



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

Wstęp	3
Cel i zakres pracy	4
1 Język migowy American Sign Language	5
1.1 Historia	5
1.2 Populacja	6
1.3 Dystrybucja geograficzna	6
1.4 Składnia języka	6
2 Konwolucyjne sieci neuronowe i rozpoznawanie obrazów	7
2.1 Sekcja C	7
2.2 Sekcja D	7
3 Implementacja aplikacji do rozpoznawania języka migowego ASL	8
4 Podsumowanie	9
Spis tabel	10
Spis rysunków	11
Spis listingów	12
Bibliografia	12

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]

1.2 Populacja

1.3 Dystrybucja geograficzna

1.4 Składnia języka

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

Spis rysunków

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; Edycja Revised ed., 1988.