

# Implementácia a využívanie modelov počítačového videnia v praxi \*

Ján Mareček

Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií

xmarecek@stuba.sk

00000000

3. november 2021

School of Hard Knocks

SOCI4568 L01 Sociology of Physics

For: Professor Y.R.U StillHere

---

## Abstrakt

Odborný článok na tému Implementácia a využívanie modelov Počítačového videnia v praxi "pozostáva z 4 kapitol. Prvá kapitola je venovaná používaniu počítačového videnia s využitím programovacích jazykov ako Python a C++ a knižnice OpenCV. Druhá kapitola je venovaná druhom a trénovaniu modelov počítačového videnia. Tretia kapitola je venovaná výhodám a nevýhodám počítačového videnia. Obsahom štvrtej kapitoli je využitie počítačového videnia v praxi.

## 1 Počítačové videnie

Počítačové videnie je automatizovaná extrakcia informácií z obrázkov. Informáciu môže predstavovať hocičo od 3D modelov, videa alebo obrázka. Niekedy sa počítačové videnie snaží napodobňovať ľudské videnie, inokedy využíva dáta alebo štatistiky. Praktické počítačové videnie je mix programovania, modelovania a matematiky. Matematická časť pomáha pochopiť špecifické algoritmy, ktoré sú používané.

---

\*Semestrálny projekt v predmete Metódy inžinierskej práce, ak. rok 2020/21, vedenie: Vladimír Mlynarovič

## 2 Programovacie jazyky a knižnice počítačového videnia

Jeden z najpopulárnejších, ak nie najpopulárnejší programovací jazyk pre prácu s počítačovým videním je Python. Python je interpretovaný, interaktívny programovací jazyk s dobrou podporou na spracovávanie obrazu. Od verzie 2.6 máme pri práci s Pythonom k dispozícii väčšinu balíkov, ktoré pri práci potrebujeme. Od verzie 3.x.x si zmenila syntax pre ale kompatibilita s balíkmi pre počítačové videnie ostala rovnaká. [Sol]

NumPy je knižnica pre programovací jazyk Python. Je užitočná na prácu s vektormi, maticami a ich operáciami na vysokej úrovni, ktoré sú potrebné na fungovanie počítačového videnia. Dané typy sú reprezentované pomocou polí.

V počítačovom videní sa stretneme aj s inými knižnicami jazyka Python. Knižnica Matplotlib je využívaná pre vizualizáciu výsledkov. Knižnica SciPy je využívaná keď potrebujeme využívať pokročilejšie matematické funkcie alebo algoritmy. [Sol]

OpenCV je open-source knižnica napísaná v C++ ,ktorá obsahuje algoritmy a funkcie zamerané na počítačové videnie. Najčastejšie sa používa s programovacími jazykmi Python, C a C++ ,ale v novších verziach už bola vydaná podpora aj na JavaScript. Má modulárnu štruktúru čo znamená, že obsahuje niekoľko hlavných knižníc.

- Hlavné knižnice OpenCv sa zaraďujú napríklad:
  1. Image Processing-modul na spracovávanie obrázkov
  2. Video Analysis-modul na spracovávanie videí
  3. Core functionality-modul definujúci základné dátové štruktúry a funkcie ,ktoré využívajú ostatné moduly
  4. Object Detection-detekcia objektov a inštancie preddefinovaných tried (napr. tvár, oči, ľudia, autá...)
  5. High-level GUI – ľahko použiteľné rozhranie na jednoduché UI funkcie.
  6. Camera Calibration and 3D Reconstruction – kalibrácia kamery, odhad polohy objektu, 3D rekonštrukcia.
  7. ... [?]

## 3 Modely a ich trénovanie

### 3.1 Typy modelov

Rôzne typy modelov nám pomáhajú riešiť rôzne typy problémov. Aké predmety sú na obrázku? Kde sú predmety na obrázku?

Klasifikácia obrázkov: Snaží sa identifikovať najvýznamnejšiu triedu objektu. Triedu môžeme chápať ako označenie, napr. pri identifikácii topánok (bežecká obuv, ...)

**Detekcia objektov:** Používa sa, keď je dôležitá poloha objektu. Vracia množstvo súradníc nazývaných ohraničujúci rámček. Príklad takéhoto modelu môže byť model na detekciu osôb (AlwaysAI/mobilenet).

**Segmentácia obrazu:** Keď je dôležitý presný tvar objektu využijeme segmentáciu obrazu. Pre presnú segmentáciu klasifikuje každý pixel.

**Detekcia orientačných bodov:** Využíva sa na určenie dôležitých bodov, ktoré zachytávajú dôležité prvky objektu. Môžeme použiť model na odhad pózy (AlwaysAI/human-pose). [LM]

### 3.2 Typy údajov pre školenie počítačového videnia a ich tréningovanie

**Generovanie množiny údajov:** Na učenie modelu počítačového videnia potrebujeme kvalitné údaje. Za kvalitný údaj sa považuje ten, čo je podobný údajom z reálneho sveta, ktoré budú použité na tréningovanie modelu.

**Typy generovania množín údajov:** Podľa toho, či chceme nech model vykonáva jednu konkrétnu úlohu alebo chceme všeobecný model, potrebujeme aj rôzne typy údajov. Ak chceme konkrétny model budeme potrebovať na jeho tréning špecifické dáta a obrázky, ktoré sú čo najviac podobné danej špecifikácii. Použitie existujúcej anotovanej množiny údajov: Výrazne môžu skrátiť čas potrebný na tréningovanie modelu ale zároveň nemáme takú kontrolu nad kvalitou údajov. Populárne verejné súbory údajov sú napr. PASCAL Visual Object Classes (VOC), ImageNet, ... Použitie vlastných údajov: vlastný súbor údajov tvorený z voľne dostupných videí a obrázkov online. Na rozdiel od existujúcej anotovanej množiny ich treba anotovať pred použitím. Použitie digitálne generovaného súboru: Používa sa, ak nie sme sami schopní zhromaždiť dostatok údajov. Vedie k lepšiemu výkonu. [LM]

Zhromaždené a anotované údaje sa použijú ako vstup pre tréning modelu. Prechádzajú a porovnávajú sa údaje, až kým nedosiahneme dobrý výkon, na základe špecifikácií pre každý typ modelu.

**Transfer Learning:** Využíva poznatky zo všeobecného tréningu a aplikuje ich na iné špecifickejšie možnosti. **Testovacia aplikácia:** Po vytvorení modelu na ňom otestujeme aj nové údaje, a zistíme či model reaguje a pracuje ako očakávame. Nejde o automatizovaný test, ale jeho výhoda je, že umožňuje rýchlo identifikovať nedostatky alebo prípady, kedy model nefunguje spoľahlivo. Potom na model použijeme situácie prípadov, kedy nefungoval úplne správne a opravíme to. [LM]

Čo je na obrázku	Predikcia modelu že obázok je...		
	Čivava	Muffin	Spolu
Čivava	450	100	550
Muffin	150	550	700
Spolu	600	650	1250
Šanca v %	0,75	0,85	



## 4 Výhody a nevýhody počítačového videnia

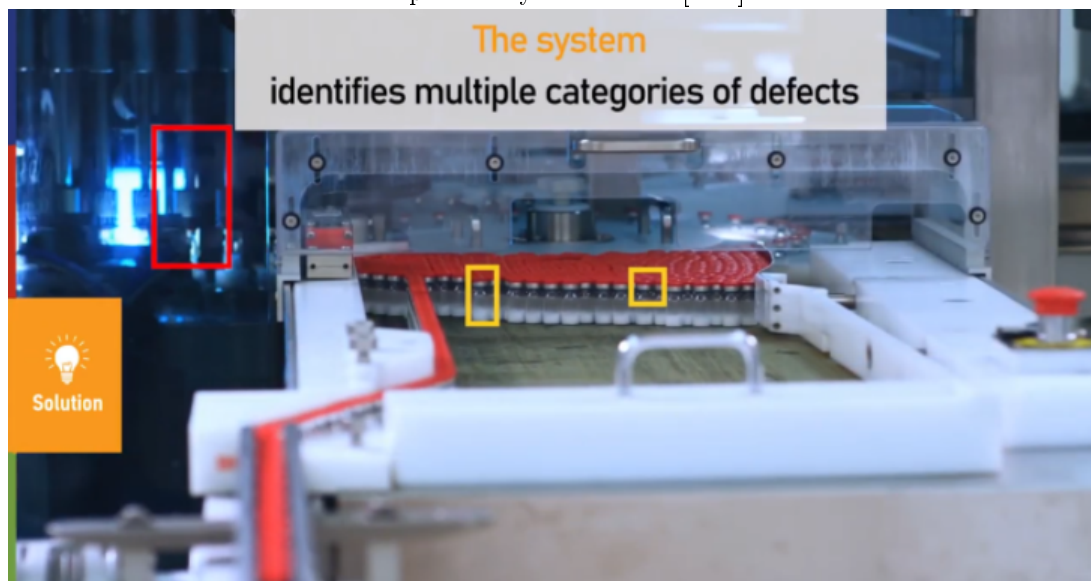
Ako každá iná technológia aj počítačové videnie má svoje výhody a nevýhody.

- Porovnanie s ľudským videním
  - ľudské videnie: zdroj obrazu → oko → kognitívny systém
  - počítačové videnie: zdroj obrazu → kamera → počítač
- Výhody
  1. Uľahčuje množstvo procesov
  2. Úplné frekvenčné spektrum pre získanie obrazov
  3. Jednoduchý a rýchly spôsob získavania údajov
  4. Generovanie presných a presných údajov
  5. Finančne efektívne
- Nevýhody
  1. Treba spracovávať veľmi veľké množstvo informácií
  2. Citlivé na osvetlenie a rôzne iné javy
  3. Reálny svet transformujeme 3D na 2D
  4. Je potrebné veľké množstvo pamäte
  5. Veľké množstvo nadbytočných údajov [Kad]

## 5 Využitie počítačového videnia

Počítačové videnie sa dnes využíva v každom odvetvy informačných technológií na svete. Jeho najväčšie využitie je v priemysle kde dokáže efektívne kontrolovať kvalitu produktov , alebo v zdravotníctve kde

zachraňuje množstvo životou kontrolou a identifikáciou problémov. Využitie má aj v rekreácií kde slúži ako zábava pre programátorov alebo nadšendov drovo , ktoré sú často využívané pri amatérkej práci s počítačovým videním . [Kad]



[?]

## 6 Záver

### Literatúra

- [Kad] Anoop Kadan. Role of computer vision in automatic inspection system. [https://www.researchgate.net/publication/281719495\\_Role\\_of\\_Computer\\_Vision\\_in\\_Automatic\\_Inspection\\_Systems/](https://www.researchgate.net/publication/281719495_Role_of_Computer_Vision_in_Automatic_Inspection_Systems/).
- [LM] CV developer Lila Mullany. Introduction to computer vision model training. <https://towardsdatascience.com/introduction-to-computer-vision-model-training-c8d22a9af22b/>.
- [Sol] Ján Erik Solem. Computer vision with python. [https://books.google.sk/books?hl=sk&lr=id=J9b\\_CH-NrycCoi=fndpg=PP2dq=computer+vision+with+python&ots=B-670WbKmxsig=VjXZ8QhGBTzgfmmKCNuZUvnZkkredir\\_esc=yv=onepageq=computer20vision20with20pythonf=false/](https://books.google.sk/books?hl=sk&lr=id=J9b_CH-NrycCoi=fndpg=PP2dq=computer+vision+with+python&ots=B-670WbKmxsig=VjXZ8QhGBTzgfmmKCNuZUvnZkkredir_esc=yv=onepageq=computer20vision20with20pythonf=false/).
- [TT] Data scientist Travis Tnag. Conditional probability. <https://towardsdatascience.com/learn-data-science-now-conditional-probability-acae6ec7b474/>.