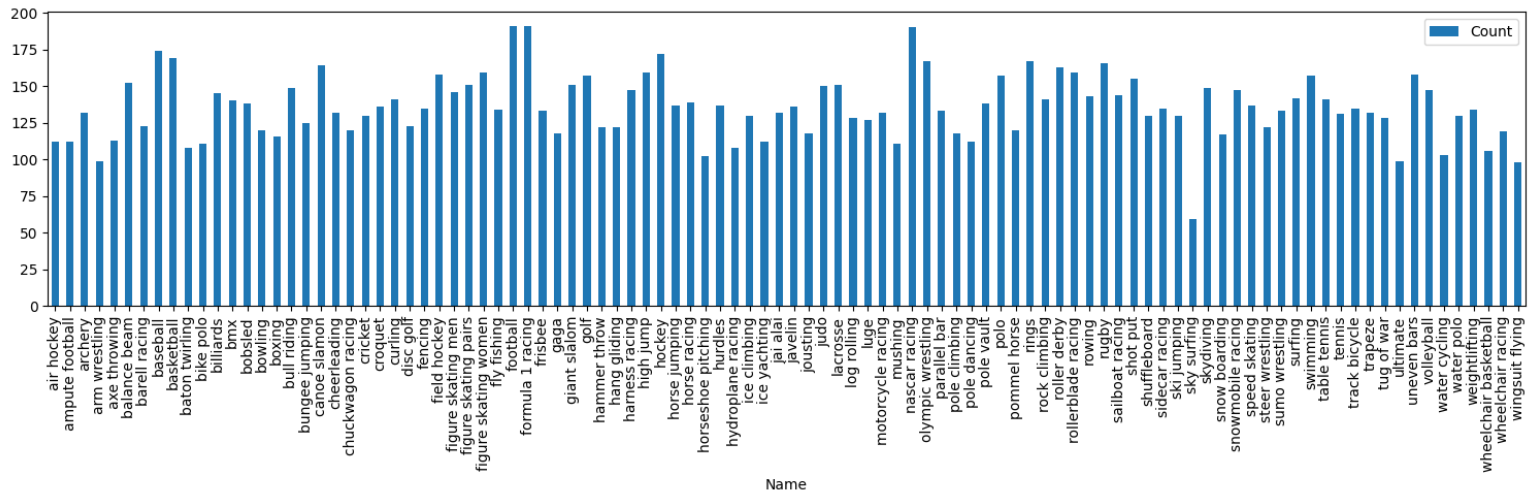


I-SUNS Zadanie 3

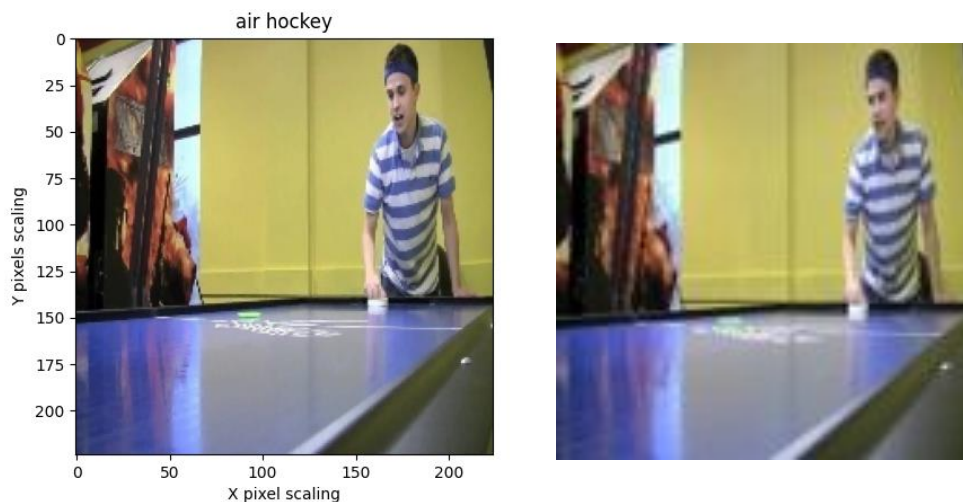
Tomáš Minárik

1. EDA

1.1. Početnosti v triedach.



- Počet všetkých obrázkov je 13572.
- Priemerný počet obrázkov na triedu je 135,7.
- Najviac početné triedy sú football(191), formula1(191) a nascar(190).
- Najmenej obrázkov obsahuje trieda sky surfing(59).



- Vľavo je zobrazený obrázok v originálnej veľkosti 224x224. Vpravo je zobrazený obrázok vo veľkosti 128x128, ktorý bude použitý na tréningovanie.

1.2. Farebnosť pixelov

- Všetky grafy sa nachádzajú v colors.pdf.
- Na grafoch je zobrazených 5 najčastejšie sa vyskytujúcich farieb. Zobrazené zľava do prava.
- Z grafov môžeme vidieť na ako povrchu sa šport odohráva

- Tráva



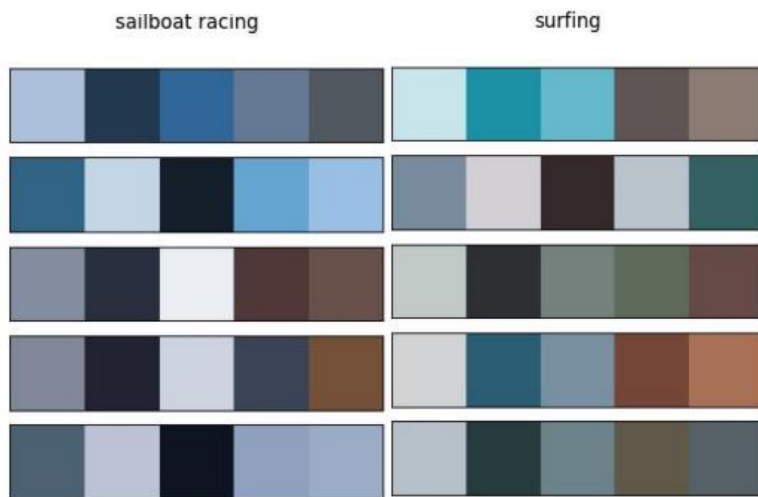
- Ľad



- Halové športy



- Vodné športy



1.3. ImageNet

- Vybrali sme si sieť NASNetMobile Keďže mala veľkosť len 23mb.
- Pre každý obrázok sme zobrali 2 výsledky s najväčšou pravdepodobnosťou a potom sme zobrali 3 príznaky, ktoré sa v triede najčastejšie vyskytovali.
- Výsledky sú v súbore results.csv.
- V 4 prípadoch sieť trafila šport: baseball, basketball, bobsled, volleyball.
- Sieť rozpoznávala hlavne predmety s ktorými sa športy vykonávajú.

archery	bow	walking_stick	rifle
hockey	puck	ski	ski_mask
rugby	rugby_ball	soccer_ball	football_helmet

- V niektorých prípadoch by sa trieda z výsledkov nedala určiť.

hammer throw	racket	volleyball	tennis_ball
olympic wrestling	horizontal_bar	balance_beam	parallel_bars

2. Trénovanie

- Načítali sme train, val a test data.
- Všetky obrázky sme zmenšili na veľkosť 128x128.
- Augmentácia tréningových dát.

```
width_shift_range=0.2,
height_shift_range=0.2,
zoom_range=0.2,
horizontal_flip=True)
```

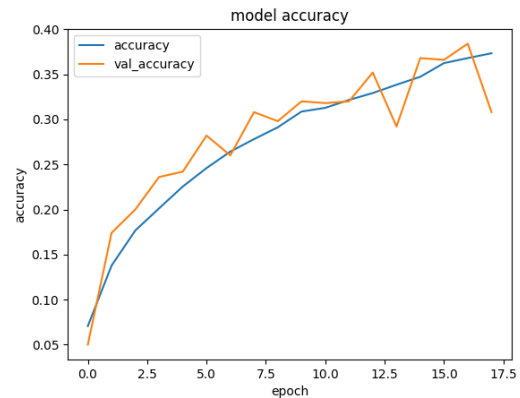
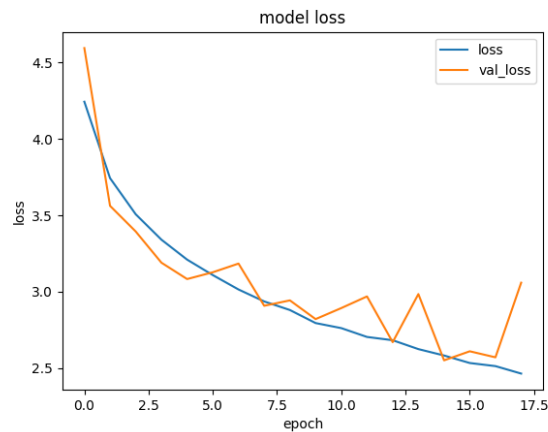
- Takisto sme pomocou generátoru dáta normalizovali.

```
ImageDataGenerator(rescale=1. / 255)
```

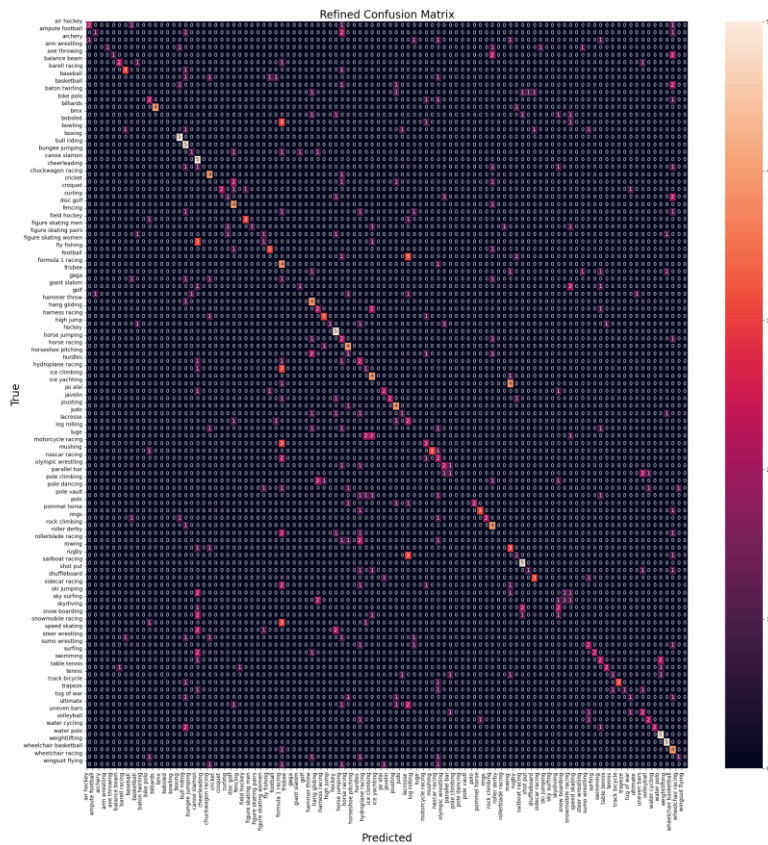
2.1. Sieť s 2 konvolučnými vrstvami.

- Obidve vrstvy majú veľkosť filtra 64 a veľkosť okna filtra 3x3.

- Aktivačnú funkciu sme použili relu.
- Výstupná vrstva má veľkosť 100 a aktivačnú funkciu softmax.
- Kriteriaľnú funkciu sme použili categorical crossentropy.
- Batch_size bol 50 a počet epoch bol 30.
- Optimizer používame SGD s learning rate=0,001.
- Použili sme earlystopping s patience=3.



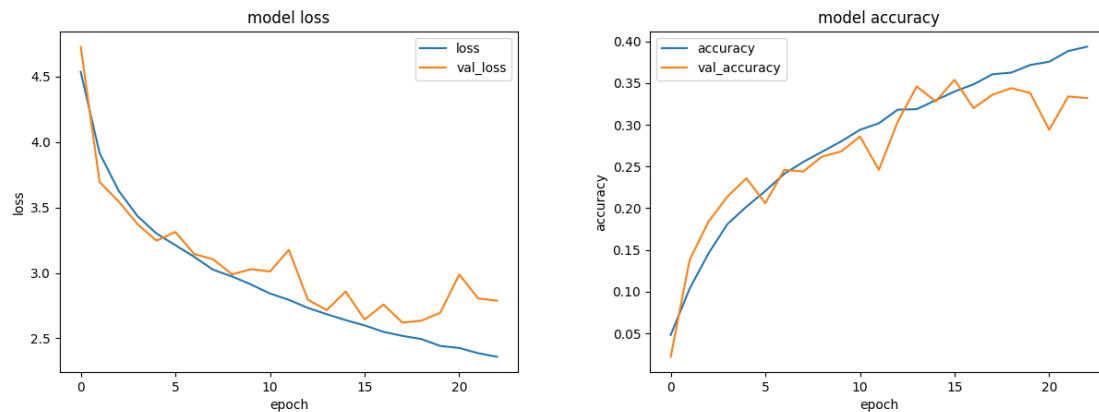
- Trénovanie bolo ukončené po 18 epoche.
- Úspešnosť na trénovacích dátach bola 37% a testovacích 33%.



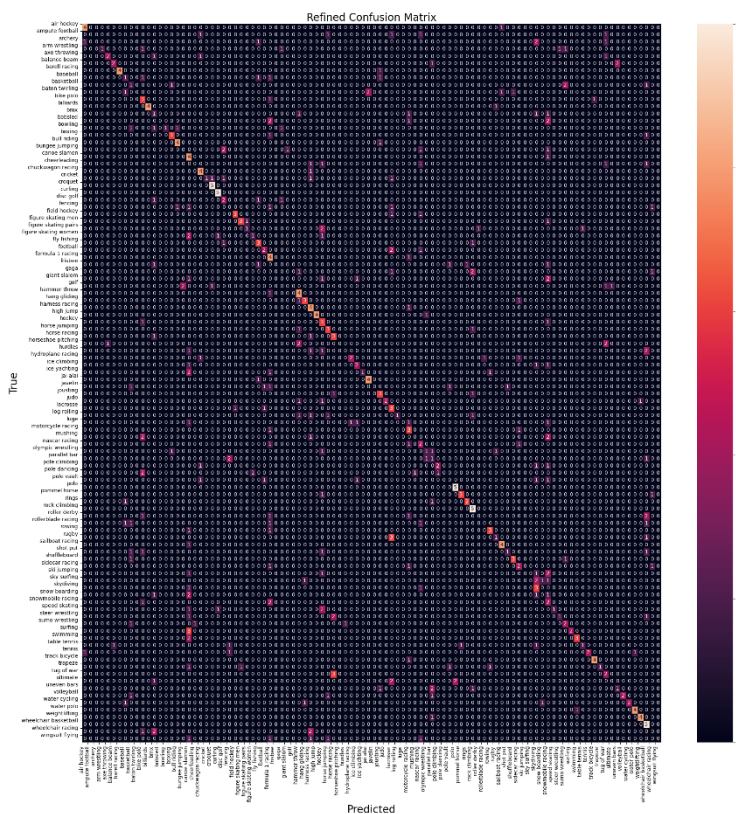
- Sieť bola najpresnejšia pre: Boxing, Bull riding, Canoe slalom, Hockey, Sailboat racing, Water polo a Weightlifting.
- Pre ice yachting sieť určila 4 vzorky ako rowing.

2.2. Sieť s 3 konvolučnými vrstvami.

- Na sieť sme použili rovnaké parametre ako na predchádzajúcu sieť s tým že 3. vrstva bola rovnaká ako predošlé 2.
- Early stopping mal pri tejto sieti patience=5.



- Trénovanie skončilo na 23. epoche.
- Úspešnosť na trénovacích dátach bola 39% a na testovacích 38%.



- Sieť bola najpresnejšia pre: Croquet, Curling, Polo, Rock climbing a Wheelchair basketball.

3. Testovanie na vlastnej databáze.



	Class	Pred1
67	rock climbing	0.698752
30	fly fishing	0.118365
24	disc golf	0.115264
20	chuckwagon racing	0.022582
37	hammer throw	0.009843



	Class	Pred2
58	nascar racing	0.348468
97	wheelchair basketball	0.283945
32	formula 1 racing	0.260637
39	harness racing	0.045970
80	snowmobile racing	0.027951



	Class	Pred3
32	formula 1 racing	0.636575
58	nascar racing	0.309429
80	snowmobile racing	0.025511
56	motorcycle racing	0.009587
6	barell racing	0.009463



	Class	Pred4
58	nascar racing	0.347958
32	formula 1 racing	0.306110
18	canoe slamon	0.084684
37	hammer throw	0.082666
80	snowmobile racing	0.043633



	Class	Pred5
67	rock climbing	0.722017
30	fly fishing	0.247735
24	disc golf	0.012852
20	chuckwagon racing	0.007882
39	harness racing	0.002822

