Rozdział 2 Biometryczna identyfikacja tożsamości na podstawie chodu

Aleksander Sawicki

Streszczenie: W artykule zawarto wyniki oryginalnych badań nad zastosowaniem splotowych sieci neuronowych w identyfikacji osób na podstawie chodu. W pracy wykorzystano ogólnodostępny korpus danych, zawierający sygnały sensorów IMU (ang. *Inertial Measurement Unit*) takie, jak przyspieszenie czy natężenia pola magnetycznego.

W przeprowadzonych badaniach nieprzetworzone wartości pomiarowe sensorów (ang. *raw data*) podano bezpośrednio na wejście klasyfikatorów CNN (ang. *Convolution Nerual Network*). Jednocześnie zastosowano technikę oddzielania zerami (ang. *zero-padding*) do przedzielenia danych pochodzących z sensorów różnego typu. Podkreślić należy, iż podejście takie stosowane było wcześniej, przede wszystkim w zagadnieniach rozpoznawania czynności (ang. *Human Activity Recognition*). Dużą wartość dodaną pracy przedstawiają operacje w dziedzinie kwaternionów, umożliwiające wyznaczanie prędkości kątowych na podstawie sygnałów orientacji. Wykorzystanie sygnałów prędkości kątowych pozwoliło na osiągnięcie bardzo dobrych rezultatów klasyfikacji w finalnym systemie biometrycznym.

Zastosowanie techniki oddzielenia zerami (ang. *zero padding*) pozwoliło osiągnąć obiecujące wskaźniki identyfikacji osób w zagadnieniu biometrii behawioralnej. Osiągnięte wyniki były porównywalne z rezultatami opisywanymi w literaturze.

Słowa kluczowe: Biometria, Klasyfikacja, IMU

Wprowadzenie

Zadaniem systemów biometrycznych jest rozróżnianie osób na podstawie wybranych cech fizjologicznych lub behawioralnych. Badania nad rozwojem i wdrażaniem tego typu rozwiązań przyciągają uwagę zarówno środowisk akademickich, jak i przemysłowych z uwagi na rosnące zapotrzebowanie identyfikacji i weryfikacji osób [1]. Systemy biometryczne bazujące na analizie chodu, w przeciwieństwie do tych wykorzystujących dane wejściowe w postaci m.in: zdjęć tęczówki oka, żył dłoni czy odcisku palca, mogą działać w sposób bardziej dyskretny [2]. W ostatnim czasie zaobserwować można rozwój tej szczególnej dziedziny biometrii behawioralnej, właśnie ze względu na brak konieczności aktywnej kooperacji z uczestnikiem badań [3].

W rozdziale przedstawiono rezultaty pracy nad systemem biometrycznej identyfikacji osób na podstawie chodu. Zaproponowana metodologia bazuje na przetwarzaniu nieprzetworzonych wartości pomiarowych sensorów ruchu IMU (ang. *Inertial Measurement Unit*) za pomocą sztucznych splotowych sieci neuronowych CNN (ang. *Convolution Neuranl Network*).

Prezentowane eksperymenty przeprowadzono z wykorzystaniem ogólnodostępnej bazy chodu ludzkiego, zgromadzonej w uniwersytecie McGill [4]. Korpus danych zawiera zestaw nagrań ruchu dwudziestu uczestników badań. Proces akwizycji danych przeprowadzony został z wykorzystaniem telefonu typu smartphone, wyposażonego w czujniki takie, jak akcelerometr i magnetometr.

Klasyfikacja chodu ludzkiego przeprowadzona została z wykorzystaniem sieci neuronowych CNN. Na wejścia klasyfikatora poddano nieprzetworzone wielkości pomiarowe sensorów. Dane te oddzielone zostały blokami zerowych wartości, aby nie doszło do ich interferencji w pierwszej warstwie splotowej. W tym celu zastosowano technikę uzupełniania zerami (ang. *zero padding*), [5]. W wyniku przeprowadzonych badań zaproponowano architekturę systemu biometrycznego, umożliwiającego identyfikację uczestników badań na poziomie około 90%.

Dużą wartość dodaną pracy stanowi zestaw procedur w dziedzinie algebry kwaternionów, umożliwiających odtworzenie sygnałów prędkości kątowej na podstawie sygnałów orientacji. Rozwiązanie tego typu zwiększyło skuteczność opracowanego systemu biometrycznego.

2.1. Stan wiedzy

Przegląd stanu wiedzy dotyczył kilku pomniejszych zagadnień i został podzielony na następujące punkty: analiza chodu z wykorzystaniem sieci CNN, opis wykorzystanej bazy danych, zestawienie skuteczności systemów biometrycznych innych zespołów naukowych, metody odtwarzania prędkości kątowej, metody analizy danych różnych grup sensorów przez sieci splotowe.

2.1.1. Metody analizy chodu z wykorzystaniem sieci CNN

Zagadnienia związane z biometrią behawioralną w znacznie większej mierze dotyczą analizy ludzkiego chodu niż jakichkolwiek innych ruchów czy też gestów. Współcześnie wyodrębnić można dwa główne nurty dotyczące zagadnień analizy ruchu z wykorzystaniem sztucznych splotowych sieci neuronowych. W pierwszym z nich autorzy wykorzystują wyniki różnorodnych transformat, doprowadzających do powstania "obrazów" dwuwymiarowych. W publikacji [3] wykorzystano nowatorską technikę nazwaną AE-GD (ang. *Angle Embedded Gait Dynamic Image*). Polegała ona na wyznaczeniu macierzy dwuwymiarowych na podstawie sygnału magnitudy akcelerometru. W przypadku publikacji [6] autorzy zastosowali transformatę czasowo-częstotliwościową w postaci Dystrubucji Wignera-Villa. Odmienne aczkolwiek bardzo podobne podejście zaprezentowali autorzy [7] wykorzystując scalogramy – wyniki transformaty falkowej.

W drugiej grupie rozwiązań zamieścić można rozwiązania bazujące bezpośrednio na wartościach pomiarowych IMU. Reprezentantem tej grupy może być publikacja [8], w której autorzy wykorzystali surowe wartości akcelerometru i żyroskopu oraz sieci splotowe w formie ekstraktora danych. Z kolei w publikacji [9] wykorzystano architekturę sieci neuronowej o wielu wejściach w celu identyfikacji osoby wykonującej wybrane gesty.

2.1.2. Baza danych chodu Uniwersytetu McGill

Opisywane w rozdziale badania dotyczą zagadnienia biometrii behawioralnej ludzkiego chodu. W przeprowadzonych eksperymentach wykorzystano ogólnodostępną bazę [4], zgromadzoną w uniwersytecie McGill. W tworzeniu korpusu danych wzięło udział 20 uczestników, odpowiednio 10 kobiet i 10 mężczyzn. Poproszeni zostali oni o udział w dwóch około 15-minutowych sesjach nagrywania chodu, które odbywały się w różnych dniach. Podkreślić należy, iż na potrzeby opisywanych w rozdziale badań wykorzystano jedynie dane zarejestrowane podczas pierwszej z nich.

Akwizycja danych przeprowadzona została z wykorzystaniem sensorów ruchu wbudowanych w telefon komórkowy typu smartphone (urządzenia HTC Nexus One). W celu odczytu danych pomiarowych czujników ruchu opracowano oraz wykorzystano autorskie oprogramowanie. Z uwagi na akwizycję danych z udziałem sensorów działających pod systemem operacyjnym, akwizycja danych przebiegała z częstotliwością próbkowania wynoszącą 27,52 \pm 4,07 Hz.

Wśród udostępnionych w korpusie danych sygnałów znajdują się::

- znacznik czasowy (ang. *timestamp*);
- wartości pomiarowe akcelerometru (odpowiednio osie OX,OY,OZ oraz magnituda);
- wartości pomiarowe magnetometru (odpowiednio osie OX,OY,OZ);
- informacje o orientacji telefonu w postaci kątów Eulera tj; Yaw, Pitch Roll.

Osoby biorące udział w eksperymencie proszone były o typowe dla nich użytkowanie telefonu komórkowego. W związku z czym w zależności od preferencji uczestnika, sensory umieszczone były w prawej lub lewej oraz przedniej lub tylnej kieszeni spodni. Jednocześnie osoby biorące udział w eksperymencie w dowolnym momencie mogły wyjąć telefon, a następnie włożyć go z powrotem, niekoniecznie zachowując oryginalne ułożenie urządzenia. Nagrania zawierają również fragmenty małej aktywności (bezruchu), w czasie których czujniki ruchu nie wykrywały znaczących zmian. Domniemywać można, że w tych okresach uczestnicy badań np. rozmawiali z osobami spotkanymi na terenie kampusu.

Finalnie zaznaczyć należy, iż poszczególne osoby posiadały różne typy obuwia, chód ich zarejestrowany został na terenie kampusu akademickiego, przy czym nie zachowywali oni stałej ścieżki spacerów. Ponadto, proces akwizycji danych odbył się z udziałem czterech różnych egzemplarzy telefonów komórkowych.

2.1.3. Porównanie rezultatów stosowania innych systemów biometrycznych

Korpus danych wykorzystywany w pracy był kilkukrotnie używany w publikacjach naukowych z uwagi na swój ogólnodostępny charakter.

W pozycji [3] autorzy opisali opracowany algorytm generowania form pseudoobrazów *GDS* (ang. *Gait Dynamics Images*) na podstawie sygnałów pomiarowych trzech osi akcelerometrów. Prezentowane metody zakładały całkowite pominięcie innych sygnałów dostępnych w bazie. Prezentowany system biometryczny wykorzystywał grafiki wygenerowane na podstawie okien przesuwnych o długości 2,4 sekund. Do oceny podobieństwa sygnałów testowych oraz wzorcowych wykorzystano sumaryczne szczytowe współczynniki korelacji. W pracy wykorzystano klasyfikator w postaci algorytmu najbliższych sąsiadów. Ocena jakości identyfikacji osób odbyła się na podstawie wyciętych bloków danych o długości 50 sekund przy wykorzystaniu okien przesuwnych o długości 2,4 sekund oraz nakładkowaniu wynoszącym 50%. Jako miarę jakości klasyfikacji wykorzystano skuteczność przy zastosowaniu prostej walidacji.

Z kolei w pracy [3] zaproponowano zmodyfikowany sposób generowania obrazów nazwany jako AE-GDI (ang. *Angle Embedded Gait Dynamic*). Przedstawiona metoda również bazowała na graficznej reprezentacji wygenerowanej na podstawie sygnałów pomiarowych akcelerometru. W przeciwieństwie do poprzedniej pozycji [4] autorzy zaimplementowali algorytm detekcji poszczególnych cykli chodu. Opracowany system biometryczny wykorzystywał grafiki wygenerowane na podstawie sygnałów pojedynczego cyklu chodu lub próbek dwóch sklejonych przebiegów. W pracy wykorzystano klasyfikator w postaci sieci splotowej CNN. Testowanie przebiegło w wysublimowanych warunkach. Z dostępnych w ciągu jednego dnia przebiegów czasowych wyznaczano blok o długości 10% całości, stanowiący zbiór treningowy. Następnie dokonywano zaawansowanej segmentacji cykli chodu oraz generowano grafiki. Zbiór testowy był generowany losowo na podstawie niewykorzystano klanych i posiadał tę samą długość co zbiór treningowy (tj. 10% całości). Proces ten powtarzany był dziesięciokrotnie. Jako miarę jakości klasyfikacji wykorzystano skuteczność. Porównanie jakości klasyfikacji poszczególnych metod zaprezentowano w tabeli 2.1.

forma danych wejściowych	Metoda	Skuteczność [%]
bloki danych	Bazowa [4] (oparta na magnitudzie akcelerometru)	67,5
bloki danych	GDI [4]	87,5
bloki danych	Zmodyfikowana metoda GDI [4] 85,5	
cykle chodu	AE-GDI (pojedynczy cykl chodu) [3] 95,4	
cykle chodu	AE-GDI (podwójny cykl chodu) [3] 96,6	

TABELA 2.1. Porównanie skuteczności systemów biometrycznych bazujących na bazie [4]

2.1.4. Odtworzenie prędkości kątowej

Współcześnie, z uwagi na wzrost aplikacji metod sztucznej inteligencji z dziedziny uczenia głębokiego, szczególnie wymagających licznych zbiorów uczących, zaobserwować można zwiększenie liczebności prac naukowych w dziedzinie augmentacji danych, jak i generowania próbek syntetycznych. Nurt ten widoczny jest również w dziedzinie przetwarzania sygnałów ruchu.

W pracach [10,11] badacze zaproponowali autorskie rozwiązania umożliwiające odtworzenie sygnałów akcelerometru oraz żyroskopu. W obu tych przypadkach zastosowano modele regresyjne umożliwiające zasymulowanie wskazań sensorów na podstawie animacji szkieletowych, wyodrębnionych z nagrań wideo lub też nagrań kamery głębi. W obu opisywanych algorytmach niezbędne było wykorzystani obu modalności w celu wcześniejszej nauki modeli regresyjnych.

Z kolei w pozycji [12] autorzy zaprezentowali analityczny sposób syntetycznego generowania wskazań magnitudy akcelerometru na postawie informacji o pozycji poszczególnych stawów animacji szkieletowej. W pracy [13] badacze zaproponowali użycie symulatorów IMU do generowania danych syntetycznych na podstawie nagrań wideo zgromadzonych na platformie YouTube. Zastosowana Aplikacja *IMUSim* [14] pozwoliła na odtworzenie serii czasowych wartości pomiarowych trójosiowych akcelerometrów, żyroskopów i magnetometrów na podstawie informacji o orientacji czujnika. Oprogramowanie to jest w stanie w sposób analityczny odtworzyć przebiegi czasowe, jak i zasymulować typowe zakłócenia poszczególnych czujników.

Czerpiąc inspirację z publikacji [13] zdecydowano się na odtworzenie brakującego sygnału z wykorzystaniem metod analitycznych, odpowiadających stosowanym w symulatorach IMU. Podejście takie nie wymagało nauki modeli regresyjnych, a więc wcześniejszej znajomości sygnałów prędkości kątowej (które nie były udostępnione w przypadku omawianej bazy danych). Nadmienić należy, iż podejście takie nie było prezentowane wcześniej dla opisywanego korpusu danych.

2.1.5. Metody łączenia danych

W literaturze podano kilka odmiennych podejść do przetwarzania bloków danych pochodzących z różnych typów sensorów przez sieci splotowe CNN.

W pozycji [8] autorzy wykorzystywali nagrania zarejestrowane za pomocą akcelerometru oraz żyroskopu, zlokalizowanych na nodze do identyfikacji osób. Do klasyfikacji wykorzystano sieć splotową, na której wejście podawano bloki danych o wymiarach $8 \times N$. Pierwszy z wymiarów wynikał z faktu wykorzystania odpowiednio sygnałów pomiarowych trójosiowych akcelerometru oraz żyroskopu oraz ich magnitud. Liczba N oznaczała liczbę próbek sygnału.

Z kolei w publikacji [5] autorzy wskazują, że mieszanie się nieprzetworzonych danych pomiarowych różnego typu (tj. pochodzących z sensorów mierzących odmienne wartości fizyczne) w pierwszej warstwie splotowej jest niepożądane. W pracy zaproponowano wstawianie bloków danych o zerowych wartościach (*tzw. Zero-padding technique*) pomiędzy zarejestrowane sygnały pomiarowe. Nadmienić należy, że wartości zerowe wstawiane były wyłącznie pomiędzy dane zarejestrowane przez sensory różnego typu. Przykładowo, pomiędzy sygnały pomiarowe dwóch akcelerometrów, nie umieszczano bloków wartości zerowych. Autorzy w publikacji argumentowali, iż prezentowane podejście doprowadzi do wyspecjalizowania filtrów splotowych do poszczególnych typów zarejestrowanych wielkości fizycznych.

Całkowicie odmienne podejście w zakresie łączenia danych z różnych typów czujników zaprezentowano w pozycji [9]. Autorzy przeprowadzili badania identyfikacji osób na podstawie wykonywanych przy pomocy dłoni w powietrzu symboli liter (ang. *Air-writing*). W swoich pracach zaproponowali wykorzystanie architektury sieci CNN o kilku wejściach (ang. *multi-input architecture*) do klasyfikacji danych zarejestrowanych z użyciem akcelerometru oraz żyroskopu. W prezentowanym podejściu dochodzi do wyspecjalizowania filtrów splotowych do poszczególnych typów danych. W przypadku architektury tego typu należy szczegółowo określić liczbę filtrów splotowych wykorzystywanych dla poszczególnych typów wartości pomiarowych.

2.2. Metodologia

Celem opracowanego systemu biometrycznego była identyfikacja osób na podstawie sygnałów pomiarowych czujników ruchu zarejestrowanych podczas chodu. Wejście systemu stanowiły sygnały ruchu oryginalnie znajdujące się w korpusie danych, wyjście zaś numery identyfikacyjne uczestników badań. Zaproponowany system posiadał następujące bloki: "*przetwarzanie sygnałów*", "*segmentacja*", "*okno przesuwne*" oraz "*decyzja*". Schemat blokowy rozwiązania zaprezentowano na rys. 2.1.

W bloku "*przetwarzanie sygnałów*" dokonywano przekształceń mających na celu zwielokrotnienie oryginalnych sygnałów ruchu. Proces ten miał na celu odnalezienie takiej formy reprezentacji danych, która zapewniłaby wysokie wskaźniki identyfikacji.



RYSUNEK 2.1. Schemat blokowy systemu biometrycznego

W etapie "*segmentacja*" dane poddane zostały procesowi prostej segmentacji. Powodował on usunięcie z nagrań fragmentów, w których uczestnik pozostawał w bezruchu. Proces ten umożliwił dalszą pracę na fragmentach rzeczywiście reprezentujących chód. W bloku "*okno przesuwne*" (ang. *sliding window*) poszczególne nagrania podzielone zostały na ramki danych o stałej długości. Umożliwiło to bezpośrednie podanie ich na wejście algorytmów klasyfikujących. Ponadto, zastosowanie okna przesuwnego z nakładkowaniem wynoszącym 50% zapewniło zwielokrotnienie danych treningowych. Finalnie w bloku "*decyzja*" z udziałem sztucznych głębokich sieci neuronowych typu CNN dokonano identyfikacji osób. Zaproponowane kroki przetwarzania omówione zostały szczegółowo w następnych sekcjach.

2.2.1. Przetwarzanie sygnałów

W prezentowanych eksperymentach zastosowano przetwarzanie sygnałów, które można podzielić na następujące grupy: przekształcenia w dziedzinie kwaternionów, implementacja techniki uzupełniania zerami (przyp. przegląd literatury) oraz filtracja częstotliwościowa.

Przekształcenia w dziedzinie kwaternionów

Wykorzystany w eksperymentach korpus danych w swojej oryginalnej formie nie zawierał informacji o prędkości kątowej. Obejmował on jednak sygnały pochodzące z: akcelerometru, magnetometru, orientacji w postaci kątów Eulera oraz znacznik czasowy. W pracach [13,14] autorzy wykorzystali informacje o orientacji w postaci kwaternionów do symulowania prędkości kątowych. W celu odtworzenia wartości pomiarowych żyroskopu z danych dostępnych w korpusie, niezbędne było dostarczenie informacji o orientacji sensora w postaci kwaternionu.

Zagadnienie to realizowane może być w dwojaki sposób: przez konwersję z reprezentacji kątów Eulera oraz przez analityczne wyznaczenie obrotu pomiędzy dwoma wektorami [15]. Konwersja z postaci kątów Eulera do postaci kwaternionu (jak i przekształcenie odwrotne) jest standardową praktyką, dzięki czemu posiada implementację w pakiecie Aerospace Toolbox środowiska MATLAB. Zagadnienie konwersji pomiędzy reprezentacjami nie zostanie więc szczegółowo opisane w tekście. Drugi stosowany w literaturze [15] sposób wyznaczenia orientacji w postaci kwaternionu bazuje na wyznaczaniu kąta obrotu pomiędzy dwoma wektorami (1):

$$\boldsymbol{a} = \frac{\left(a_{x}, a_{y}, a_{z}\right)}{\|\boldsymbol{a}\|}$$
$$\boldsymbol{b} = \frac{\left(b_{x}, b_{y}, b_{z}\right)}{\|\boldsymbol{b}\|} , \qquad (1)$$
$$\boldsymbol{\alpha} = \operatorname{acos}\left(\langle \boldsymbol{a} \mid \boldsymbol{b} \rangle\right)$$

gdzie:

a, *b* – znormalizowane trójelementowe wektory;

 $\langle a|b\rangle$ – iloczyn skalarny wektorów *a* i *b*;

α – kąt obrotu pomiędzy wektorami.

Oś obrotu wyznaczona jest na podstawie iloczynu wektorowego zgodnie z równaniem (2):

$$\boldsymbol{n} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\sin(\alpha)},\tag{2}$$

gdzie:

 $a \times b$ – iloczyn wektorowy wektorów *a* i *b*.

Finalnie kwaternion reprezentujący obrót pomiędzy wektorami wyznaczony jest zgodnie z równaniem (3):

$$\boldsymbol{q} = \begin{bmatrix} \cos(\alpha/2) & n_x \cdot \sin(\alpha/2) & n_y \cdot \sin(\alpha/2) & n_z \cdot \sin(\alpha/2) \end{bmatrix}, \quad (3)$$

gdzie:

q – kwaternion reprezentujący obrót pomiędzy dwoma wektorami; n_x , n_y , n_z – poszczególne składowe wektora n (osie obrotu), wyznaczone w równaniu (2).

W przypadku przeprowadzanych badań wyznaczano orientację pomiędzy wskazaniami trójosiowego akcelerometru w chwili *t*, oraz globalnym wektorem orientację **G** – (0,0,1). Opisywany sposób pozwolił na wyznaczenie kwaternionu, reprezentującego orientację wyznaczoną jedynie na podstawie sensora akcelerometrycznego.

Seria czasowa kwaternionów reprezentujących orientację wraz ze znacznikiem czasowym posłużyć mogą do odtworzenia sygnałów prędkości kątowej. Proces ten rozpoczęty jest od wyznaczenia różniczki kwaternionu zgodnie z równaniem (4):

$$\dot{\boldsymbol{q}}_{t} = (\boldsymbol{q}_{t+1} - \boldsymbol{q}_{t}) / \Delta T, \qquad (4)$$

gdzie:

 $q_{(t+1)}, q_t$ – orientacja w postaci kwaternionu w chwili *t*+1 oraz *t*; ΔT – różnica w czasie akwizycji próbki *t*+1 oraz *t*;

Kolejnym krokiem jest odtworzenie prędkości kątowej zgodnie z równaniem (5):

$$\omega_t \left(\boldsymbol{q}_t, \boldsymbol{q}_t \right) = 2 \cdot \boldsymbol{W} \left(\boldsymbol{q}_t \right) \cdot \boldsymbol{q}_t, \qquad (5)$$

gdzie:

 ω_t – wektor prędkości kątowych ($\omega_x, \omega_y, \omega_z$) w chwili t;

W – macierz mapująca kwaternion q_t oraz jego różniczkę, do prędkości kątowych wyrażonych we współrzędnych świata.

Wartość macierzy W zależy od wartości kwaternionu q w chwili t. Związek ten opisany jest równaniem (6):

$$\boldsymbol{W}(\boldsymbol{q}_{t}) = \begin{bmatrix} -q_{x} & q_{w} & -q_{z} & q_{y} \\ -q_{y} & q_{z} & q_{w} & -q_{x} \\ -q_{z} & -q_{y} & q_{x} & q_{w} \end{bmatrix},$$
(6)

gdzie:

 q_w, q_x, q_y, q_z – wartość składowych *w*, *x*, *y*, *z* kwaternionu w chwili *t*;

Dokonując właściwych eksperymentów skorzystano z własnoręcznie przygotowanej implementacji procedury wyznaczania prędkości kątowej. Skorzystano w tym celu z otwartych źródeł oprogramowania T-SIM [16]. Oryginalny kod źródłowy nie umożliwiał wyznaczenia prędkości kątowej dla nierównomiernie próbkowanego sygnału orientacji (co ma miejsce w wykorzystywanym korpusie danych).

Technika uzupełniania zerami

Idea jaka przyświeca stosowaniu techniki uzupełniania zerami (ang. Zero-padding technique) [5] jest oddzielenie danych pochodzących z sensorów mierzących odmienne wielkości fizyczne, tak aby nie doszło do ich mieszania w pierwszej warstwie sploto-wej sieci neuronowej CNN.

Podczas prac wykorzystano bloki danych o wartościach "0" oraz wymiarze $3 \times N$, gdzie N oznacza długość ramki danych. Pierwszy z wymiarów wynika z faktu, że dane pomiarowe poszczególnych trójosiowych sensorów mają wymiar $3 \times N$, a zastosowane filtry splotowe wymiar 4×4 . Na rys. 2.2 zaprezentowano porównanie mieszania się danych dla bloków danych o wymiarze A) $3 \times N$, oraz B) $2 \times N$.



RYSUNEK 2.2. Technika uzupełniania zerami o wymiarze a) 3×N, b) 2×N, oraz filtach splotowych o wymiarach 4×4

Na osi Y rys. 2.2 zaznaczono: dane zarejestrowane z użyciem trójosiowego akcelerometru (tj. a_x , a_y , a_z), blok wartości zerowych oraz blok odtworzonych danych prędkości kątowych (g_x , g_y , g_z). Dla filtrów o wymiarze 4×4, w przypadku a) nie dochodzi do przetwarzania przez kernel danych z różnego typu sensorów. W przypadku b) dla ostatniego z kerneli nastąpi przetwarzanie z użyciem danych a_z oraz g_x . W przypadku stosowania trójosiowych sensorów wymiar wstawionych bloków danych powinien więc wynosić co najmniej 3 (dla filtrów splotowych o wymiarze 4×4).

Filtracja częstotliwościowa

W celu wyeliminowania szybkozmiennych zakłóceń (szum biały) w sygnałach pomiarowych akcelerometru [17] zastosowano cyfrową filtrację dolnoprzepustową. Zaimplementowano filtr Butterwortha czwartego rzędu o częstotliwości odcięcia wynoszącej 10 Hz. Podkreślić należy, że wartość ta dobrana była eksperymentalnie i ograniczona była częstotliwością Nyquista.

2.2.2. Segmentacja danych

Proces segmentacji danych może zachodzić w dwojaki sposób. Z jednej strony polegać może na wyselekcjonowaniu pojedynczych cykli chodu [3]. Z drugiej zaś strony dotyczyć może rozróżnienia chodu od sytuacji statycznych [18].

W pracach prezentowanych w rozdziale zastosowano drugą z omawianych metod. Proces wycinania fragmentów bezruchu, czy też małej aktywności, bazowano na inspekcji ręcznej. Zaznaczyć jednak należy, że kody źródłowe oprogramowania dostępne są na platformie Github [19]. Szczególnie istotne z punktu widzenia powtarzalności eksperymentów jest umieszczenie numerów próbek, które uwzględnione zostały w finalnych eksperymentach.

W literaturze popularne jest wykorzystanie informacji o magnitudzie sygnału akcelerometru w celu oceny aktywności uczestnika badań. O ile wskazania poszczególnych osi sensora zależą od jego orientacji [4,17], o tyle wskazania ich normy niosą jednoznaczną informację o ruchu czujnika. Na rys 2.3 zaprezentowano sygnały zarejestrowane podczas krótkotrwałej przerwy w chodzie. Na rys 2.3 a) wykreślono sygnał magnitudy akcelerometru. Za pomocą szarego tła zaznaczono fragmenty uznane za aktywny chód. Fragment sygnału o wartościach około 10 m/s² uznany został za przerwę w aktywności. Wartość ta wynika z rejestracji przyśpieszenia grawitacyjnego przez czujnik [17] i jest typowa dla sytuacji statycznych pozbawionych ruchu. Z kolei na rys 2.3 b) wykreślono sygnały zarejestrowane przez kolejne trzy osie akcelerometru. W przypadku przerwy w chodzie zauważyć można brak znaczących powtarzalnych zmian.



RYSUNEK 2.3. Sygnały zarejestrowane podczas krótkiej przerwy w chodzie: a) sygnał magnitudy akcelerometru; za pomocą szarego tła zaznaczono okresy uznawane za aktywny chód b) sygnały zarejestrowane przez poszczególne osie czujnika

Podkreślić należy, że w bazie danych przed zastosowaniem procesu segmentacji czas zarejestrowanych sygnałów przez każdego uczestnika badań wynosił co najmniej 900 sekund (tj. 15 min). Jednakże, dla uczestnika badań o nr. identyfikacyjnym 7, długość zgromadzonych danych przekraczała 1500 sekund.

W wyniku przeprowadzonego procesu usuwania pauz w nagraniach dla poszczególnych uczestników zgromadzono co najmniej 800 sekund nagrań. Na rys. 2.4 zaznaczono liczebność nagrań z podziałem na poszczególnych uczestników. Dysproporcja w długości nagrań uczestnika #7 została zmniejszona.



RYSUNEK 2.4. Długość nagrań zgromadzonych dla poszczególnych uczestników badań przed oraz po zastosowaniu segmentacji

2.2.3. Podział danych na zbiór treningowy i testowy

Ewaluacja algorytmów klasyfikacji osób przeprowadzona została z wykorzystaniem prostej walidacji. Zbiór treningowy utworzony został na podstawie 60% czasu trwania nagrań, podczas gdy zbiór testowy stanowił 40% pozostałych danych. Zaznaczyć jednak należy, iż zarówno z początków, jak i końców poszczególnych nagrań, zostały usunięte bloki danych o długości 250 próbek (~8 sekund). Podejście takie wynikało z faktu, iż podczas ruszania, jak i hamowania, cykl chodu mógł być zaburzony.

Podczas generowania zbiorów uczących oraz treningowych dane podzielone zostały na okna o czasie trwania 64 próbek oraz nakładkowaniu wynoszącym 50%. Przy częstotliwości próbkowania około 30 Hz długość okna wynosiła około 2,14 sekund, co jest zbliżone do wartości literaturowych, tj. 2,4 sekund [4] oraz 2,14 sekund [3].

2.2.4. Klasyfikacja chodu

Proces klasyfikacji chodu odbył się z wykorzystaniem splotowych sieci neuronowych przy wykorzystaniu biblioteki programistycznej PyTorch oraz języka programowania Python. Przeprowadzono szereg eksperymentów, w których dokonywano m.in selekcji danych wejściowych sieci, jak i zmian architektury sieci neuronowych. Wszystkie z opisywanych klasyfikatorów bazowały na architekturze typu AlexNet [21] i składały się z naprzemiennych warstw "*Convoluton*" oraz "*MaxPooling*". Do nauki sieci wykorzystano funkcję kosztu entropii krzyżowej (ang. *cross entropy*), optymalizator SGD (ang. *stochastic gradient descent*) oraz współczynnik nauki (ang. *learning rate*) o wartości 0,01.

System biometryczny I

W ramach pierwszego z eksperymentów sprawdzono możliwość podania na wejście sieci neuronowych dwóch typów sygnałów: przyspieszenia oraz orientacji. W tym przypadku dane wejściowe składały się z sygnałów przyspieszenia (w formie oryginalnej i filtrowanej) oraz sygnałów orientacji wyznaczonej na podstawie wartości pomiarowych przyspieszenia (w formie oryginalnej i filtrowanej). Poszczególne typy danych oddzielone zostały za pomocą bloków danych o wartości zerowej (o wymiarze 3×64). Przebieg łączenia danych wejściowych zaprezentowano na rys. 2.5. Orientacja wyznaczona została z wykorzystaniem wzorów (1), (2), (3). Wektorem *G* zaznaczono globalny wektor grawitacji.



RYSUNEK 2.5. Struktura danych wejściowych Systemu I

Szczegóły dotyczące sieci splotowej przedstawiono w tabeli 2.2. Zastosowanie sieci neuronowej w opisanej konfiguracji pozwoliło na osiągnięcie skuteczności klasyfikacji równej **51,4%**.

TABELA 2.2. Architektura Systemu I oraz Systemu II, in/out – wejściowa/wyjściowa liczba filtrów, ks- kernel size, p-padding, s-stride

Warstwy	Тур	Szczegóły
1	Conv2D	in=1, out=3, ks=(4,4), p=0
2	MaxPool2d	ks=(3,3), s=(1,1)
3	Conv2D	in=3, out=6, ks=(3,3), p=0
4	MaxPool2d	ks=(2,4), s=(1,1)
5	Conv2D	in=6, out=6, ks=(5,5), p=0
6	MaxPool2d	ks=(2,5), s=(1,2)
7	Conv2D	in=6, out=9, ks=(2,2), p=0
8	MaxPool2d	ks=(1,2), s=(1,2)
9	Flatten	
10	Fully Connected	99×20
11	Dropout	0,5
12	Softmax	

System biometryczny II

W przypadku kolejnego badanego modelu wykorzystano zmodyfikowany zestaw danych wejściowych. W tej konfiguracji bloki danych składały się z sygnałów przyspieszenia (w formie oryginalnej i filtrowanej) oraz sygnały prędkości kątowych odtworzonych na podstawie wartości pomiarowych przyspieszenia (w formie oryginalnej i filtrowanej). Przebieg łączenia danych wejściowych zaprezentowano na rys 2.6. Prędkość kątowa została wyznaczona z wykorzystaniem wzorów (4), (5), (6).



RYSUNEK 2.6. Struktura danych wejściowych Systemu II

Na wejście sztucznej sieci neuronowej podano dane o wymiarze 15×64. W porównaniu do *Modelu I*, w danych wejściowych zamiast sygnałów orientacji (w postaci kątów Eulera) znalazły się sygnały prędkości kątowej. Zastosowano dokładnie tę samą architekturę sieci, co w punkcie poprzednim. Zastosowanie sieci neuronowej w opisanej konfiguracji pozwoliło na osiągnięcie skuteczności klasyfikacji równej **54,9**%.

System biometryczny III

W drodze przeprowadzonych eksperymentów najlepszy wyniki klasyfikacji uzyskano przy wykorzystaniu danych wejściowych w postaci bloków składających się z sygnałów: filtrowanego przyspieszenia, prędkości kątowej, natężenia pola magnetycznego, oraz orientacji rys. 2.7.



RYSUNEK 2.7. Przebieg łączenia danych wejściowych dla Systemu III

Szczegóły dotyczące sieci splotowej przedstawiono w tabeli 2.3. Zastosowanie sieci neuronowej o takiej architekturze pozwoliło na osiągnięcie skuteczności identyfikacji równej **89,3%.**

TABELA 2.3. Architektura Systemu III

Warstwy	Тур	Szczegóły
1	Conv2D	in=1, out=2, ks=(4,4), p=0
2	MaxPool2d	<i>ks</i> =(3,6), <i>s</i> =(1,2)
3	Conv2D	in=3, out=6, ks=(3,3), p=0
4	MaxPool2d	ks=(2,4), s=(1,2)
5	Conv2D	in=6, out=6, ks=(5,5), p=0
6	MaxPool2d	ks=(2,5), s=(1,1)
7	Flatten	

Warstwy	Тур	Szczegóły
8	Fully Connected	192×40
9	Sigmoid	
10	Fully Connected	40×20
11	Softmax	

Zestawienie wyników eksperymentów

W tabeli 2.4 zaprezentowano zestawienie wyników badań dotyczących identyfikacji osób na postawie chodu. *System I* oraz *II* posiadały tożsamą strukturę sieci splotowej, a różnica w klasyfikacji wynikała z uwzględnienia różnego zestawu sygnałów wejściowych. Najlepsze wyniki identyfikacji wynoszące około 90% zaobserwowano dla ostatniego z systemów biometrycznych.

TABELA 2.4. Zestawienie skuteczności opraco	wanych systemów biometrycznych
---	--------------------------------

System biometryczny	Skuteczność [%]
1	51,4
	54,9
	89,3

Podsumowanie

W niniejszej pracy przedstawiono wyniki eksperymentów dotyczących identyfikacji osób na podstawie chodu. W przeprowadzonych testach zbadano możliwość rozróżniania tożsamości na podstawie danych pomiarowych czujników ruchu.

Przyjęta metodologia badań zakładała zastosowanie prostej segmentacji, usuwającej z przebiegów jedynie okresy bezruchu. W prezentowanym podejściu brak jest algorytmów wyznaczania cykli chodu czy ich początków. Uzyskane wyniki identyfikacji osób (89,3%) zbliżone są wartościami do rezultatów stosowanych w literaturze [4], (87,5%), w której stosowano dane wejściowe w podobnej formie (tj. brak podziału na poszczególne cykle chodu). W przeciwieństwie do pozycji [3, 4], w toku których przetwarzanie wstępne danych oparto na generowaniu pseudo- obrazów, w niniejszej pracy zaproponowano bezpośrednie podanie sygnałów pomiarowych na wejście sieci splotowych CNN.

W przeprowadzonych badaniach dokonywano licznych eksperymentów z topologią sieci i strukturą danych wejściowych. Zastosowanie odtworzonych, a nie uwzględnionych w oryginalnym korpusie sygnałów prędkości kątowych, pozwoliło na zwiększenie skuteczności identyfikacji. Zaobserwowano wzrost o około 3,5 punktów procentowych dla *systemu II* względem *systemu I* (tabela 2.4). O ile wartość ta jest stosunkowo niewielka, należy pamiętać, że o tyle wykorzystanie sygnałów prędkości kątowych pozwoliło na osiągnięcie skuteczności klasyfikacji równej 89,3% w finalnym systemie identyfikacji.

Zaznaczyć należy, że identyfikacja przeprowadzona została z wykorzystaniem danych zgromadzonych w ciągu jednego dnia. Prezentowane podejście charakteryzują pewne ułatwienia, np. każdy z uczestników badań posiada stały rodzaj obuwia. Naturalną kontynuacją rozpoczętych badań jest ich rozszerzenie tak, aby obejmowały one dane zgromadzone w ciągu kilku dni.

Badania zostały zrealizowane w ramach pracy nr WI/WI/1/2019, sfinansowanej ze środków Ministerstwa Nauki i Szkolnictwa Wyższego w Polsce.

Bibliografia

- [1] Zou Q., Wang Y., Wang Q., Zhao Y., Li Q., Deep Learning-Based Gait Recognition Using Smartphones in the Wild, arXiv:1811.00338, 2020
- [2] Wang S.; Liu J., *Biometrics on Mobile Phone* [in] *Recent Applications Biometrics*; InTech: Rijeka, Croatia, 2011
- [3] Zhao Y., Zhou S., Wearable device-based gait recognition using angle embedded gait dynamic images and a convolutional neural network, Sensors, 2017, vol. 17, no. 3
- [4] Zhong, Y., Deng Y., Sensor orientation invariant mobile gait biometrics, Proceedings of the IEEE International Joint Conference on Biometrics (IJCB 2014), Clearwater, FL, USA, 2014
- [5] Ha S., Yun J.-M., Choi S., *Multi-modal convolutional neural networks for activity recognition*, IEEE International Conference on Systems, Man, and Cybernetics, 2015
- [6] Dehzangi O., Taherisadr M., ChangalVala R., *IMU-based gait recognition using convolutio*nal neural networks and multi-sensor fusion, Sensors, 2017, vol. 17, no. 12
- [7] Wong W. Y., Wong M. S., Lo K., *Clinical Applications of Sensors for Human Posture and Movement Analysis: A Review*, Prosthetics and orthotics international, 2017, 31
- [8] Gadaleta M., Merelli L., Rossi M., *Human authentication from ankle motion data using convolutional neural networks*, Proc. IEEE Stat. Signal Process. Workshop, 2016
- [9] Ding Y.,Xue Y., A Deep Learning Approach to Writer Identification Using Inertial Sensor Data of Air-Handwriting, IEICE Transactions on Information and Systems, 2019
- [10] Banos O., Calatroni A., Damas M., et al., *Kinect=IMU? Learning MIMO signal mappings to automatically translate activity recognition systems across sensor modalities*, Proceedings of the International Symposium on Wearable Computers (ISWC), Newcastle, UK, 2012
- [11] Rey V. F., Hevesi P., Kovalenko O., Lukowicz P., Let there be IMU data: generating training data for wearable, motion sensor based activity recognition from monocular RGB videos. Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, ACM. 2019,
- [12] Lago P, Takeda S., Okita T., Inoue S., MEASURed: Evaluating Sensor-Based Activity Recognition Scenarios by Simulating Accelerometer Measures from Motion Capture [in] Kawaguchi N. (ed.) Human Activity Sensing Corpus and Applications, Springer Series in Adaptive Environments, 2019
- [13] Kwon H., Tong C., Haresamudram H., Gao Y., et al., *IMUTube: Automatic extrac-tion of virtual on-body accelerometry from video for human activity recognition*, arXiv:2006.05675, 2020

- [14] Young A. D., Ling M. J., Arvind D.K., *IMUSim: A Simulation Environment for Inertial Sensing AlgorithmDesign and Evaluation*, Proceedings of the 10th International Conference on Information Processing in Sensor Networks, Chicago, IL, USA, 2011
- [15] Sawicki A., Walendziuk W., Idzkowski A., *The gravitational acceleration components elimi*nation from the accelerometer measurement data, Proc. SPIE 10031,2016
- [16] Stanley M, Simulate Motion for Sensor Fusion, materialy promocyjne firmy NXP, https:// www.nxp.com/company/blog/simulate-motion-for-sensor-fusion:BL-SIMULATE-MO-TION-SENSOR-FUSION [dostęp: 1.12.2020]
- [17] Bieda R., Grygiel, R., Galuszka, A., Naive Kalman filtering for estimation of spatial object orientation, Methods and Models in Automation and Robotics (MMAR), 20th International Conference, 2015
- [18] Barralon P., Vuillerme N., Noury N., Walk Detection With a Kinematic Sensor: Frequencyand Wavelet Comparison, Proceedings of IEEE International Conference of Engineering in Medicine and Biology Society (EMBS), 2006
- [19] Sawicki A. Repozytorium projektu Github https://github.com/asawicki-pb/gait-biometry [dostęp: 1.12.2020]
- [20] Krizhevsky A., Sutskever I., Hinton G.E., *ImageNet classification with deep convolutional neural networks*. Commun. ACM 60(6), 2017