



**UMCS**

UNIwersytet Marii Curie-Skłodowskiej  
w Lublinie

Wydział Matematyki, Fizyki i Informatyki

Kierunek: **informatyka**

Specjalność: –

**Rafał Hrabia**

nr albumu: 296583

**Rozpoznawanie liter języka migowego z  
zastosowaniem technik uczenia  
maszynowego**

Sign Language Letter Classification using Deep Learning

Praca licencjacka  
napisana w Katedrze Cyberbezpieczeństwa  
pod kierunkiem dr hab. Michała Wydry

**Lublin 2021**

# Spis treści

|                                                                       |           |
|-----------------------------------------------------------------------|-----------|
| <b>Wstęp</b>                                                          | <b>3</b>  |
| <b>Cel i zakres pracy</b>                                             | <b>4</b>  |
| <b>1 Język migowy American Sign Language</b>                          | <b>5</b>  |
| 1.1 Historia . . . . .                                                | 5         |
| 1.2 Populacja . . . . .                                               | 6         |
| 1.3 American Manual Alphabet . . . . .                                | 7         |
| <b>2 Konwolucyjne sieci neuronowe i rozpoznawanie obrazów</b>         | <b>9</b>  |
| 2.1 Sekcja C . . . . .                                                | 9         |
| 2.2 Sekcja D . . . . .                                                | 9         |
| <b>3 Implementacja aplikacji do rozpoznawania języka migowego ASL</b> | <b>10</b> |
| <b>4 Podsumowanie</b>                                                 | <b>11</b> |
| <b>Spis tabel</b>                                                     | <b>12</b> |
| <b>Spis rysunków</b>                                                  | <b>13</b> |
| <b>Spis listingów</b>                                                 | <b>14</b> |
| <b>Bibliografia</b>                                                   | <b>14</b> |

# Wstęp

Sieci neuronowe mają za zadanie naśladowanie zachowań sieci neuronów znajdujących się w mózgu człowieka. Zostały stworzone do rozwiązywania zadań trudnych lub prawie niemożliwych do opisanego za pomocą reguł, wyrażeń logicznych i innych narzędzi programistycznych. Wraz z rozwojem sieci neuronowych powstało wiele wariantów, które ze względu na swoją budowę lepiej lub gorzej sprawdzają się w różnych problemach. W przypadku rozpoznawania obrazów w postaci dwu-wymiarowej macierzy dla danych monochromatycznych lub trój-wymiarowej macierzy dla zdjęć kolorowych jednym z najlepszych wyborów będą sieci konwolucyjne. Sieci te rozpoznają wzorce, poczynając od linii horyzontalnych i wertykalnych, a w dalszych warstwach kończąc na skomplikowanych strukturach. Budowa i działanie sieci konwolucyjnych daje wielki potencjał do klasyfikacji obrazów. Ta praca przedstawi przykład takiej klasyfikacji wieloklasowej z użyciem sieci konwolucyjnych na przykładzie alfabetu Amerykańskiego Języka Migowego.

# Cel i zakres pracy

Celem pracy jest stworzenie programu wyposażonego w wytrenowany model do rozpoznawania obrazu, który w czasie rzeczywistym używając kamery internetowej będzie w stanie odczytać i wyświetlić na ekranie transkrypcję znaków języka migowego pokazywanych przez osobę znajdującą się w polu widzenia kamery. Dodatkowo do pracy będą składać się: utworzenie zbioru danych składającego się z około 50000 zdjęć zawierających wszystkie litery alfabetu ASL, utworzenie modelu z warstw konwolucyjnych i gęstych oraz wytrenowanie modelu i tuning parametrów. W części teoretycznej pracy zostanie przybliżony temat języka migowego American Sign Language, jak również temat konwolucyjnych sieci neuronowych i sposobu działania modeli.

# Rozdział 1

## Język migowy American Sign Language

American Sign Language (Amerykański Język Migowy, ASL) to złożony język wizualno-przestrzenny używany przez ludzi głuchoniemych w Stanach Zjednoczonych Ameryki oraz anglojęzycznych częściach Kanady. Jest to w pełni kompletny język naturalny. ASL jest językiem natywnym dla wielu głuchoniemych mężczyzn, kobiet i dzieci, a także niektórych słyszących dzieci w rodzinach, gdzie opiekunowie prawni są niesłyszący.[1]

### 1.1 Historia

Pochodzenie dzisiejszej społeczności osób głuchoniemych w Stanach Zjednoczonych jest powszechnie utożsamiane z założeniem pierwszej szkoły dla niesłyszących - American School for the Deaf (ASD), założonej w 1817 roku w Hartford, Connecticut. Przed założeniem szkoły ASD na terenie USA działało wiele niezależnych społeczności, poczynając od małych grup o wielkości pojedynczej rodziny do większych - całych wsi. W małych społecznościach uformowały się niezależne znaki i systemy języka migowego, które są obecne po dziś dzień w tych środowiskach. Istnieją dowody na to, że głuchonieme dzieci kształtowały swoje własne systemy języków migowych, które były o wiele bardziej wyrafinowane od tych, używanych w społeczności, w których się znajdowały.[2]

Istnieją również doniesienia o innym niezależnie uformowanym systemie języka migowego Martha's Vineyard Sign Language (MVSL), który istniał przed założeniem American School for the Deaf. Język ten był głównie używany w wsi Chilmark na wybrzeżu Massachusetts. Powstanie MVSL zapoczątkował fakt, że wspomniana społeczność Chilmark miała wysoki odsetek mutacji genetycznych prowadzących do głuchoty. W skali miasteczka około 4 procent mieszkańców było głuchoniemych. Wynikało to z wysokiego odsetka mieszanych małżeństw od wielu pokoleń, począwszy od hrabstwa Kent w Anglii, zanim wieś Chilmark została założona w roku 1690.[3] Mieszkańcy, którzy nie byli

głuchoniemi również posługiwali się językiem MVSL, wówczas gdy znajdowali się w towarzystwie osób z niepełnosprawnością ale również, gdy w gronie rozmówców nie było osoby głuchoniemej. Język był używany do czasu założenia szkoły dla osób głuchoniemych ASD w Hartford. Dzieci z wsi Chilmark zaczęto wysyłać do szkoły American School for the Deaf we wczesnych latach dwudziestych XIX wieku. Skutkowało to zatarciem się języków MVSL i nowego języka migowego ewoluującego w dzisiejszy język ASL.[2]

Kiedy szkoła dla głuchoniemych została założona, stała się miejscem, gdzie wiele pomniejszych systemów migowych stykało i mieszało się ze sobą przez 175 lat. Z mieszanki tych języków powstał dzisiejszy język ASL. Głuchoniemy nauczyciel - Laurent Clerc, pochodzący z Francji, był pierwszym nauczycielem w wspomnianej placówce. Z kraju swojego pochodzenia przywiózł wiedzę o Francuskim Języku Migowym, którego znaków nauczał w amerykańskiej szkole. Ta sytuacja spowodowała bardzo mocne doprawienie wówczas powstającego języka ASL o aspekty zaciągnięte z Francuskiego Języka Migowego. Na dzień dzisiejszy około 60 procent współczesnego systemu ASL opiera się na starym migowym języku Francuskim.[2]

## 1.2 Populacja

Informacje dostępne w internecie nie dają jednoznacznej odpowiedzi na pytanie - ile osób używa języka ASL. Wyszukane rezultaty są bardzo niejednoznaczne.[4] Znalezione szacunkowe (Tabela 1.1) liczby użytkowników języka ASL wahają się począwszy od 100000 do nawet 15000000. Warto zwrócić uwagę, że prawdopodobnie szacunki Aetna IntelliHealth zostały omyłkowo utożsamione z użytkownikami języka ASL, a tak naprawdę mogą przedstawiać liczbę osób ze znacznym i całkowitym ubytkiem słuchu.[4]

Starając się uogólnić wszystkie szacunki płynące ze źródeł internetowych, kształtują się dwa główne twierdzenia:

- Jest mniej niż dwa miliony użytkowników ASL, ale bardziej prawdopodobne, że ta liczba jest mniejsza i wynosi mniej niż półtora miliona ludzi w Stanach Zjednoczonych Ameryki;
- ASL może być trzecim najczęściej występującym językiem w Stanach Zjednoczonych.[4]

Trzeba zauważyć jednak, że publikacje i twierdzenia, w których został poruszony temat populacji użytkowników języka ASL pochodzą głównie z lat dwutysięcznych i powinny być brane pod uwagę z dystansem. Obecna liczba ludzi używających tego języka może być mniejsza lub znacznie bardziej prawdopodobnie - większa.

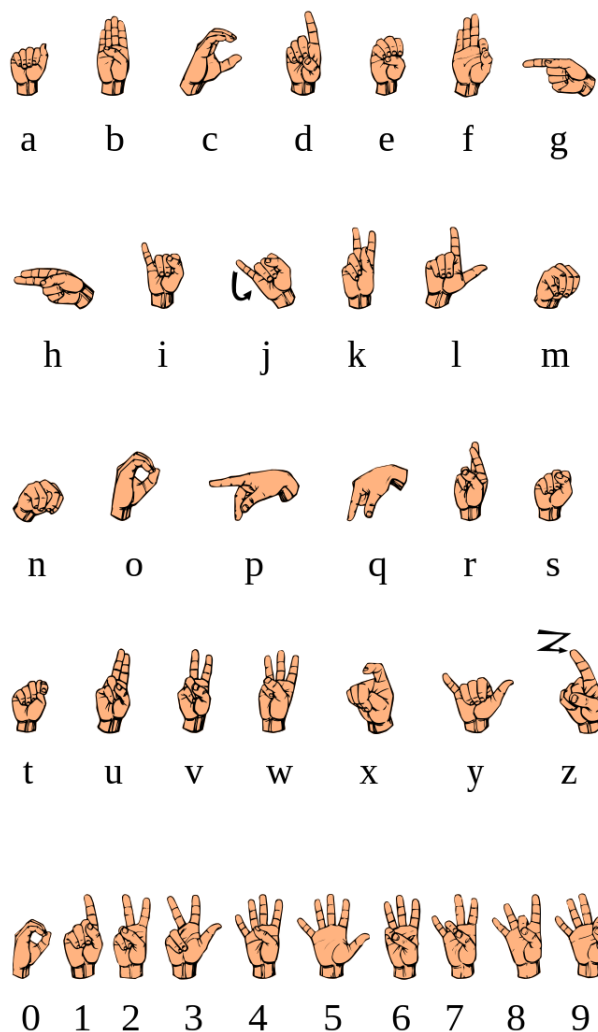
| Oszacowany ranking                         | Źródło                                                                                                                                                                                                                                                                 |
|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 100,000 – 500,000                          | ERIC Digests (Wilcox & Peyton, 1999)<br>MSN Encarta (Wilcox, 2004)<br>Ethnologue.com (Ethnologue, 2004)                                                                                                                                                                |
| 250,000 – 500,000                          | American Sign Language Program @ The University of Iowa (Department of Speech Pathology and Audiology, 2004)<br>ASLTA (NC ASLTA and NCAD Ad Hoc Committee, 2004)<br>Colorado Department of Human Services (Colorado Commission for the Deaf and Hard of Hearing, n.d.) |
| 300,000 – 500,000                          | Barnes&Noble.com (Costello, 1994)<br>SignWriting.org (Rosenberg, 1999)                                                                                                                                                                                                 |
| 500,000                                    | American Academy of Family Physicians (CDGAP, 1997)<br>ASLinfo.com (ASLinfo.com, n.d.)<br>DEAF C.A.N.! (Deaf Community Advocacy Network, n.d.)                                                                                                                         |
| 500,000 – 2,000,000                        | Brenda Schick, Ph.D. (Schick, 1998)<br>DawnSignPress (DawnSignPress, 2003)<br>Gallaudet University Library (Harrington, 2004)                                                                                                                                          |
| 15,000,000                                 | Aetna IntelliHealth (Gordon, 2001)                                                                                                                                                                                                                                     |
| 3rd most used language in the U.S.         | HandSpeak (HandSpeak.com, n.d.)<br>Health Literacy Consulting (Osborne, 2003)<br>Missouri Office of State Courts Administrator (Office of State Courts Administrator, n.d.)                                                                                            |
| 4th most used language in the U.S.         | The ASHA Leader Online (Scott & Lee, 2003)<br>Deaf Resource Library (Nakamura, 2002)<br>NIDCD (National Institute on Deafness and Other Communication Disorders, 2000)                                                                                                 |
| 3rd to 10th most used language in the U.S. | Wikipedia (Wikimedia, 2003)                                                                                                                                                                                                                                            |

Tabela 1.1: Szacowane rankingi użytkowników ASL według różnych źródeł

## 1.3 American Manual Alphabet

American Manual Alphabet to zbiór różnych gestów dłonią reprezentujący wszystkie litery angielskiego alfabetu. Są one ważnym elementem komunikacji i uzupełniają American Sign Language o zestaw znaków pozwalających na wyrażenie słów, dla których rozmówca nie zna konkretnego znaku lub nazw własnych, takich jak imiona czy nazwy firm. Często używane są przez użytkowników ASL dla upewnienia się, że idea tego co chcieli powiedzieć została dokładnie i zrozumiale zakomunikowana drugiej stronie rozmowy. Początkujący użytkownicy języka ASL, z racji ograniczonego zbioru znaków często posiłkują się alfabetem migowym do komunikowania słów, dla których nie znają

odpowiadających znaków.[5]



Rysunek 1.1: American Manual Alphabet [Wikipedia]

W alfabecie znajduje się 26 znaków odpowiadających literom języka angielskiego. Na odpowiedniki liter składa się 19 ułożeń dłoni. Niektóre ułożenia występują kilkakrotnie tak jak litery "i" i "j" oraz "p" i "k". Różnica pomiędzy tymi znakami polega na ruchu dłonią (przy literze "j" występuje ruch w kształcie półkola) lub pozycji względem ciała (przy literze "k" dłoń jest skierowana palcami do góry, a przy literze "p" w dół lub w bok).[5]



## Rozdział 2

# Konwolucyjne sieci neuronowe i rozpoznawanie obrazów

### 2.1 Sekcja C

### 2.2 Sekcja D

## Rozdział 3

# Implementacja aplikacji do rozpoznawania języka migowego ASL

## Rozdział 4

## Podsumowanie

# Spis tabel

|     |                                                                     |   |
|-----|---------------------------------------------------------------------|---|
| 1.1 | Szacowane rankingi użytkowników ASL według różnych źródeł . . . . . | 7 |
|-----|---------------------------------------------------------------------|---|

# Spis rysunków

|     |                                      |   |
|-----|--------------------------------------|---|
| 1.1 | American Manual Alphabet [Wikipedia] | 8 |
|-----|--------------------------------------|---|

## Spis listingów

# Bibliografia

- [1] K. Nakamura, “About american sign language,” *Deaf Research Library*, 1995.
- [2] B. Bahan, *Non-Manual Realization of Agreement in American Sign Language*. PhD thesis, Boston University, 1996.
- [3] N. E. Groce, *Everyone Here Spoke Sign Language: Hereditary Deafness on Martha’s Vineyard*. Harvard University Press Revised ed., 1988.
- [4] R. E. Mitchell, T. A. Young, B. Bachleda, and M. A. Karchmer, “How many people use asl in the united states?: Why estimates need updating,” *Sign Language Studies*, 2006.
- [5] E. Costello, *American Sign Language Dictionary*. Random House, 2012.