

**UNIWERSYTET RZESZOWSKI**

**Kolegium Nauk Przyrodniczych**

Damian Jamroży

Nr albumu: 113729

Kierunek: Informatyka

**Prototyp i analiza platformy do rozpoznawania gestów polskiego  
języka migowego w czasie rzeczywistym z zastosowaniem  
metod uczenia maszynowego**

Praca magisterska

Praca wykonana pod kierunkiem

Dr. Inż. Bogusława Twaroga

Rzeszów, 2024

Spis treści

[Wstęp 5](#_Toc175220205)

[1. Przegląd literatury i analiza istniejących rozwiązań 7](#_Toc175220206)

[1.1. Przegląd technologii rozpoznawania gestów 7](#_Toc175220207)

[1.2. Przegląd metod uczenia maszynowego 7](#_Toc175220208)

[1.3. Wyzwania w rozpoznawaniu gestów języka migowego 7](#_Toc175220209)

[1.4. Analiza istniejących platform do rozpoznawania języka migowego 7](#_Toc175220210)

[2. Projekt platformy 8](#_Toc175220211)

[2.1. Założenia funkcjonalne i niefunkcjonalne 8](#_Toc175220212)

[2.2. Architektura systemu 8](#_Toc175220213)

[2.3. Warstwa frontendowa (interfejs użytkownika) 8](#_Toc175220214)

[2.4. Warstwa backendowa (przetwarzanie danych) 8](#_Toc175220215)

[2.5. Warstwa AI (model uczenia maszynowego) 8](#_Toc175220216)

[3. Budowa modelu rozpoznawania gestów 9](#_Toc175220217)

[3.1. Przygotowanie danych wejściowych 9](#_Toc175220218)

[3.2. Preprocessing i augmentacja danych 9](#_Toc175220219)

[3.3. Wybór architektury modelu 9](#_Toc175220220)

[3.4. Sieci neuronowe i ich zastosowanie 9](#_Toc175220221)

[3.5. Algorytmy klasyfikacji 9](#_Toc175220222)

[3.6. Trening i walidacja modelu 9](#_Toc175220223)

[3.7. Proces uczenia modelu 9](#_Toc175220224)

[3.8. Ocena skuteczności modelu 9](#_Toc175220225)

[4. Implementacja platformy 10](#_Toc175220226)

[4.1. Integracja modelu AI z aplikacją webową 10](#_Toc175220227)

[4.2. Mechanizmy komunikacji między frontendem a backendem 10](#_Toc175220228)

[4.3. Widoki aplikacji 10](#_Toc175220229)

[5. Ocena działania platformy 11](#_Toc175220230)

[5.1. Testy w rzeczywistych warunkach 11](#_Toc175220231)

[5.2. Skuteczność i dokładność rozpoznawania gestów 11](#_Toc175220232)

[5.3. Analiza błędów i propozycje ulepszeń 11](#_Toc175220233)

[6. Dyskusja 12](#_Toc175220234)

[6.1. Porównanie wyników z innymi badaniami 12](#_Toc175220235)

[6.2. Omówienie wyników w kontekście postawionych hipotez 12](#_Toc175220236)

[6.3. Możliwości zastosowania platformy w praktyce 12](#_Toc175220237)

[Podsumowanie i wnioski 13](#_Toc175220238)

[Bibliografia oraz Netografia 14](#_Toc175220239)

[Spis ilustracji, tabel oraz wykresów 15](#_Toc175220240)

# Wstęp

Obecny rozwój technologii pozwala na automatyzację wielu procesów, w tym procesów finansowych, administracyjnych oraz sprzedażowych. Wymienione elementy są związane stricte z obsługą biznesów, które umożliwiają komercyjny zarobek dla twórców oprogramowania. Algorytmy mają w tym przypadku ułatwić pracę lub całkowicie zastąpić osoby fizyczne w ich pracy. Dzięki takim praktykom pracodawcy czy też różnego rodzaju organizacje zwiększają swoje dochody poprzez przyspieszeniu pracy lub w gorszym przypadku, oszczędzają fundusze poprzez zredukowanie etatów. Szerokie zastosowanie sztucznej inteligencji jest widoczne w każdym aspekcie naszego życia. Coraz większa ilość sklepów, banków czy też producentów wszelkiego rodzaju produktów decyduje się na wdrażanie sztucznej inteligencji. Tendencja wzrostowa popularności rozwiązań opartych na SI jest stale obserwowana przez specjalistów, którzy stawiają różnego rodzaju hipotezy dotyczące przyszłości rozwoju sztucznej inteligencji. Zgodnie z opinią specjalistów z Wyższej Szkoły Biznesu National-Louis University, AI może powodować utratę miejsc pracy oraz restrukturyzację zawodów, jednakże tym samym może zwiększać zapotrzebowanie na specjalistów w branży kreatywnej oraz IT. Autor artykułu zwraca uwagę na problematykę dotyczącą etyki związanej ze sztuczną inteligencją oraz potrzebę nieustannej nauki oraz rozwoju.[9]

Istnieją również aspekty SI, które są niezaprzeczalnie pozytywne, chociażby zastosowane w strefach pożytku publicznego czy też w rozwiązaniach dla osób z dysfunkcjami. Wszelkiego rodzaju protezy, pojazdy, syntezatory mowy, algorytmy analizujące tekst, dźwięk czy też obraz pozwalają na łatwiejsze funkcjonowanie osób niepełnosprawnych. Niestety obszary te są często pomijane, ze względu na stosunkowo niewielkie grono odbiorców.

Podejmując dalszą próbę analizy problemu możemy zaobserwować wysokie zainteresowanie wadami wzroku oraz problemami ruchowymi, natomiast niskie zainteresowanie dysfunkcją głosową w odniesieniu do osób głuchoniemych. Istnieje wiele rozwiązań, które ułatwiają kontakt wzrokowy, chociażby takie jak regulacja wielkości czcionek we wszelkiego rodzaju aplikacjach, asystenci głosowi czy też operacyjne korekty wzroku. Funkcje ruchowe wspierane są przez różnorodne protezy, pojazdy, a także specjalne miejsca dostosowane do ich potrzeb np. miejsca parkingowe. Niestety w stosunku do problematyki osób głuchoniemych nie ma zbyt wielu rozwiązań technologicznych, które mogą ułatwić ich życie w sposób nieinwazyjny.

Kontekst i znaczenie tematu

Opis ogólny tematu: Wprowadzenie do problematyki rozpoznawania gestów, szczególnie w kontekście polskiego języka migowego.

Znaczenie praktyczne i teoretyczne: Omówienie, dlaczego temat jest ważny, zarówno z perspektywy naukowej, jak i praktycznej, np. ułatwienia komunikacji dla osób niesłyszących.

Tło historyczne i rozwój technologii: Krótkie omówienie, jak rozwijały się technologie rozpoznawania gestów i jaką rolę pełnią obecnie.

Problem badawczy

Definicja problemu: Jasne sformułowanie, jaki konkretny problem badawczy podejmowany jest w pracy.

Znaczenie problemu: Wyjaśnienie, dlaczego rozwiązanie tego problemu jest ważne i jakie są jego potencjalne implikacje.

Ograniczenia i wyzwania: Zidentyfikowanie głównych trudności i ograniczeń związanych z tematem, np. ograniczenia technologiczne, brak dostępnych danych itp.

Hipotezy i cele badawcze

Hipotezy badawcze: Sformułowanie hipotez, które będą testowane w pracy. Może to obejmować przypuszczenia dotyczące skuteczności modelu, jego wydajności w czasie rzeczywistym, czy też trafności rozpoznawania gestów.

Cele badawcze: Prezentacja głównych celów pracy, zarówno ogólnych, jak i szczegółowych. Mogą one obejmować stworzenie prototypu platformy, przetestowanie różnych modeli uczenia maszynowego, integrację modelu z aplikacją webową, itp.

Oczekiwane wyniki: Krótkie przedstawienie, jakie wyniki autor pracy spodziewa się uzyskać.

Zakres pracy

Zakres teoretyczny: Określenie, które aspekty teoretyczne zostaną omówione w pracy, np. teorie związane z rozpoznawaniem obrazów, metody uczenia maszynowego, specyfika polskiego języka migowego.

Zakres praktyczny: Określenie, jakie aspekty praktyczne zostaną zrealizowane, np. stworzenie aplikacji, implementacja modelu AI, testowanie platformy.

Ograniczenia badawcze: Omówienie, co nie zostało objęte badaniami i dlaczego, np. ograniczenia czasowe, technologiczne, dostępność danych.

# Przegląd literatury i analiza istniejących rozwiązań

## Przegląd technologii rozpoznawania gestów

## Przegląd metod uczenia maszynowego

## Wyzwania w rozpoznawaniu gestów języka migowego

## Analiza istniejących platform do rozpoznawania języka migowego

# Projekt platformy

## Założenia funkcjonalne i niefunkcjonalne

## Architektura systemu

## Warstwa frontendowa (interfejs użytkownika)

Wybór technologii i narzędzi

## Warstwa backendowa (przetwarzanie danych)

## Warstwa AI (model uczenia maszynowego)

Biblioteki do uczenia maszynowego

# Budowa modelu rozpoznawania gestów

## Przygotowanie danych wejściowych

Źródła danych

## Preprocessing i augmentacja danych

## Wybór architektury modelu

## Sieci neuronowe i ich zastosowanie

## Algorytmy klasyfikacji

## Trening i walidacja modelu

## Proces uczenia modelu

## Ocena skuteczności modelu

# Implementacja platformy

## Integracja modelu AI z aplikacją webową

## Mechanizmy komunikacji między frontendem a backendem

## Widoki aplikacji

# Ocena działania platformy

## Testy w rzeczywistych warunkach

Wydajność w czasie rzeczywistym

## Skuteczność i dokładność rozpoznawania gestów

## Analiza błędów i propozycje ulepszeń

# Dyskusja

## Porównanie wyników z innymi badaniami

## Omówienie wyników w kontekście postawionych hipotez

## Możliwości zastosowania platformy w praktyce

# Podsumowanie i wnioski

Główne osiągnięcia pracy

Wnioski z przeprowadzonych badań

Propozycje dalszych badań i rozwoju platformy

# Bibliografia oraz Netografia

1. Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2017). SubUNets: End-to-End Hand Shape and Continuous Sign Language Recognition. IEEE International Conference on Computer Vision (ICCV).
2. Huang, J., Zhou, W., Zhang, Q., Li, H., & Li, W. (2018). Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition. IEEE Transactions on Circuits and Systems for Video Technology, 29(9), 2822-2832.
3. Koller, O., Zargaran, S., Ney, H., & Bowden, R. (2015). Deep Sign: Hybrid CNN-HMM for Continuous Sign Language Recognition. In Proceedings of the British Machine Vision Conference (BMVC).
4. Pandey, A., Mishra, M., & Verma, A. K. (2022). Real-Time Sign Language Recognition Using Machine Learning and Neural Networks. IEEE Access.
5. Patel, K., Gil-González, A.-B., & Corchado, J. M. (2022). Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics, 11(11), 1780.
6. Patel, M., & Shah, N. (2021). Real-Time Gesture-Based Sign Language Recognition System. IEEE International Conference on Information Technology and Engineering (ICITE).
7. Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2015). Sign Language Recognition Using Convolutional Neural Networks. European Conference on Computer Vision Workshops (ECCVW).
8. Zhang, Z., & Liu, C. (2020). Skeleton-Based Sign Language Recognition Using Whole-Hand Features. IEEE Access, 8, 68827-68837.
9. <https://www.wsb-nlu.edu.pl/pl/wpisy/wplyw-sztucznej-inteligencji-na-przyszlosc-pracy-nowe-perspektywy-i-wyzwania> (22.08.2024)

# Spis ilustracji, tabel oraz wykresów

**Nie można odnaleźć pozycji dla spisu ilustracji.**

**Nie można odnaleźć pozycji dla spisu ilustracji.**

**Streszczenie:**

**Prototyp i analiza platformy do rozpoznawania gestów polskiego języka migowego w czasie rzeczywistym z zastosowaniem metod uczenia maszynowego.**

**Abstract:**

**Prototype and Analysis of a Real-time Polish Sign Language Gesture Recognition Platform Using Machine Learning Method.**

****