



# Kő-papír-olló

Képfeldolgozás haladóknak MSc gyakorlat

Hergert Lea, Balogh András, Cservenák Bence



# Feladat

Kő-papír-olló játék

- Egyszerű, intuitív felület
- Automatikus detekció
- Több játékmód

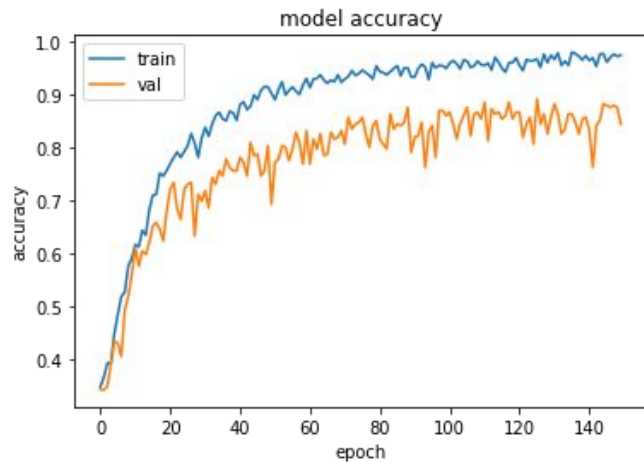


# Megvalósításhoz használt eszközök

- PyQt
- OpenCV
- Keras
- Mediapipe
- scikit-learn

# Detektálás neurális hálóval: “homemade” háló

- Saját konvolúciós hálózat
  - 5 réteg: 3 konvolúciós- és pooling-, 2 teljesen csatolt réteg
  - 3 osztály
- >1300 tanító / teszt kép
- Augmentáció
  - Zoom
  - Eltolás
  - Forgatás
  - Tükrözés
- Detekció pontossága: ~85%



# Detektálás neurális hálóval: “transfer” háló

- Transfer learning
  - MobileNet architektúra [1]
  - ImageNet adatbázison tanítva
- Legfelső réteg kicserélése
- 1300 kép, augmentáció
- Detekció pontossága: ~95%

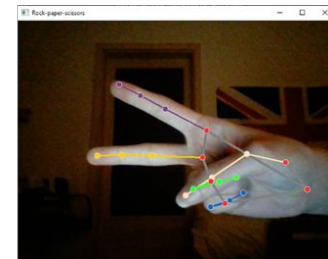
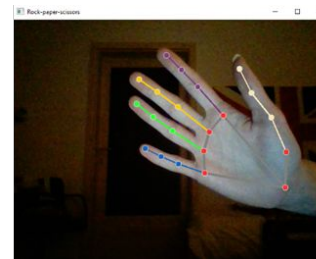
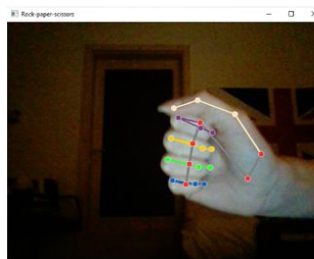
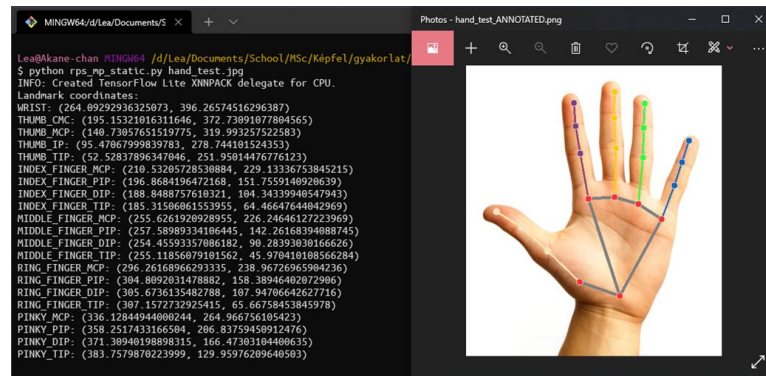
Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32 \text{ dw}$	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64 \text{ dw}$	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128 \text{ dw}$	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128 \text{ dw}$	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256 \text{ dw}$	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256 \text{ dw}$	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5×	Conv dw / s1	$3 \times 3 \times 512 \text{ dw}$
	Conv / s1	$1 \times 1 \times 512 \times 512$
Conv dw / s2	$3 \times 3 \times 512 \text{ dw}$	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024 \text{ dw}$	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool $7 \times 7$	$7 \times 7 \times 1024$
FC / s1	$1024 \times 1000$	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

[1]: Howard, Andrew & Zhu, Menglong & Chen, Bo & Kalenichenko, Dmitry & Wang, Weijun & Weyand, Tobias & Andreetto, Marco & Adam, Hartwig. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

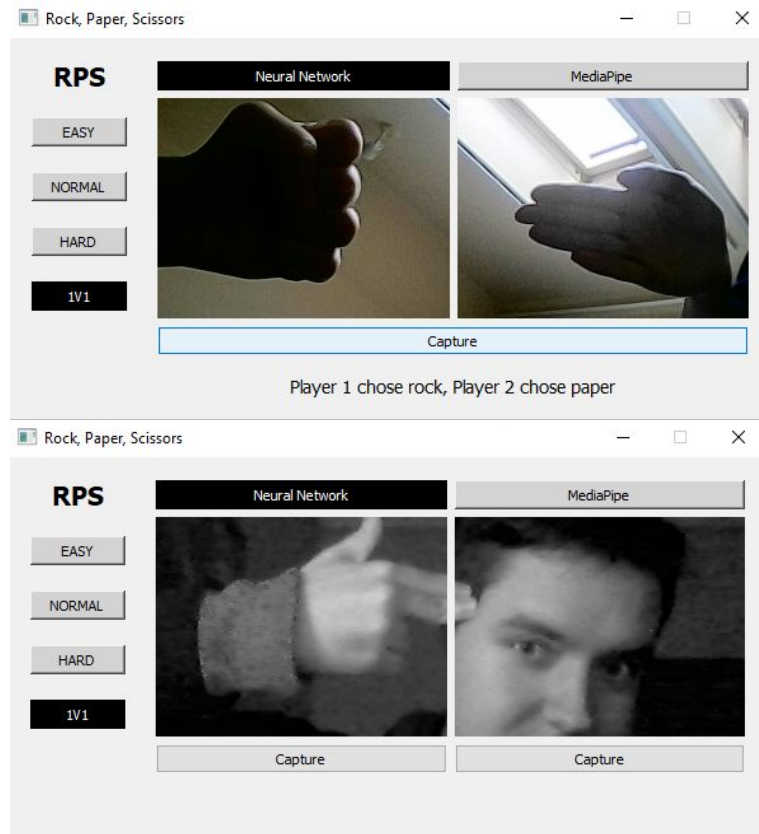
# Detektálás MediaPipe keretrendszerrel

- MediaPipe: kézdetektálás
- Detektált kulcspontokból jellemzők származtatása
  - Kulcspontok páronkénti távolsága
- Osztályozás
  - 4 osztály
  - Random Forest osztályozó
- Detekció pontossága: ~98% videófolyamon, ~75% statikus képen

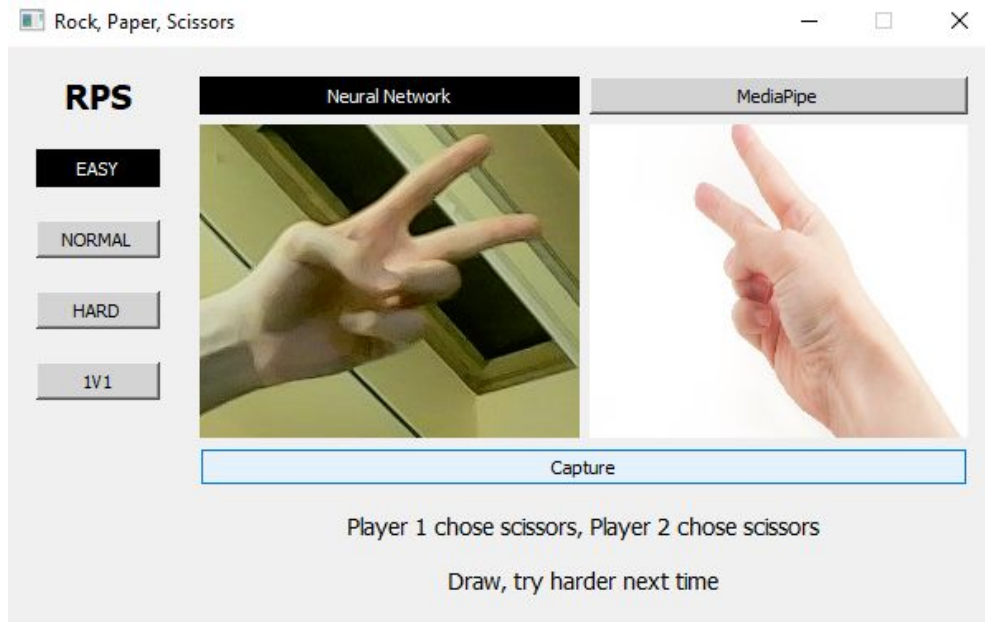


# Játék

- Számítógép vagy másik játékos ellen
- 256x256-os kamerakép
- Játékmódok:
  - Easy: véletlenszerű
  - Normal: utolsó két lépés alapján
  - Hard: játékos mindig veszít
  - 1v1: két játékos



# Játékelmény







# Demo

Videó: <https://youtu.be/5qZuU7YMEGQ>



**Köszönjük a figyelmet!**