

Sledovanie osoby vo viacerých videozáznamoch

1/22

Meno študenta: Andrej Zbín

Vedúci práce: RNDr. Zuzana Černeková, PhD.

Motivácia

- Bezpečnosť
- Autentifikácia
- Autorizácia
- Automatizácia
- ...

Ciele práce

- ▶ Návrh a implementácia systému
- ▶ Metódy detekcie osôb a ich porovnanie
- ▶ Techniky rozpoznávania osôb a ich porovnanie

Proces

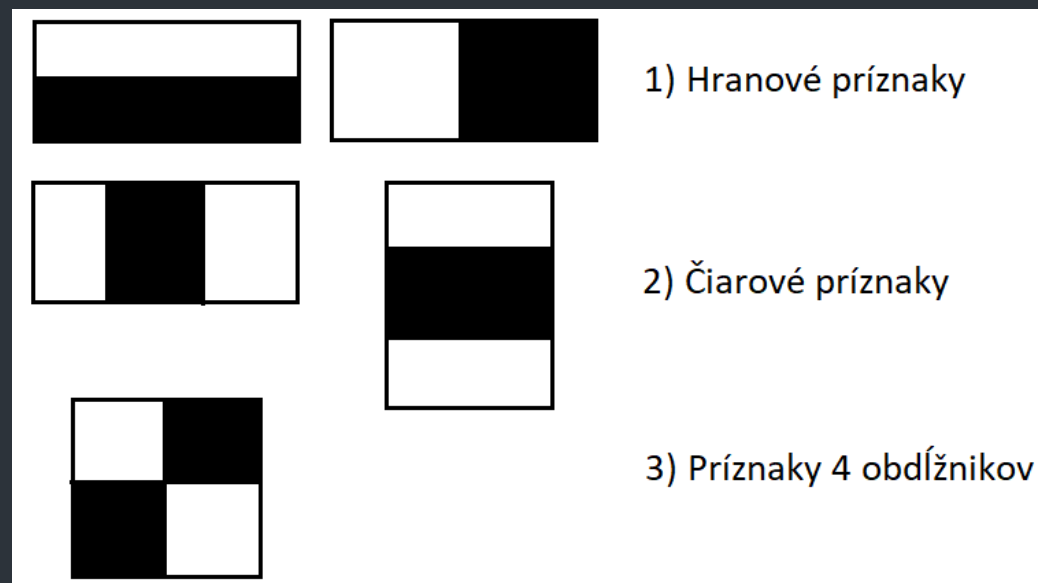
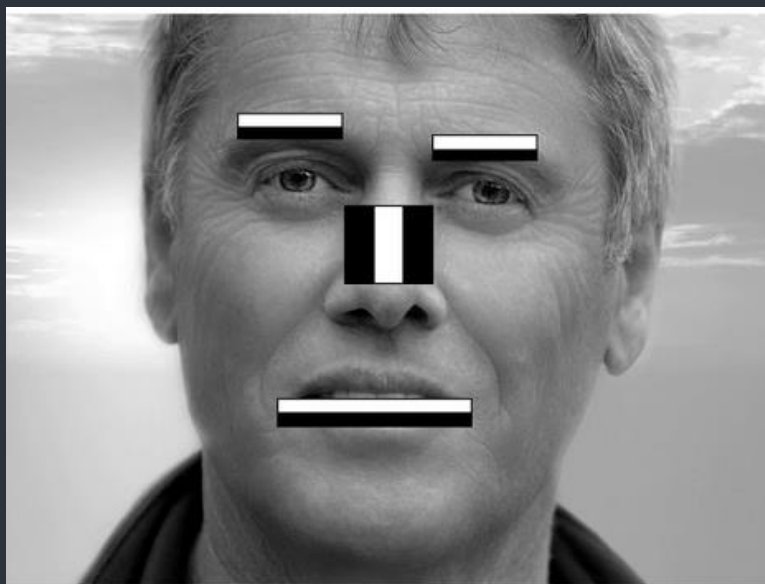
- Vstup
- Predspracovanie
- Normalizácia
- Detekcia
- Extrakcia príznakov
- Klasifikácia
- Výstup

Detekcia osôb

- ▶ Haarové príznaky
- ▶ Histogram orientovaných gradientov
- ▶ Konvolučné neurónové siete
- ▶ ...

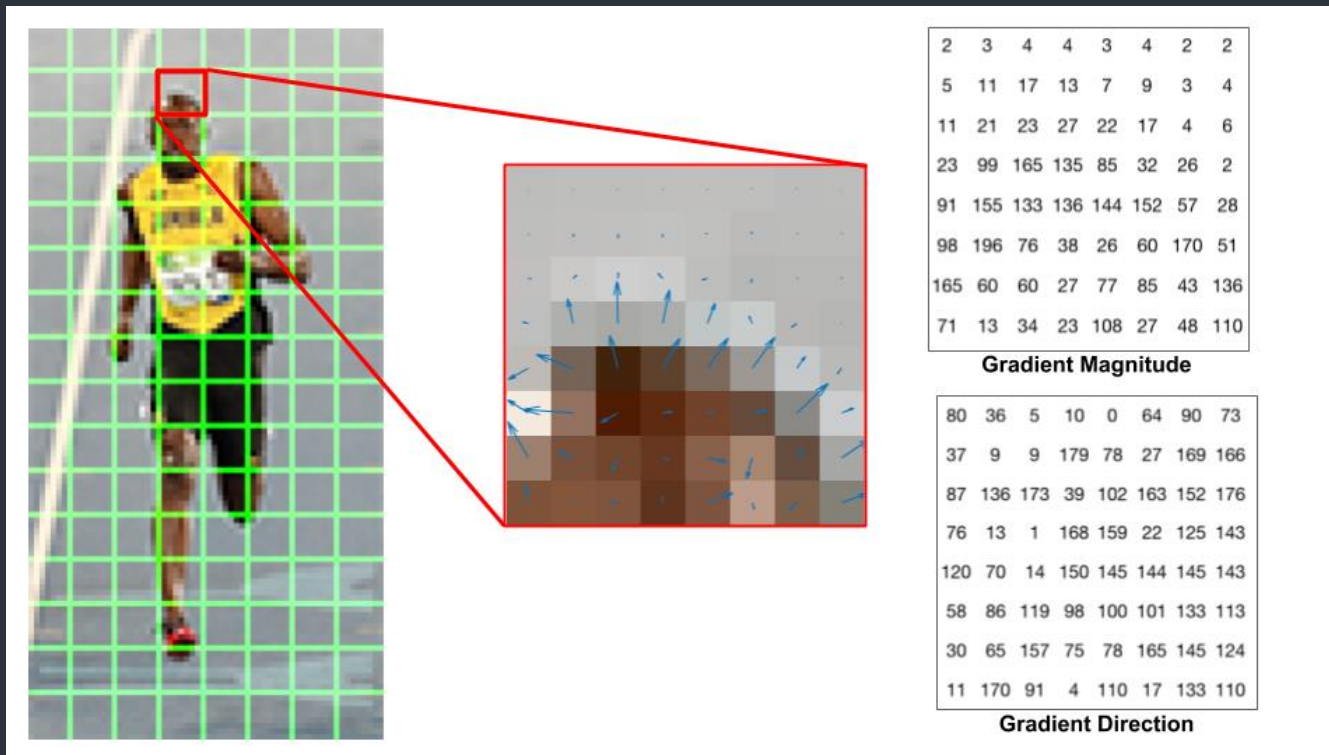
Haarové príznaky

- ▶ Viola-Jones
- ▶ Rýchlosť
- ▶ Kontrolované prostredie



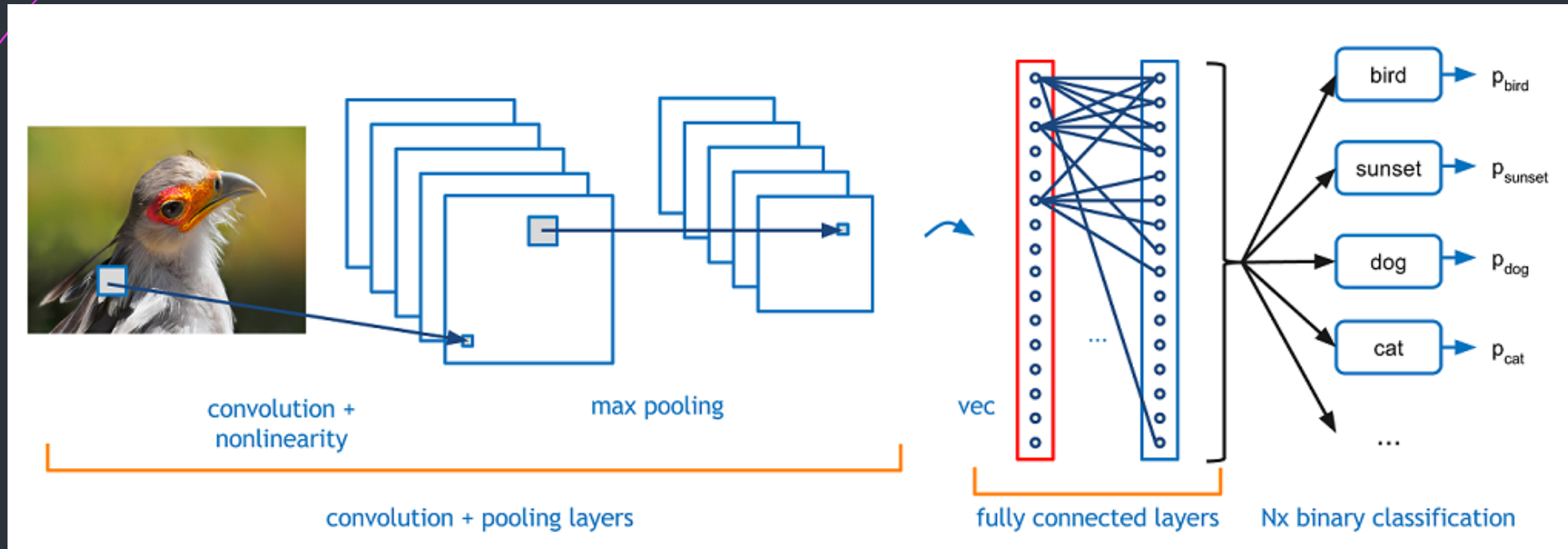
Histogram orientovaných gradientov

- Rozdelenie na regióny
- Výpočet histogramu



Konvolučné neurónové siete

- Klasifikácia vstupu
- Učenie s učiteľom
- Konvolučné vrstvy

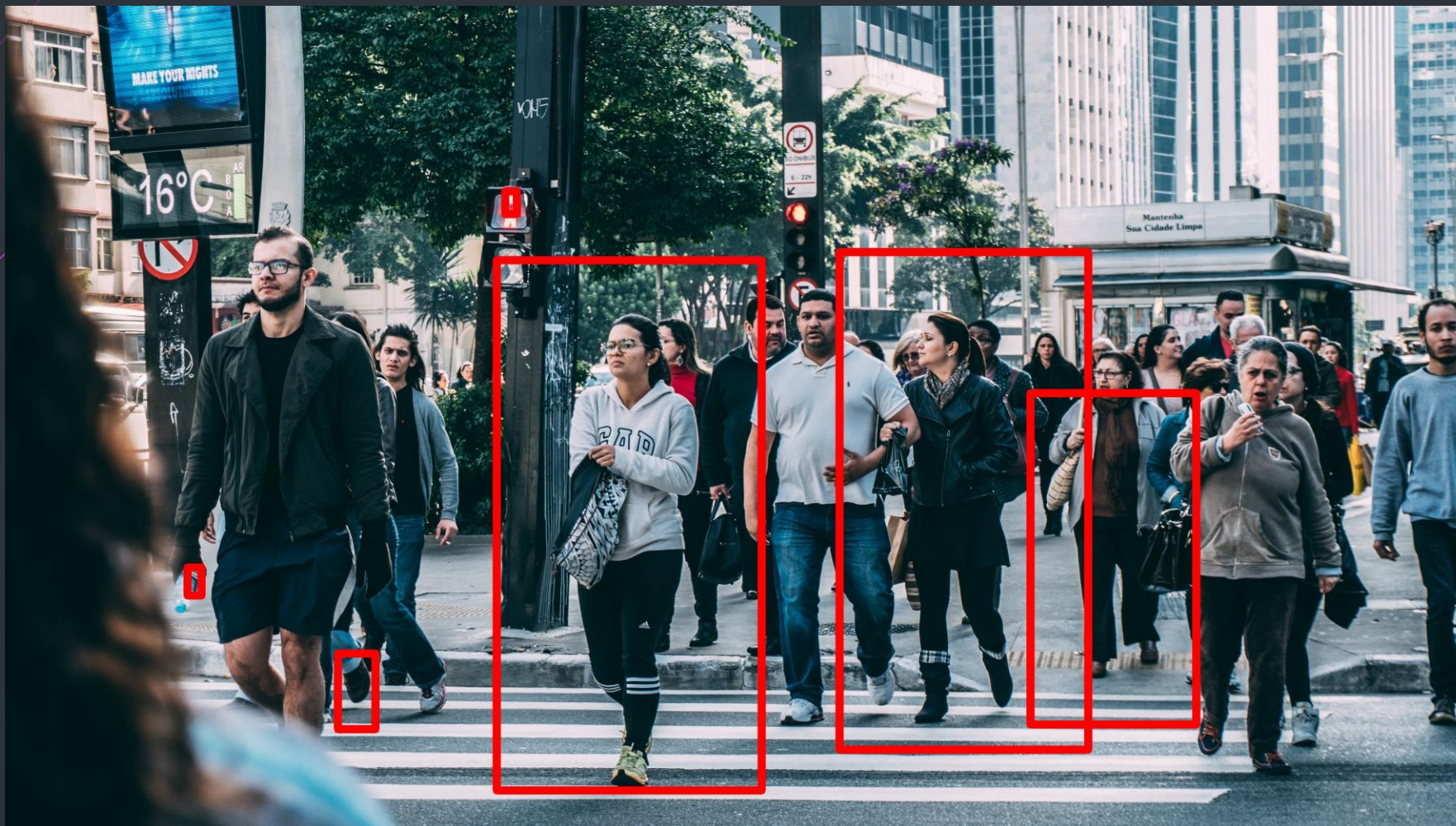


Porovnanie

- ▶ OpenCV
- ▶ MobileNet-SSD

Metóda	Čas potrebný na detekciu
Haarové príznaky	168ms
Histogram orientovaných gradientov	8 188ms
Neurónová sieť	4 901ms

Haarové příznaky



Histogram orientovaných gradientov



Konvolučná neurónová sieť



Extrakcia príznakov

- ▶ HSV histogram (Hue, Saturation, Value)
 - ▶ Pixel $(h, s, v) \rightarrow \text{histogram}[h][s][v] += 1$
 - ▶ Rozdelenie na regióny
 - ▶ Spojenie histogramov
- ▶ Škálovo invariantný lokálny ternárny vzor
 - ▶ Porovnanie pixelu so susedmi
 - ▶ Parametre p – počet susedov, r – polomer kružnice, c - threshold

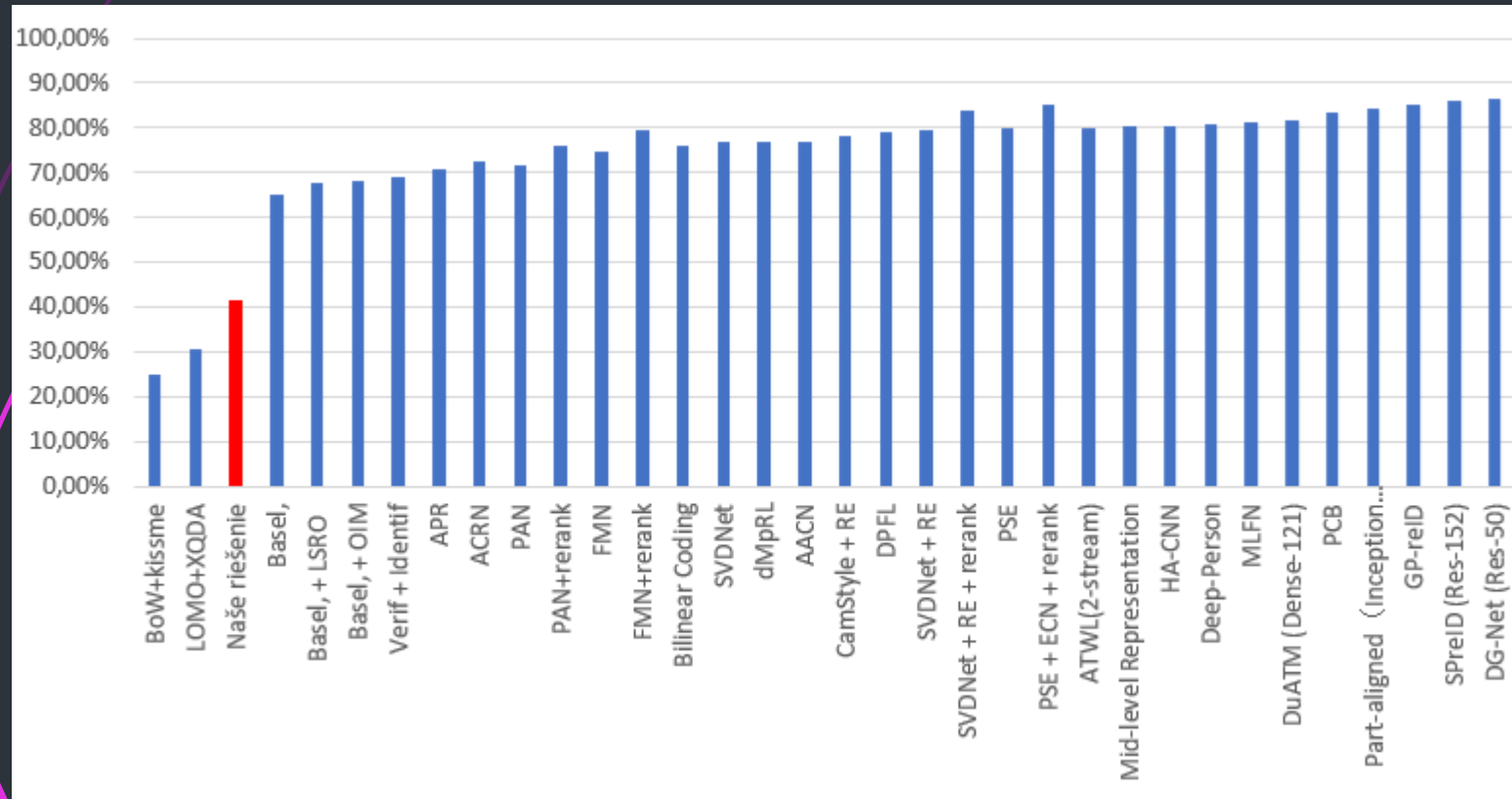
$$SILTP(I) = \sum_{n=0}^{P-1} s(V_I, V_{I_n}) * 3^n$$
$$s(c, p) = \begin{cases} 2 & \text{ak } (1 + T) * c < p \\ 1 & \text{ak } (1 - T) * c > p \\ 0 & \text{inak} \end{cases}$$

Klasifikácia

- ▶ Metóda k-najbližších susedov
 - ▶ Vzdialenosti medzi vektormi príznakov
- ▶ Metóda podporných vektorov (SVM)
 - ▶ Hľadanie nadroviny rozdelujúcej priestor vektorov
- ▶ Dataset - DukeMTMC-reID
- ▶ Redukcia dimenzie - PCA

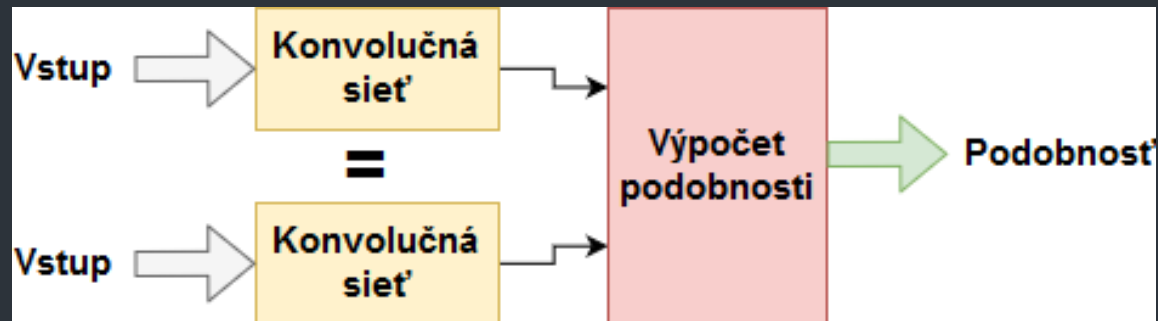
Porovnanie

- ▶ SVM dosahuje lepšie výsledky
- ▶ Porovnanie s ďalšími riešeniami:



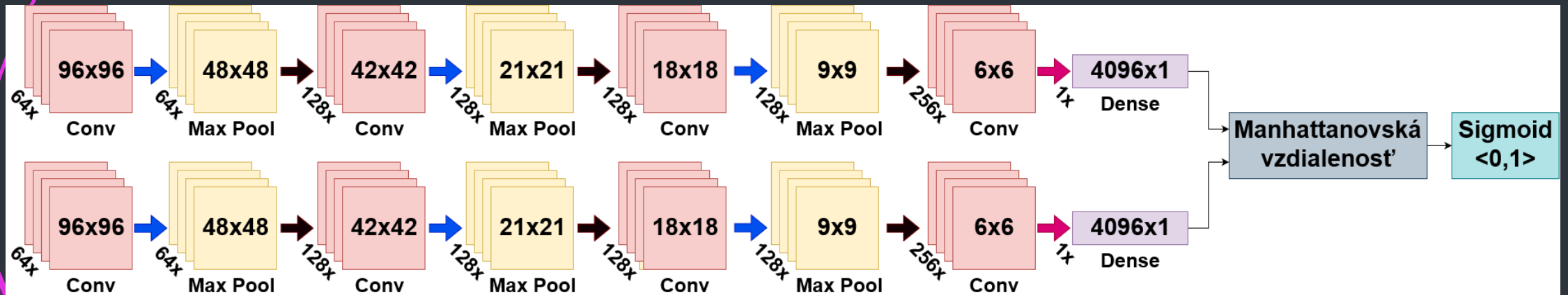
Siamská konvolučná neurónová sieť

- ▶ Hľadanie podobnosti 2 vstupov
- ▶ Výpočet vzdialenosti:
 - ▶ Manhattanovská vzdialenosť
 - ▶ Euklidovská vzdialenosť
- ▶ Klasifikácia
 - ▶ podobnosť so záberom z databázy



Architektúra siete

- ▶ 4 kovuľčné vrstvy
- ▶ ReLU vrstvy
- ▶ 3 Max pool vrstvy
- ▶ Manhattanovská vzdialenosť
- ▶ Sigmoid



Trénovanie

- ▶ 2 siete
 - ▶ Celá postava
 - ▶ Tvár
- ▶ Dataset tvárí - ChokePoint
- ▶ Chybová funkcia – Binárna entropia

$$L(p, q) = t(p, q) * \log(s(p, q)) + (1 - t(p, q)) * \log(1 - s(p, q))$$

$$t(p, q) = \begin{cases} 1 & \text{ak } p \text{ a } q \text{ patria do rovnakej triedy} \\ 0 & \text{inak} \end{cases}$$

$$s(p, q) = \textit{predikcia podobnosti}$$

Výsledky

Klasifikácia tvárí

		Predikcia	
		Zhoda	Nezhodna
Realita	Zhoda	5084	3601
	Nezhoda	420	8265

Presnosť	0,769
Senzitivita	0,585
Špecificita	0,952
Precíznosť	0,924
Negatívna predikčná hodnota	0,697

Klasifikácia postáv

		Predikcia	
		Zhoda	Nezhodna
Realita	Zhoda	1661	540
	Nezhoda	158	2043

Presnosť	0,841
Senzitivita	0,755
Špecificita	0,928
Precíznosť	0,913
Negatívna predikčná hodnota	0,791

Implementácia

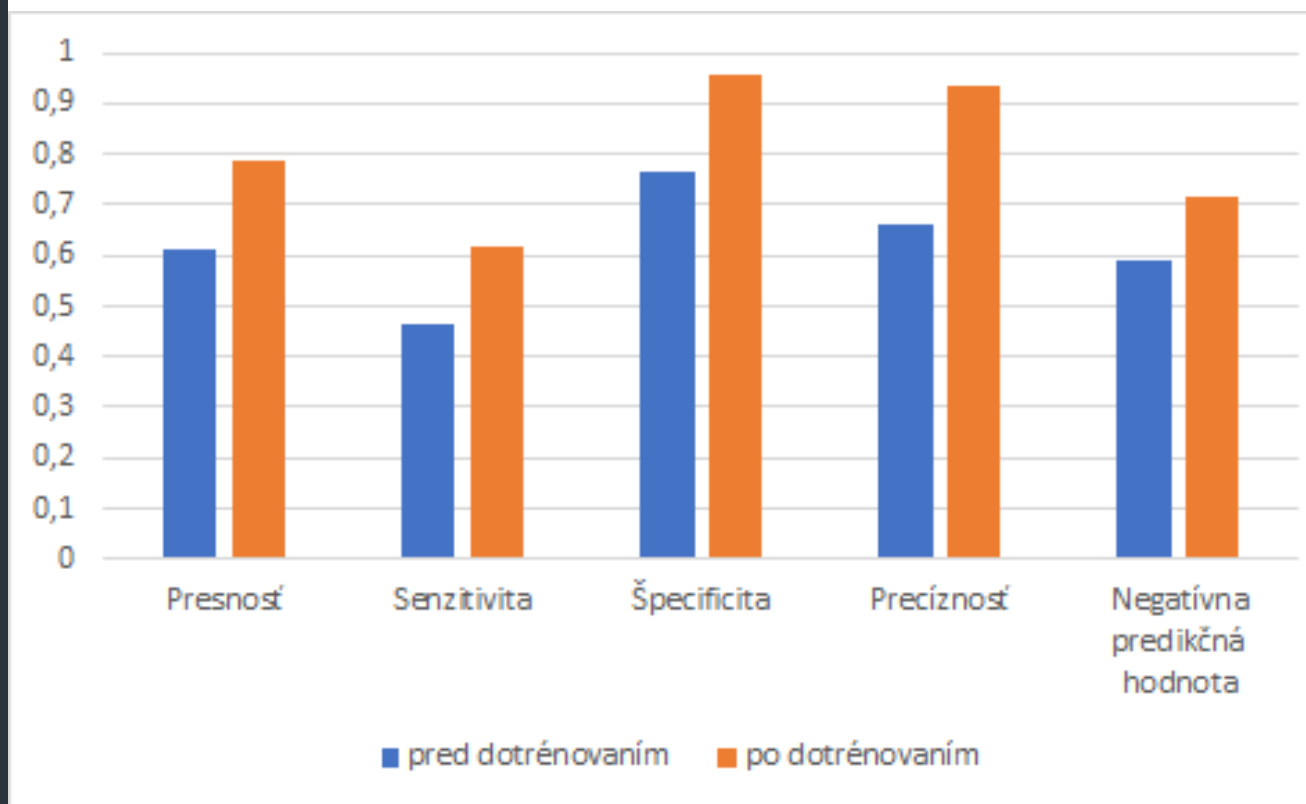
- ▶ Python3, OpenCV, MobileNet-SSD, Siamská konvolučná neurónová sieť
- ▶ Rozpoznanie uloženej osoby, sledovanie a rozpoznávanie neznámych osôb
- ▶ Dotrénovanie
- ▶ Dataset ChokePoint

Implementácia

- ▶ Uložené predchádzajúce snímky pre každú osobu
- ▶ Detekcia každých 30 snímok
- ▶ Klasifikácia: Porovnanie so všetkými
- ▶ Korelačný tracker (dlib) – každú snímku
- ▶ Centroid tracker – pri detekcii
 - ▶ Priradenie detegovaných osôb k osobám detegovaným v predchádzajúcej fáze
 - ▶ Kontrola: porovnanie snímky s predtým uloženými
- ▶ Dotrénovanie: potvrdzovanie zhody

Porovnanie výsledkov (podľa postavy)

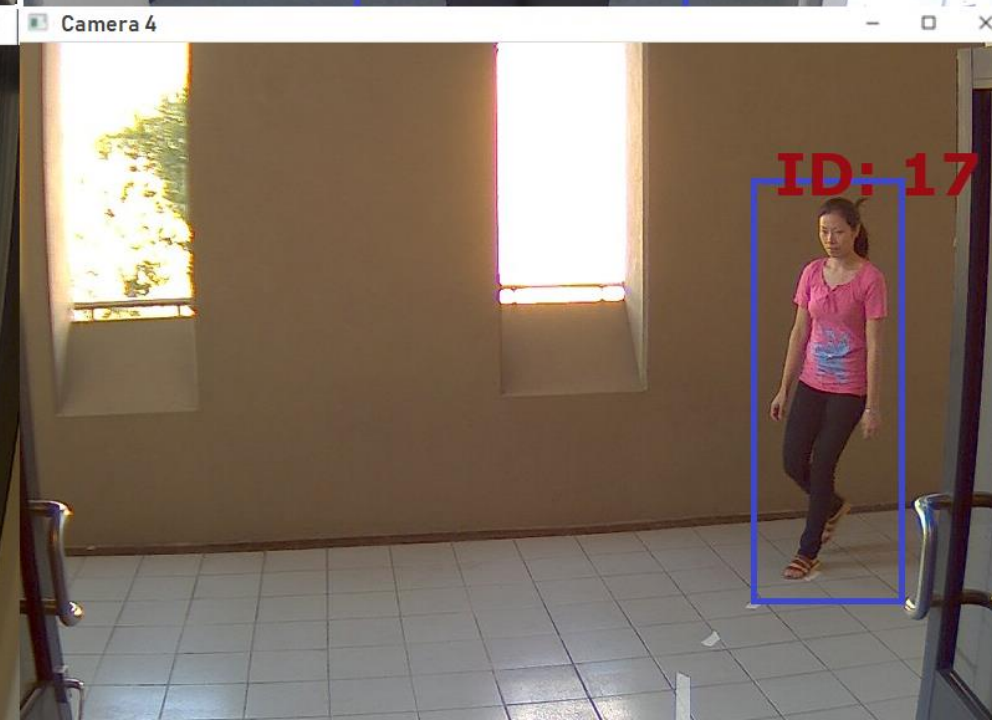
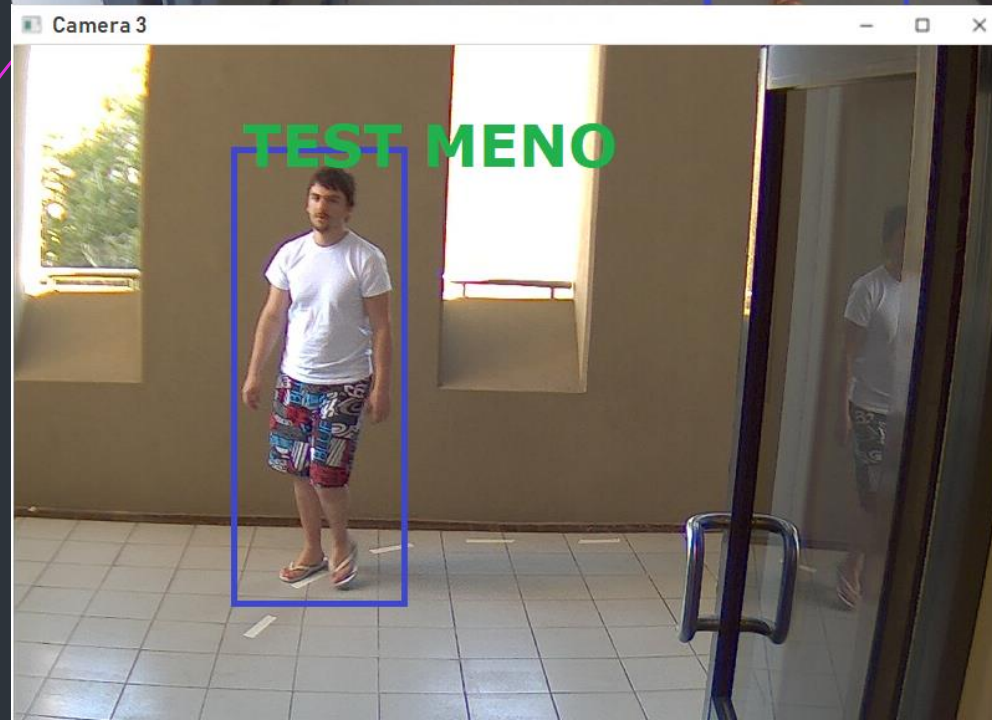
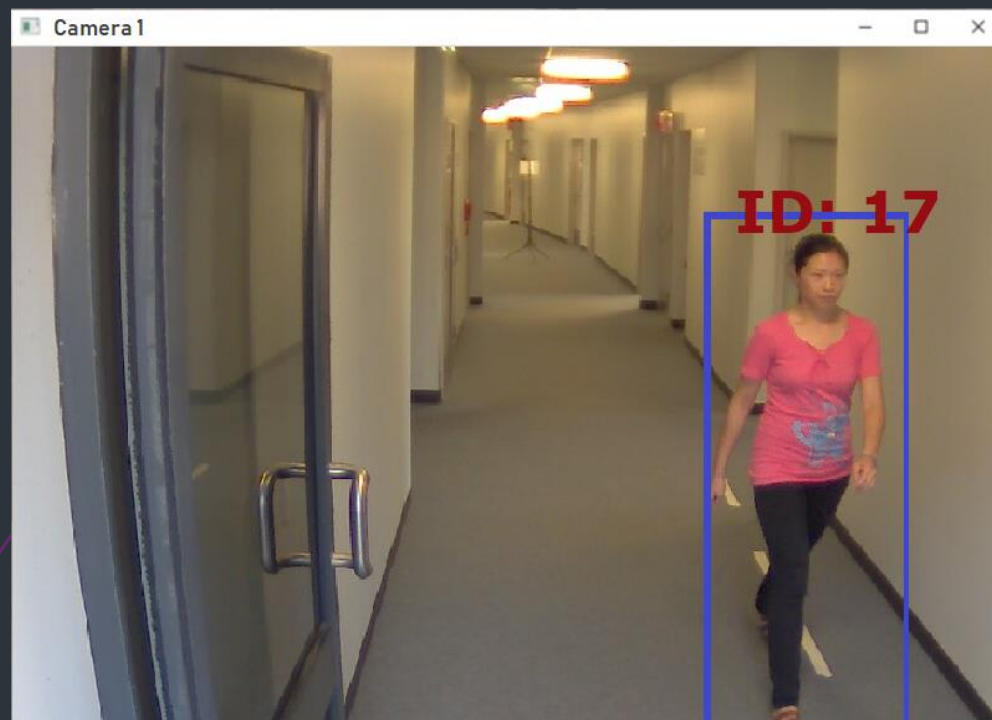
		Predikcia pred dotrénovaním		Predikcia po dotrénovaní	
		Zhoda	Nezhodna	Zhoda	Nezhodna
Realita	Zhoda	696	804	925	575
	Nezhoda	352	1148	66	1434



Výsledky

- ▶ V praxi nepresné vyhodnotenie veľmi nevedí
 - ▶ Kombinácia pomocou tváre a postavy
 - ▶ Viacero uložených snímok

21/22



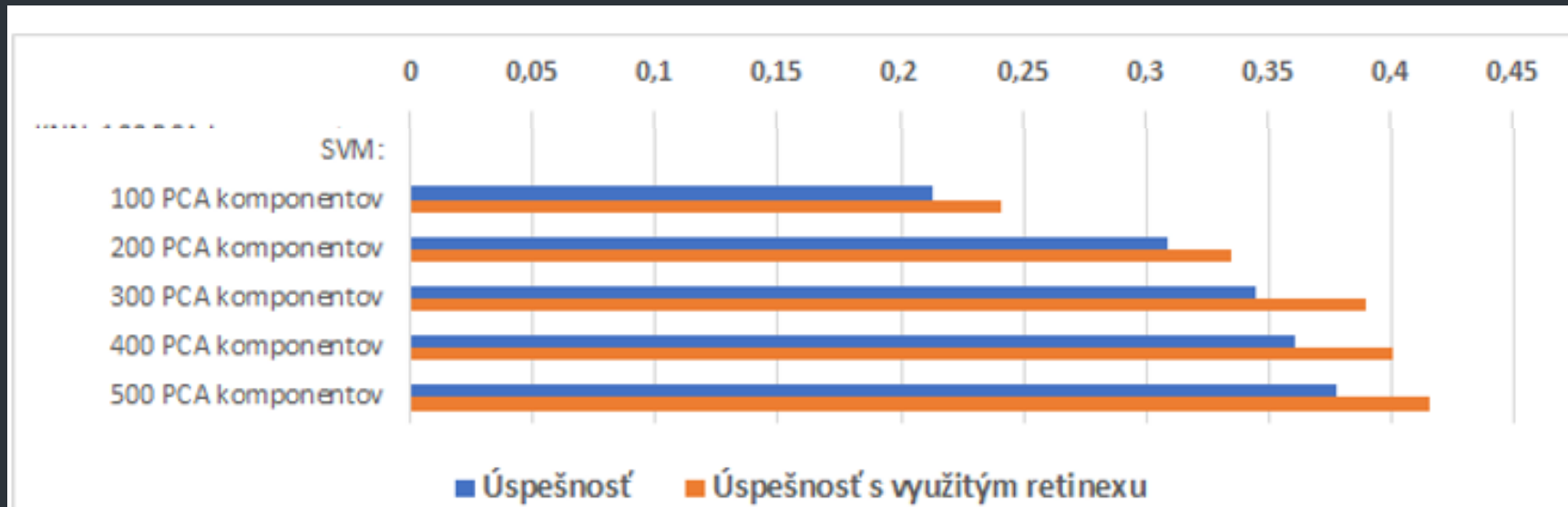
Možnosti ďalšej práce

- ▶ Spoľahlivosť
- ▶ Rýchlosť
- ▶ Sledovanie v dave



Ďakujem za pozornosť

- ▶ Na Obr. 2.3 kde autor píše, že porovnáva výsledky popísaných riešení s rôznymi parametrami nie je jasné čo je na x-ovej osi. Ktoré príznaky boli využité pri klasifikácii?
- ▶ Pomer správnej klasifikácie k celkovému počtu
 - 0 = žiadny testovaný príklad správne klasifikovaný
 - 1 = všetky testované príklady správne klasifikované
- ▶ Kombinácia 2x škálovo invariantného lokálneho ternárneho vzoru s rôznymi parametrami a HSV histogramu



- ▶ Vyššie uvedená formulácia cieľov má aspoň dve rôzne interpretácie: 1. na základe fotografie osoby určiť vo videu časové intervaly, kedy (prípadne aj kde) je daná osoba v obraze; 2. v každom momente určí, ktoré osoby z danej databázy sa nachádzajú v obraze. V prvom rade bolo potrebné upresniť, ktorý z týchto dvoch problémov sa rieši - a potom tiež to, či sa rieši offline alebo v reálnom čase. Prirodzená sa mi zdá prvá interpretácia, avšak z textu práce postupne vyplýva, že sa rieši ten druhý, všeobecnejší problém.
- ▶ Implementácia rieši všeobecnejší problém a teda rieši oboje interpretácie
- ▶ Riešenie funguje offline – načíta záznam z disku, aj online – záznam načíta z webkamery alebo internetového streamu
- ▶ Offline riešenie je spoľahlivejšie, nevynecháva snímky

- ▶ V časti 2.4 na str. 15 sa píše, že počet tried KNN klasifikátora musí byť vopred určený. Tento problém sa rieši použitím siamskej neurónovej siete. Pozná diplomant nejaké alternatívne riešenie? (V práci žiadne neuvádza.) Ak áno, aké výhody resp. nevýhody má voľba siamskej neurónovej siete v porovnaní s iným riešením či riešeniami?
- ▶ Počítanie vzdialenosti medzi extrahovanými príznakmi
 - ▶ Sieť pravdepodobne nájde lepšie príznaky, vylepšuje sa na základe vzdialenosti