Rozpoznávanie objektov vo videách

Výskumníci a vývojári v oblasti informačných technológií sa inšpirovali funkciou ľudského zraku a rozhodli sa vytvoriť aplikácie na detekciu videoobjektov, ktoré poskytujú strojom schopnosť analyzovať obrázky a zisťovať objekty, ktoré sa v nich nachádzajú. Najprv vyvinuli protokoly a postupy určené na fungovanie len na obrázkoch. Dnes sa však veci posunuli k video-obrazom. Cieľom takéhoto nástroja je umožniť stroju lokalizovať, identifikovať a klasifikovať objekty, ktoré možno vidieť na vstupných pohyblivých obrázkoch. Stroje a ich postupy neberú do úvahy obrazy ako celok. Na ich analýzu musia snímky rozdeliť a pracovať s pixelmi a ich vlastnosťami. Zvyčajne kombinujú detekciu obrazu a videostopy, aby prišli so svojimi výsledkami.

Cieľom záverečnej práce je predstaviť rôzne prístupy k rozpoznávaniu objektov vo videách a vyhodnotiť ich úspešnosť. V teoretickej časti je žiadúce sa zamerať na rôzne algoritmy rozpoznávania objektov vo videách. V praktickej časti je žiadúce vybrať vhodnú testovaciu sadu videí, na ktorých sa porovnajú a vyhodnotia rôzne prístupy k rozpoznávaniu objektov.

Charakter práce:

Výskumný – stanovenie predpokladov/hypotéz, metodika výskumu, výsledky výskumu (štatistická interpretácia), interpretácia výsledkov výskumu (vecná interpretácia).

Predmetové prerekvizity:

Úvod do strojového učenia (1., mgr);

Neurónové siete (1., mgr);

Hĺbková analýza dát (2., mgr).

Najdôležitejšie kompetentnosti získané spracovaním témy:

vykonávať vedecký výskum;

princípy umelej inteligencie;

poskytnúť vizuálnu prezentáciu údajov;

referovať o výsledkoch analýzy;

vykonať analýzu údajov.

# LINKS

<https://www.sciencedirect.com>

<https://link.springer.com>

<https://dl.acm.org>

<https://ieeexplore.ieee.org/Xplore/home.jsp>

­­­­­­­­­­­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

<https://labelyourdata.com/articles/ai-video-recognition>

Video Object Recognition with TensorFlow API

YOLO (You Only Look Once)

SSD Multibox (Single-Shot Detector)

ImageAI

TorchVision

<https://www.ridgerun.com/video-based-ai>

<https://www.ict.eu/en/projects/ai-video-recognition-assessing-video-footage-machine-learning-algorithm>

<https://www.sciencedirect.com/science/article/pii/S0031320302002303>

<https://www.sciencedirect.com/science/article/pii/S0925231213006085>

<https://www.sciencedirect.com/science/article/pii/S0926580511000781>

<https://link.springer.com/article/10.1007/s00464-017-5878-1>

<https://www.sciencedirect.com/science/article/pii/S0167865501001350>

<https://www.mdpi.com/2073-431X/2/2/88>

<https://www.sciencedirect.com/science/article/pii/S0952197612001030>

<https://link.springer.com/chapter/10.1007/978-3-319-16811-1_2>

<https://www.sciencedirect.com/science/article/pii/S0031320310005649>

static -> basic objects like flames -> comlex things like faces

<https://paperswithcode.com/task/video-classification>

<https://blog.coast.ai/five-video-classification-methods-implemented-in-keras-and-tensorflow-99cad29cc0b5>

Video classification - Mask R-CNN, YOLOv3, YOLOR, and YOLOv7

<https://viso.ai/computer-vision/video-analytics-ultimate-overview/#:~:text=Real-time%20Object%20Detection%20in,datasets%20such%20as%20MS%20COCO>.

<https://www.researchgate.net/publication/348363084_A_review_on_Video_Classification_with_Methods_Findings_Performance_Challenges_Limitations_and_Future_Work>

<https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn>

<https://www.apriorit.com/dev-blog/609-ai-long-short-term-memory-video-classification>

<https://learnopencv.com/introduction-to-video-classification-and-human-activity-recognition/#heading3>

# OBSAH

1. Úvod
2. Popis riešeného problému / Analýza súčasného stavu
   1. Úvod do AI?
      1. Úvod
      2. História – oheň v obrázkoch
      3. Ako AI spracováva obraz
   2. Kde sa využíva rozpoznávanie vo videách
   3. Budúcnosť rozpoznávania
   4. Machine learning
      1. Netflix, Facebook
   5. Neurónové siete
   6. Deep learning
   7. Modely na výber - **MoViNet, gluon-cv, GoogLeNet, VideoMAE, 3D CNNs, Two-Stream CNNs, Siamese networks**
   8. Knižnice
   9. Model 1 – **CNN**
      1. Ako pracuje + popis + zdôvodnenie výberu
   10. Model 2 – **RNN**
       1. Ako pracuje + popis + zdôvodnenie výberu
   11. Model 3 – **Long Short-Term Memory (LSTM)/3D CNN**
       1. Ako pracuje + popis + zdôvodnenie výberu
   12. Videá – objekty, aktivity, niečo ako programy v tv(správy, filmy...), gestá
       1. Ako prebehol výber
       2. Základné parametre
       3. Video 1
          1. Popis, prečo a na čo sa upriamiť
       4. Video 2
          1. Popis, prečo a na čo sa upriamiť
       5. Video 3
          1. Popis, prečo a na čo sa upriamiť
3. Ciele záverečnej práce
4. Metodika výskumu
   1. Modely použité na videách + nejaký kód
5. Výsledky
   1. Štatistické výsledky – porovnanie modelov na grafoch a použitie štatistických metód
6. Záver

## CNN - Convolutional Neural Network

<https://cs.nju.edu.cn/wujx/paper/CNN.pdf>

<https://onlinelibrary.wiley.com/doi/full/10.1002/advs.202204723>

Čo je to CNN?

Konvolučná neurónová sieť alebo CNN je sieťová architektúra, ktorá používa niektoré vlastnosti vstupných dát tak, aby redukovala množstvo prepojení potrebných na spojenie vrstiev neurónovej siete(1).

Čo ale znamená slovo „konvolučná“?

<https://www.sciencedirect.com/science/article/pii/S1047320316302267?casa_token=vG4FO0tSm5YAAAAA:kSEw-aSE2mvdJUhHePtQ_aL5k_1d8TpdO1xPYJws4hNpQkZQIe5ONj5cZk5r7kbMWfoWTPykd7w>

<https://theaisummer.com/cnn-architectures/>

<https://poloclub.github.io/cnn-explainer/>

Klady a zápory

<https://onlinelibrary.wiley.com/doi/full/10.1002/advs.202204723>

Obrázok, na ktorom je text

Automaticky generovaný popis

Ako funguje pri videách?

Ako funguje CNN?

1 Leonardo De Marchi, Laura Mitchell - Hands-On Neural Networks\_ Learn how to build and train your first neural network model using Python (2019, Packt Publishing)

## Deep Learning – hĺbkové učenie

Veľmi jednoducho popisuje hĺbkové učenie John D. Kelleher vo svojej knihe a to nasledovne: „Hĺbkové učenie je podobor umelej inteligencie, ktorý sa zameriava na vyváranie veľkých modelov neurónových sietí, ktoré sú schopné vytvárať presné rozhodnutia na základe dát. Najlepšie sa využíva v prípadoch, kde máme komplexné dáta alebo obrovské datasety.“ Taktiež spomína, že hĺbkové učenie používajú v dnešnej dobe obrovské firmy ako Facebook, Google alebo aj Microsoft. (<https://books.google.sk/books?hl=en&lr=&id=b06qDwAAQBAJ&oi=fnd&pg=PP9&dq=deep+learning&ots=_oHTQKl_XN&sig=_Bw5JqJ17zT1w4mD-Gd82zft4HM&redir_esc=y#v=onepage&q=deep%20learning&f=false>)

Čo sú to big data

Prečo je ale hĺbkové učenie tak populárne a prečo sa považuje za tak silný nástroj?

A big advantage with deep learning, and a key part in understanding why it’s becoming popular, is that it’s powered by massive amounts of data. The “Big Data Era” of technology will provide huge amounts of opportunities for new innovations in deep learning. As per Andrew Ng, the chief scientist of China’s major search engine Baidu and one of the leaders of the Google Brain Project, “The analogy to deep learning is that the rocket engine is the deep learning models and the fuel is the huge amounts of data we can feed to these algorithms.” <https://towardsdatascience.com/why-deep-learning-is-needed-over-traditional-machine-learning-1b6a99177063>

With the use of proper neural network architecture (number of layers, number of neurons, non-linear function, etc.) along with large enough data, a deep learning network can learn any mapping from one vector space to another vector space. That’s what makes deep learning such a powerful tool for any machine learning task. <https://venturebeat.com/datadecisionmakers/this-is-what-makes-deep-learning-so-powerful/>

História

A deep learning pioneer, Yann LeCun (creator of the convolutional neural network and ACM Turing award winner) <https://venturebeat.com/datadecisionmakers/this-is-what-makes-deep-learning-so-powerful/>

Negatíva/limity

Zdroje: <https://medium.com/@Ambarish_224/the-power-of-deep-learning-how-its-changing-the-world-13780515be31>

<https://www.ise.ncsu.edu/fuzzy-neural/wp-content/uploads/sites/9/2022/08/neuralnetworksanddeeplearning.pdf>

<https://books.google.sk/books?hl=en&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=deep+learning&ots=MNU-bopHRS&sig=q1bMvFrwP7c_3EhUmHruRIxBGCE&redir_esc=y#v=onepage&q&f=false>

## Úvod do AI

* 1. Úvod do AI?
     1. Úvod
     2. História
     3. Ako AI spracováva obraz

### História

História umelej inteligencie(UI) obsahuje mnoho fantázie, ukážok a prísľubov. Už Homér v ôsmom storočí pred Kristom spomínal mechanické trojnožky, ktoré sa vyznačovali inteligenciou. Až v poslednej dobe ale komunita okolo UI začala budovať stroje, ktoré by overovali hypotézy, aké dovtedy existovali len teoreticky. Avšak stále sa bojíme a očakávame budúcnosť, v ktorej nás čaká úplná umelá inteligencia. (A (Very) Brief History of Artificial Intelligence)

Medzi týmito udalosťami je celá história, ktorú by sme radi popísali v tejto kapitole.

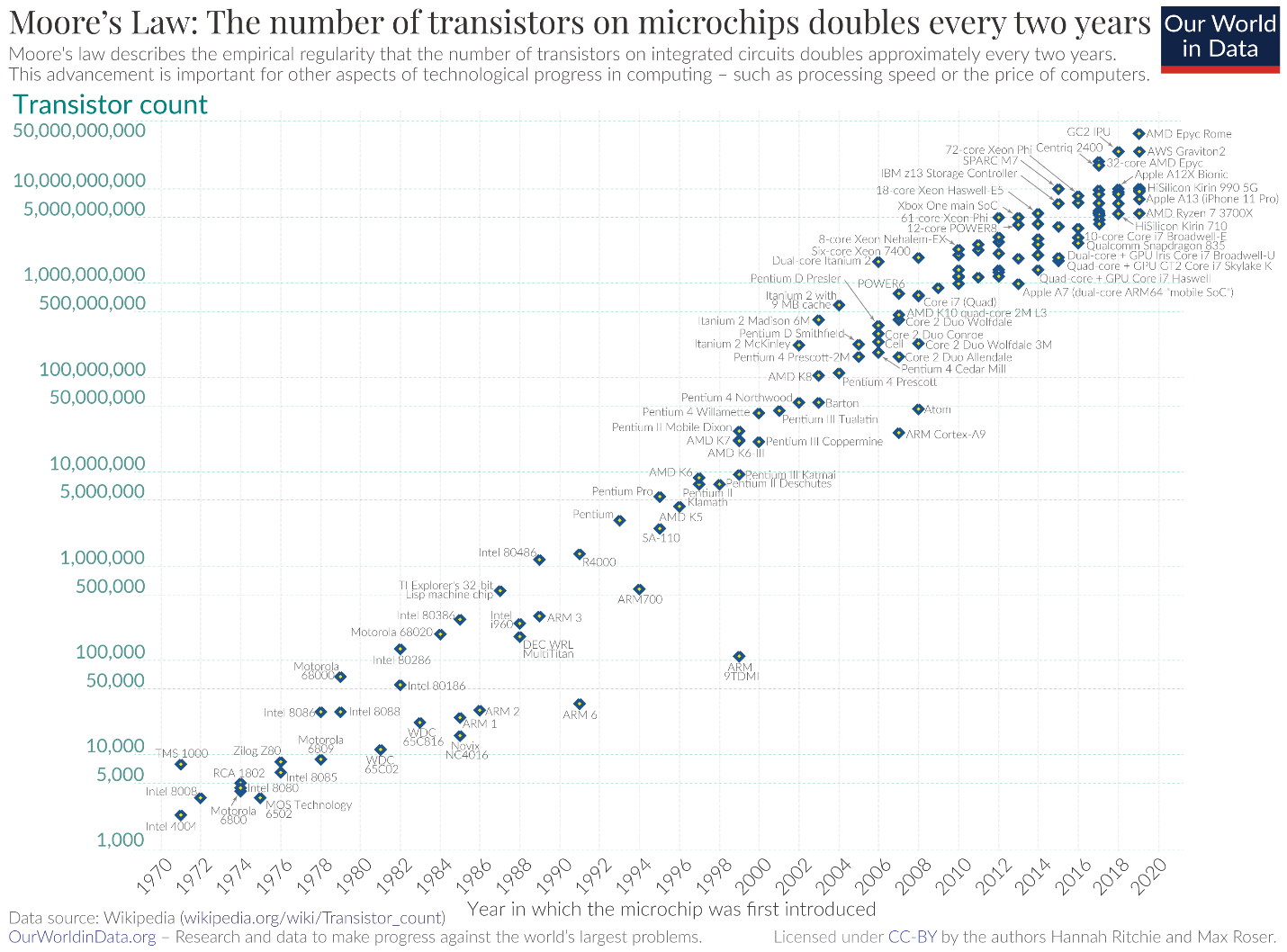
Alana Turing je veľmi známe meno. Vytvoril stroj, ktorý predstavil teoretický model počítača už v roku 1937. Bolo to v období druhej svetovej vojny a na opačnej strane, za nepriateľskou líniou, vytvoril Konrad Zuse prvý digitálny počítač. Zaoberal sa taktiež myšlienkou inteligencie počítača a či by vedel vyhrať v šachu. Vytvoril taktiež svoj programovací jazyk: Plankalkul, ktorý okrem riešenia matematických problémov mal byť použitý na vytvorenie umelej inteligencie. (Beranek et al. - HISTORY OF ARTIFICIAL INTELLIGENCE P. McCorduck ) Jeho predpoveď, že UI je jednu alebo dve generácie ďaleko, sa mu nevyplnila ako môžeme vidieť po približne 80. rokoch.

Mediztým sa vedecká sféra začala zaoberať do väčšej mieri kybernetickému chápaniu informácií a vyjadrovaním psychologických a biologických fenoménov matematickými formulami. Čo viedlo Johna McCarthyho aby zorganizoval s ďalšími vedcami v roku 1956 spoločné štúdium umelej inteligencie na Darthmoutskej univerzite. (Beranek et al. - HISTORY OF ARTIFICIAL INTELLIGENCE P. McCorduck )

John McCarthy, ktorý sa považuje za otca umelej inteligencie, vtedy prvý krát použil pojem „umelá inteligenica“ v oficiálnej práci a to priamo na spomínanej univerzite. (Johnmccarthy — Father of Artificial Intelligence)

Vývoj počítačov priamo ovplyvňoval pohľad na UI. Už známe programovateľné počítače ENIAC predstavili cestu, ktorá avšak neviedla priamo k UI a potrebovali sa prístupy a metódy vo vývoji. V 50. a 60. rokoch sa výskumy zameriavali na symbolickú reprezentáciu znalostí, ale tie neboli postačujúce a prešlo sa na pravdepodobnostné a štatistické modely.( Reconciling deep learning with symbolic artificial intelligence: representing objects and relations) Tu taktiež pribudli aj neurónové siete. (Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modeling)

V posledných desaťročiach vidíme rýchli pokrok UI, dôvodom je exponenciálny rast výpočtovej kapacity a dostupnosť dát. Vytvárajú sa nové modeli strojového učenia a hĺkového učenia, ktoré dosahujú doteraz nevídané výsledky. (Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions)



Odkaz na obrázok a zlepšiť obrázok Moorovho pravidla

### Ako AI spracováva obraz

#### Teória videí

Podobne ako fyzický film a analógové video sa digitálne video správa rovnako. Ide o stream obrazov, ktoré sú usporiadané v presnom časovom horizonte a to štandardne 24 snímkov za sekundu. Avšak stretneme sa aj s iným počtom od 12 až do 30. Záleží podľa spôsobu vytvorenia videa.

K obrazu môžeme mať aj zvuk. Ten sa ukladá ako samostatný stream a spája sa s obrazom v presne časovej synchronizácii.

V digitálnom videu existuje avšak špeciálny prvok a to pixel. Keďže sa tu uplatňuje rovnaká stratégia ako pri analógovej televízii: rozdeľ a panuj, tak sa obraz delí na mriežku. Najmenším prvkom mriežky je teda pixel, ktorý nesie informáciu o intenzite a farbe v číslach. (celý text doteraz: Digital Video 101: Understanding How Digital Video Works)

RELEVANT AT ALL?

Kniha: 978-3-319-61455-7.pdf str13

Obrázok, na ktorom je text, snímka obrazovky, písmo, dokument

Automaticky generovaný popis

Video môže mať rôznu kvalitu, ktorá záleží od spôsobu snímania obrazu –

<https://www.unite.ai/applying-ai-to-real-time-video-processing-the-basics-and-more/>

<https://www.cameralyze.co/blog/how-deep-learning-helps-with-real-time-video-processing>

(last one already used in Kde sa využíva rozpoznávanie vo videách)

<https://www.kdnuggets.com/2021/02/deep-learning-based-real-time-video-processing.html>

(if needed more content: https://www.newworldencyclopedia.org/entry/Digital\_video)

#### Prečo spracovávať videá ako datasety?

Úvod napojiť možno na predošlý odstavec

V konečnom dôsledku sú to len obrázky, nie?

Keďže vo video processingu ide o to aby sa automaticky generovali vety v prirodzenom jazyku, ktoré opisujú čo sa deje vo videu, tak sa dajú využiť vo veľa prípadoch, kedy obrázky nestačia.( Video Description: A Survey of Methods, Datasets, and Evaluation Metrics)

* Dáta, ktoré obsahujú veľké množstvo údajov a teda sú hodnotné a diverse
* Reprezentácia zo života, ktorá je výborná pre trénovanie lebo sa približuje realite
* Ide stále o sekvencie obrazov
* Môžeme využiť učenie na jednej úlohe a potom len presunúť údaje na druhú a maličko ich len pomeniť
* Videá vieme ľahko upraviť nástrojmi ako orezať ho, škálovať, pridávať noise, predchádame tak overfittingu a zlepšujeme generaizáciu
* Veľká accessibilita: s príbudajúcim množstvom videí na internete, surveillence cameras...
* Videá môžu obsahovať aj zvuk a text

Klady oproti obrázkom:

* Obsahujú dočastné informácie, ako sa menia snímky
* Dynamické scény: kedy môžu sledovať objekt, predikovať pohyb a rozpoznávať udalosti, napríklad pri surveillance, autonómnych vozidlách, robotoch
* Rozumieť kontextu: rozumieť komplexným scénam
* Long range dependencies: udalosti môžu trvať viacero snímkov alebo sekúnd a zachytávanie týchto prepojení je kľučové
* V niektorých prípadoch vedia videá predstavi´t viac informácií za rovnaký obsah dát ako pri obrazoch
* Práca s dátami: napríklad náhodný snímok a speed perturbations, and temporal jittering môžu zlepšiť generalizáciu
* Multi modal fusion: text, zvuk
* Spracovanie v reálnom ćase: niektpré aplikácie vyźadujú
* Predikcia udalostí

Rozpoznávanie vo videách v reálnom čase je kľúčové pre rýchle a spoľahlivé zaznamenávanie a spracovávanie dát. Keďže modely pre deep learning nevyžadujú ľudí aby ich učili, tak sú veľmi populárne vďaka ich samostatnosti. Ak sú dobre nastavené, tak sa ich popularita ešte viac prehlbuje. V nasledujúcich riadkoch popíšeme pár základných smerov.( How Deep Learning Helps With Real-Time Video Processing)

Prvý dôležitý smer je detekcia pohybu. Používa sa na určenie prítomnosti relevantného pohybu v pozorovanej sfére. Cieľom analýzy je:

* zistiť pohyb v rámci sekvencie videa
* sledovať tento pohyb objektu v čase
* zlúčiť objekty, ktoré sa hýbu spoločne
* identifikovať smer pohybu

(Video Processing Using Deep Learning Techniques: A Systematic Literature Review)

Ďalší smer je rozpoznávanie ľudských akcií(Human Action Recognition - HAR), ktorý má za cieľ automaticky skúmať a rozpoznávať povahu akcie z neznámych sekvencií vo videu. (Vision-based human action recognition: An overview and real world challenges)

Podobne ako rozpoznávanie ľudských akcií fungujú aj ďalšie smery, spomenieme napríklad rozpoznávanie a detekcia objektov, sledovanie objektov, analýza správania a mnohé iné.

? did I use this? https://link.springer.com/article/10.1007/s11042-020-08976-6

## Kde sa využíva rozpoznávanie vo videách

<https://www.sciencedirect.com/science/article/pii/S092523122030494X>

Obrázok, na ktorom je text, snímka obrazovky, písmo, dokument

Automaticky generovaný popis

* Surveillance and Security
  + A piece of luggage having been abandoned or a person moving in a restricted access area are typical examples of primitive events. <https://link.springer.com/article/10.1057/sj.2012.48>
  + Face identification: Hallucination and superresolution, <https://link.springer.com/article/10.1057/sj.2012.48> chceck for even more
* Autonómne vozidlá
  + a traffic monitoring
  + https://www.mdpi.com/1424-8220/22/4/1663
  + <https://www.sciencedirect.com/science/article/abs/pii/S0262885603000040>
  + https://www.sciencedirect.com/science/article/pii/S221501612100265X
* Human-Computer Interactions HCI
  + Rozpoznávanie gést a pohybov
    - <https://www.sciencedirect.com/science/article/abs/pii/S0167865515000616>
    - https://mecs-press.org/ijisa/ijisa-v4-n5/IJISA-V4-N5-8.pdf
* Video content understanding
* Športová analýza
* Medicína: medical scans, operácie, detekovanie abnormalít
  + https://link.springer.com/article/10.1007/s11554-010-0185-2
* Analýza zhromaždených ľudí
* Groceries
  + https://link.springer.com/article/10.1007/s12161-022-02384-2
* Industriálna automatizácia: kontrola kvailty, monitorovanie manufaktury, zabezpečenie bezpečnostných politík
  + <https://www.researchgate.net/profile/Joel-Jung/publication/268297183_No-Reference_Metric_for_a_Video_Quality_Control_Loop/links/553d14780cf245bdd76a814f/No-Reference-Metric-for-a-Video-Quality-Control-Loop.pdf>
* Healthcare: detekcia pádu, monitorovaie aktivít, asistovanie starším a postihnutým ľuďom
  + <https://www.sciencedirect.com/science/article/abs/pii/S0925231212003153>
  + https://link.springer.com/article/10.1007/s11760-013-0493-7
* Školstvo: zapojenie študentov a vytváranie osobných výučbových plánov
  + https://ieeexplore.ieee.org/abstract/document/9138372
* Monitorovanie prostredia: wildlife, natural distater monitoring.
  + <https://books.google.sk/books?hl=en&lr=&id=X7syQInA1aoC&oi=fnd&pg=PA15&dq=environment+video+processing&ots=4ODYPbolBv&sig=hzHnUkc6EbFoMqlVFuv1CF_fegw&redir_esc=y#v=onepage&q=environment%20video%20processing&f=false>
  + <https://ieeexplore.ieee.org/abstract/document/6787292>
  + https://books.google.sk/books?hl=en&lr=&id=Nlg9I-jGccIC&oi=fnd&pg=PR8&dq=video+processing+in+ecology&ots=f1jXBuq4LT&sig=87vM9zadAglD\_inSfLuYN9fyWNo&redir\_esc=y#v=onepage&q=video%20processing%20in%20ecology&f=false

Veľa smerov prepája vyššie popísané využívané smery v jeden celok. Napríklad autonómne autá potrebujú detekovať objekty, ich smer, ľudí, aktivity a predvídať ich pohyb a možno aj úmysly. Všetko sa to musí diať v reálnom čase a reakcia musí byť okamžitá. Čiže aj výpočty musia byť najrýchlejšie.

Najväčšie prekážky a problémy sú analýza, spracovanie a prenos dát, najmä potom limity výpočtovej techniky, uskladnenie dát a preťaženie prenosových limitov pre dáta. Všetky tieto procesy stoja veľa energie a preto už samotný prenos a ukladanie dát je potrebný kontrolovať. (Video Surveillance Processing Algorithms utilizing Artificial Intelligent (AI) for Unmanned Autonomous Vehicles (UAVs))

V podstate vieme rozdeliť tieto senzory do dvoch skupín: analýza statického prostredia(napríklad automatické hľadanie čiar na ceste) a detekcia statických alebo hýbajúcich sa prekážok(detekcia objektov) v relevantnom priestore k vozidlu. (A survey of video processing techniques for traffic applications)

Detekcia pre autonómne vozidlá obsahuje namä: detekciu objektov, sledovanie objektov, predikcia trajektórie a predpovedanie kolízie. Keďže na cestách sú kolízne objekty vysoko nepredvídateľné (chodci, cyklisti a aj vo v+seobecnosti objekty môžu kedykoľvek meniť smery) tak sú senzory a k nim používané algoritmy nesmierne dôležité. Najmä aby fungovali sto percentne keďže ide o ľudské životy. (Evaluation of 3D Vulnerable Objects’ Detection Using a Multi-Sensors System for Autonomous Vehicles)

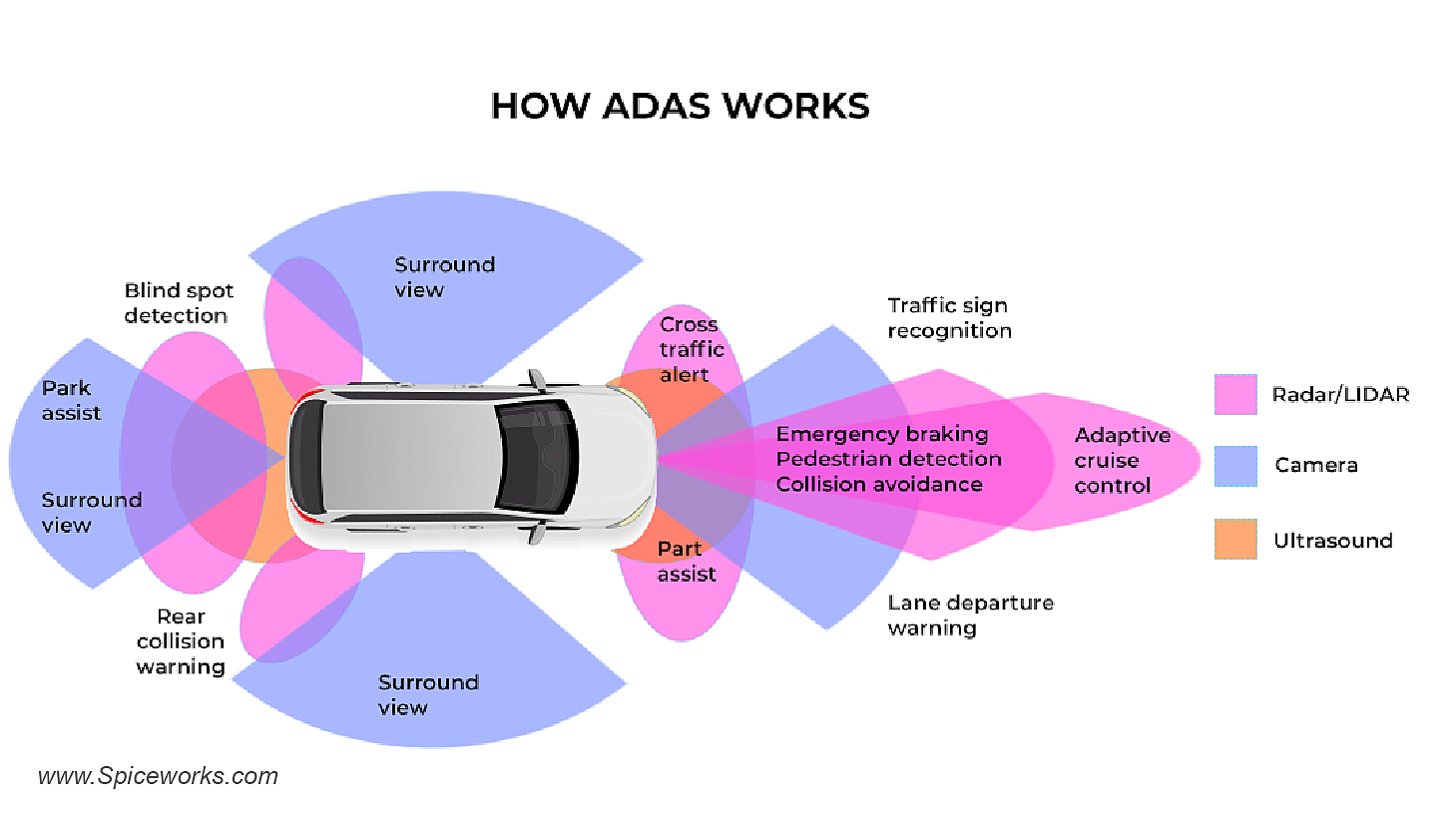


Figure 1https://www.mobileappdaily.com/ai-and-autonomous-vehicles

Ide o veľmi náročnú sieť algoritmov a UI, ktorá musí k tomu pracovať stopercentne, keďže ide o ľudské životy. Podobne sa využíva aj monitorovanie dopravy. Tieto systémy sú rýchle, ľahké na sprevádzkovanie, nízko nákladové a ľahko udržateľné a taktiež sú schopné pokryť širokú oblasť. (A survey of video processing techniques for traffic applications)

Ďalším zaujímavým odvetím, kde sa rozširuje využitie video proseccing je Human-Computer Interactions HCI. Ide o smer, kde sa vedci snažia priviesť robotov zo scifi do reality. Keďže rozpoznávanie gést a mimiky je u ľudí pomerne automatická vec a ľahko sa ju učíme pri dospievaní, tak o to ťažšie sa nám to popisuje pre počítačový program.

Medicína je ďalšie odvetie, kde sa video processing dá využívať, ale v dnešnej dobe sa umelá inteligencia skôr podiela na rozpoznávaní z obrazových materiálov. Ide hlavne o to, že vie rozpoznávať z prístrojov, ktorých výstup sú obázky. Napríklad röntgenové snímky, MRI, CT skeny a rôzne iné snímky napríklad z márnice(Machine learning and artificial intelligence in the service of medicine: Necessity or potentiality?). Avšak aj spracovanie videí sa rozširuje v posledných rokoch. Dá sa využiť vo veľa prípadoch, napríklad pri operáciach, kde je čas veľmi kľúčový.

Video processing sa využíva skôr vo výskumnej sfére medicíny. Našli sme rôzne výskumy, ktoré by mali v budúcnosti dopomôcť k zlepšeniu medicíny a starostlivosti o ľudí. Napríklad

TODO kniha: <https://books.google.sk/books?hl=en&lr=&id=bb0sCgAAQBAJ&oi=fnd&pg=PR11&dq=video+processing+medicine&ots=_R_0kIJXqk&sig=QWIBMawlp4hibBax-M9-ezkEBNs&redir_esc=y#v=onepage&q=video%20processing%20medicine&f=false>

TODO fall detection:

<https://link.springer.com/article/10.1007/s11760-013-0493-7>

<https://www.sciencedirect.com/science/article/abs/pii/S0167865515000616>

<https://mecs-press.org/ijisa/ijisa-v4-n5/IJISA-V4-N5-8.pdf>

<https://www.cameralyze.co/blog/how-deep-learning-helps-with-real-time-video-processing>

<https://ieeexplore.ieee.org/abstract/document/9563948>

## Firmy, ktoré využívajú video processing

* Google
* Facebook
* Amazon
* Microsoft
* IBM
* NVIDIA
* Intel
* Tesla
* Adobe
  + **Social Media and Content Platforms**: Companies like Google (YouTube) and Facebook use AI-based video processing to enhance user experience, recommend personalized content, and facilitate content recognition and tagging.
  + **E-commerce and Retail**: Amazon leverages AI video processing in its products and services, including Amazon Prime Video and Amazon Rekognition, for content analysis and video-based applications.
  + **Technology and Software Development**: Microsoft, IBM, NVIDIA, and Intel are involved in various aspects of AI research and video processing, providing hardware and software solutions for video analytics, security, and accelerating video processing tasks.
  + **Autonomous Vehicles and Transportation**: Tesla uses AI-powered video processing for its Autopilot feature, enabling advanced driver-assistance capabilities.
  + **Multimedia and Creativity**: Adobe incorporates AI and machine learning in its video editing and creative software tools to enhance video processing capabilities.
  + **Messaging and Augmented Reality**: Snap Inc. (parent company of Snapchat) utilizes AI-based video processing for interactive filters and augmented reality features in their messaging app.

## Negatíva video processing

Jedna z najväčších prekážok, s ktorými s apri video processingu stretneme je rovnaká ako pri práci s modelmi v celej umelej inteligencii a to pri falošne negatívnych výsledkoch modelu.

Napríklad ako popisuje článok: , využíva sa video processing pri detekcii zbraní na bezpečnostných kamerových záznamoch. Rozlišuje sa, či je na obraze zbraň, aká je(strelná alebo nie) a dokonca je užitočná informácia aj či ju drží človek. A tu už veľmi zaváži, či je jednotka rýchleho nasadenia poslaná zbytočne na miesto alebo nie. (Human pose estimation for mitigating false negatives in weapon detection in video-surveillance)

Ďalším problémom je aj problém v spracovávaní v reálnom čase. Zvyčajne techniky robia komprosim medzi presnosťou a efektivitou. Zvyčajne dochádza k tomu to kompromisu kvôôi vyššej náročnosti na výpočtovú techniku. Napríklad častý problem odhadu pohybu je matematicky náročný, veľakrát nemá presne definované riešenie a trpí nepresnými výsledkami. (Real-time motion estimation for image and video processing applications)

Problémov je viacero aj pri samotných datasetoch. Keďže Od analógovej televízie po digitálnu televíziu, od VHS videokaziet k DVD, od mobilných telefónov, ktoré sa používali len na volanie a posielanie textových správ, k mobilným telefónom, ktoré fungujú ako fotoaparáty, webové prehliadače, navigačné systémy, sociálne siete a takmer sa nepoužívajú na volanie, došlo k veľkej revolúcii v spôsobe, ako používatelia vytvárajú, zdieľajú a sledujú videá za posledné roky. (Video Coding Standards and Video Formats), čo na jednu stranu sťažuje prácu s datasetmi, ale na druhej ak natrénujeme model správne, tak máme možnosť využiť ho pri viacero prípadoch bez ohľadu na kvalitu videí.

Samozrejme obrovským problémom môže byť aj bezpečnosť dát. V tvorbe záznamov sa vytávarajú už dlhú dobu zabezpečenia, ktoré napríklad rozpixelovávajú obraz na určitých miestach(tvár), aby sa predišlo problémom s GDPR protokolmi, ale práve to má aj druhú stránku a takéto videá môžu mať tento efekt nenávratný a teda video často krát nevieme použiť pre výskumné ciele. Avšak sa naskytne otázka, či to nie je vlastne dobré pre naše bezpečie. (The future of video analytics for surveillance and its ethical implications). Ovšem toto sa týka len oficiálne vytváraných záznamov napríklad z bezpečnostných kamier, osobne vytvorené videá, ktoré sú na internete sa nekontrolujú do takej miery a vedia byť aj zneužité.

<https://ieeexplore.ieee.org/abstract/document/5947693>

* Large data size
  + storage capacity, computation resources => cost and time consuming, training time and model complexity
* overfitting – many redundancies

Negatívum celej AI: berou nám práci

<http://heb-nic.in/cass/admin/freePDF/oegeu7swbtkfcij72sr3.pdf>

## Budúcnosť video processingu

Spracovávanie videí pomocou umelej inteligencie má obrovský potenciál a už teraz vznikajú užitočné nápady a projekty ktoré nám pomôžu v bežnom živote. Okrem autonómnych vozidiel sú nápady na pozorovanie detí v malom veku u ktorých sa na základe ich pohybu môžu detekovať problémy a dokonca aj choroby.( Monitoring infants by automatic video processing: A unified approach to motion analysis)

Taktiež aj mnohé odvetia, ktoré by vedeli využiť dronové zábery môžu zužitkovať spracovávanie z videí, napríklad agriculture. Farmári by vedeli detekovať ľahšie choroby pri plodinách na rozsiahlych poliach, taktiež vytvárať mapy polí na obrábanie a taktiež samotné monitorovanie plodín. (AI in Agriculture). Podobne sa dajú dronové zábery využiť aj pri kontrolovaní statiky budov a mostov, stavu parkov a ciest, čo sú pravidelné kontroly vykonávané ľudmi.

Možností je priam nekonečno, záleží aké nápady budú zrealizovateľné a akým smerom sa bude smerovať video processing.

<https://www.sciencedirect.com/science/article/pii/S0010482516303031>

## Prepojenie s datasetom:

Dobrá kniha na image processing:

<https://books.google.sk/books?hl=en&lr=&id=UM_GCfJe88sC&oi=fnd&pg=PP1&dq=video+processing&ots=-iqDVnXWJ5&sig=RmnL4j7xBGJSvdQk1ccSsbBbEmg&redir_esc=y#v=onepage&q=video%20processing&f=false>

<https://openaccess.thecvf.com/content_cvpr_2018/papers/Huang_What_Makes_a_CVPR_2018_paper.pdf>

<https://dspace.mit.edu/bitstream/handle/1721.1/77452/826647817-MIT.pdf?sequence=2&isAllowed=y>

## Dataset

### <https://www.kaggle.com/datasets/mdaspace/canadarm3-collision-avoidance-day-1-part-1?resource=download>

<https://github.com/xiaobai1217/Awesome-Video-Datasets#Video-Classification>

<https://lcmou.github.io/ERA_Dataset/>

<https://www.kaggle.com/code/michaelapers/the-plasticc-astronomy-starter-kit/notebook>

<https://catalog.data.gov/dataset/?_res_format_limit=0&res_format=ZIP>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

<https://lcmou.github.io/ERA_Dataset/>

<https://www.eso.org/public/videos/archive/category/fulldome/list/3/>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## Modely

* Two-Stream CNNs
* Convolutional 3D – C3D
* Inflated 3D ConvNet – I3D
* Temporal Shift Module – TSM
* Spatio-Temporal 3D ConvNet – S3D
* SlowFast Networks
* R(2+1)D (R(2+1)D ConvNet)
* I3D with Non-Local Blocks- I3D-NL
* Attention in RNNs – ARTNet
* Efficient Convolutional Network – ECO

Obrázok, na ktorom je text, snímka obrazovky, písmo, číslo

Automaticky generovaný popis

## TODO

|  |  |  |
| --- | --- | --- |
| 1 | Začať s prácou na datasete |  |
| 1.1 | Pozrieť modely, ktoré použijem |  |
| 1.2 | pozrieť si presný postup na prácu s datasetom |  |
| 1.3 | dokumentovať všetko |  |
| 1.4 | použiť dataset https://lcmou.github.io/ERA\_Dataset/ |  |
| 1.5 | overiť modely, ktoré sú použité na datsete |  |
| 1.6 | pridať vlastné modely na datasete |  |
| 1.7 | interpretovať výsledky + štatistika |  |
| 1.8 | pozrieť ONNX a možné uplatnenie na práci |  |
| 2 | Písať analýzu |  |
| 2.1 | pokračovať v kde sa využíva rozpoznávanie v sekcii medicína |  |
| 3 | ODOVZDAŤ 10.11.: |  |
| 3.1 | Analýzu súčasného stavu, v rozsahu minimálne 10 strán |  |
| 3.2 | Ciele záverečnej práce |  |
| 3.3 | Metodiku/postup na dosiahnutie cieľa práce |  |
| 3.4 | Zoznam bibliografických odkazov |  |
| 4 | Poster |  |

- Ciele záverečnej práce, tzn. hlavný cieľ a z neho vyplývajúce chronologicky usporiadané úlohy (čiastkové ciele/podciele), ktoré vedú k naplneniu hlavného cieľa záverečnej práce

- Metodiku/postup na dosiahnutie cieľa práce, ktorá by mala korešpondovať s požiadavkami na obsah z hľadiska charakteru práce

- Zoznam bibliografických odkazov, minimálne 20 zdrojov, kde hlavne používajte digitálne knižnice akceptovaných vydavateľstiev ako je Elsevier (https://www.sciencedirect.com/) a Springer (https://link.springer.com/), ako aj spoločností ACM (https://dl.acm.org/) a IEEE (https://ieeexplore.ieee.org/Xplore/home.jsp).

<https://www.eso.org/public/videos/>

<https://gea.esac.esa.int/archive/>

<https://www.eso.org/public/videos/archive/category/fulldome/list/3/>

<http://simbad.cds.unistra.fr/simbad/>