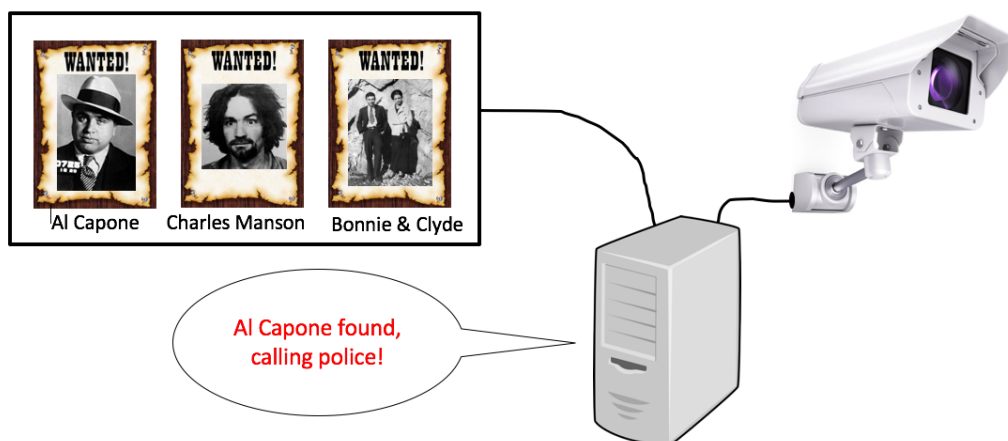


Detekce hledaných osob ve videu

David Bažout



Abstract

Cílem této práce je vytvořit nástroj umožňující vyhledávání podezřelých osob ve videozáznamu pocházejícího z dohledových kamer. Hledané osoby jsou systému určeny pomocí několika fotografií obličeje. Výstup tvoří informace o výskytu hledaných osob na konkrétních snímcích.

Úloha je řešena rozdělením problému na úlohu detekce tváře a její následnou identifikaci. Experimenty s existujícími přístupy na vhodných datových sadách poskytují relevantní porovnání úspěšnosti metod v různých situacích z reálného provozu. Výstupy testů poskytují informace o vhodných metodách a jejich optimálním nastavení pro tuto konkrétní úlohu.

Práce se zabývá i návrhem vhodné architektury, průzkumem existujících knihoven implementujících zkoumané metody a dalšími způsoby optimalizace výpočtu. Výsledkem je implementace modulu, který splňuje zadané parametry.

Keywords: detekce tváře — rozpoznávání tváře — dohledový systém

Supplementary Material: N/A

*xbazou00@stud.fit.vutbr.cz, Faculty of Information Technology, Brno University of Technology

1. Úvod

Někdy i vteřiny mohou při hledání nebezpečných osob znamenat rozdíl mezi životem o smrtí civilistů. Rychlé a efektivní vyhledávání v množství kamerových záznamů nelze provádět pouze manuálně. Moderní výpočetní technika a pokročilé technologie počítačového vidění nabízejí prostředky pro sledování budov, objektů nebo veřejných prostranství v reálném čase.

Cílem této práce je vytvoření modulu, který umožňuje efektivní vyhledávání osob podle obličeje ve videu

a splňuje následující požadavky. Hledané osoby jsou systému zadávány pomocí sady fotografií jejich obličeje. Fotografie mohou být různého stáří a v nepříliš dobré kvalitě. Video může pocházet z IP kamery v reálném čase, nebo je zde možnost načítání ze souboru. Záznam IP kamery může dosahovat rozlišení Full-HD. Modul by měl být schopen pracovat i s videem pořízeným v nepříznivých světelných podmínkách. Tváře ve videu mohou být vzdálené od objektivu kamery, je zde i možnost jejich značného odklonu od kolmé-

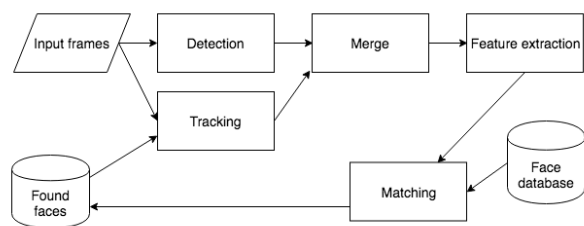


Figure 1. Schéma logických komponent modulu.

ho pohledu a podmínkou není ani dobré rozlišení. Systém je schopen určit, na kterých snímcích videa a ve které části se hledaná osoba nachází. Je nutné dosahovat únosné výpočetní náročnosti, aby mohla být data zpracovávána i na přenosném počítači v reálném čase.

Úloha lze rozdělit na problém detekce tváře a její následnou identifikaci. V současnosti je dostupné množství metod poskytující řešení pro oba typy úloh [1] [2]. Úspěšnost těchto metod může být ovlivňována kvalitou osvětlení na videozáznamech, jejich rozlišením, natočením tváře od objektivu kamery, vzdáleností tváře od objektivu kamery, objekty zakrývající části tváře a mnoha dalšími vlivy. Důležitou vlastností metod je i jejich výpočetní náročnost. Je nutné zvážit i možnosti úprav algoritmu zpracování dat, využití sledovacích algoritmů nebo jiných komponent pro zlepšení úspěšnosti prohledávání.

Cílem testování metod detekce a identifikace je objektivně porovnat jejich úspěšnost pomocí vhodné statistické metriky a na základě výsledků určit nejvhodnější metody pro tuto konkrétní úlohu. Je nutné zvážit i možnosti nastavení parametrů metod a jejich dopad na výpočetní náročnost. Součástí testování je i průzkum existujících datových sad, porovnání jejich dat, informací zanesených v anotacích a výběr vhodné datové sady k otestování zkoumané vlastnosti. Výsledky testů umožňují provést relevantní porovnání existujících přístupů z výše nastíněných hledisek a lze tedy vyhodnotit nejvhodnější metody a jejich optimální nastavení pro zpracování dat z dohledového systému.

Za hlavní přínos práce považují průzkum existujících přístupů a pečlivý přístup k jejich testování existujících z hledisek, která jsou pro realizaci dohledového systému klíčová.

2. Popis systému a přehled testovaných metod

Na obrázku 1 je schématicky zobrazena základní struktura systému pro detekci a identifikaci hledaných osob na videu. Zpracování videosekvence v dohledových systémech probíhá obvykle po jednotlivých snímcích.

Prvním krokem je provedení detekce tváří nacházejících se na snímku. Toho je docíleno pomocí logického celku, který je označován jako detektor. Úloha detekce má za úkol určit, kde se ve fotografii obličej nachází. Existuje množství detektorů [1] [3] [4] [2], které s různou úspěšností a výpočetními nároky dokáží vyřešit tuto úlohu. Funkce detektorů je negativně ovlivňována mnoha nepříznivými vlivy. Mezi tyto vlivy patří různé rozlišení fotografií, barevná hloubka fotografie nebo nízká kvalita osvětlení tváře. Dalším negativním faktorem jsou různá natočení obličejů, zakryté části, sluneční brýle nebo pokrývky hlavy. Je nutné zajistit, aby v nejmenší možné míře docházelo k situaci, kdy systém tvář nedetekuje, za tolerance občasného výskytu falešných detekcí [5].

Využití sledovacího algoritmu [6] není pro funkci systému bezpodmínečně nutné. Jeho úlohou je určit polohu objektu na základě navazujících snímků videozáznamu. Motivací využití sledování může být optimalizace výpočetní náročnosti detekce. Sledování rovněž napomáhá identifikovat a odstranit její chyby.

Cílem úlohy rozpoznávání je vytvoření strojového popisu (tzv. příznakového vektoru) detekované osoby a jeho porovnání s příznaky v databázi. Příznakový vektor je množina skalárních hodnot, které popisují určité vlastnosti zkoumaného objektu. Porovnání příznakových vektorů spočívá v aplikaci metriky vyjadřující jejich vzdálenost. Důležitým výstupem testování je vhodná prahová hodnota vzdálenosti pro shodu dvou příznakových vektorů. Stejně jako při úloze detekce je zde množství faktorů negativně ovlivňujících výsledky identifikace [6]. Je třeba zajistit invariantnost extrahovaných příznakových vektorů proti různému osvětlení, natočení tváře, špatné kvalitě snímků a geometrickým deformacím. Metoda by měla být schopna určit relevantní příznakový vektor i pro nízká rozlišení vstupních dat [2].

3. Metody detekce, testy a vyhodnocení

Mezi nejúspěšnější technologie pro detekci obličejů ve fotografii patří klasické metody Viola-Jones [1], HOG [3] a LBP [1] a metody založené na konvolučních neuronových sítích [4]. V současné době jsou dosahovány nejlepší výsledky v oblasti detekce obličejů aplikací neuronových sítí a existuje množství detektorů obličejů na nich založených.

Mezi nejznámější datové sady z oblasti detekce tváře patří datová sada FDDB (Face Detection Data Set and Benchmark) [7] nebo datová sada Wider face [8]. Tyto datové sady příliš neodpovídají záznamům z dohledového systému.

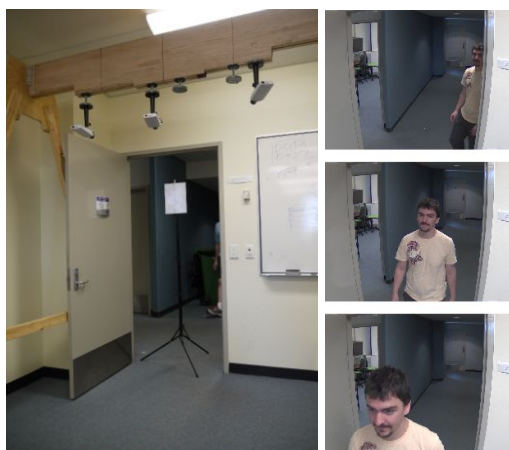


Figure 2. Ukázka dat z datové sady ChokePoint [9].

K porovnání metod pro detekci tváře jsem z tohoto důvodu použil modifikovanou datovou sadu ChokePoint [9]. Datová sada ChokePoint[9] byla vytvořena pro experimenty v oblasti reidentifikace osob za podmínek odpovídajících dohledovým systémům reálného světa. Obsahuje snímky s náročnými světelnými podmínkami, různým natočením tváří a ostroty obrazu. Je tvořena 48-mi sekvencemi videozáznamů, které byly pořizovány třemi kamerami umístěnými nad dveřmi v odlišných úhlech.

V datové sadě se vyskytuje celkem 25 různých osob (19 mužů a 6 žen) v první části a 29 různých osob (23 mužů a 6 žen) v části druhé. Snímky byly pořízeny v rozlišení 800x600px s počtem 30 snímků za vteřinu. Na prvních 100 snímcích každé sekvence je pouze prázdné pozadí. Celkový počet snímků sekvence je 100 000 a na 64 000 snímcích se vyskytuje obličej. Osoby přicházejí ke dveřím z různých úhlů a procházejí směrem dovnitř. Data se podobají záznamům z interního kamerového systému. Anotace nese pro každou osobu souřadnice levého a pravého oka a její identifikátor. Ukázka dat je na obrázku 2.

Tato datová sada je určena primárně pro úlohu rozpoznávání, ale lze ji s drobnými odchylkami aplikovat i na úlohu detekce tváře. Výsledkem metod detektorů tváře je obdélník ohraničující oblast detekce. S ohledem na anotace datové sady ChokePoint [9] lze testovat, že se souřadnice očí nacházejí uvnitř obdélníku detekce. Problémem je, že anotace jsou dostupné pouze pro osoby, které procházejí dveřmi. Na snímcích dochází ale i k výskytu jiných osob, což negativně ovlivňuje relevanci výsledků testování. Problém jsem vyřešil poloautomatickým určením inkriminovaných snímků a jejich vyřazení z testovacích dat. Při porovnání metod (obrázek 3) byla využita metrika ROC křivek [5] na základě různých prahových hodnot pro skóre detekcí.

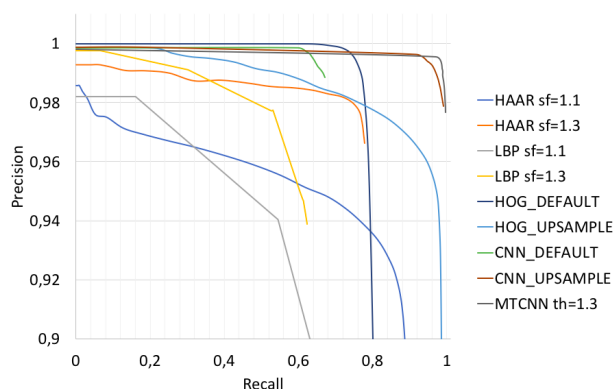


Figure 3. Precision - recall křivky pro metody detekce [1] [3] [4] [2].

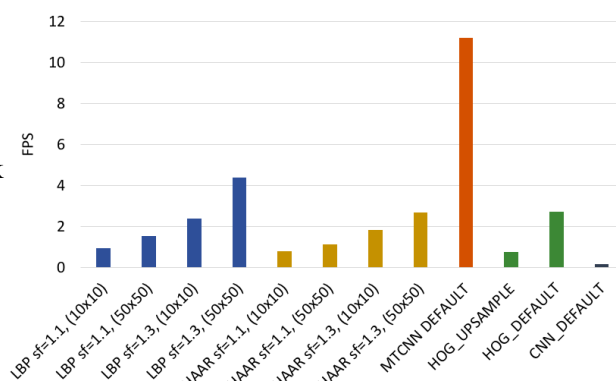


Figure 4. Výkonnostní srovnání metod detekce[1] [3] [4] [2].

Důležitou vlastností testovaných metod je i jejich výpočetní náročnost kvůli podmínce na zpracování dat v reálném čase. Rychlost zpracování snímků odpovídá průměrné hodnotě pro vstupní data o rozlišení Full-HD na 1 ze 4 virtualizovaných jader procesoru Intel i5@1.8GHz. Z výsledků na obrázku 4 je patrné, že je pro všechny testované metody klíčová minimální velikost detekovaných tváří. Metody Viola-Jones a LBP byly testovány pro minimální velikost 10x10px a 50x50px, výchozí nastavení MTCNN detektoru odpovídá 60x60px, výchozí nastavení metod HOG a CNN odpovídá nastavení velikosti 80x80px a varianta metody HOG_UPSAMPLE provádí detekce do velikosti 40x40px. Pro optimalizaci výpočetních nároků je nutné zvážit minimální velikost detekce potřebnou k extrakci relevantního příznakového vektoru.

V praxi se lze setkat se situací, kdy osoby ve videozáznamu nesledují objektiv kamery a jejich tvář s objektivem svírá určitý úhel. Natočení tváře výrazně ovlivňuje úspěšnosti metod detekce a z tohoto důvodu je důležité porovnat testované metody s ohledem na invarianci k tomuto jevu. Pro účely testování invariance metod detekce k natočení tváře jsem použil upravenou datovou sadu Head Pose Image Dataset[10], která je podrobněji popsána v následující kapitole. Úprava

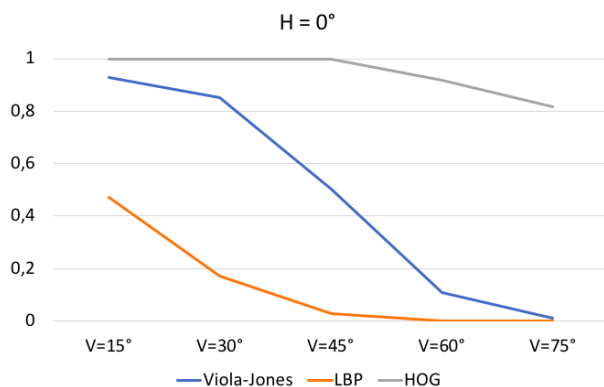


Figure 5. Testy invariance klasických detektorů[1] [3].

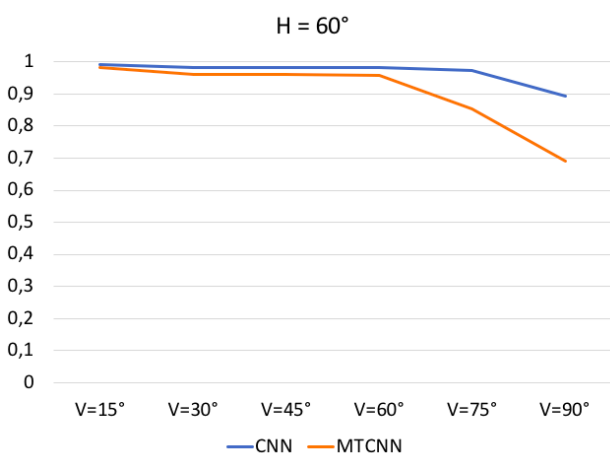


Figure 6. Testy invariance detektorů založených na neuronových sítích [4] [2].

datové sady spočívala v doplnění 200 fotografií bez výskytu obličejů, aby zde byl větší prostor pro falešné detekce. Doplnující fotografie byly získány z datové sady ChokePoint [9]. Na několika prvních vteřinách těchto záznamů se vyskytují pouze prázdné záběry vnitřních prostor.

Pro nově vzniklou datovou sadu byla vypočítána matice. Každé fotografii datové sady odpovídal řádek matice s hodnotami vertikálního a horizontálního úhlu natočení, hodnotami IoU (intersection over union) pro každou detekci a jejich skóre. Hodnota IoU byla určena porovnáním detekce s anotací a skóre je poskytováno metodou detekce. Metrikou pro porovnání metod byla křivka popisující plochu pod ROC křivkou vzhledem k odlišným úhlům natočení. ROC křivka, jejíž plocha je v grafech zanesena, byla sestavena[11] na základě odlišných prahových hodnot skóre a bylo využito pevné hodnoty $\text{IoU} \geq 0.5$ pro kladně hodnocenou detekci. Vzhledem k velkým rozdílům v úspěšnosti klasických metod a metod založených na neuronových sítích jsem výsledný graf rozdělil do obrázků 5 a 6.

Vysokého pokrytí s nízkým počtem falešných detekcí podle výsledků testování dosahuje MTCNN[4] detektor. Jeho implementace zahrnuje i velice rychlý korelační tracker. Vynikajících výsledků dosahuje i CNN detektor implementovaný v knihovně Dlib[2]. Mezi jeho negativa patří vysoká výpočetní náročnost a nutnost akcelerace na grafické kartě. Z tohoto důvodu je pro výslednou aplikaci vhodnější fakultní implementace detektoru MTCNN s přihlédnutím na dostačující invarianci oproti natočení tváře v porovnání s robustností příznaku identifikace.

4. Metody identifikace, testy a vyhodnocení

Proces extrakce příznakového vektoru z fotografie tváře a jeho porovnávání s databází je důležitou součástí celého systému. Cílem jeho testování je získat informace o chování metod a jejich výpočetní náročnosti za použití různých nastavení a rozlišení vstupních dat. Tyto informace poslouží nejen pro určení optimálního nastavení pro účely vyvíjeného systému, ale i pro získání informací o použitelnosti metody v praxi. Lze podle nich specifikovat i vhodný hardware pro realizaci výpočtu (CPU, grafická karta), ale i pro pořizování záznamů (minimální rozlišení IP kamery). Stejně jako u metod detekce se zde vyskytují klasické přístupy mezi které patří metody Eigenfaces[12] nebo Fisherfaces[6], ale nejlepších výsledků je dosažováno metodami[2] [4], které využívají konvoluční neuronové sítě.

Mezi nejznámější datové sady spadající do oblasti rozpoznávání tváře patří datová sada LFW (Labeled Faces in the Wild)[11]. Fotografie této datové sady jsou v příliš dobrém rozlišení a kvalitě oproti datům, pro které je aplikace určena. Vhodná datová sada pro účely testování by měla být podobná datům z reálného provozu. Měly by se zde vyskytovat náročné světelné podmínky a nízká kvalita záznamů nejlépe z dohledových kamer. Osoby na záznamech by měly být různých národností a jejich oblečení by mělo zahrnovat i prvky, které zakrývají části obličeje. Z dostupných datových sad předcházejícímu popisu velice dobře odpovídá datová sada ChokePoint [9], jejíž specifikace je uvedena v předchozí kapitole.

Jednou z nejlepších volně dostupných metod pro extrakci příznakového vektoru je model neuronové sítě typu ResNet-34 implementovaný v knihovně Dlib [2]. Konkrétně se jedná o modifikaci residuální neuronové sítě z dokumentu [13] s nižším počtem vrstev a počtem filtrů pro každou vrstvu snížených na polovinu. Tato síť byla natrénována na datové sadě[2], která byla vytvořena spojením datových sad The face scrub [14],

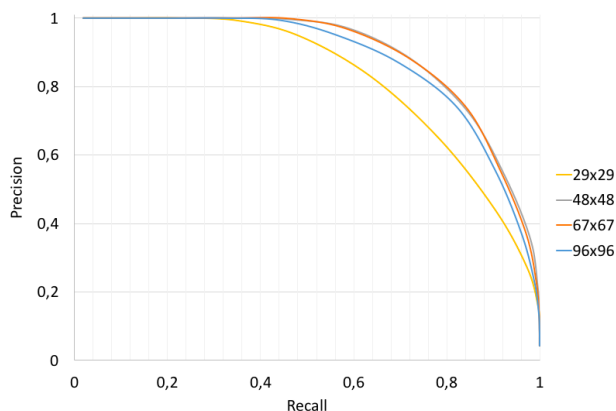


Figure 7. Výsledky testování metody identifikace z knihovny Dlib na datové sadě ChokePoint[9].

VGG dataset[15] a dalším množstvím fotografií tváří z internetu. Celkový počet fotografií tváře vzniklé datové sady dosahuje počtu 3 milionů. Výsledný model dosahuje podle dokumentace[2] úspěšnosti 99,38% na známé srovnávací datové sadě LFW (Labeled Faces in Wild) [11].

Vhodnou metrikou pro porovnání metod identifikace osob je ROC křivka. Lze ji sestavit podle různých prahových hodnot euklidovské vzdálenosti mezi příznaky. Podle ní lze rovněž zjistit optimální prahovou hodnotu euklidovské vzdálenosti pro shodu mezi příznaky[5]. Výsledky jsou zobrazeny v grafu 7.

Úhel natočení tváře ovlivňuje úspěšnost detekce i následné identifikace. Metody detekce i identifikace dosahují různých úrovní invariance k tomuto jevu. V případě, kdy je použita neinvariantní metoda detekce s invariantním identifikátorem, odolnost identifikátoru proti natočení tváře zůstává nevyužita. Stejná situace vzniká i v případě opačném. Metody detekce i identifikace byly z tohoto důvodu podrobeny testování jejich invariance oproti natočení tváře vzhledem ke kameře.

Srovnání lze provést na datových sadách vytvořených speciálně pro tento problém. Datových sad s anotacemi, které nesou informace o souřadnicích, identifikátoru a úhlu natočení tváře, je dostupných několik. Při testování byla využita datová sada Head Pose Image Dataset [10], která se skládá z 2790 fotografií 15 různých osob. Ke každé osobě existují 2 série fotografií po 93 snímcích. Varianty natočení pokrývají horizontální i vertikální úhly v rozpětí od -90° do 90° po krocích 15° pro vertikální úhel a 30° pro horizontální úhel (obrázek 10). Série fotografií stejné osoby se odlišují různým účesem, oblečením nebo brýlemi. Anotace nesou informace o souřadnicích tváře, identifikátoru osoby, číslu série, pořadovém číslu fotografie a horizontálním a vertikálním úhlu natočení. Výsledky testování invariance pro metodu identifikace z knihovny Dlib jsou na obrázcích 8 a 9.

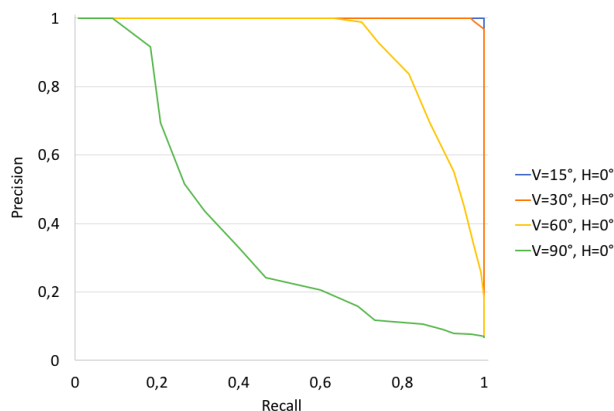


Figure 8. Výsledky testování invariance metody identifikace z knihovny Dlib[2] na datové sadě Head Pose Image Dataset[10] vzhledem k horizontálnímu natočení tváře.

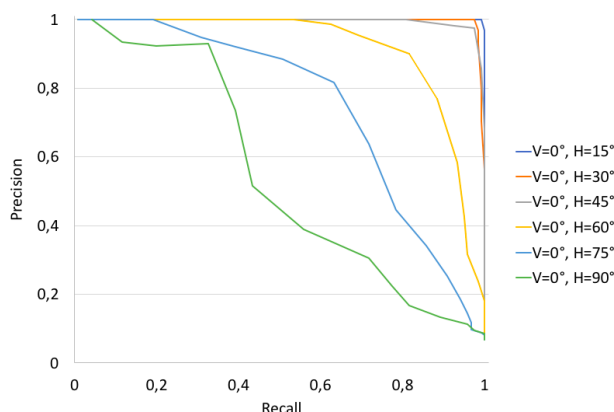


Figure 9. Výsledky testování invariance metody identifikace z knihovny Dlib[2] na datové sadě Head Pose Image Dataset[10] vzhledem k vertikálnímu natočení tváře.



Figure 10. Ukázka dat z datové sady Head Pose Image Dataset [10].

Z důvodu vysoké invariance oproti natočení tváře a dalším negativním vlivům je pro implementaci modulu vhodná volně dostupná implementace extraktoru příznakových vektorů z knihovny Dlib[2]. Pro lepší využití výpočetních zdrojů a rozložení zátěže výpočtu detekce a identifikace je vhodné přenést výpočet extraktoru na GPU.

5. Popis aplikace

Výsledný systém by měl mít podobu samostatného modulu implementovaného v jazyce C++. Je požadována síťová komunikace se vzdáleným serverem pro ukládání modulem získaných dat. Tuto komunikaci zajišťuje fakultní implementace komunikačního rozhraní propojující větší množství modulů.

Vstupem pro modul je IP adresa kamery zadávaná parametrem příkazové řádky. Modul by měl být schopen komunikovat po Wi-Fi nebo Ethernetovém připojení s běžně používanými IP kamerami s maximálním rozlišením 1080p. Druhou možností je zadání názvu souboru s videosekvencí. Modul by měl být schopen pracovat se všemi běžnými formáty video souborů.

Dalším důležitým parametrem je cesta ke složce obsahující fotografie hledaných osob. Fotografie mohou být ve všech běžně používaných formátech. Jejich rozlišení může být od několika desítek pixelů po několik megapixelů. Název fotografie bez přípony systému specifikuje identifikátor osoby. Na fotografii by se měla vždy vyskytovat jediná osoba. V případě výskytu většího množství osob může systém využít fotografii k identifikaci libovolné osoby ze snímku.

Výstupem modulu je množina hodnot specifikujících každou detekci na snímku videa. Množina hodnot specifikujících detekce nese informace o časové známce snímku, souřadnicemi specifikující ohraničení detekce, identifikátor osoby v případě, že se podařilo identifikovat shodu s databází, a příznakový vektor extrahovaný z detekce.

Modul by měl být schopen zpracování dat v reálném čase, což odpovídá alespoň hodnotě 5 FPS. Nižší hodnota FPS je opodstatněna nepříliš rychlými změnami ve videu. Modul by měl být připraven na využití v terénu a měl by si vystačit s hardwarovou konfigurací výkonnějšího notebooku. Lze uvažovat minimální procesor Intel-i5 a možnost využití dedikované grafické karty.

Implementace modulu v jazyce C++ využívá známé knihovny OpenCV [1] pro načítání a práci se snímky videosekvence, modelu neuronové sítě z knihovny Dlib[2] pro úlohu extrakce příznakového vektoru a rozpoznávání a fakultní implementace MTCNN detektoru. Kód je pro zvýšení přehlednosti rozdělen do několika logických modulů.

Modul byl spuštěn na záznamu o rozlišení FullHD z mobilního zařízení umístěného na mostu spojujícím budovy fakulty (obrázek 11). Na referenčním stroji MacBook Air 2017 s procesorem Intel-i5 se podařilo dosáhnout rychlosti zpracování 8 FPS bez akcelerace výpočtu na grafické kartě. Hledané osoby byly



Figure 11. Ukázka dat z mobilního zařízení umístěného na mostu spojující budovy fakulty.

spolehlivě identifikovány do vzdálenosti přibližně 50 metrů od kamery. Na podrobnějších výsledcích testování na vlastní datové sadě se pracuje.

6. Závěr

Cílem této práce byl návrh a implementace systému schopného efektivně vyhledávat osoby identifikované sadou fotografií ve videu. Podařilo se identifikovat základní komponenty, ze kterých se tento systém skládá. Průzkum existujících metod a otestování jejich klíčových vlastností na vhodných datových sadách umožnil jejich výběr a volbu jejich optimálního nastavení pro výslednou implementace modulu v jazyce C++ splňujícího předem stanovené požadavky.

Na datech z reálného provozu se podařilo s detektorem využívajícím konvolučních neuronových sítí v kombinaci s extraktorem příznakového vektoru z knihovny Dlib dosáhnout na videozáznamu o rozlišení FullHD spolehlivě detekovat hledané osoby do vzdálenosti přibližně 50 metrů. Zpracování snímků probíhalo rychlostí 8 FPS na procesoru Intel-i5 @ 1.8GHz bez akcelerace na grafické kartě.

Úloha detekce a identifikace je v současné době řešena s nejlepšími výsledky s využitím neuronových sítí. Vývojem složitějších architektur, výkonnější počítačů pro trénování modelů a vytvářením větších a náročnějších datových sad bude pravděpodobně dosahováno stále působivějších výsledků.

Poděkování

Úplným závěrem bych chtěl poděkovat panu Ing. Vítězslavu Beranovi Ph.D. za jeho podnětné a konstruktivní rady, které významnou měrou dopomohly k tvorbě této práce.

References

- [1] Gary R Bradski. *Learning OpenCV*. Computer Programming. Robotics. O'Reilly, Sebastopol, 2008.

- [2] Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.
- [3] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *In CVPR*, pages 886–893, 2005.
- [4] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, Oct 2016.
- [5] Mala Sundaram and Ambika Mani. Face recognition: Demystification of multifarious aspect in evaluation metrics, Jul 2016.
- [6] Milan Šonka. *Image processing, analysis, and machine vision*. Thomson, Toronto, 3rd ed. edition, 2008.
- [7] Vidit Jain and Erik Learned-Miller. Fddb: A benchmark for face detection in unconstrained settings. Technical Report UM-CS-2010-009, University of Massachusetts, Amherst, 2010.
- [8] Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Wider face: A face detection benchmark. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [9] Yongkang Wong, Shaokang Chen, Sandra Mau, Conrad Sanderson, and Brian C. Lovell. Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition. In *IEEE Biometrics Workshop, Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 81–88. IEEE, June 2011.
- [10] Nicolas Gourier, Daniela Hall, and James L. Crowley. Estimating face orientation from robust detection of salient facial structures. In *FG NET WORKSHOP ON VISUAL OBSERVATION OF DEICTIC GESTURES*, 2004.
- [11] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.
- [12] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [14] Hong-Wei Ng and Stefan Winkler. A data-driven approach to cleaning large face datasets. In *2014 IEEE International Conference on Image Processing (ICIP)*, pages 343–347. IEEE, 2014.
- [15] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, 2015.