



Projekt č. VI20172020068

Nástroje a metody zpracování videa a obrazu pro zvýšení efektivity operací bezpečnostních a záchranných složek (VRASSEO)

Technická zpráva 2018

Detektor hledané osoby ve video-streamu podle  
obrazu obličeje

David Bažout, Vítězslav Beran

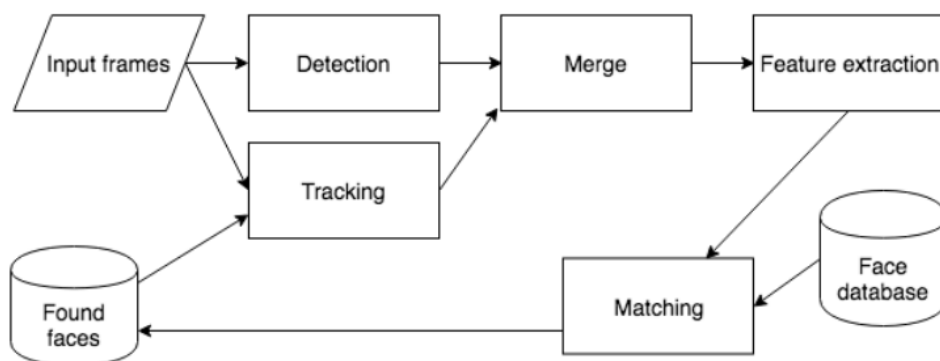
Fakulta informačních technologií  
Vysokého učení technického v Brně  
Božetěchova 1/2  
612 66 Brno, Česko

Prosinec 2020

## Technická dokumentace

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 kolmého 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.

Architektura typického systému pro reidentifikaci hledaných osob je ilustrováno obrázkem 1. Celý problém lze rozdělit na úlohu detekce tváře a její následnou identifikaci. V současnosti je dostupné množství metod poskytující řešení pro oba typy úloh. Úspěšnost těchto metod může být ovlivňována kvalitou osvětlení na videozáznamech, jejich rozlišením, natočení 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.



Obr. 1. Architektura typického systému pro reidentifikaci hledaných osob.

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 považuji průzkum existujících přístupů a pečlivý přístup k jejich testování existujících z hledisek, která jsou pro realizaci dohledového systému klíčová.

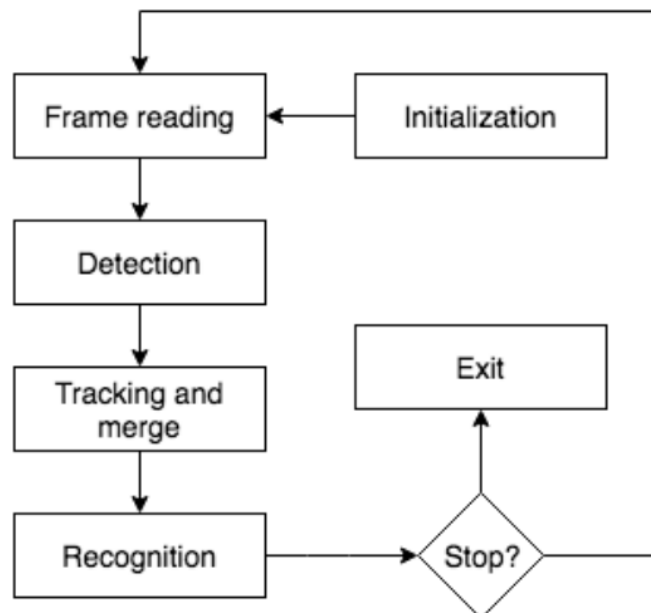
Podrobné informace k postupu návrhu systému a vyhodnocení metod jsou uvedeny v bakalářské práci Detekce hledaných osob ve videu (David Bažout, 2019, FIT VUT).

## Programátorský manuál

Výsledkem podrobného návrhu a experimentů je terminálová aplikace implementovaná v jazyce Python 3. Zdrojový kód je logicky rozčleněn do tříd implementujících podúkoly uvedené v seznamu níže.

- **class Database** – Reprezentuje databázi hledaných osob, implementuje načítání fotografií ze složky a optimalizované vyhledávání.
- **class FeatureExtractor** – Realizuje normalizaci fotografie tváře a následnou extrakci příznakového vektoru.
- **class Face** – Reprezentuje detekovanou tvář, realizuje její sledování v rámci videosekvence.
- **class Monitor** – Realizuje kompletní algoritmus reidentifikace osob, využívá ostatních tříd.
- **class Vian** – Zajišťuje odesílání detekcí na ViAn server.
- **class Recorder** – Ukládá video s rendery detekovaných osob do souboru.

Po spuštění je v první fázi provedeno zpracování argumentů příkazové řádky, inicializace spojení s ViAn serverem a extrakce příznaků pro fotografie hledaných osob ze zadané složky. Poté probíhá algoritmus reidentifikace hledaných osob ilustrovaný na obrázku 2.



Obr. 2. Algoritmus reidentifikace hledaných osob.

Vstupní videosekvence je po snímcích načítána pomocí knihovny **OpenCV**<sup>1</sup>. Pro každý snímek je provedena detekce tváří pomocí metody **MTCNN**<sup>2</sup>. Každá tvář je od okamžiku detekce

<sup>1</sup> <https://opencv.org/>

<sup>2</sup> [https://kpzhang93.github.io/MTCNN\\_face\\_detection\\_alignment/](https://kpzhang93.github.io/MTCNN_face_detection_alignment/)

sledována během pohybu mezi snímky. Ke sledování bylo využito korelačního trackeru z knihovny **Dlib**<sup>3</sup>. S každým novým snímkem je provedena extrakce příznakového vektoru pro všechny prozatím nerozpoznané tváře a tento příznak je porovnáván s databází. Extrakce příznakového vektoru je realizována modifikovanou neuronovou sítí **ResNet-34**<sup>4</sup> rovněž z knihovny **Dlib**. Porovnávání s databází tváří je paralelizováno pomocí knihovny **Numpy**<sup>5</sup>.

## Uživatelský manuál

Před použitím nástroje je nutné aktivovat předpřipravené virtuální prostředí jazyka Python. Poté je nástroj spustitelný s parametry blíže specifikovanými v tabulce níže.

Název parametru	Význam
--source	Řetězec definující zdroj videa podle standardu OpenCV.
--window	Umožňuje spuštění aplikace bez vizualizace. (True/False)
--save	Název souboru pro uložení vizualizovaného výstupu.
--database	Cesta ke složce s databází fotografií.
--export	Cesta ke složce pro export snímků detekcí.
--tracking-threshold	Práh úspěšnosti pro trackovací algoritmus. (5.0-20.0)
--match-threshold	Práh pro porovnání s příznaky databáze. (0.56)
--min-detection	Minimální velikost detekovaného obličeje v px. Vyšší hodnota přináší rychlejší vyhodnocení.
--detection-confidence	Práh pro úspěšnou detekci obličeje. (0.993)
--detection-step	Detekce obličejů pouze pro každý n-tý snímek.
--extraction-step	Extrakce příznaků u nerozpoznané tvare je prováděna každý n-tý snímek. Zrychlení v případě vysokého počtu tváří. (1)
--vian	URL adresa ViAn serveru.
--access-token	Autentizační token.
--dataset	Jméno příslušného datasetu na ViAn serveru.
--rmq-uri	RabbitMQ URI.

## Instalační manuál

Instalace nástroje vyžaduje interpret jazyka Python (verze 3.7.3). Dále je nutné vytvořit virtuální prostředí jazyka Python a nainstalovat do něj závislosti uvedené v souboru **requirement.txt**. Po dokončení instalace závislostí je nástroj připraven k použití. Konkrétní příkazy jsou uvedeny v souboru **Readme.md**.

<sup>3</sup> <http://dlib.net/>

<sup>4</sup> <http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html>

<sup>5</sup> <https://numpy.org/>

## Reference

BAŽOUT David a BERAN Vítězslav. Detekce a identifikace hledaných osob ve videu. Technická zpráva k projektu č.:VI20172020068, 2018. Dostupné z: <https://www.fit.vut.cz/research/publication/12393/>

BAŽOUT, David. Detekce hledaných osob ve videu. Brno, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. 2019-06-12. Vedoucí práce Beran Vítězslav. Dostupné z: <https://www.fit.vut.cz/study/thesis/22160/>