



Aplikacja do nauki Polskiego Języka Migowego

M.M. Czajka, D. Kubacka, M. Grudziński, M.R. Janiszewski

Politechnika Poznańska, Wydział Automatyki, Robotyki i Elektrotechniki

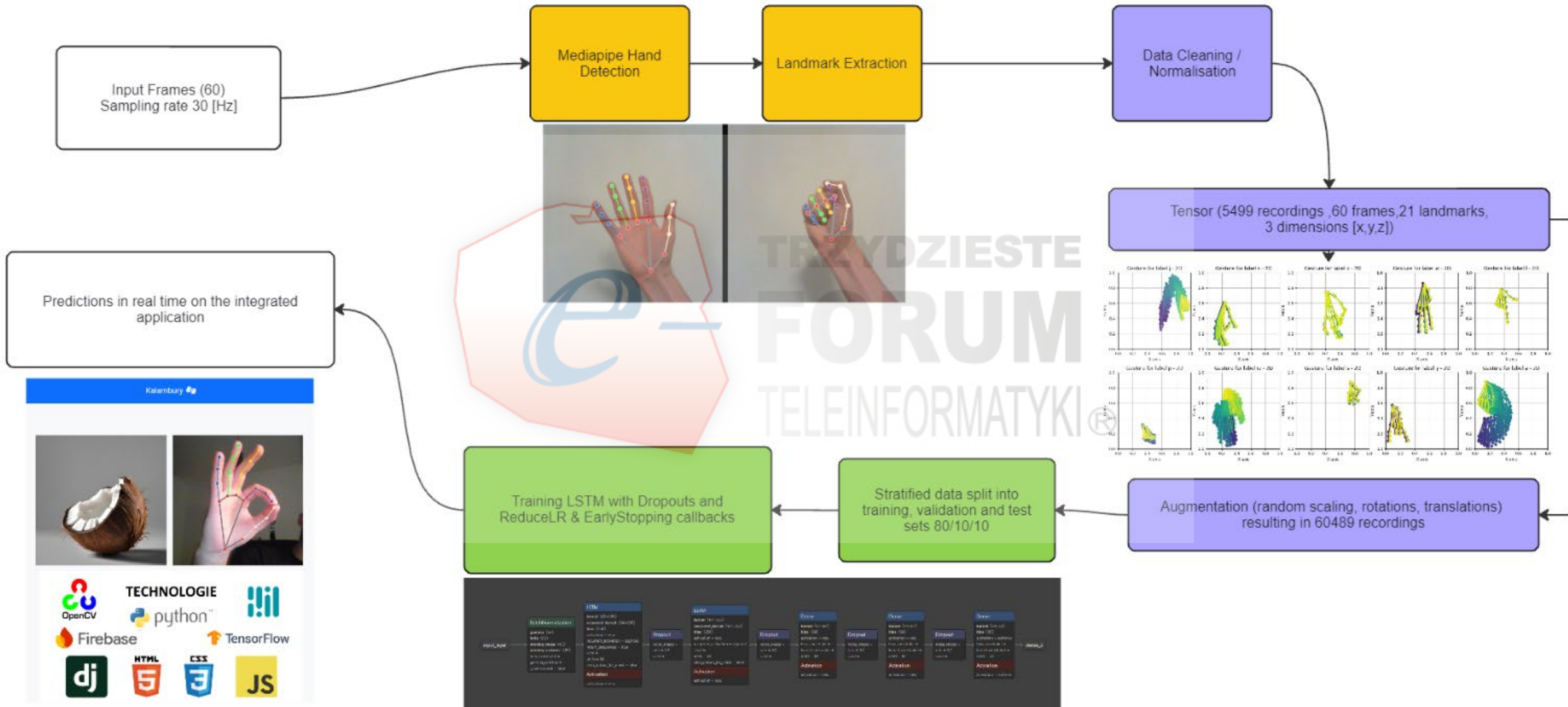


W Polsce żyje około 50 000 osób Głuchych oraz od 800 000 do 900 000 osób niedosłyszących¹. Osoby te często napotykają na liczne trudności, zarówno wynikające ze stereotypów, jak i z niedostosowanego do ich potrzeb systemu edukacji oraz barier na rynku pracy.

Wartość społeczną interaktywnej aplikacji do nauki Polskiego Języka Migowego (PJM):

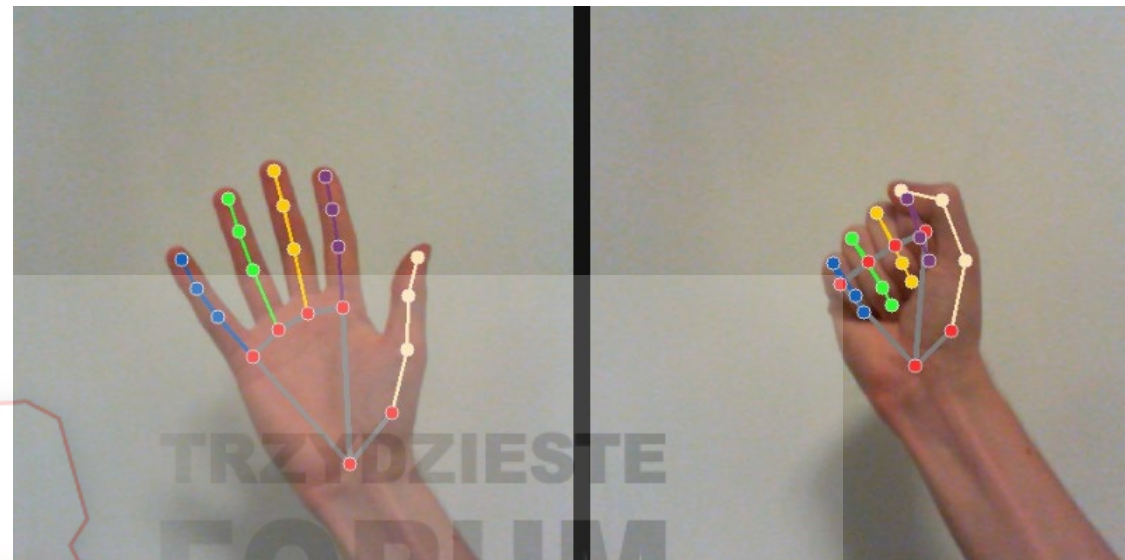
- Zwiększenie świadomości społecznej,
- Eliminacja barier komunikacyjnych,
- Promowanie wzajemnego zrozumienia i akceptacji w społeczeństwie.

¹ [https://orka.sejm.gov.pl/opinie8.nsf/nazwa/408_20190305_2/\\$file/408_20190305_2.pdf](https://orka.sejm.gov.pl/opinie8.nsf/nazwa/408_20190305_2/$file/408_20190305_2.pdf)

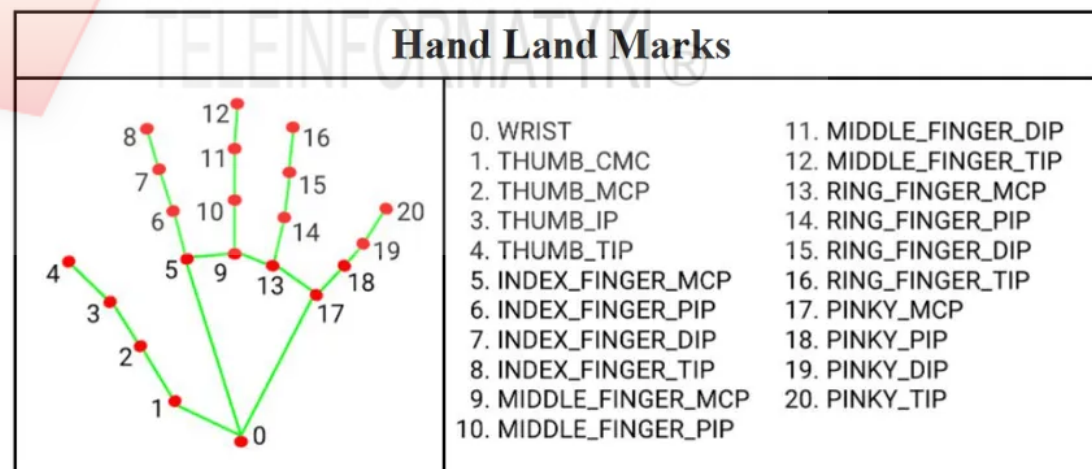


Rysunek: Schemat blokowy przedstawiający strukturę projektu

Każda sekwencja składa się z 60 klatek zawierających 21 punktów w 3 wymiarach. Punkty charakterystyczne dłoni są ekstraktowane przy użyciu biblioteki MediaPipe. Łącznie zebrano ok. 5500 tensorów, które dodatkowo poddano augmentacji (10 razy dodano losową rotację, translację i skalowanie).



Rysunek: Przykład wykrytych punktów charakterystycznych



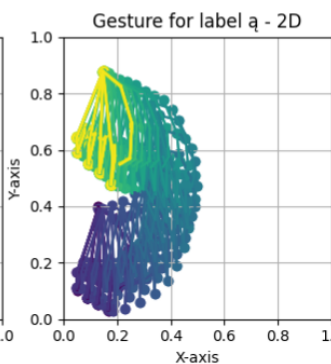
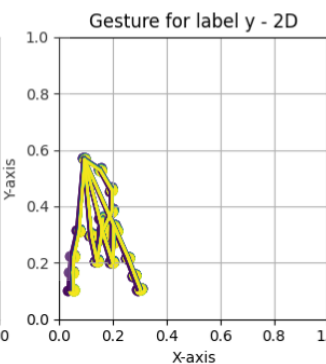
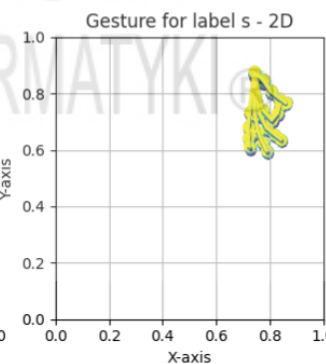
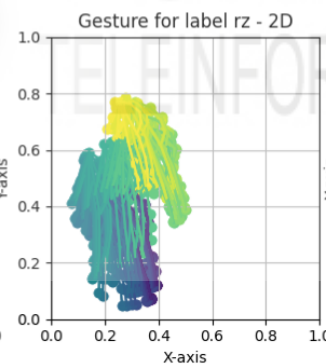
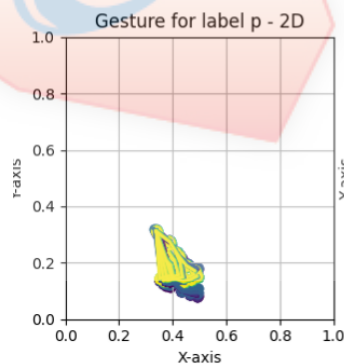
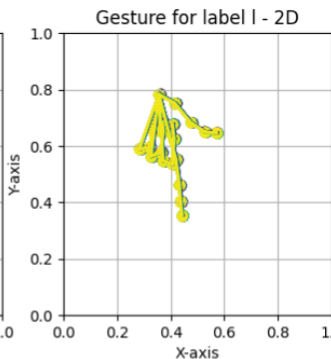
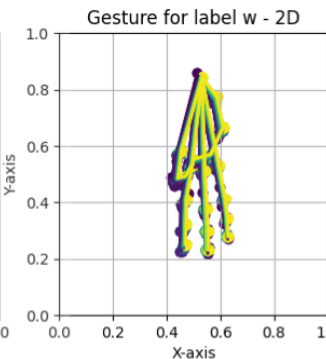
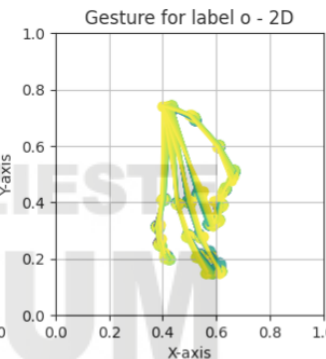
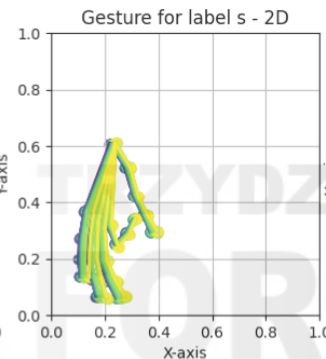
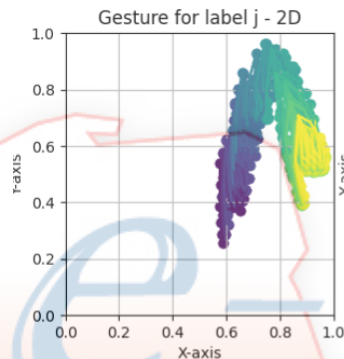
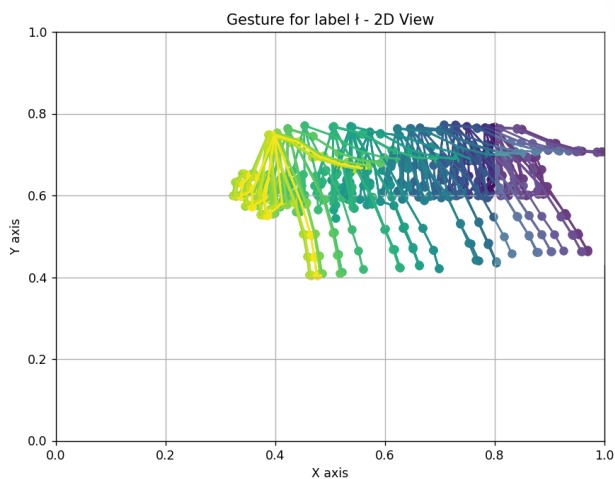
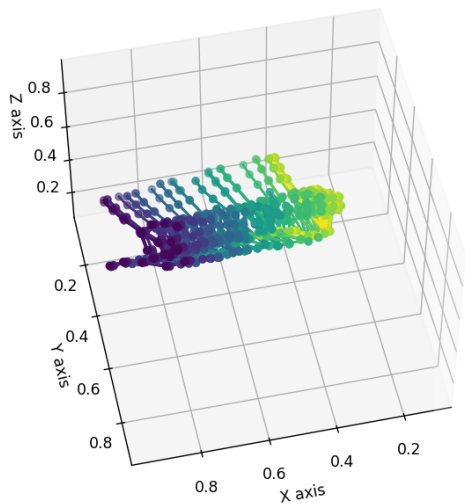
Rysunek: Opis poszczególnych punktów charakterystycznych.

Źródło: https://ai.google.dev/edge/mediapipe/solutions/vision/gesture_recognizer?hl=pl

Time

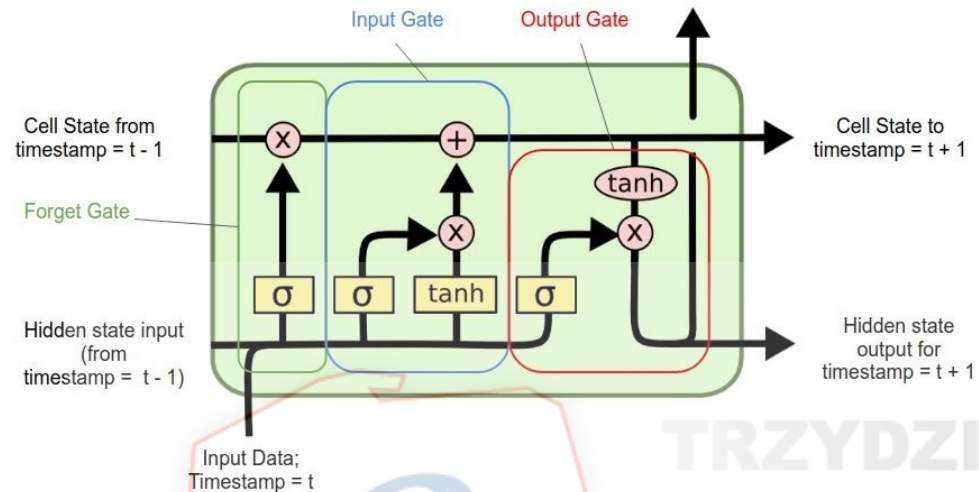


Gesture for label t - 3D View



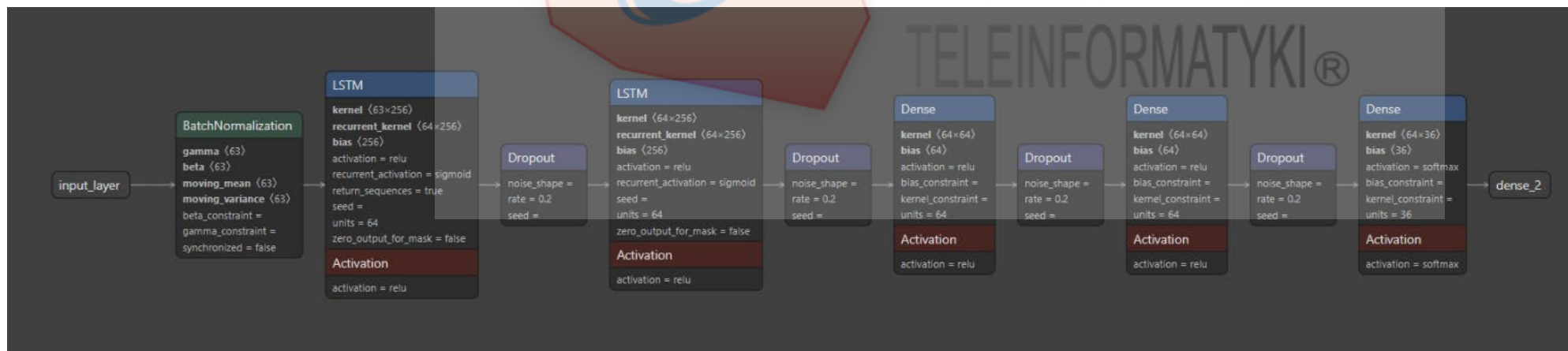
Rysunek: Wizualizacja przykładowych sekwencji ze zbioru danych treningowych

Struktura zbudowanego modelu (LSTM)

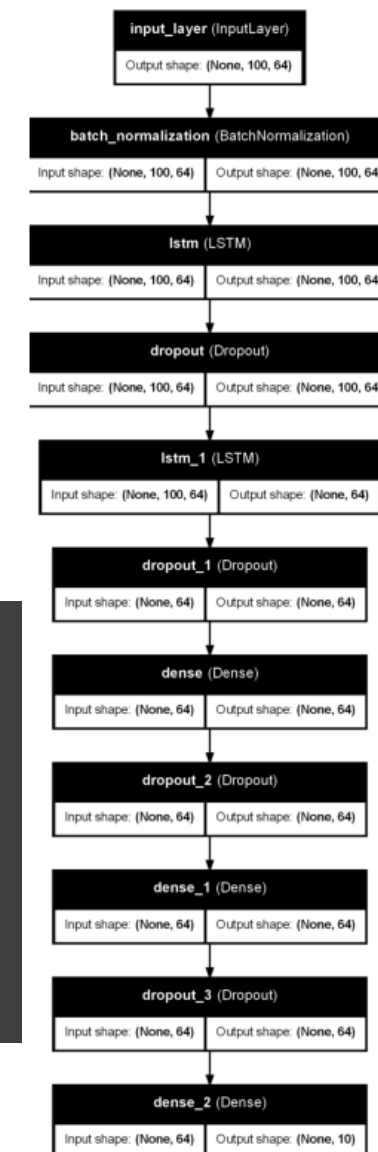


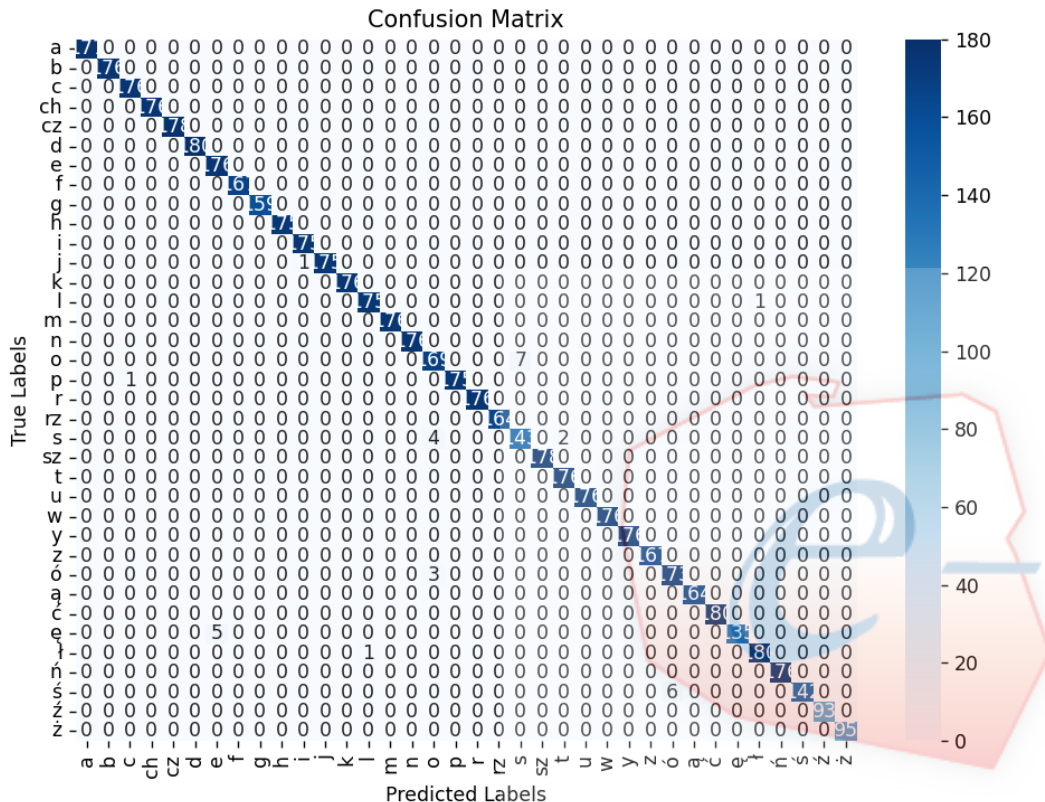
Rysunek: Wizualizacja bloku LSTM:

Źródło: <https://medium.com/@tm.nidheesh95/lstm-making-neural-networks-remember-what-matters-98a8d8eca844>



Rysunek: Wizualizacja struktury wykorzystanej sieci neuronowej (wykonana przy pomocy Tensorboard)



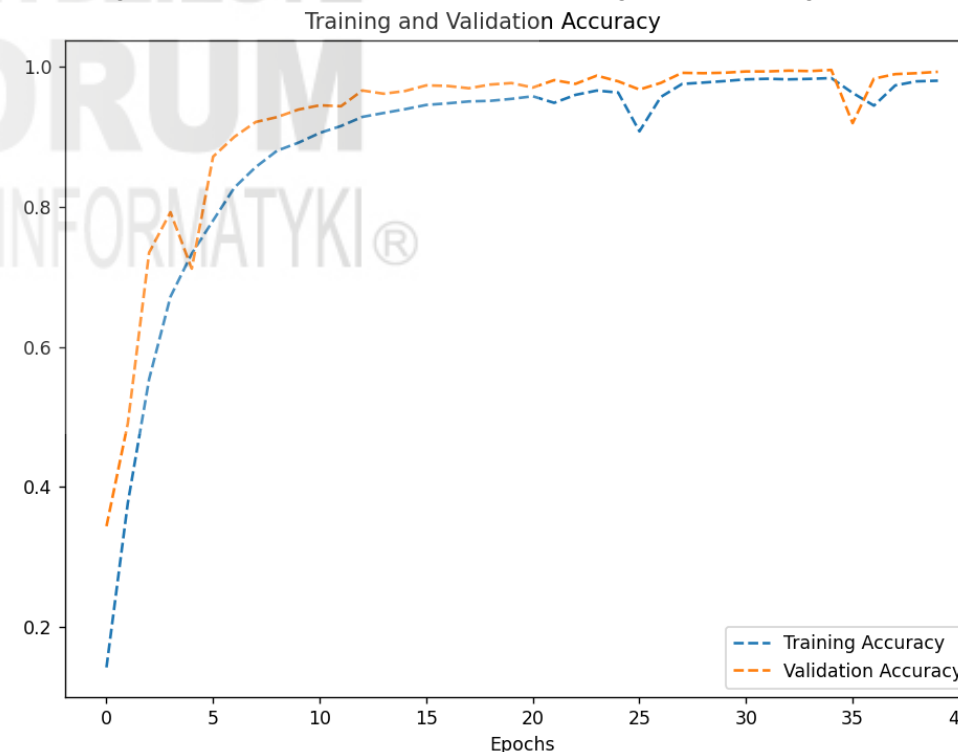


Rysunek: Macierz pomyłek wytrenowanego modelu

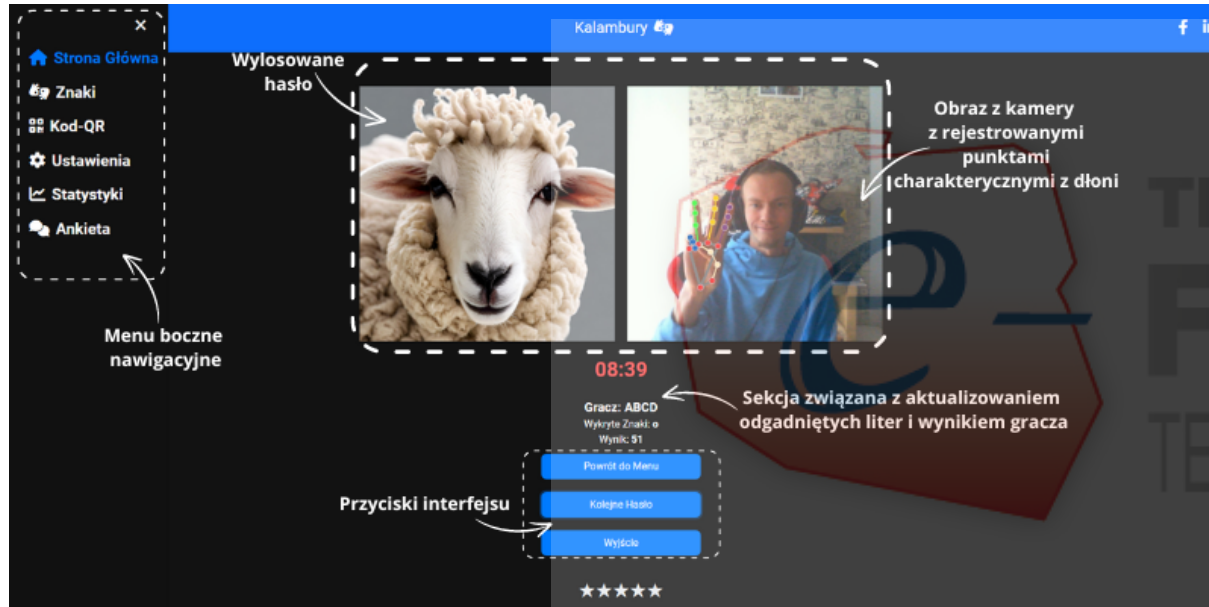
Po wielu cyklach treningu i ewaluacji na losowo inicjalizowanych podzbiorach można zauważyć bardzo wysoką dokładność predykcji. Po zwiększeniu bazy danych z nagraniami, przeprowadzeniu augmentacji, zastosowaniu warstw dropout i wspomnianych wywołań zwrotnych, macierz pomyłek i testy modelu na zintegrowanej aplikacji sugerują jedynie okazjne pomyłki modelu w przypadku podobnych znaków jak "o" i "s" czy "i" i "j".

Do treningu sieci neuronowej wykorzystano:

- **Podział danych** : w stosunku 80/10/10. Zastosowano rozwarstwienie, aby uwzględnić wszystkie etykiety w każdym zestawie i zapobiec wpływowi nierównowagi klas na wyniki.
- **Optymalizator** : Adaptive Moment Estimation (Adam).
- **Funkcja kosztu** : entropia krzyżowa.
- **Szybkość uczenia**: 0,001 z dodatkowymi wywołaniami zwrotnymi, takimi jak wczesne zatrzymanie i zmniejszenie stałej uczenia.



Rysunek: Wykres przedstawiający przebieg treningu

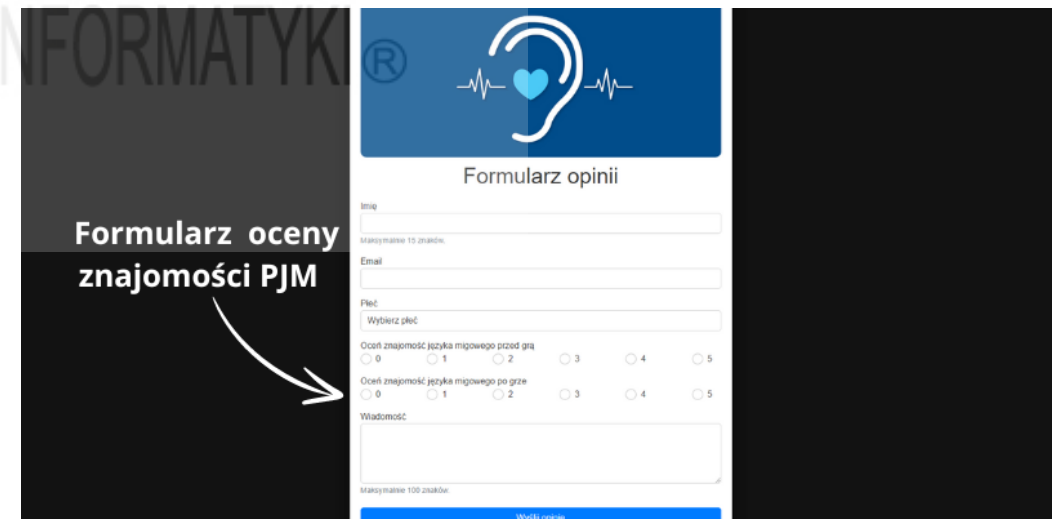


Rysunek: Zrzut ekranu z modułu z rozgrywką

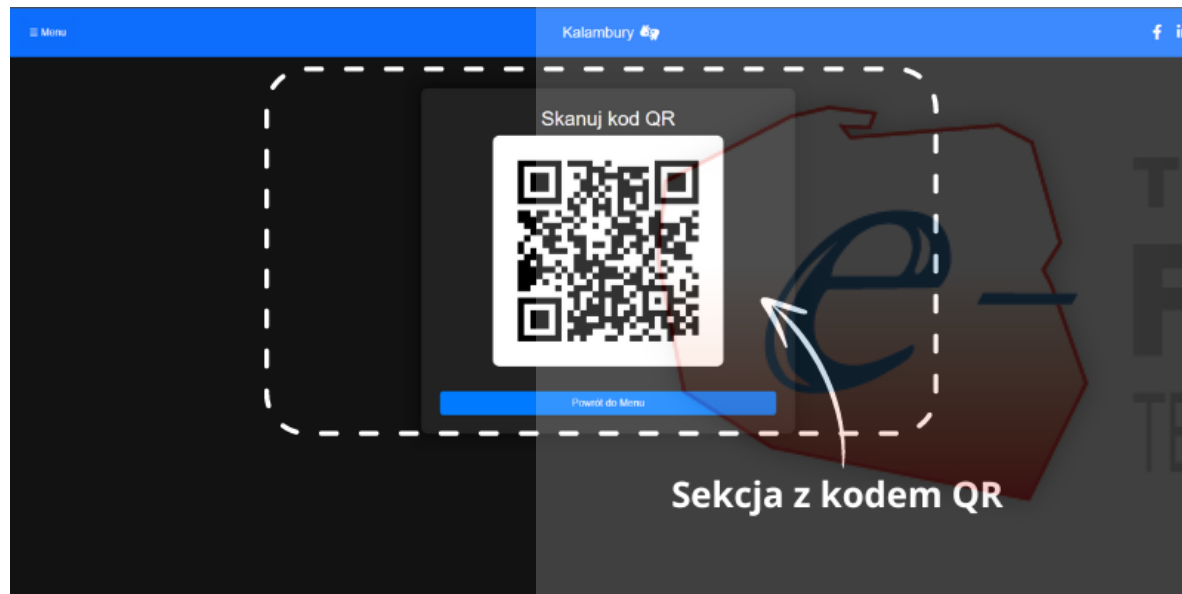
Interfejs aplikacji jest responsywny na wielu urządzeniach. Boczne menu nawigacyjne zapewnia szybki dostęp do wielu funkcji, takich jak ustawienia i ankieta oceny znajomości PJM. Sekcja gry zawiera obraz z kamery, który rejestruje punkty charakterystyczne z dłoni użytkownika, wyświetla wylosowane hasło, pozostały czas oraz wyniki, umożliwiając interaktywną naukę i śledzenie postępów w czasie rzeczywistym.



Rysunek: Zrzut ekranu z sekcji ustawień

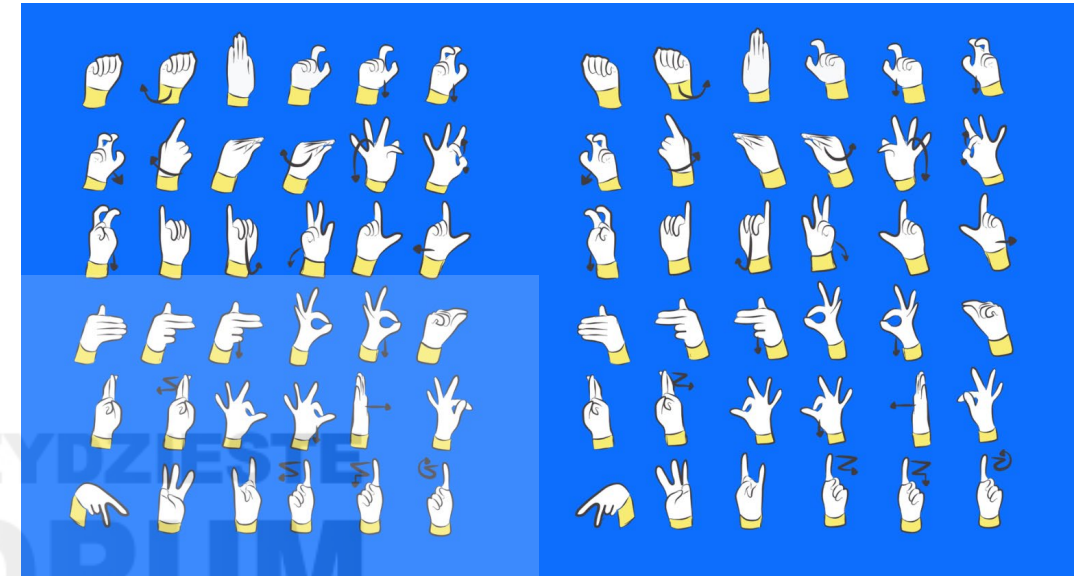


Rysunek: Zrzut ekranu z formularza

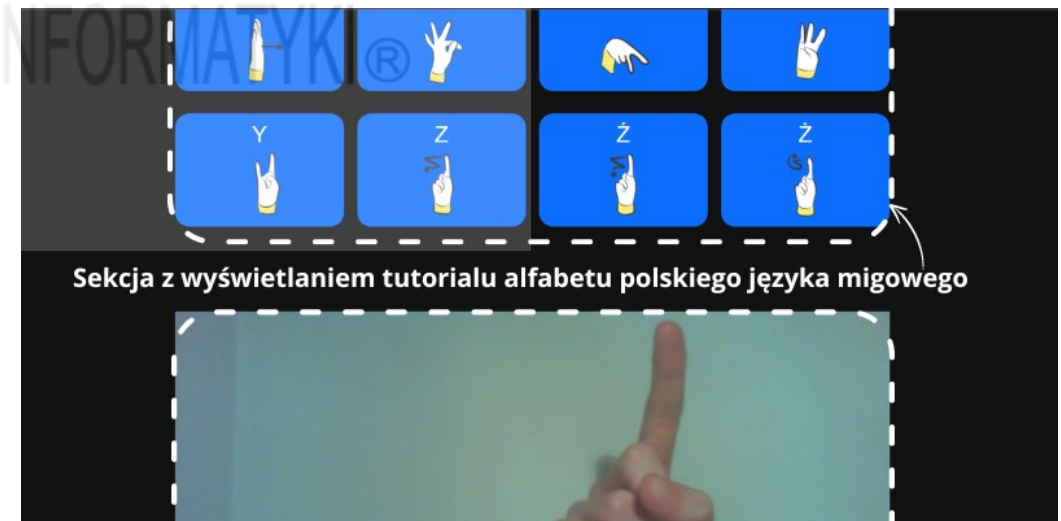


Rysunek: Zrzut ekranu z podziękowania w formie kodu QR

W celu ułatwienia użytkownikowi nauki przygotowano sekcję z nagraniami przedstawiającymi każdą literę alfabetu PJM. Dzięki temu użytkownicy mają możliwość zapoznania się z tymi gestami zarówno przed rozpoczęciem gry, jak i w trakcie jej trwania.



Rysunek: Grafiki przedstawiające litery PJM



Rysunek: Zrzut ekranu z sekcji tutoriali

Aplikacja (we wcześniejszej formie) została przedstawiona podczas Festynu Naukowego organizowanego przez Szkołę Podstawową nr 3 w Środzie Wielkopolskiej. W wydarzeniu tym brały udział osoby dorosłe oraz dzieci, co zapewniło dużą różnorodność rozmiarów dłoni i umiejętności obsługi komputera. Podczas Festynu zauważyliśmy, że projekt spełnia swój cel zachęcania poznania Polskiego Języka Migowego, a jej forma jest przyjazna, zarówno dzieciom, jak i dorosłym.

Możliwe kierunki rozwoju:

- utworzenie strony internetowej z przedstawioną aplikacją,
- zaimplementowanie wersji na urządzenia mobilne,
- powiększenie zbioru danych lub sprawdzenie innych struktur sieci neuronowej, aby sprawdzić czy możliwa jest poprawa klasyfikacji pokazywanych gestów,
- rozszerzenie aplikacji na słowa Polskiego Języka Migowego.



Rysunek: Kod QR z linkiem do repozytorium na GitHub



Rysunek: Kod QR z linkiem do filmu prezentującego aplikację



Rysunek: Grafika przedstawiająca wykorzystane w projekcie technologie