

DW Poznań - #10 Augmentacja danych

2019-06-04

Czyli jak polepszyć model generując nowe dane

Agenda

01. **Dlaczego człowiek rozpoznaje obiekty**
02. **Historia Augmentacji - modele i ich jakość przed i po co.**
03. **Augmentacja tekstu**
<https://towardsdatascience.com/data-augmentation-in-nlp-2801a34dfc28>
<https://towardsdatascience.com/data-augmentation-library-for-text-9661736b13ff>
04. **Augmentacja obrazu.**
<https://www.tensorflow.org/tutorials/images/classification?hl=da>
<https://towardsdatascience.com/exploring-image-data-augmentation-with-keras-and-tensorflow-a8162d89b844>

Dlaczego człowiek rozpoznaje rzeczy

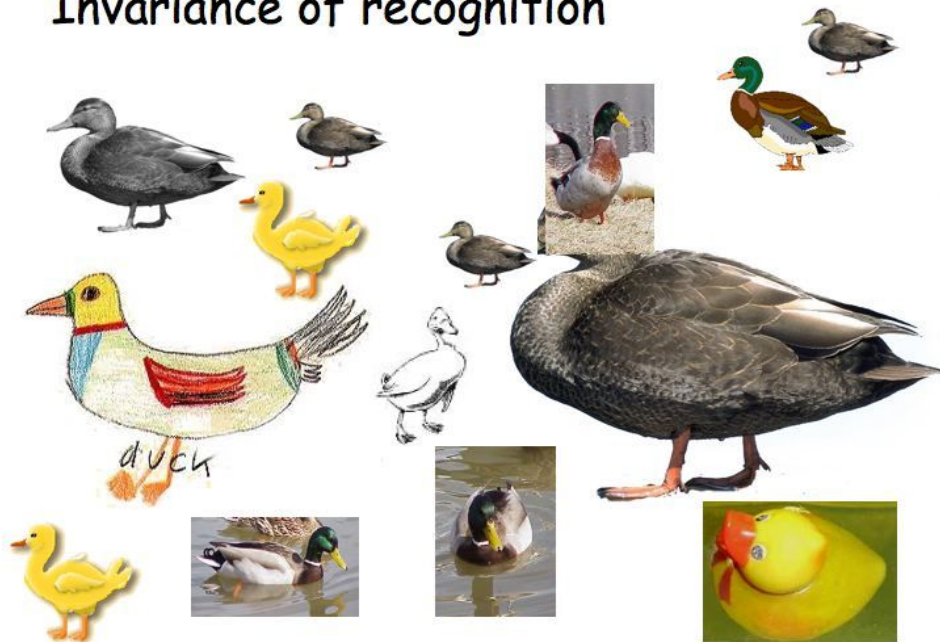
- Gdy człowiek raz zobaczy kota to potem rozpozna kota niezależnie od tego jak jest ustawiony, rozciągnięty, powiększony czy przesunięty.



Dlaczego człowiek rozpoznaje rzeczy

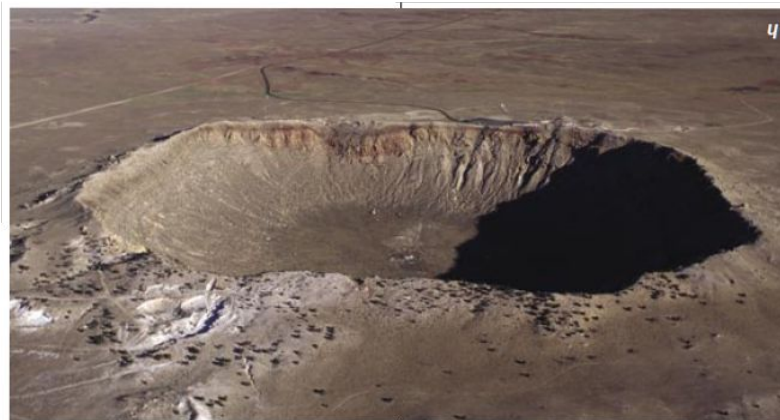
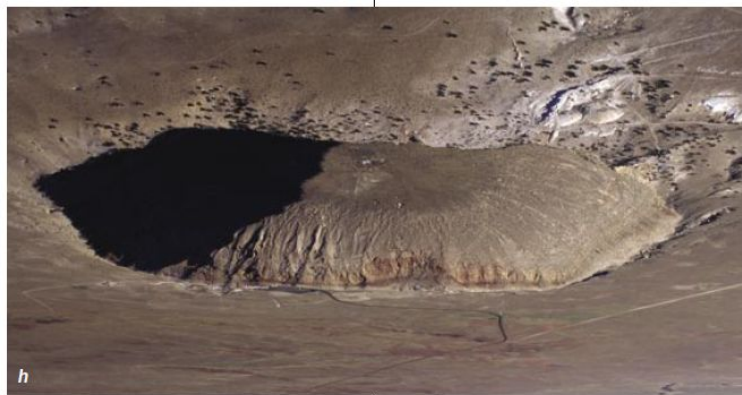
- Za każdym razem, gdy patrzymy na jakiś przedmiot, nasz mózg wyodrębnia jego cechy i w taki sposób, że wielkość, orientacja, oświetlenie, perspektywa itp. nie mają znaczenia. Pamiętasz obiekt przez jego kształt i nieodłączne cechy. Nie ma znaczenia, jak obiekt jest umieszczony, jak duży lub mały jest, ani z jakiej strony jest dla Ciebie widoczny.

Invariance of recognition

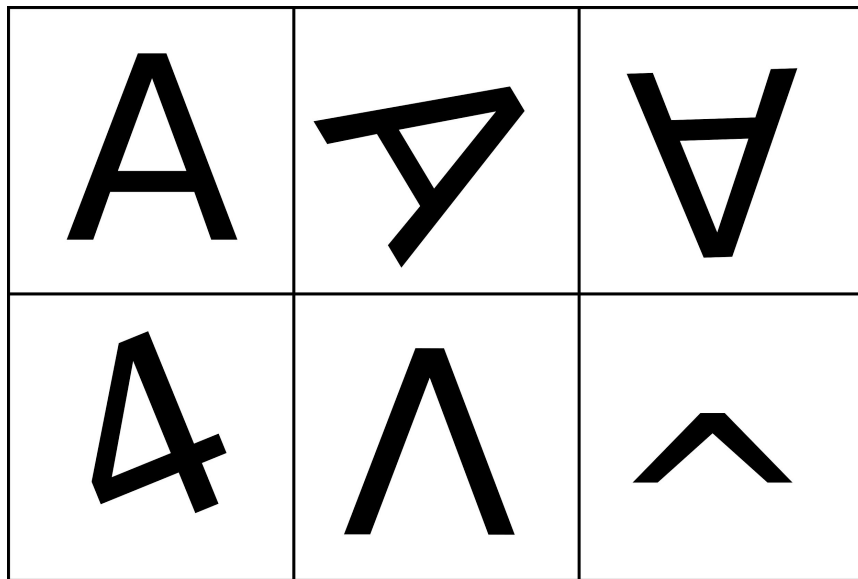


Dlaczego człowiek rozpoznaje rzeczy

- Nasz mózg wie (nauczył się), że światło pada zawsze z góry. W konsekwencji na lewym obrazku musi być góra a na prawym dolina.



Augmentacja - wstęp



Augmentacja - wstęp

Co to?

Zwiększenie liczby danych testowych przez zastosowanie różnorodnych transformat na pierwotnych danych

Zapobiega przeuczeniu

Warunki na których działa sieć mogą być bardziej zróżnicowane (kąt, oświetlenie itp)

Zapobiega uczeniu się niewłaściwych wzorców (np pies jest zawsze czarny, a kot biały)

Augmentacja tekstu

- Losowe usuwanie słów / liter
- Losowe dodawanie słów / liter
- Podmiana synonimów
- Losowa zmiana kolejności

Augmentacja tekstu

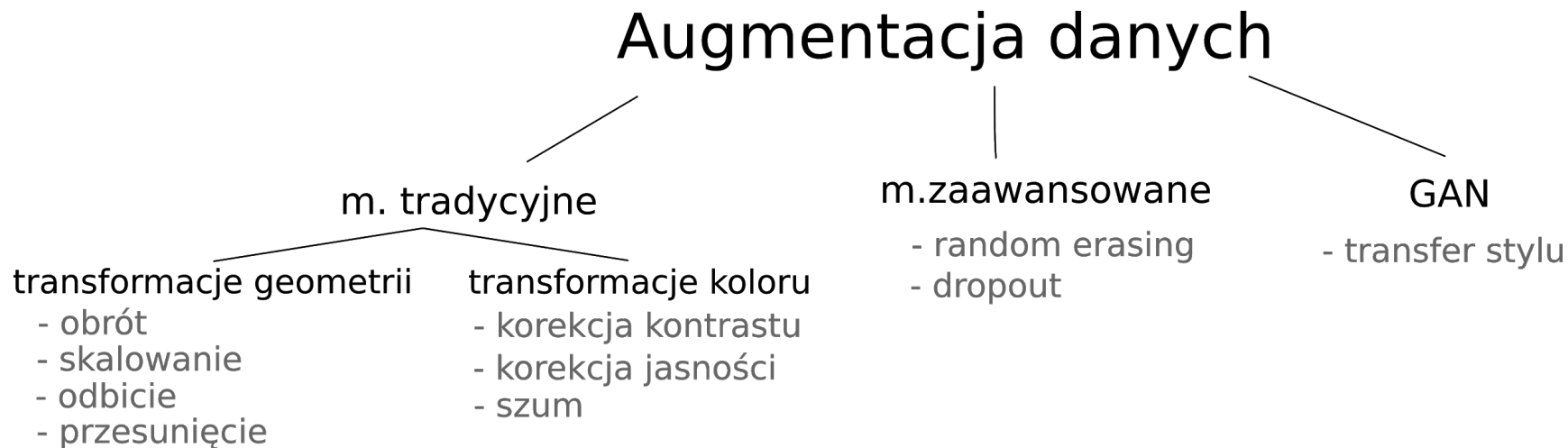
01. TextAugment

<https://github.com/dsfsi/textaugment>

02. Nlpaug

<https://pypi.org/project/nlpaug/>

Augmentacja obrazu



Augmentacja obrazu

01. Augmentor

<https://github.com/mdbloice/Augmentor>

02. Imgaug

<https://github.com/aleju/imgaug>

03. Albumentations

<https://albumentations.readthedocs.io/en/latest/#>

04. Keras ImageDataGenerators

<https://keras.io/api/preprocessing/image/#imagedatagenerator-class>

Kolejne kroki

- Spotkanie za 2 tygodnie

<https://github.com/dataworkshop/dw-poznan-project>

Dziękuję