

Rozpoznawanie liter języka migowego z zastosowaniem technik uczenia maszynowego

Rafał Hrabia

Uniwersytet Marii Curie Skłodowskiej

30 stycznia 2022

Cel pracy

- Przybliżenie zagadnienia języka migowego American Sign Language oraz sieci konwolucyjnych CNN
- Stworzenie zbioru danych zawierających obrazy, na których znajdują się znaki języka migowego American Sign Language
- Stworzenie oraz wytrenowanie modelu rozpoznającego znaki języka migowego
- Implementacja programu (PoC) umożliwiającego rozpoznanie języka migowego w czasie rzeczywistym

Istota pracy

- Możliwe w przyszłości ułatwienie komunikacji osób nie znających języka migowego z osobami głuchoniemyymi

Język migowy American Sign Language

Jest to wariant języka migowego używanego w krajach obu Ameryk, gdzie natywnym językiem jest angielski.[1]

Posiada 26 znaków (American manual alphabet), używanych do wyrażania słów.[2]

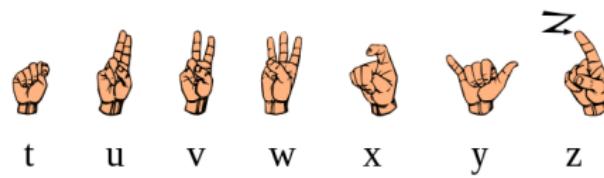
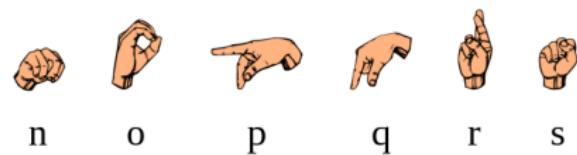
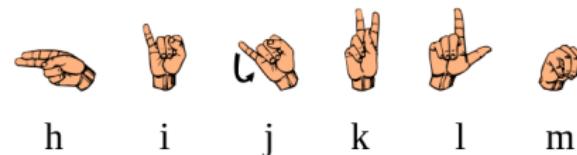
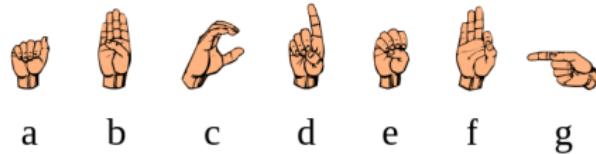
Składnia języka opiera się na subject–verb–object (SVO) czyli w zdaniu występuje podmiot, orzeczenie i obiekt, do którego skierowany jest czasownik.[3] Przykładowo:

FATHER LOVE CHILD

co będzie oznaczało

"The father loves the child."

Język migowy American Sign Language



Treningowy zbiór danych

Treningowy zbiór danych zawiera około 50 000 autorskich obrazów poszczególnych liter alfabetu ASL.

Litera	A	B	F
Obraz			

Model i architektura

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 200, 200, 32)	896
conv2d_1 (Conv2D)	(None, 200, 200, 32)	9248
max_pooling2d (MaxPooling2D)	(None, 66, 66, 32)	0
)		
conv2d_2 (Conv2D)	(None, 66, 66, 64)	18496
conv2d_3 (Conv2D)	(None, 66, 66, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 22, 22, 64)	0
conv2d_4 (Conv2D)	(None, 22, 22, 128)	73856
conv2d_5 (Conv2D)	(None, 22, 22, 128)	147584
max_pooling2d_2 (MaxPooling2D)	(None, 7, 7, 128)	0
conv2d_6 (Conv2D)	(None, 7, 7, 256)	295168
conv2d_7 (Conv2D)	(None, 7, 7, 256)	590080
flatten (Flatten)	(None, 12544)	0
dense (Dense)	(None, 1500)	18817500
dropout (Dropout)	(None, 1500)	0
dense_1 (Dense)	(None, 26)	39026

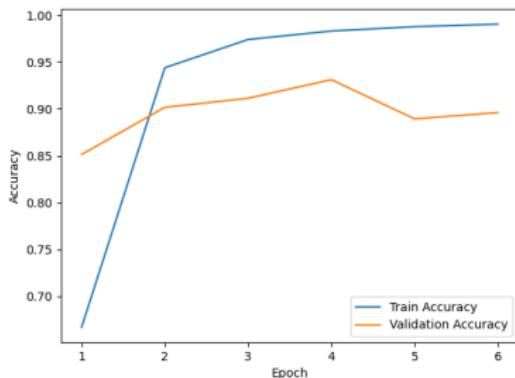
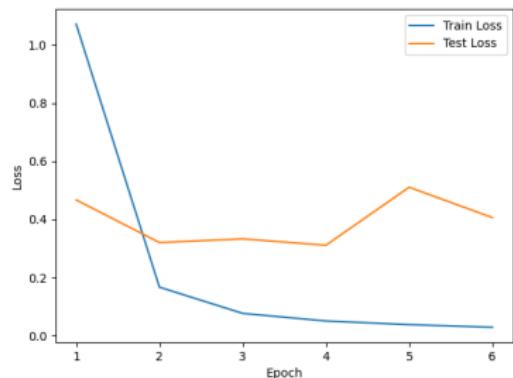
Total params: 20,028,782

Trainable params: 20,028,782

Non-trainable params: 0

Trening modelu

Model osiąga przyzwoite wyniki rzędu 85-90% dokładności klasyfikacji na zbiorze walidacyjnym.



Program do rozpoznawania języka migowego w czasie rzeczywistym



- [1] Carol Padden. *Sign Language Geography*. 2010.
- [2] Elaine Costello. *American Sign Language Dictionary*. Random House, 2008.
- [3] Carol Neidle. *The Syntax of American Sign Language: Functional Categories and Hierarchical Structures*. Cambridge, MA: The MIT Press., 2000.