

# Klasifikace pomocí CNN

Hledání klávesnice ve videu

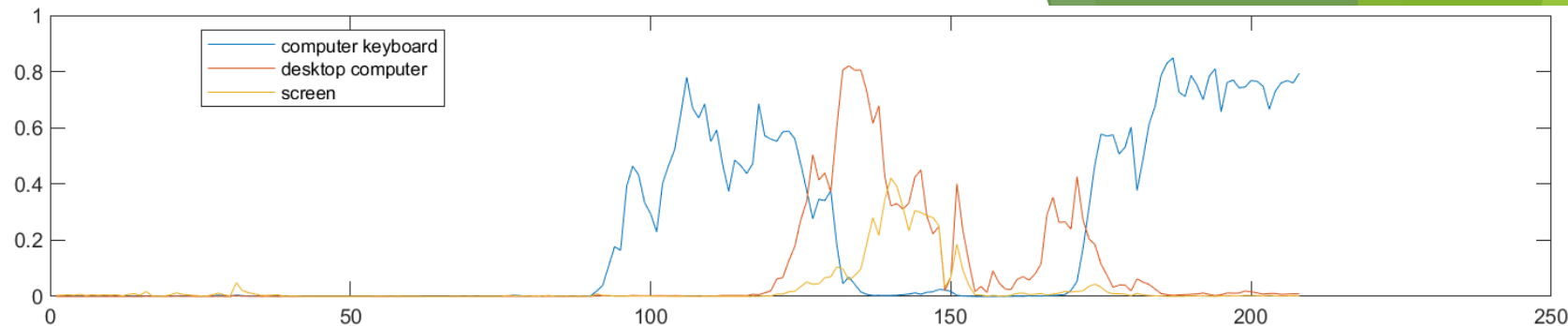
# Úvod

- ▶ Minulou hodinu jsme sestavili jednoduchou CNN (konvoluční neuronovou síť) a klasifikovali obrázky čísel, což je klasická úloha klasifikace.
- ▶ Klasifikace nám prozradí, co se na obrázku nachází. Už nám neřekne kde, k tomu slouží algoritmy pro detekci objektů (pochopitelně jsou založené na podobných principech)
- ▶ ImageNet -> Je rozsáhlá databáze obrázků a soutěž v jejich rozpoznávání. Z této soutěže pochází konvoluční neuronová síť Alexnet - V roce 2012 s výrazným náskokem zvítězila.
- ▶ Alexnet (a jiné CNN) jsou natrénovány na rozpoznávání obrázku do 1000 tříd.
  - ▶ Jejich seznam: [https://zodoc.tech/posts/en/imagenet\\_classes](https://zodoc.tech/posts/en/imagenet_classes)
- ▶ Alexnet a jiné natrénované sítě můžete velmi snadno v matlabu:
  - ▶ <https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html>

# Klasifikace

- ▶ Ke klasifikaci libovolného obrázku je potřeba neuronová síť a obrázek ve správném formátu (rozměr a počet kanálů). Konkrétně pro Alexnet je to 227: [https://zodoc.tech/posts/en/classify\\_by\\_alexnet](https://zodoc.tech/posts/en/classify_by_alexnet)
- ▶ Výstupem funkce classify je popisek (label) o kterém se síť domnívá, že je na obrázku. Druhým výstupním parametrem je vektor pravděpodobností všech tříd. [https://zodoc.tech/posts/en/classification\\_results\\_in\\_barplot](https://zodoc.tech/posts/en/classification_results_in_barplot). V součtu dává 1. Neuronová síť vždycky něco vymyslí, vždycky obrázek vyhodnotí jako „něco“. U jistějších klasifikací mívá hodnotu skóre vyšší (viz květina a jejich 98 %).
- ▶ Do kterých labelů síť klasifikuje se dá vytáhnout z poslední vrstvy sítě.

# Úloha



Výsledek pro 3 zvolené třídy

- ▶ Nalezněte místa ve videu (soubor kancl.mp4), kdy je zabraná klávesnice.
  - ▶ Klasifikujete každý snímek.
    - ▶ Využijte libovolnou neuronovou síť ke klasifikaci. Výsledky se u různých sítí liší.
      - ▶ Zde je jejich seznam, včetně předpokládané přesnosti a velikost: <https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html>
      - ▶ Skript koncipujte tak, aby se jedním příkazem dala síť přehodit a všechno fungovalo stejně (jedná se především o vstupní velikost sítě)
- ▶ Výstupem bude:
  - ▶ Graf, na jehož xové ose bude číslo snímku. Na y ose bude skóre které dosáhla klávesnice v daném snímku.
  - ▶ V grafu bude společně s klávesnicí i třída, která má celkově největší součet skóre přes všechny snímky + napište proč si myslíte, že tomu tak je. Jestli je ta třída správná, atd..
  - ▶ Tedy i nějaká diskuze nad výsledkem. Je to použitelné, proč to funguje/nefunguje. Jak je to s rozdílem mezi jednotlivými sítěmi. Jak by se to daly chyby napravit, atd..  
Diskutujte též, jak by

# Rady a poznámky

- ▶ Video si převed'te do 4 rozměrné struktury
  - ▶ První 2 rozměry jsou velikosti obrázků, tak aby odpovídali vstupní vrstvě sítě.
  - ▶ Třetí rozměr je 3, tedy počet barevných kanálů
  - ▶ 4. rozměr je počet snímků ve videu
  - ▶ Funkce predict, či classify (jsou vlastně stejné) dokáže požrat tuto strukturu a rovnou vám v matici vyhodnotit skóre pro každý snímek a každou třídu.
- ▶ Video je naschvál v otřesné kvalitě, tak aby zabíralo málo místa a bylo snadné s ním operovat.