Klasyfikator ruchów czujnika IMU na podstawie rekurencyjnej sieci neuronowej

Mateusz Woźniak
wozniakmat@student.agh.edu.pl

Maciej Pawłowski
maciekp@student.agh.edu.pl

Abstrakt

Artykuł przedstawia realizację klasyfikatora ruchów człowieka na podstawie danych z czujnika IMU (Inertial Measurement Unit) za pomocą rekurencyjnej sieci neuronowej. Czujnik IMU, mierzący przyśpieszenia postępowe i kątowe, jest powszechnie stosowany w wielu dziedzinach, takich jak lotnictwo, robotyka, wirtualna rzeczywistość i medycyna. W artykule omówiono proces zbierania danych z urządzenia Samsung Galaxy S10e 2019, a następnie wykorzystanie tych danych do trenowania sieci neuronowej w celu klasyfikacji pięciu z góry ustalonych ruchów. Wyniki mogą mieć zastosowanie w różnych dziedzinach, w tym w detekcji przeciągnięcia samolotu, kontroli stabilności pojazdów i zaawansowanych systemach wspomagania kierowcy. Zaproponowana architektura rekurencyjnej sieci neuronowej osiąga 96% precyzji w zadaniu klasyfikacji ruchów.

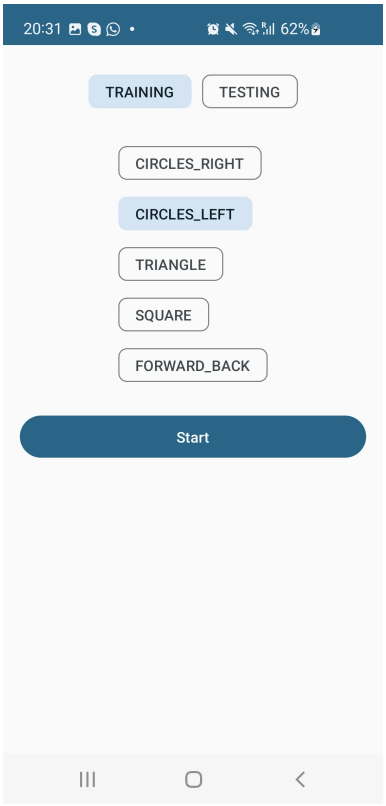
1 Wstęp

Czujnik IMU mierzy przyśpieszenia postępowe i kątowe używając żyroskopu, akcelerometru i magnetometru. IMU jest powszechnie stosowane w lotnictwie, robotyce, wirtualnej rzeczywistości i medycynie. W lotnictwie umożliwia precyzyjne sterowanie statkami powietrznymi, natomiast w robotyce wspomaga autonomiczne poruszanie się robotów. Jednym z zastosowań klasyfikatora ruchów może być detekcja przeciągnięcia samolotu. Z kolei w motoryzacji, IMU znajduje użycie w systemach kontroli stabilności pojazdów oraz w zaawansowanych systemach wspomagania kierowcy, które poprawiają bezpieczeństwo. Celem projektu była realizacja klasyfikatora pięciu z góry ustalonych ruchów przy użyciu rekurencyjnej sieci neuronowej. Zakres projektu obejmował zebranie zbioru danych, implementację modelu, trening i ewaluację modelu.

2 Zbiór danych

Dane zostały zebrane z urządzenia Samsung Galaxy S10e 2019. W procesie uczestniczyły dwie osoby, które wykonywały ruch. Każdy ruch został zebrany ręcznie, a seria danych z czujników była zapisywana do pliku csv o ilości wierszy takiej jak ilość kroków czasowych. Interwał próbkowania pomiaru z czujnika został ustawiony na $50ms$. To oznacza, że w ciągu każdej sekundy trwania ruchu następowało 20 odczytów. Ustaloną listę ruchów pokazuje tabela 1.

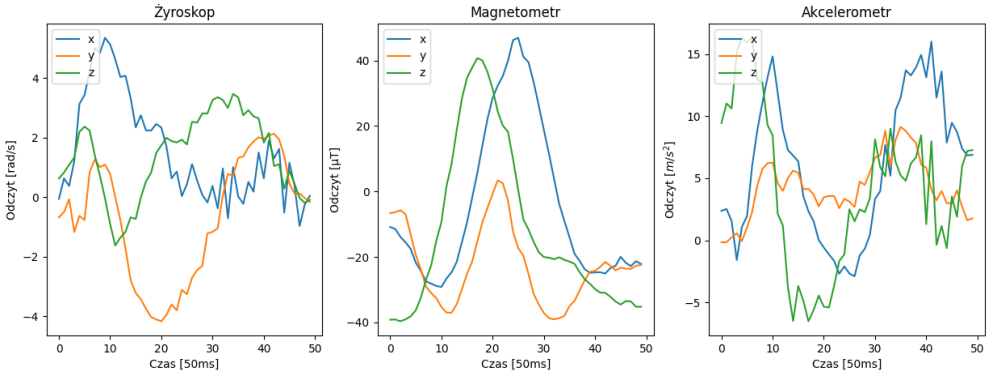
Ruchy były wykonywane przy włączonej aplikacji mobilnej napisanej w Kotlinie (rys. 1).



Rysunek 1: Aplikacja mobilna do zbierania danych z czujnika IMU

Dane wysyłano na zdalny serwer, na którym został uruchomiony mikroserwis HTTP napisany w Go. Mikroserwis zbierał dane i zapisywał je na dysk. Dzięki temu zbieranie danych odbywało się sprawnie, a mikroserwis monitorował ilość zebranych danych dla każdej z klas. Oto przykładowy fragment pliku csv (oraz jego wizualizacja na rys. 2):

```
gyro_x;gyro_y;gyro_z;magnetometer_x;...
-0.062384613;-0.15294538;-0.32650748;-43.62;...
-0.87361366;0.41210496;-0.49510628;-43.5;...
-1.7758616;0.20685424;-1.4266758;-43.92;...
-1.8662697;-0.46876273;-1.4040737;-44.399998;...
-0.7575493;-0.8077929;-1.488984;-44.28;...
0.06895141;-1.5170075;-2.1273382;-46.14;...
```



Rysunek 2: Przykładowy odczyt ruchu dla klasy CIRCLES_RIGHT

Tabela 1: Tabela dostępnych ruchów

Klasa	Opis ruchu	Ilość próbek
SQUARE	Ruch imitujący rysowanie kwadratu	203
TRIANGLE	Ruch imitujący rysowanie trójkąta	179
CIRCLES_LEFT	Rysowanie okręgu przeciwnie z ruchem wskazówek zegara	193
CIRCLES_RIGHT	Rysowanie okręgu zgodnie z ruchem wskazówek zegara	191
FORWARD_BACK	Ruch "od siebie - do siebie"	187

Dane zostały załadowane do tensora o kształcie (B, T, C), gdzie

- B - wsad (32)
- T - oś czasu
- C - cechy (9)

W przypadku, gdy próbki były różnej długości, tensor dopełniany był zerami, by T było równe 100. Dzięki temu kształt tensora odpowiadał wejściu sieci neuronowej. Zbiór treningowy stanowił 80% całego zbioru danych, a zbiór walidacyjny - 20%.

3 Model

Ze względu na to, że zadanie klasyfikacji ruchów miało być niezależne od czasu, tzn. że ruch mógł zawierać dowolną ilość próbek zadecydowano o użyciu rekurencyjnej sieci neuronowej z komórką LSTM [1]. Long Short-Term Memory (LSTM) to rodzaj architektury sieci neuronowej rekurencyjnej (RNN), zaprojektowanej w celu przezwyciężenia ograniczeń tradycyjnych RNN w przechwytywaniu i uczeniu się długoterminowych zależności w danych sekwencyjnych. LSTMy zostały wprowadzone przez Seppa Hochreitera i Jürgena Schmidhubera w 1997 roku [2]. Założeniem było, aby sieć neuronowa nauczyła się zależności w czasie pomiędzy zmianami w przyśpieszeniach czujnika [3] [4]. W związku z tym w pierwszym podejściu zdecydowano sprawdzić architekturę składającą się z 32 komórek LSTM oraz dwóch warstw gęstych o wielkości 24 i 5. Na ostatniej warstwie została zastosowana funkcja aktywacji $softmax$ by uzyskać dystrybucję prawdopodobieństwa klas (realizuje to $CrossEntropyLoss$)

Po 5 próbach zostały znalezione hiperparametry modelu, które dawały najlepsze zachowanie sieci. Warstwa LSTM musiała mieć 22 komórki, a warstwa gęsta - 32 neurony. Przeszukiwanie przestrzeni hiperparametrów zostało zrealizowane metodyką gridsearch. Finalny model ma następującą architekturę:

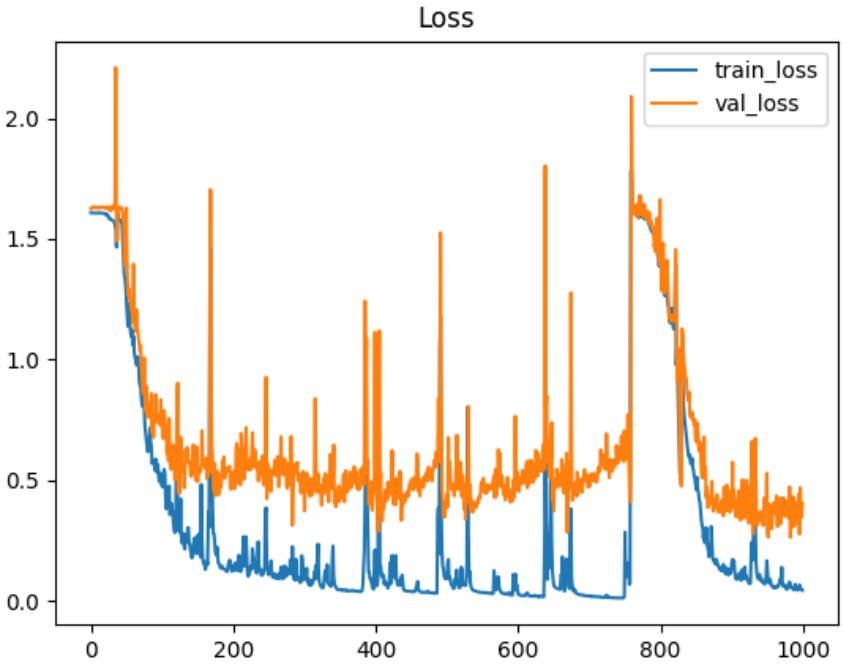
```
import torch
import torch.nn as nn

class Net(nn.Module):
  def __init__(self, input_size):
      super(Net, self).__init__()
      self.lstm = nn.LSTM(input_size, 22, batch_first=True)
      self.fc1 = nn.Linear(22, 32)
      self.fc2 = nn.Linear(32, 5)

  def forward(self, x):
      _, (h_n, _) = self.lstm(x)
      x = h_n[-1, :, :]
      x = self.fc1(x)
      x = self.fc2(x)
      return x
```

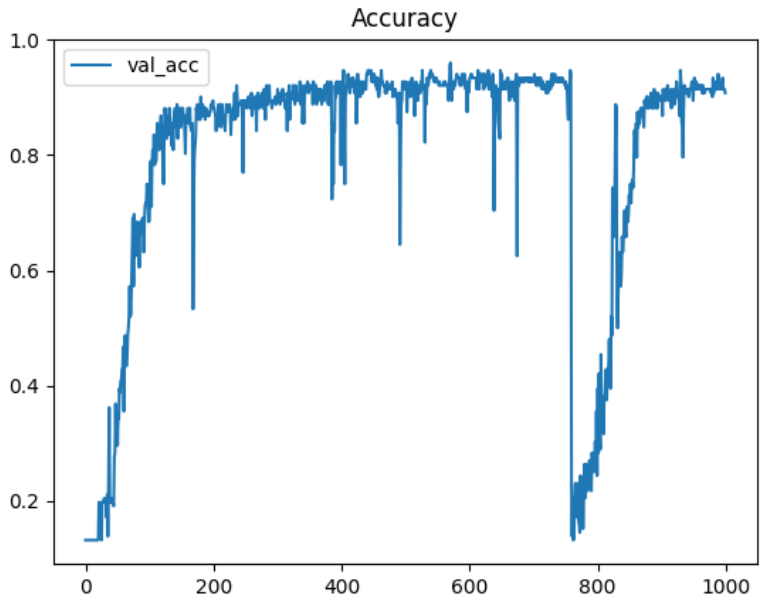
4 Obserwacje

Model był w stanie rozwiązać problem klasyfikacji i dostarczał zadowalających rezultatów. Znajdowanie daleko-zasięgowych zależności w danych często stanowi trudność dla sieci gęstych. W przypadku rekurencyjnej sieci neuronowej optymalizator był w stanie znaleźć odpowiedni kierunek by dostosować wagi sieci (rys. 3).

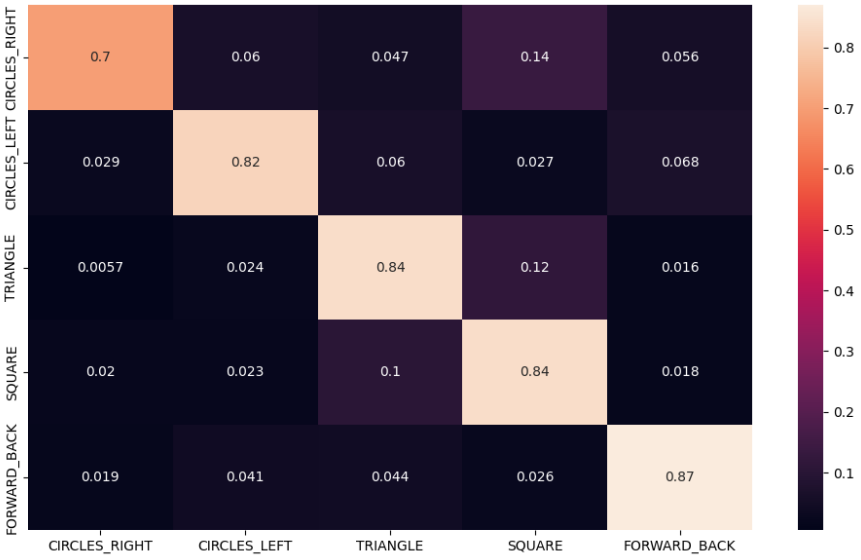


Rysunek 3: Wykres funkcji straty

Użyty optymalizator to $Adam$ ze współczynnikiem uczenia 0.001 Spośród 5 treningów sieci, najwyższą precyzję klasyfikacji, jaką dostarczył model, odnotowano na poziomie 96% (rys. 4). Sieć wykorzystywała wszystkie 9 cech (3 osie żyroskopu, 3 osie akcelerometru, 3 osie magnetometru). Macierz błędów jest widoczna na rys. 5.



Rysunek 4: Wykres precyzji na zbiorze walidacyjnym



Rysunek 5: Wykres precyzji na zbiorze walidacyjnym

W trakcie opracowywania modelu, odnotowano znaczne różnice w zakresach odczytów w zależności od producenta i modelu smartfonu. W związku z tym całe badanie zostało oparte na jednym urządzeniu Samsung Galaxy S10e.

5 Ewaluacja modelu

Ze względu na to, że ewaluacja na smartfonie w przypadku rekurencyjnych sieci neuronowych jest trudna, sieć neuronowa została przeniesiona na serwer VPS. Zastosowana została iteracyjna metodologia prac nad projektem:

1. Wytrenowanie model na lokalnym sprzęcie (laptop).
2. Budowa obrazu Dockera z wagami zawartymi w środku i wysłanie go do rejestru
3. Na serwerze zdalnym: pobranie obrazu i uruchomienie serwera Flask

Użycie takiej metodologii pozwoliło szybko systematycznie ulepszać jakość modelu. Aplikacja mobilna wykonuje żądanie HTTP POST do serwera VPS, by ten zwrócił wyniki.

6 Wnioski

Na podstawie wykresu funkcji straty i wykresu funkcji precyzji modelu można stwierdzić, że decyzja o wykorzystaniu rekurencyjnej sieci neuronowej była trafna. Zaproponowana sieć neuronowa cechuje się wysoką skutecznością w zadaniu klasyfikacji ruchów z urządzenia IMU.

7 Sprzęt

Do pomiarów został wykorzystany smartfon Samsung Galaxy S10e 2019 wyposażony w czujnik IMU. Implementację modelu została wykonana we frameworku PyTorch. Trening sieci neuronowej odbywał się na Apple Macbook M1 16GB.

Literatura

[1] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3] P. Rivera, E. Valarezo, M.-T. Choi, and T.-S. Kim, "Recognition of human hand activities based on a single wrist imu using recurrent neural networks," *Int. J. Pharma Med. Biol. Sci*, vol. 6, no. 4, pp. 114–118, 2017.

[4] S. Ashry, R. Elbasiony, and W. Gomaa, "An lstm-based descriptor for human activities recognition using imu sensors," in *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics, ICINCO*, vol. 1, 2018, pp. 494–501.