

„ROZPOZNAWANIE KOTKÓW OD PIESKÓW”

Michał Skibiński (260352)

30.05.2022

spis treści

1	Wstęp	2
1.1	cel projektu	2
1.2	narzędzia	2
2	Model	3
2.1	przygotowanie danych	3
2.2	budowa modelu	3
2.3	przebieg treningu	4
3	Wyniki	5
3.1	kryterium jakości modelu	5
3.2	jakość modelu	5
3.3	przyczyny takiej a nie innej jakości modelu	7
4	przypisy	7

1 Wstęp

1.1 cel projektu

Przedmiotem projektu jest zadanie klasyfikacji binarnej (klasyfikacji kotów oraz psów).

1.2 narzędzia

- problem rozwiązany został za pomocą **konwolucyjnej sieci neuronowej (CNN)**. Ponieważ sieci neuronowe operują na macierzach, są one szczególnie wydajne przy pracy z obrazami, które tak naprawdę są macierzami pikseli.
- Rozwiązanie zaprojektowane zostało w języku **python**, ze względu na jego prostotę, niezawodność oraz wielorakość bibliotek ogólnodostępnych.
- biblioteka użyta do rozwiązania problemu to **tensorflow**. Została ona użyta ze względu na możliwość wykonywania złożonych operacji na sieciach neuronowych, przy użyciu stosunkowo niewielkiej ilości kodu. Oprócz tego biblioteka ta działa bardzo efektywnie, oraz została do niej napisana bogata dokumentacja.

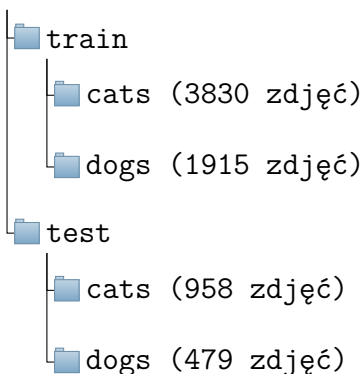
2 Model

2.1 przygotowanie danych

aby przygotować dane kolejno:

1. zaprojektowany został algorytm do odczytywania zdjęć z zadanego folderu i tworzeniu folderu plików z podziałem na **dane testowe i treningowe** oraz klasy, tak by struktura wyglądała następująco:

input for model



ilość danych testowych zdefiniowana została na 20 procent wszystkich danych.

2. usunięta została różnica w ilości fotografii kotów i psów (ważne aby sieć neuronowa miała tyle samo danych z każdej z klas)
3. dokonana została **normalizacji danych** - wartości pikseli zostały przeskalowane, tak by pochodziły z zakresu $[0,1]$
4. dokonana została standaryzacja danych - wszystkie zdjęcia przeskalowane zostały do rozmiarów $[224 \times 224]$ px.
5. redukcja wymiarów danych przy użyciu **PCA** nie została zastosowana, ponieważ wiązała się ze zbyt dużą utratą danych.

2.2 budowa modelu

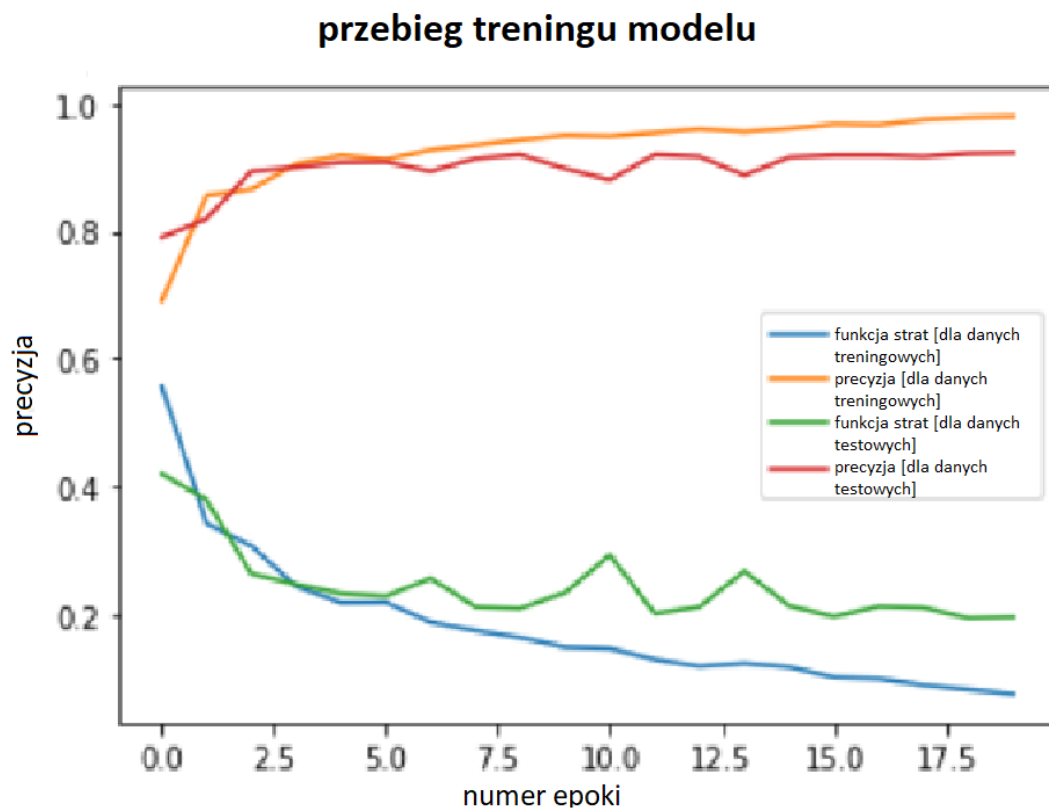
model inspirowany jest dotychczasowymi modelami do rozwiązania tego typu problemów. Zanim pojawił się finałowy model stworzone zostało ok 10 innych modeli, każdy był usprawnieniem poprzedniego. przy udoskonalaniu modelu, nacisk kładziony był na zamianie **funkcji aktywujących** dla poszczególnych warstw oraz liczbie ich liczbie.

własności modelu:

- wejście dla modelu: macierz o rozmiarach $(224 \times 224 \times 3)$ - gdzie 3 wynika z reprezentacji pikseli poprzez RGB

- liczba warstw ukrytych: 20
- wyjście: wartość 0 lub 1 (0 - dla psa, 1 - dla kota)
- metoda optymalizacji: **stochastyczny spadek gradientu(SGD)** z prędkością nauki = 0.0001
- funkcja strat: binarna entropia krzyżowa
- ilość generacji(epok): 20
- populacja w każdej generacji: 60
- funkcja regularyzacji: l2 (bez regularyzacji występowało **przetrenowanie modelu**)

2.3 przebieg treningu



3 Wyniki

3.1 kryterium jakości modelu

jakość modelu mierzona została ze względu na **precyzję**. kryterium to zostało wybrane, ze względu na jego prostotę, precyzja jest dobrą metryką do oceny wydajności modelu w prostych modelach, takich jak ten.

$$\text{precyzja} = \frac{\text{ilość poprawnie sklasyfikowanych zdjęć}}{\text{ilość wszystkich zdjęć}} * 100\%$$

$\text{precyzja} \in [0\%, 100\%]$ - im wyższa tym lepsza

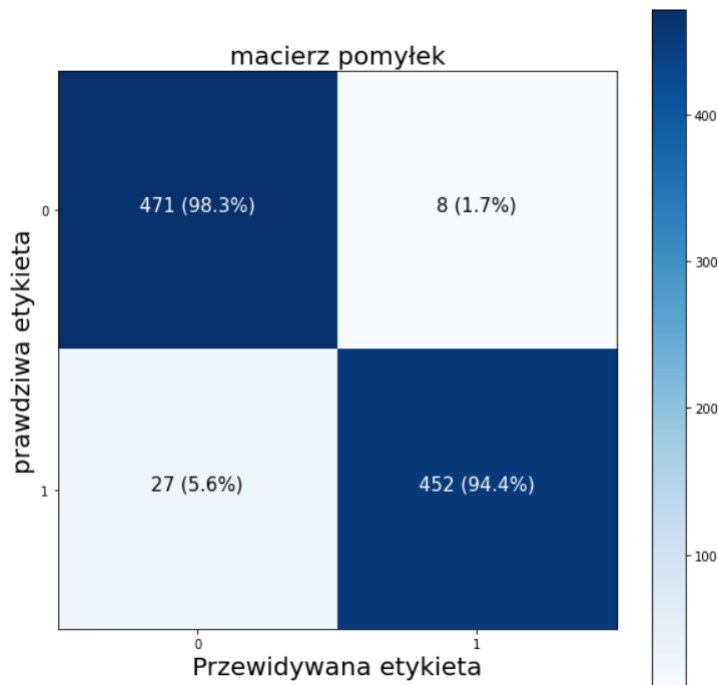
3.2 jakość modelu

po skończonym treningu precyzja:

- dla danych treningowych wyniosła: **98 procent**
- dla danych testowych: **93 procent**

Do przedstawienia graficznie precyzji dla danych testowych modelu użyta została macierz pomyłek (confusion matrix).
gdzie:

- 0 - etykieta dla *psa*
- 1 - etykieta dla *kota*



Jak widać model ma większe trudności z poprawnym zaklasyfikowaniem kotów niż psów.

Prawie 6 % kotów zostało sklasyfikowanych jako psy - a tylko niecałe 2 % psów zostało błędnie sklasyfikowanych jako koty

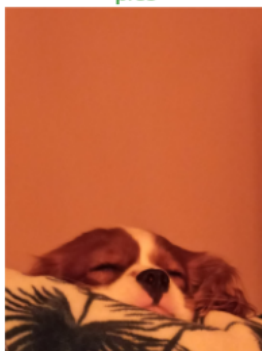
Aby zweryfikować poprawność modelu uruchomiony on został dla zdjęć zrobionych przeze mnie oraz moich znajomych, tak aby mieć pewność że zdjęcia są różnych formatów, rozmiarów, różnej jakości i nie były wcześniej w żaden sposób przerobione

przykładowe predykcje modelu:

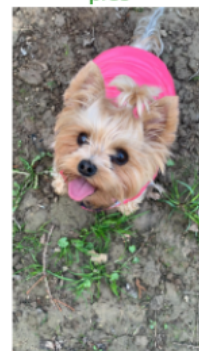
pies



pies



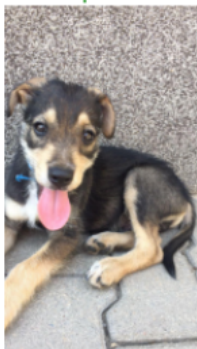
pies



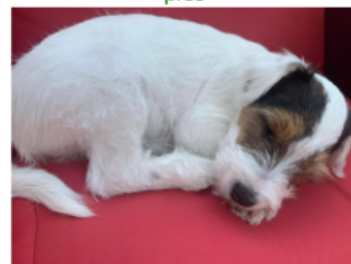
kot



pies



pies



3.3 przyczyny takiej a nie innej jakości modelu

model okazał się stosunkowo skuteczny, przy klasyfikacji zdjęć nie spreparowanych skuteczność na poziomie 93 procent, możemy uznać za zadowalającą.

dłaczego model nie ma skuteczności na poziomie 100 procent?

1. ponieważ urządzenie na którym prowadzone było trenowanie modelu dysponowało niewielką mocą obliczeniową, a operacje na zdjęciach, są skomplikowane - trenowanie sieci zakończone zostało gdy jej dokładność miała dalej tendencje rosnące.
2. niektóre ze zdjęć mogły być nieostre, lub przedstawiać zwierzę które jest mieszanką genetyczną psa i kota. przykład takich zdjęć:



4 przypisy

zdjęcia użyte do budowy dla modelu:

<https://www.kaggle.com/datasets/zippyz/cats-and-dogs-breeds-classification-oxford-dataset>

kod rozwiązania:

<https://github.com/michalskibinski109/projekt/blob/main/PROJEKT.ipynb>