



Inżynieria Oprogramowania 2020

Temat projektu: Rozpoznawanie znaków języka migowego

Kryptonim projektu: Sign

Lider:

Przemysław Jabłecki

Członkowie zespołu:

Anna Świadek

Grzegorz Sroka

Filip Ślęzyk

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji
Kierunek Informatyka

Spis treści

| | |
|---|----------|
| Opis problemu | 3 |
| Wstęp | 3 |
| Kontekst problemu i potrzeby użytkowników | 3 |
| Istniejące możliwości | 4 |
| Podsumowanie | 4 |
| Wizja rozwiązania | 5 |
| Wizja projektu | 5 |
| Perspektywa biznesowa | 5 |
| Podsumowanie | 6 |
| Koncepcja systemu | 7 |
| Koncepcja architektury | 7 |
| Opis przebiegu działania programu | 7 |
| Przewidywane technologie | 9 |
| Zarys funkcjonalności systemu | 9 |

1. Opis problemu

1.1. Wstęp

Język migowy służy do porozumiewania się bez używania narządu słuchu, zazwyczaj przez same osoby głuche lub głuchonieme. Choć badania nie podają dokładnej liczby posługujących się tym językiem, to szacuje się, iż w samych Stanach Zjednoczonych takich osób może być nawet 0,5 mln ¹. Oznacza to, iż większość populacji nie potrafi rozpoznawać znaków języka migowego, który dla pewnej jej grupy stanowi naturalny sposób komunikacji.

1.2. Kontekst problemu i potrzeby użytkowników

Osoby głuche i głuchonieme, posługujące się językiem migowym, w codziennym życiu mogą spotkać się z barierą komunikacyjną. Na przyjęciu, w sklepie, czy nawet pośród członków rodziny mogą znaleźć się ludzie nie znający języka migowego. Życie osób migających byłoby znacznie ułatwione, gdyby w naturalny dla nich sposób, używając języka migowego, mogli przekazywać informacje osobom, dla których język ten nie jest znany.

Istnieją też ludzie nie będący głuchoniemi, czy głusi, którzy chcą nauczyć się języka migowego oraz osoby, które nagle w skutek wypadku straciły zmysł mowy czy słuchu i są zmuszone do nauki tego języka. Takie osoby nie zawsze mają możliwość sprawdzenia, czy pokazywane przez nich znaki są rozpoznawalne dla innych. Dostarczane przez różne instruktaże (jak np. poradniki, tutoriale internetowe) informacje nie zawsze dają możliwość skonfrontowania poczynionych w nauce języka migowego postępów z rzeczywistością.

Kursy nastawione na naukę języka są często bardzo drogie, a język migowy tym się różni od innych języków, że nie ma formy pisemnej. Przez to, żeby się go nauczyć i przekonać, czy wykonujemy poprawnie musimy mieć kogoś, kto nas oceni. Cena godziny nauki tego języka z wykwalifikowanym tutorem kosztuje około 100zł. Jest to wydatek, który może być zbyt wysoki dla wielu biedniejszych osób. Takie osoby potrzebują tańszego zamiennika, bądź narzędzia, które pozwoli im na ćwiczenie znaków w wolnym czasie. Takie lekcje nie mogą się też odbywać w dowolnym momencie, a brak dostępu do nauczyciela języka migowego może opóźnić naukę.

Nauczyciele języka migowego potrzebują również możliwości zadania swoim uczniom zadania domowego i ocenienia jak im poszło. W tym przypadku przydałoby się narzędzie mogące to zweryfikować nabyte przez ucznia umiejętności.

¹ Mitchell, Ross; Young, Travas; Bachleda, Bellamie; Karchmer, Michael (2006). *"How Many People Use ASL in the United States?: Why Estimates Need Updating"*. *Sign Language Studies*. 6 ISSN 0302-1475.

Nauka języka migowego (jak każdej innej umiejętności) bywa trudna i łatwo się przy niej zniechęcić. Osoby, którym na tym zależy mogą potrzebować narzędzia, które będzie ich motywować do dalszej nauki. Widząc swoje postępy na przestrzeni czasu nabiera się nowych chęci i sił, żeby iść naprzód.

Ponadto, osoby głuchonieme są częściowo pozbawione możliwości korzystania z wygodnego interfejsu komunikacji z komputerem i innymi urządzeniami elektronicznymi, dostępnego dla osób pełnosprawnych. Przy większości filmów, bądź utworów muzycznych umieszczonych w Internecie nie ma nawet możliwości włączenia napisów, a tłumacz języka migowego w rogu ekranu jest prawdziwą rzadkością². Z tej przyczyny osoby głuchonieme pozostają częściowo wykluczone z grona użytkowników najnowszych rozwiązań technologicznych.

1.3. Istniejące możliwości

Można wskazać kilka istniejących rozwiązań, które potencjalnie mogłyby zaspokoić wymienione wyżej potrzeby.

Osoby głuche i głuchonieme, mogą oglądać w Internecie filmy instruktażowe w których osoby znające język migowy uczą jak pokazać konkretną literę. Taka forma instruktażu jest dość rozpowszechniona, m.in. na takich platformach jak YouTube, czy Coursera. Niestety filmy instruktażowe nie umożliwiają jednoznacznej weryfikacji w czasie rzeczywistym, czy udało nam się poprawnie pokazać literę, czy nie.

Grupa osób, które chcą nauczyć się języka migowego, może korzystać z usług tłumaczy i trenerów. Niewątpliwie jednak, biorąc pod uwagę małą popularność języka migowego, takich nauczycieli jest bardzo niewielu, przez co efektywna nauka migania może być utrudniona przez brak zasobów ludzkich. Samodzielna nauka również nie jest idealnym rozwiązaniem. Po pierwsze, nie zawsze znak, który jest czytelny dla osoby uczącej się, będzie czytelny dla innych. Po drugie, osoba ucząca się może szybko stracić motywację do nauki, jeśli nie będzie mogła sprawdzić jej efektów. Nie ma dostępnym programów, które mogłyby na żywo pokazywać postępy czynione w nauce języka migowego.

1.4. Podsumowanie

Życie osób głuchych i głuchoniemych jest utrudnione. Pomimo, że świat ciągle dąży do tworzenia takim ludziom jak najlepszych warunków życia, przykładowo, zapewniając w programach telewizyjnych tłumaczy języka migowego, wciąż istnieje znaczna część populacji, z którymi osoby migające nie będą mogły się łatwo skontaktować w bezpośredniej rozmowie. Dlatego bardzo ważne jest kształcenie ludzi w tym kierunku. Potrzebne są narzędzia, które mogą naukę języka migowego ułatwić. Jest to szczególnie ważne teraz w XXI wieku. Trzeba uczynić wszystko co tylko możliwe, żeby każda grupa

² <https://www.w3.org/2007/08/video/positions/Meraka.pdf>

ludzi, która potrzebuje i chce się czegoś nauczyć mogła dla osiągnięcia tego celu korzystać z dobrodziejstw dzisiejszej technologii.

2. Wizja rozwiązania

2.1. Wizja projektu

Ważnym aspektem proponowanego projektu jest jego łatwość obsługi oraz efektywność. Jako możliwy wariant rozwiązania proponujemy stworzenie aplikacji desktopowej, spełniającej powyższe kryteria, oferującą bardzo intuicyjny interfejs użytkownika oraz formę nie wymagającą dużej ilości zasobów. Przed rozpoczęciem pracy użytkownik musi siedzieć na wprost kamery i jego dłoń musi być w niej widoczna. Gdy aplikacja zostanie uruchomiona, następuje przejście do widoku głównej funkcjonalności projektu. W drugim okienku widoku będzie można zobaczyć obraz z kamery internetowej w czasie rzeczywistym.

Użytkownikowi zostanie wyświetlona litera języka angielskiego, którą będzie aktualnie ćwiczył. Następnie przy pomocy kamery będzie mógł zaprezentować znak, który zostanie przetworzony przez aplikację i przetłumaczony na odpowiadający mu symbol z języka angielskiego. Symbol ten będzie porównywany z docelową literą. Jeśli program stwierdzi zgodność na Front-edzie zostanie wyświetlona informacja informująca o sukcesie. Następnie program przejdzie do kolejnej losowo wybranej litery. Baza danych zawierająca informacje o skuteczności użytkownika pokazywaniu danej litery języka migowego będzie dostępna z poziomu Front-endu.

2.2. Perspektywa biznesowa

W powyższych punktach zostało wykazane, iż osób posługujących się językiem migowym w sposób natywny jest niewiele względem całej populacji, zatem również nie powstało znacząco dużo aplikacji ułatwiających naukę języka migowego przez takie osoby. Rozwiązaniem konkurencyjnym do prezentowanej aplikacji jest Five App³ zaimplementowana przez zespół Mateusza Macha. Rozwiązanie umożliwia wprowadzanie znaków alfabetu łacińskiego, które są przetwarzane w GIF-y przedstawiające język migowy. Spektakularny sukces tej aplikacji udowadnia, że jest to nisza rynkowa, która nie została jeszcze nasycona i projekt może przyczynić się do pomocy osobom głuchoniemym oraz głuchym. Wymienione konkurencyjne przedsięwzięcie różni się znacząco od przedstawionego w tej dokumentacji rozwiązania, ponieważ nie pozwala na samodzielną weryfikację postępów w nauce języka migowego, zapisywanie historii postępów w nauce oraz w ogólności nie jest oparta o rozpoznawanie obrazu.

³ <https://businessinsider.com.pl/firmy/five-app-komunikator-dla-gluchych-tlumacz-jezyka-migowego/kcd079d>

Aplikacja stworzona w ramach projektu o znaku kodowym Sign będzie działać niezależnie od systemu operacyjnego, a zatem może być dystrybuowana poprzez Microsoft Store, czy też Mac App Store.

2.3. Podsumowanie

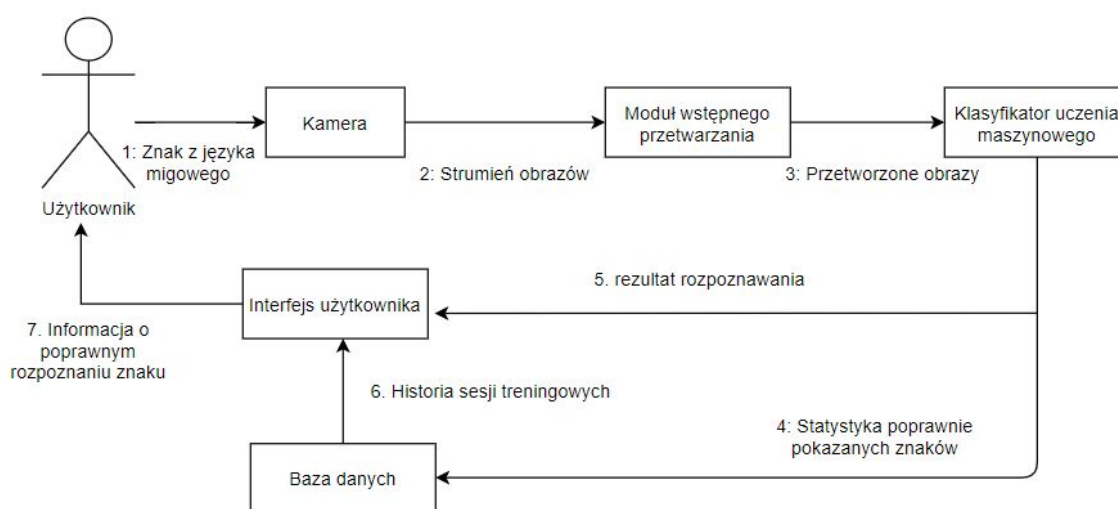
W powyższych punktach wykazano, że ten innowacyjny projekt ma potencjalny rynek zbytu i może się znacząco przyczynić do ułatwiania ludziom nauki języka migowego. Tworzona aplikacja jest oparta o technikę uczenia maszynowego, a zatem w trakcie implementacji precyzja rozpoznawania obrazu będzie sukcesywnie umacniana. Ponadto układ widoku będzie przyjazny dla użytkownika i ergonomiczny.

3. Koncepcja systemu

3.1. Koncepcja architektury

W ramach tworzonego systemu można wyróżnić kilka głównych komponentów logicznych. Pierwszym z nich jest kamera wbudowana w komputer, na którym użytkownik będzie korzystał z aplikacji. Będzie ona rejestrowała znaki migowe pokazywane przez użytkownika. Kolejnym komponentem systemu jest moduł wstępnego przetwarzania obrazu, odpowiedzialny za ekstrakcję istotnych dla dalszej detekcji cech dostarczonego strumienia obrazów (przykładowo, wyodrębnienie krawędzi, korekta kolorów). Następnym komponentem jest klasyfikator uczenia maszynowego, który będzie wykorzystywał wytrenowany wcześniej model do detekcji znaku na podstawie danych dostarczonych z modułu wstępnego przetwarzania. Pozostałymi komponentami systemu będzie baza danych, służąca do zapisywania postępów użytkownika w pokazywaniu literek, jak i moduł interfejsu użytkownika, dostarczający między innymi podgląd obrazu z kamery, aktualnie ćwiczoną przez użytkownika literę alfabetu, oraz informacje o postępach w nauce.

3.2. Opis przebiegu działania programu



Rysunek 1: Diagram komunikacji w przypadku standardowej pracy z systemem

1. Znak języka migowego

Użytkownik prezentuje wybrany znak alfabetu migowego do kamery internetowej za pomocą

2. Strumień obrazów

Kamera odczytuje i przesyła kolejne klatki obrazu do modułu wstępnego przetwarzania.

3. Przetworzone obrazy

Strumień obrazów z kamery trafia do modułu wstępnego przetwarzania, który odpowiednio przetwarza odebrane klatki. Znak jest następnie rozpoznany za pomocą klasyfikatora.

4. Statystyka poprawnie pokazanych znaków

Do bazy danych zapisywane są informacje o poprawności pokazania znaku przez użytkownika.

5. Rezultat rozpoznawania

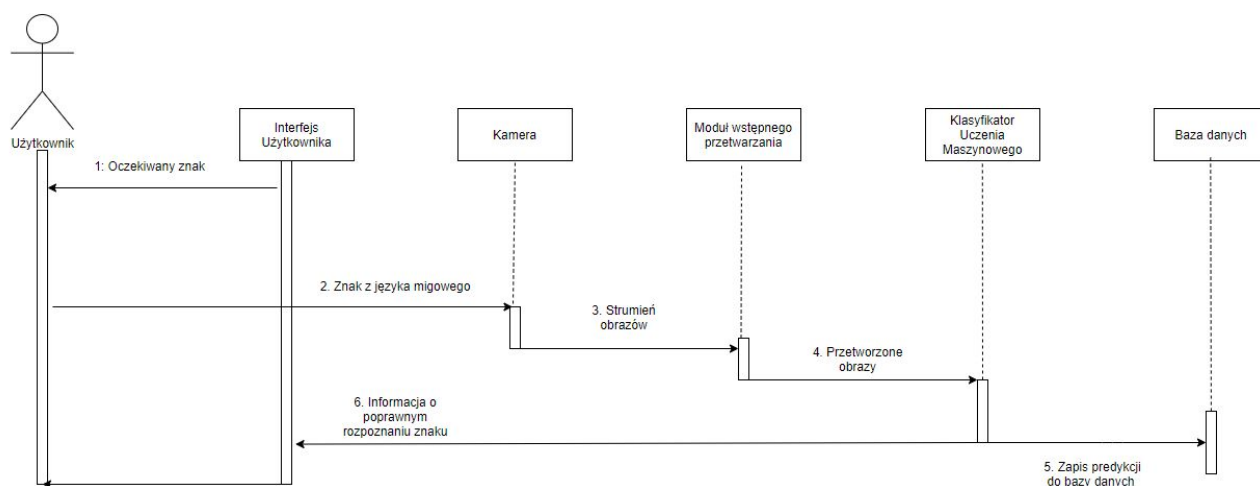
Jednocześnie z zapisem danych informacja o sukcesie bądź porażce rozpoznania znaku jest przesyłana do graficznego interfejsu użytkownika.

6. Historia sesji treningowych

Użytkownik może za pomocą interfejsu poprosić o ukazanie historii poprzednich sesji treningowych.

7. Informacja o poprawnym rozpoznaniu znaku

Jeśli użytkownik poprawnie pokazał aktualnie ćwiczony znak języka migowego zostaje on o tym poinformowany.



Rysunek 2: Diagram sekwencji

3.3. Przewidywane technologie

W projekcie zdecydowaliśmy się użyć języka Python w wersji 3, gdyż jest to obecnie najbardziej popularny język programowania wykorzystywany w uczeniu maszynowym. Z pewnością mogą się dla nas okazać przydatne biblioteki numpy i scipy. Rozważamy też wykorzystanie OpenCV⁴ - biblioteki służącej do obróbki obrazu. Baza danych będzie wykonana w technologii SQLite. Cechuje się ona wysoką niezawodnością, prostotą obsługi. Jest to obecnie najczęściej używany silnik bazodanowy na świecie⁵. Front-end aplikacji zostanie zaimplementowany przy pomocy frameworka Angular⁶.

3.4. Zarys funkcjonalności systemu

System umożliwia użytkownikowi rozpoznanie znaku z języka migowego. Dzięki wykorzystaniu bazy danych użytkownik będzie w stanie odtworzyć swoje wcześniejsze wywołania. Program z działaniem czasu będzie ciągle ulepszany, żeby osiągnąć coraz większą dokładność w rozpoznawaniu znaków. Co więcej, zapewni on użytkownikowi wysoki poziom bezpieczeństwa i niezawodności. Interfejs będzie zbudowany zgodnie z najlepszymi zaletami user experience, żeby był łatwy w obsłudze i funkcjonalny.

⁴ <https://opencv.org/>

⁵ <https://www.sqlite.org/index.html>

⁶ <https://angular.io/>