpornych metod biometrycznych, które nie opierają się wyłącznie na analizie skuteczności klasyfikatora, lecz również na strukturze danych wejściowych. Przedstawiony w pracy problem naukowy jest ujęty w sposób ścisły i został poparty formalną analizą matematyczną oraz empiryczną weryfikacją, co uzasadnia wagę prowadzonych badań.

Ocena wyników przedstawionych w rozprawie

Autor rozprawy, Pan Aleksander Sawicki, konsekwentnie wykazuje, że odpowiednie uzupełnianie danych — przez algebraiczną augmentację oraz generowanie syntetyków — podnosi skuteczność identyfikacji osób w realistycznym układzie cross-day (trening i test w różne dni). Kluczowe są też dwie decyzje inżynierskie: transformacja do globalnego układu odniesienia oraz dobór filtracji, które same w sobie mocno poprawiają bazę danych pomiarowych przed dodaniem danych syntetycznych. Transformacja globalna ustawia dane tak, jakby telefon był zawsze tak samo obrócony. Dzięki temu porównania dotyczą zawsze to samego. Użycie filtrów dolnoprzepustowych wygładza sygnał i usuwa przypadkowe drgania, pozostawiając tylko ślad (rytm) chodu.

Rozpoznanie osoby przeprowadzane jest z wykorzystaniem konwolucyjnych sieci neuronowych (CNN), które są właściwym klasyfikatorem w pracy.

Autor wykorzystuje trzy warianty: bazowy (Vanilla) CNN, CNN z mechanizmem uwagi oraz Multi-Input CNN. Autor zaproponował architekturę CNN w konfiguracji 6×128 (3 osie akcelerometru + 3 osie żyroskopu) i 128 kroków czasowych. Warstwa wyjściowa ma liczbę neuronów równą liczbie osób, więc sieć uczy się identyfikacji tożsamości. W części VR porównywano te same trzy klasyfikatory; najlepsze wyniki zwykle dawał Multi-Input CNN, a w analizie długoterminowej użyto też syjamskiej sieci CNN (wariant do porównań par próbek). Autor podkreśla laboratoryjny charakter autorskiego zbioru danych, co jest uczciwym zastrzeżeniem.

Ze szczegółowej obserwacji eksperymentów wynika, że:

1. Augmentacja algebraiczna (IMU) na autorskim zbiorze 100-osobowym poprawia skuteczność klasyfikatora CNN z ok. $F1=0,68$ do $F1=0,86$ w układzie lokalnym; po przejściu do układu globalnego wynik osiąga wartość $F1=0,97$ (*cross-day*). To bardzo duży, praktyczny skok.
2. Model LSTM (Long Short-Term Memory) — rekurencyjna sieć neuronowa do analizy szeregów czasowych — z modułem wyjściowym MDN (Mixture Density Network), który wyznacza parametry mieszaniny rozkładów normalnych (wagi, średnie, odchylenia), służy tu do generowania losowych, lecz realistycznych próbek IMU. MDN najlepiej wypada przy 1–2 komponentach (dla jednego komponentu $F1=0,789$), podczas gdy 3–4 prowadzą do przeuczenia. Po dopasowaniu filtracji sygnału (dolnoprzepustowy filtr Butterwortha trzeciego rzędu) i strategii losowania próbek, autor uzyskuje wartości współczynnika $F1=0,975$, co jest najlepszym wynikiem w scenariuszu *cross-day* na zbiorze autorskim.
3. Zbadano także autoenkoder wariacyjny (VAE) do generowania sztucznych danych; przyniósł on jedynie niewielką poprawę.
4. Dla strategii VR (z różnymi kontrolerami), zamiana „pozycja+orientacja” na „przyspieszenia + prędkości kątowe” (pseudo-IMU) wyraźnie pomaga: np. $0,80 \rightarrow 0,82$ $F1$ (Cosmos), $0,77 \rightarrow 0,82$ (Quest), $0,83 \rightarrow 0,92$ (Vive). Interpretacja jest spójna i unika się kłopotów z reprezentacją kątów Eulera.

Przedstawione w rozprawie wyniki są spójne, uzasadnione i świadczą o wysokim poziomie dojrzałości badawczej. Przeprowadzone eksperymenty potwierdzają, że proponowane rozwiązania umożliwiają inteligentne dostosowanie modelu do zmieniających się warunków rzeczywistych w sposób efektywny obliczeniowo. Tym samym wyniki uzyskane w rozprawie stanowią istotny wkład w rozwój metod biometrycznych, w tym metod związanych z biometrią chodu.

Znaczenie dokonań Autora dla rozwoju dyscypliny

Znaczenie dokonań mgra inż. Aleksandra Sawickiego dla rozwoju dyscypliny naukowej, jaką jest Informatyka Techniczna i Telekomunikacja, można ocenić jako istotne i wielowymiarowe, zwłaszcza w kontekście analizy danych obarczonych szumem i zmiennością, pochodzących z czujników takich jak akcelerometr czy żyroskop. Autor zaproponował kilka oryginalnych metod analizy takich danych, które poszerzają dotychczas poznany aparat badawczy. Proponowane przez Autora metody mogą znaleźć zastosowanie w systemach medycznych jako dodatkowe elementy diagnozy i monitorowania pacjentów czy w systemach biometrycznego rozpoznawania osób.

Krytyczna ocena wyników

Chociaż sama praca jest wyraźnym osiągnięciem naukowym, to wyniki są często prezentowane jako średnie lub mediany z wielokrotnych powtórzeń oraz na wykresach pudełkowych. Brak jednak formalnych testów istotności i przedziałów ufności. Dobrym uzasadnieniem tej opinii są wyniki dla filtrów 3 Hz vs 6 Hz. Autor raportuje m.in. średnią miarę F1 i 20-krotne powtórzenia oraz pokazuje wykresy pudełkowe, ale nie przeprowadził formalnego, twardego wnioskowania statystycznego np. testem dla par Wilcozona lub t -testem, jeśli spełnia warunki. Wtedy nie tylko widać ile wynosi F1, ale też na ile można temu wynikowi ufać, czy różnice są prawdziwe, ile to kosztuje obliczeniowo i kiedy model zawodzi. To sprawia, że wnioski byłyby bardziej wiarygodne dla ewentualnych działań wdrożeniowych. W tego typu pracach agreguje się wyniki do poziomu osoby (np. średnia, mediana) i testuje się na N osobach. To najczęstsza, metoda analizy wyników w biometrii behawioralnej.

W biometrii warto dodatkowo raportować EER (Equal Error Rate), czyli punkt, w którym odsetek fałszywych akceptacji równa się odsetkowi fałszywych odrzuceń i krzywe ROC (Receiver Operating Characteristic), czułość vs. fałszywe alarmy.

W pracy pominięto obliczenia dla wskaźników ROC, AUC, DET i EER. Dlaczego? To ważne parametry we wszelkiego typu analizach biometrycznych. Trzon rozprawy to rozpoznawanie (identyfikacja) osób na podstawie chodu z sygnałów IMU (akcelerometr + żyroskop). Autor buduje następujący schemat postępowania: wyrównanie do globalnego układu odniesienia + filtracja + klasyfikator CNN uczony na oknach 6×128 . Model LSTM-MDN pojawia się jako generator próbek syntetycznych do wzmocnienia treningu. Dodatkowo sprawdzono to w modelu VR, przeliczając pozycję i orientację kontrolerów na „pseudo-IMU” i znów wykonując identyfikację osób.

Można było policzyć także dokładność ACC (Accuracy) dla klasyfikatora CNN. Dla każdego okna 6×128 sieć CNN zwraca przecież wektor P o długości c , gdzie c to liczba osób w zbiorze. Element $p_c \in P$ to szacowane prawdopodobieństwo, że okno należy do osoby c . Biorąc $\argmax$ wektora P uzyskujemy informację o najbardziej prawdopodobnej osobie. W biometrii najuczciwiej liczyć ACC *per osoba*, czyli dla każdej osoby agregować decyzję z

wielu okien (np. głosowanie większościowe lub średnia prawdopodobieństw), a potem policzyć średni odsetek trafień po osobach. To najlepsze podejście w biometrii behawioralnej.

Te zabiegi, chociaż są pracochłonne, znacznie podniosłyby walory pracy. Dlaczego tego nie zrobiono? Liczę na dyskusję na ten temat w trakcie obrony.

Nie wiem też dlaczego Autor nie pokusił się na wykonanie eksperymentu porównującego metodę VR (z pseudo IMU) z metodą LSTM-MDN na tym samym klasyfikatorze CNN z tymi samymi hiperparametrami. Można to było zrobić w dwóch wersjach (1) LSTM-MDN vs VR oraz (2) LSTM vs VR czyli bez danych syntetycznych, żeby sprawdzić, czy różnice pochodzą z jakości syntetyków, czy nie. VR był przecież tylko poligonem doświadczalnym dla tych samych idei co w IMU z telefonu komórkowego. Taki eksperyment zapewniłby czyste porównanie: identyczny CNN, te same pseudo-IMU i metoda; różniłyby się tylko próbki syntetyczne z LSTM-MDN. Pozwoliłby też oszacować, ile próbek syntetycznych na osobę potrzeba przy danym czasie nagrania i częstotliwości próbkowania.

Jeśli okno ma długość L , s jest krokiem przesunięcia okna, a sygnał ma częstotliwość próbkowania F_s [Hz] i trwa T [s], to liczba rzeczywistych (nie-syntetycznych) próbek treningowych przypadających na jedną osobę wynosi:

$$N_{\text{real}} = \left\lfloor \frac{F_s \cdot T - L}{s} \right\rfloor + 1.$$

Sugerowane, przykładowe warianty augmentacji MDN:

$$r \in \{0.5, 1, 2\},$$

gdzie r to proporcje liczby danych syntetycznych do liczby realnych próbek na osobę.

Wtedy liczba syntetycznych próbek na osobę wyniesie:

$$N_{\text{synth}} = \text{round}(r \cdot N_{\text{real}}).$$

Ta informacja pozwala na obserwację, czy dane rzeczywiste są mniej liczne niż dane syntetyczne dla jakiejś osoby. Taki pomiar ma oczywiście sens tylko wtedy, kiedy wystąpi niejednakowa liczba realnych okien lub różne czasy nagrań.

Dla IMU autor raportuje głównie liczbę próbek(cykli) na dzień (rzędu 3,8–4,4 tys.), a nie łączny czas nagrań.

Wmienione powyżej uwagi mają wyłączny charakter polemiczny, zachęcający do dyskusji w trakcie obrony. Moja wysoka ocena merytorycznego wkładu Autora w rozwój dyscypliny Informatyki Technicznej i Telekomunikacji pozostaje niezmieniona, co podkreślałem także w innych partiach mojej recenzji.

Ocena bibliografii

Bibliografia jest w przeważającej mierze trafnie dobrana i dobrze osadza rozprawę w literaturze przedmiotu. Autor czerpie zarówno z klasycznych jak i przeglądowych prac o biometrii chodu (m.in. przeglądy i ujęcia „smartphone - IMU”) jak i z pozycji nowszych — to daje szeroką i aktualną perspektywę problemu rozpoznawania osób na podstawie sygnałów z czujników inercyjnych. Na plus zaliczam szczególnie obecność przekrojowych opracowań o biometrii chodu i IMU [1–4], włączenie prac o biometrii w środowiskach VR [5], uwzględnienie literatury dotyczącej orientacji/układu odniesienia i modelowania sygnałów [8] oraz rzetelne odwołania do publicznych opisów zbiorów danych (Signet, 200GaitData itd.) wraz z adresami i datami dostępu [31–36]. W bibliografii są też cytowania najnowszych prac aplikacyjnych o rozpoznawaniu chodu (np. z EAAI 2024) i lekkich architekturach CNN dla IMU [38–39], a także pozycje opisujące komponent LSTM-MDN wykorzystywany w rozprawie (np. dla generowania pisma odręcznego) [17].

Wydzielono również osobno publikacje autorskie Doktoranta, co przejrzysto odcina własne dorobki od literatury kontekstowej. Publikacje autorskie są merytoryczne i dobrze wspierają tezę rozprawy, a konferencje dokumentują drogę do wyników końcowych.

Jednocześnie zauważam miejsca dla ewentualnego przyszłego wzmocnienia: Rozprawa cytuje zastosowania LSTM-MDN (Graves) [17], ale nie widać klasycznej pozycji wprowadzającej same Mixture Density Networks (Bishop, 1994). W części o VAE pojawia się wpis popularno-naukowy (blog) [43], przydatny ilustracyjnie, lecz brakuje cytowania pracy źródłowej nt. Auto-Encoding Variational Bayes (Kingma & Welling). Uzupełnienie tych pozycji zwiększyłoby rygor metodologiczny wywodu.

Odwołania do VGG/ResNet są prowadzone przez dokumentację PyTorch [40–41]; warto dodać artykuły źródłowe Simonyan & Zisserman (VGG) oraz He et al. (ResNet), aby odwołania nie ograniczały się do dokumentacji narzędziowej.

Zagadnienie przesunięcia dystrybucji danych między dniami jest szeroko omawiane w treści (m.in. w kontekście cross-day), ale nie widać klasycznych odnośników do literatury o dataset i covariate shift, np. monografie lub prace przeglądowe. Wskazanie kanonicznych źródeł ugruntowałoby tę oś argumentacji.

Przy opisie przejścia do układu globalnego i filtracji częstotliwościowej dobrze byłoby dołożyć rozpoznawalne w środowisku źródła o estymacji orientacji i filtrach stosowanych w czujnikach inercyjnych (obok pozycji polskojęzycznej [8]). Obecnie ten fragment bazuje głównie na opisie własnym Autora.

W kilku miejscach bibliografia odwołuje się do materiałów mniej ważnych (blog VAE [43], repozytorium projektu [45]) lub dokumentacji frameworków [40–41]. To nie jest błąd, ale dobrze byłoby, gdzie to możliwe, dołączyć również recenzowane publikacje odpowiadające tym treściom.

Punktowe uzupełnienia mają charakter wzmacniający, bez wpływu na

ogólnie pozytywną ocenę doboru literatury.

Wnioski końcowe

Podsumowując przedstawioną powyżej szczegółową charakterystykę rozprawy doktorskiej Pana mgr inż. Aleksandra Sawickiego, stwierdzam, że praca spełnia wymagania stawiane pracom doktorskim z obszaru nauk technicznych w myśl odpowiednich przepisów ustawowych, tj. Ustawy prawo o szkolnictwie wyższym i nauce, Dz. U. 2024, poz. 1571.

Wnoszę o dopuszczenie opiniowanej rozprawy doktorskiej do dalszych etapów przewodu doktorskiego.

Kierując się równocześnie dojrzałą koncepcją rozwiązań zaproponowanych w pracy doktorskiej, szerokim zestawem eksperymentów i dużą liczbą dobrej jakości opublikowanych prac zarówno w czasopismach jak i materiałach konferencyjnych, **wnioskuję o wyróżnienie pracy doktorskiej Pana mgr inż. Aleksandra Sawickiego.**