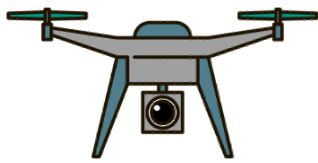
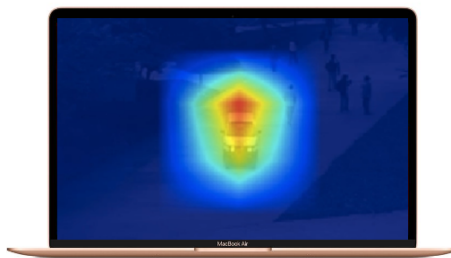


Detekce anomálií v chování davu na videodatech z dronu

David Bažout*



Video source



AI analysis



Security

Abstrakt

V posledních letech dochází k rychlému rozvoji využití dronů v mnoha odvětvích. Své uplatnění nachází i v oblasti národních bezpečnostních složek. Cílem této práce je návrh a implementace nástroje provádějícího analýzu davových scén. Tento nástroj zajišťuje včasnou identifikaci podezřelého chování osob a usnadňuje jeho lokalizaci.

Zkoumaná scéna je rozdělena do mřížky složené z buněk, které uchovávají model obvyklého pohybu v dané oblasti. Na základě videodat je vypočten hustý optický tok, který je dále zakódován do příznakového vektoru. Porovnáním příznaku s modelem obvyklého pohybu je určena úroveň anomálie jednotlivých buněk.

V rámci práce byl navržen a implementován kompletní systém umožňující detekci anomálie v davu a její vizualizaci do mapového podkladu. Mezi hlavní přínosy patří návrh vhodného algoritmu stabilizace videa, vývoj konvolučního autoenkodéru extrahujícího relevantní příznakové vektory a implementace modelu pozadí navrženého pro zpracování on-line dat.

Klíčová slova: stabilizace videa – detekce anomálie – davové scény – drony

Přiložené materiály: [Ukázka vizualizační webové aplikace](#)

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

1. Úvod

Úkolem bezpečnostních složek v rámci společenských událostí s vysokou návštěvností je monitorování mnohdy i rozsáhlých oblastí s omezeným počtem lidských zdrojů. Rychlost zásahu v případě výskytu nebezpečné situace může rozhodovat o záchraně lidských životů.

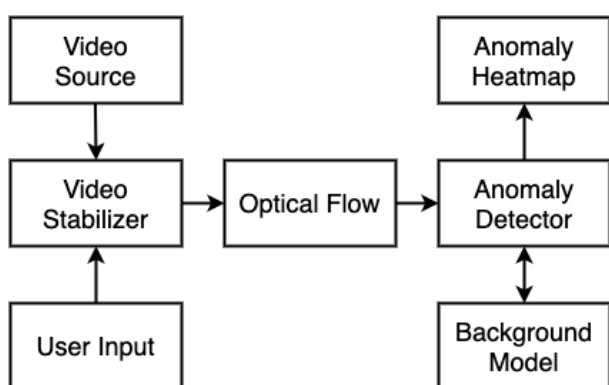
S monitorováním rozsáhlých oblastí již v současnosti pomáhají drony. Dron je po svém vzletu navigován do stacionární pozice vhodné pro sledování dané scény, kde zaujímá stacionární polohu, nicméně získaná videodata mohou být ovlivněna mnoha nepříznivými vlivy.

Uživatel potřebuje být upozorněn na neobvyklé chování v davové scéně. Při řešení úlohy je nutné vzít v úvahu nedostupnost jakýchkoliv trénovacích dat specifických pro danou scénu. V rámci reakce na nebezpečnou událost je nutná intuitivní identifikace místa jejího výskytu.

Detekce anomálie v davu je stále výrazně otevřeným problémem a to nejen v oblasti vývoje metod, ale i v pouhé interpretaci problému. V odborných článcích je za netypické chování davu obvykle považována změna rychlosti nebo směru pohybu, změnu vzhledu nebo výskyt netypického objektu [1]. Problémem existujících řešení [2, 3, 4] je jejich závislost na konkrétní scéně, jejíž změna vyžaduje opakování procesu trénování. Existující řešení dále neintegrují stabilizační metodu umožňující odstínění pohybů kamery oproti scéně.

Schéma navrhovaného řešení je na obrázku 1.

V první fázi probíhá registrace přicházejícího videa do mapového podkladu za pomoci uživatelského vstupu. Tato registrace je v průběhu analýzy udržována pomocí metody stabilizace připravené na specifika těchto videodat. Zkoumaná scéna je rozdělena do mřížky složené z buněk, které uchovávají model obvyklého pohybu v dané oblasti. Z přicházejících videodat je vypočten hustý optický tok. Optický tok je kódován do příznakového vektoru pomocí konvolučního autoenkodéru [3]. Navrhovaná metoda nevyžaduje žádné dotrénování pro konkrétní scénu. Model obvyklého pohybu je inicializován prázdný a průběžně aktualizován. Výsledkem analýzy je teplotní mapa reprezentující úroveň anomálie napříč scénou.



Obrázek 1. Schéma navrhovaného řešení.

Navrhovaný detektor anomálie byl porovnán s existujícími přístupy v rámci třech datových sad [5, 6, 7]. Ukázalo se, že dosahuje rozumných výsledků ve srovnání s ostatními metodami, ale navíc nabízí i možnost on-line trénování.

2. Specifikace úlohy

Cílem uživatele je sledování mnohdy i rozsáhlého prostoru pomocí dronu s IP kamerou. Dron zaujímá po úvodním manévrování stacionární polohu ve výšce 80 - 150 metrů nad terénem. IP kamera poskytuje RGB videozáznam scény o typických rozměrech 300x300 metrů s rozlišením 1920x1080 pixelů. Ukázka snímků z IP kamery je na obrázku 2. Scéna může být snímána za nízkého osvětlení během ranních nebo večerních hodin. Dále je zde riziko narušení viditelnosti vlivem nepříznivého počasí.



Obrázek 2. Ukázka typických videodat z dronu.

Videodata snímaná dronem mají svoje specifika. Jeho stacionární poloha může být narušena náhlým poryvem větru, korekcí trajektorie letu nebo dalšími nežádoucími okolnostmi. Částečně je tento problém řešen pomocí stabilizačního mechanismu úchyty kamery, nicméně i ten má svoje omezení. Je nutné počítat i s možností krátkodobého výpadku v přenosu dat.

Uživatel potřebuje mít přehled o situaci celkového proudění davu osob a potřebuje být upozorněn zejména na jeho zrychlení, zastavení, rozdělení nebo sročení. Změna v proudění davu nemusí být vždy způsobena výskytem nebezpečné situace, ale zároveň představuje její indikátor. Vyhodnocení potenciálního nebezpečí je v odpovědnosti uživatele a i z tohoto důvodu je zde tolerován výskyt falešných upozornění. Při řešení úlohy je nutné vzít v úvahu nedostupnost jakýchkoliv trénovacích dat specifických pro danou scénu.

V rámci reakce na nebezpečnou událost je nutná intuitivní identifikace místa jejího výskytu. Z tohoto důvodu je klíčové udržovat mapování obrazových dat na mapový podklad.

3. Stabilizace videa

Z popisu charakteru videodat v sekci 2 vyplývá nutnost rozlišování mezi pohybem snímacího zařízení a mezi samotným pohybem ve scéně. Vzhledem ke snímání

poměrně rozsáhlé oblasti představuje i drobný záchvěv kamery významně rušivý element při zkoumání pohybu v porovnání s pomalým prouděním davu.

Stabilizace obrazu je nezbytná i z důvodu udržování kontextu přicházejících videodat vzhledem k mapovému podkladu. Před začátkem analýzy optického toku lze od uživatele manuálně získat odpovídající body ve videodatech a mapovém podkladu a následně vypočítat i odpovídající transformační matici. K udržení kontextu přicházejících videodat k mapovému podkladu musí být udržována ještě transformační matice mezi počátečním a aktuálním snímkem, kterou lze určit skládáním dílčích transformací mezi jednotlivými snímky.

Z pohledu stabilizace je nutné řešení rozdělit na stabilizaci drobných záchvěvů a na situaci extrémních pohybu (nebo výpadků přenosu), kdy se může celá scéna posunout mimo záběr kamery. Drobné korekce nestability záznamu mohou být poměrně spolehlivě řešeny detekcí klíčových bodů a sledováním jejich pohybu do následujícího snímku pomocí metody řídkého optického toku [8]. Řešením této podúlohy je nalezení transformační matice popisující homografii mezi dvěma po sobě jdoucími snímky. V rámci této práce byly klíčové body detekovány pomocí Harrisova detektoru [9] a pro výpočet řídkého optického toku bylo využito Lucas-Kanade metody [8].

Problémem navrhovaného přístupu je pohyb davu nebo jiných objektů ve scéně. V případě detekce klíčových bodů na pohybujících objektech a jejich sledování do následujícího snímku je do homografie [10] započítán pohyb nesouvisející se záchvěvy kamery. To lze vyřešit uživatelským vstupem v podobě binární masky o rozlišení vstupního videa. Pomocí této masky uživatel specifikuje oblasti, ve kterých není žádný pohyb očekáván a klíčové body z těchto oblastí mohou být bezpečně použity pro určení homografie.

Výpočet homografie dále uvažuje transformaci pouze mezi dvěma 2D souřadnicovými systémy - souřadnicemi vstupního videa a mapového podkladu. Reálná scéna snímaná dronem je ale 3D charakteru a z tohoto důvodu by při hledání homografie mohlo dojít k dalším nepřesnostem [10]. Vzniká tedy další požadavek na binární masku zadávanou uživatelem - měla by zahrnovat pouze plochy ležící ve stejné rovině.

V případě silného záchvěvu nebo úplné ztráty původní scény není možné využít předchozí přístup s dostatečnou přesností. Namísto toho lze ve scéně pravidelně detekovat vhodné klíčové body společně s jejich deskriptory. K tomu lze využít výpočetně efektivních klíčových bodů ORB [8]. V případě selhání dříve popsaného přístupu lze provést detekci a následně

hledání odpovídajících klíčových bodů mezi původním a aktuálním snímkem a tím určit hledanou homografii [10]. Tento přístup je více robustní, ale zároveň i výpočetně náročnější.

4. Metody detekce anomálie

Častým způsobem detekce anomálie v chování davu je zkoumání změn optického toku. Existuje mnoho metod pro jeho výpočet, nicméně nutnost zpracování v reálném čase výrazně limituje možnost využití některé z metod založené na konvolučních neuronových sítích [4]. Dvojice po sobě jdoucích snímků musí být nejprve transformována do stejného souřadnicového systému, aby byl odstíněn vliv pohybu kamery vůči scéně.

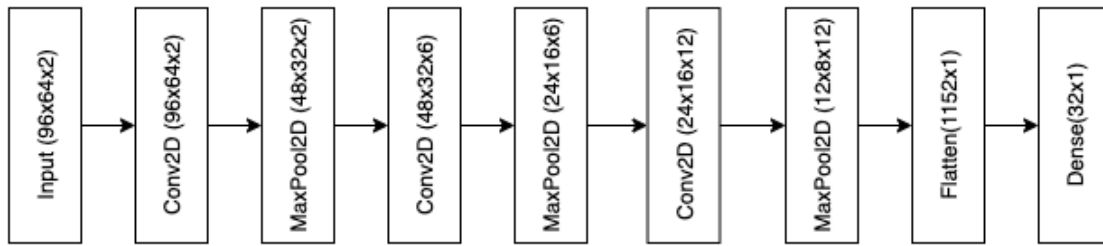
Vypočtený optický tok je vzhledem k zachování prostorového kontextu nejprve transformován do prostoru mapového podkladu za použití transformačních matic získaných při řešení úlohy stabilizace obrazu. Prostor mapového podkladu je rozdělen do mřížky tvořené buňkami o souhlasné velikosti. Základním principem výpočtu úrovně anomálie pro danou buňku je modelování optického toku v daném místě. Porovnáním aktuálního optického toku s jeho modelem je pomocí vhodné metriky vypočtena hodnota anomálie.

Specifikem řešené úlohy je nedostupnost jakýchkoliv trénovacích dat pro scénu, ve které bude detekce anomálie prováděna. Metoda detekce anomálie tedy musí být schopna vytvářet a průběžně aktualizovat model pohybu ve scéně pouze na základě on-line dat přicházejících z dronu, a to bez dostupnosti anotací.

Hlavní myšlenka navrhované metody pro detekci anomálie spočívá v popisu optického toku dané buňky pomocí relevantního příznakového vektoru [3]. Na základě série příznakových vektorů je průběžně aktualizován generativní model jejich pravděpodobnostního rozložení. Úroveň anomálie dané buňky je následně určena porovnáním aktuálního příznaku s modelem.

Extrakce příznakového vektoru

Pro extrakci příznakového vektoru bylo využito neuronové sítě s architekturou typu Denoising Convolutional Autoencoder [3]. Tato neuronová síť se rozděluje na část kodéru a dekodéru. Kodér nejprve provádí kompresi vstupního optického toku dané buňky do příznakového vektoru o výrazně nižší dimenzionalitě oproti vstupním datům. Na základě tohoto vektoru usiluje symetricky odpovídající dekodér o rekonstrukci původního vstupu. V průběhu trénování je tedy identifikována klíčová informace, která musí být do příznaku zakódována pro úspěšnou rekonstrukci původního vstupu.



Obrázek 3. Navrhovaná architektura neuronové sítě pro extrakci příznakového vektoru z výřezu optického toku.

Schéma navrhované architektury neuronové sítě inspirované článkem [3] je na obrázku 3. Vstupem neuronové sítě je výřez optického toku, který náleží dané buňce modelu. Následuje postupné snižování rozlišení a zvyšování počtu kanálů vstupního vzorku obvyklým způsobem pomocí konvolučních a pooling vrstev. V poslední fázi je tato informace zakódována pomocí plně propojené vrstvy pouze do nízkodimenzionálního příznaku, pomocí kterého lze na on-line datech vytvářet model pozadí.

Úspěšnost navrhované architektury neuronové sítě lze dále optimalizovat změnou domény vstupních dat. Anomalie v davu nemusí být vždy patrná pouze z optického toku. Z pohledu řešené úlohy považuji za vhodné vyhodnotit vliv využití obrazové informace nebo optického toku několika posledních snímků.

Statistický model pozadí

Vhodným a velmi často používaným přístupem pro modelování pravděpodobnostního rozložení multidimenzionálních dat je využití směsi Gaussovských rozložení [11]. Nicméně průzkum existujících implementací poukázal na nedostupnost řešení s možností inkrementálního trénování na online datech. Z tohoto důvodu jsem se rozhodl navrhnout vlastní implementaci.

Navrhovaný algoritmus pracuje s množinou Gaussovských rozložení, které jsou definovány počtem prvků, lineárním součtem a kvadratickým součtem. Ve výchozím stavu je tato množina prázdná. Algoritmus umožňuje průběžnou aktualizaci modelu pozadí pomocí příznakového vektoru. V prvním kroku aktualizace je nalezena vzdálenost příznakového vektoru a středu nejbližší Gaussovy křivky z dané množiny. Pokud je tato vzdálenost menší než stanovený práh, dojde k aktualizaci statistik nejbližší Gaussovy křivky. Pokud je množina Gaussovských rozložení prázdná nebo pokud vzdálenost překračuje stanovený práh, je inicializována nová Gaussova křivka se středem odpovídajícím danému příznaku.

Na operaci aktualizace modelu navazuje výpočet pravděpodobnosti výskytu daného příznakového vektoru vzhledem k modelu pozadí. V první fázi je pro

každou Gaussovu křivku vypočtena úroveň odpovědnosti za jeho vygenerování podle vztahu 1. Proměnná w_z reprezentuje odpovědnost Gaussovy křivky s identifikátorem z za vygenerování příznaku x . Proměnné μ a σ^2 reprezentují parametry GMM modelu.

$$w_z = \frac{\mathcal{N}(x | \mu_z, \sigma_z^2)}{\sum_k \mathcal{N}(x | \mu_k, \sigma_k^2)} \quad (1)$$

Každá Gaussova křivka reprezentuje určitý cluster pohybu davu. Podobnost příznakového vektoru s tímto clusterem je vyjádřena pomocí váhy odpovědnosti w_z . Kromě podobnosti je zde ale důležité započítat i pravděpodobnost výskytu daného clusteru π_z . Pravděpodobnost příznakového vektoru vzhledem k modelu pozadí je tedy vypočtena podle vztahu 2.

$$p(x | \eta) = \sum_z w_z \pi_z \quad (2)$$

5. Datové sady

Pro vývoj metod strojového učení obecně je zcela zásadní dostupnost relevantních datových sad pro trénování a následné vyhodnocení jejich úspěšnosti. Problémem úloh detekce anomálie je přílišná subjektivnost specifikace daného problému. Různé metody definují pojem anomálie odlišným způsobem, což komplikuje jejich relevantní srovnávání. V následující kapitole bude proveden průzkum existujících datových sad s cílem identifikace vhodné datové sady relevantní k uživatelské úloze blíže specifikované v kapitole 2.

UCSD Dataset

Datová sada UCSD Dataset [5] je tvořena krátkými videosekvencemi o rozlišení 238x158 pixelů skládajících se z 200 snímků. Byla pořízena pomocí stacionární kamery snímající chodník z vyvýšené pozice. Hustota osob na záznamech kolísá od silně zalidněných scén po výskyt pouhých několika chodců. Dataset je rozdělený do dvou nezávislých částí, které odpovídají různým scénám. Ukázka dat je na obrázku 4.

Pro každou scénu byla pořízena trénovací i testovací podčást. Trénovací podčást obsahuje záznamy pouze bez výskytu anomálií. Anomálie je v rámci tohoto datasetu definována jako výskyt nežádoucích objektů (cyklista, skateboardista, vozíčkář nebo vůz) nebo neobvyklý vzor pohybu. K veškerým trénovacím snímkům je dostupná anotace o globálním výskytu anomálie. K několika vybraným sekvencím jsou dále dostupné masky definující lokalitu anomálie.



Obrázek 4. Ukázka dat z UCSD datasetu.

Avenue Dataset

Tento dataset [6] tvoří 16 trénovacích a 21 testovacích videosekvencí o celkovém počtu 30652 snímků s rozlišením 640x360px. Trénovací data zahrnují pouze situace s obvyklým chováním a anomálie se vyskytují pouze v testovacích datech. Neobvyklé situace jsou v anotacích označeny pomocí ohraničujících rámečků.



Obrázek 5. Ukázka dat z Avenue datasetu.

Videa byla pořízena na veřejném prostranství s kamerou ve stacionární poloze. Občasné dochází k záchvěvům kamery nebo řídkému výskytu některého z druhů běžného chování. Anomálií v testovacích datech představuje celkem 14 typů neobvyklých situací (běh, házení objektů, nezvykle pomalá chůze a další). Ukázka dat tohoto datasetu je na obrázku 5. V záznamech se objevují anomálie v podobě netypických pohybových vzorů, nicméně malá vzdálenost od scény

a umístění kamery v malé výšce snižují relevantnost těchto dat.

UMN Crowd Dataset

Jedná se o další veřejně dostupnou datovou sadu [7] zaměřenou na detekci anomálie v chování davu. Celkově se skládá ze 3 videosekvencí o celkové délce 4 minuty a 17 vteřin. Scéna je vždy snímána pomocí stacionární kamery umístěné na vyvýšené pozici. Každá videosekvence se rozděluje na části s obvyklým a netypickým chováním. Ukázka dat je na obrázcích 6.



Obrázek 6. Ukázka dat z UMN datasetu.

Train Station Dataset

Datová sada [12] tvořená 33 minut dlouhým záznamem nádražní haly o rozlišení 720x480 pixelů. Záznam je pořízený z vyvýšeného místa pomocí stacionární kamery. Po celou dobu záznamu se zde vyskytuje poměrně velké množství osob vztažených k danému prostoru. Anotace obsahují trajektorie pohybu chodců ve scéně, což umožňuje provádění analýzy davového chování. Ukázka dat je na obrázku 7. Tento videozáznam by mohl být vzhledem k podobnosti dat se záznamy pořízenými dronem využit zejména pro trénování metody pro extrakci relevantního příznakového vektoru bez přítomnosti učitele.

6. Vyhodnocení metody stabilizace

V posledních letech bylo navrženo mnoho metod stabilizace videa. Nicméně univerzální metrika pro jejich srovnání stále nebyla přesně specifikována [13]. Pro vyhodnocení navrhované stabilizační metody bylo využito metriky MDE (Mean displacement error) [14] z důvodu její souvislosti z řešenou úlohou. Tato metrika hodnotí přímo kvalitu nalezené transformace a vyjadřuje průměrnou odchylku v posunu jednotlivých pixelů oproti anotaci.



Obrázek 7. Ukázka dat z Train Station Datasetu.

Pro vyhodnocení metody stabilizace byla vytvořena vzhledem k její specifčnosti vlastní datová sada. Tato datová sada se skládá z vlastních videí pořízených kamerou umístěnou na palubě dronu společně se záznamy veřejného prostranství z datové sady Multidrone Public Dataset [15]. Celkem se jedná o 9 videosekvencí opatřených manuální anotací s maskou definující stabilní oblasti ve videu a transformační maticí specifikovanou člověkem.

Scénáře vlastní datové sady zahrnují přelety dronu konstantním směrem a rychlostí, drobné záchvěvy, kompletní ztrátu scény nebo výskyt pohybu v maskou definovaných stabilních oblastech. Průměrná odchylka jednotlivých pixelů je v porovnání s ručně pořízenou anotací 28 pixelů při použití stabilizační metody a 132 pixelů bez stabilizace. Část chyby zde připadá na nepřesnost ručně pořízené anotace.

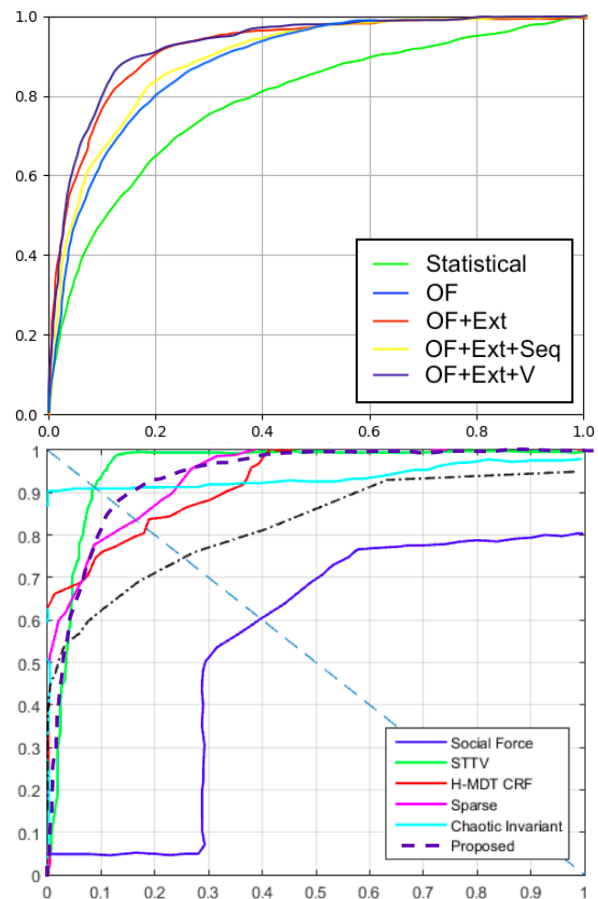
7. Experimenty s detektorem anomálie

Původní konvoluční autoenkodér navržený v kapitole 4 posloužil jako výchozí bod veškerých experimentů. Na jeho základě bylo navrženo několik dalších variant a ty byly poté vyhodnoceny na datových sadách UCSD [5], UMN [7] a Avenue [6]. Úspěšnost modelů byla ověřena v off-line i on-line režimu a bylo provedeno porovnání jejich úspěšnosti s existujícími přístupy.

Výchozí architektura konvolučního autoenkodéru (varianta OF) byla dále doplněna o další plně propojenou vrstvu o celkovém počtu 1152 neuronů (varianta OF + Ext). Se vzrůstajícím počtem parametrů modelu se zde otevírá možnost zachycení komplexnější souvislosti ve vstupních datech. Poslední navrhovanou variantou založenou pouze na analýze optického toku byla architektura, jejíž vstup tvoří trojice sousedních snímků optického toku (varianta OF + Ext + Seq). Příznakový vektor by mohl být v tomto případě robustnější vůči nestabilitě vstupního optického toku a zároveň by zde mohla být zachycena i změna pohybu.

V některých situacích nemusí být výskyt anomálie patrný pouze na základě optického toku, ale důležitou roli zde může sehrát i vizuální informace. Například v datové sadě UCSD je za jednu z anomálií považován průjezd skateboardisty. Nicméně skateboardista se od běžného chodce liší pouze svým vzhledem a v doméně optického toku je při zachování souhlasné rychlosti a směru pohybu nerozpoznatelný. Poslední návrh konvolučního autoenkodéru (varianta OF + Ext + V) má tedy na vstupu kromě jednoho vzorku optického toku i obrazová data v odstínech šedé.

Pro účely trénování byla nejprve využita datová sada Train Station Dataset [12]. V této datové sadě se vyskytuje pohyb mnoha osob snímaný z kamery umístěné vysoko nad scénou, což velmi dobře odpovídá situacím řešené úlohy. Tato datová sada byla později rozšířena o videosekvence datové sady UMN, UCF [16] a UCSD, aby mohl být zkoumán vliv rozmanitějších dat na úspěšnost detekce. V tomto případě byla ale trénovací data natolik rozmanitá, že se nedařilo rozumným způsobem minimalizovat chybovou funkci.

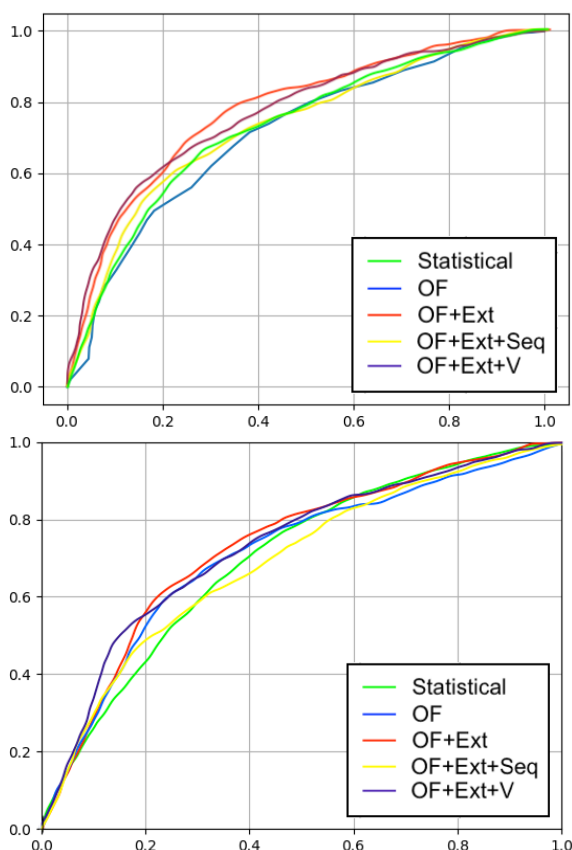


Obrázek 8. Výsledky vyhodnocení autoenkodéru a statistické metody na UMN datasetu a jejich porovnání s existujícími přístupy. (Převzato z [17])

Ukázalo se, že zcela zásadní vliv na výsledky trénování má volba aktivačních funkcí. Pro aktivaci výstupní vrstvy enkodéru byla zvolena aktivační funkce

sigmoid vzhledem ke vhodnosti normalizace hodnot příznakového vektoru do intervalu $\langle 0; 1 \rangle$. Aktivace výstupní vrstvy byla ponechána lineární, protože se v rekonstruovaném optickém toku mohou vyskytovat kladné i záporné hodnoty z teoreticky neomezeného rozsahu. U ostatních vrstev byla po několika neúspěšných experimentech ponechána lineární aktivace.

Za výchozí datovou sadu pro vyhodnocení veškerých variant konvolučního autoenkodéru jsem zvolil UMN dataset vzhledem k jednoduchosti její interpretace pojmu anomálie. Vzhledem k nedostupnosti trénovacích dat bez výskytu anomálie byly metody vyhodnoceny pouze v on-line režimu. Pro každou videosekvenci byl inicializován prázdný model pozadí, který byl během testování průběžně aktualizován. Scéna byla rozdělena do mřížky o velikosti 5x5 buněk. Vyhodnocení bylo provedeno z pohledu globálního výskytu anomálie a za skóre pro sestavení ROC křivky byla považována maximální hodnota z pole buněk. Výsledky vyhodnocení jsou včetně porovnání s původní statistickou metodou a existujícími přístupy na obrázku 8.

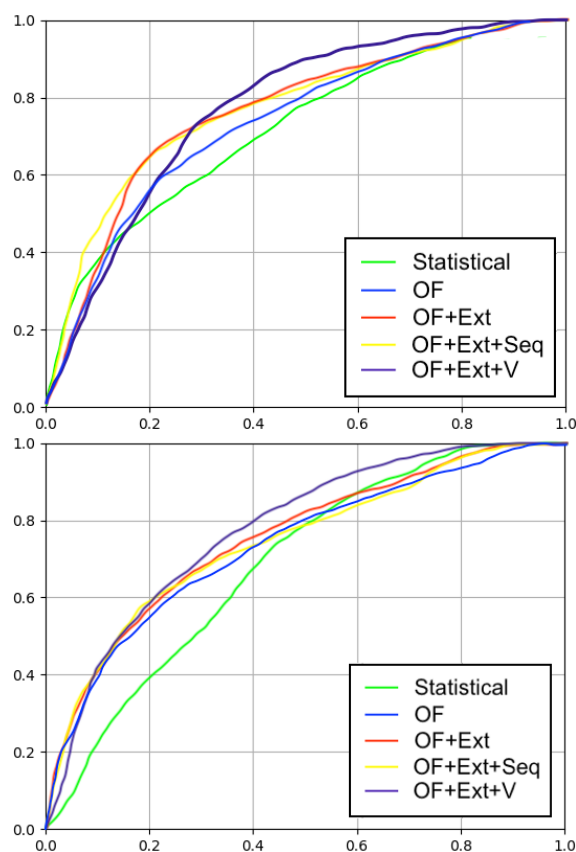


Obrázek 9. Vyhodnocení navržených metod pro on-line i off-line režim na UCSD datasetu.

V dalším experimentu bylo provedeno vyhodnocení navržených metod na datové sadě UCSD. Tato datová sada již není triviální a vyskytují se zde i anomálie s projevem pouze ve vizuální doméně. Scéna

byla opět rozdělena do mřížky o rozměrech 5x5 buněk a metody byly vyhodnoceny z pohledu výskytu globální anomálie. V off-line režimu byly modely pozadí pro jednotlivé buňky inicializovány na základě trénovacích sekvencí. V on-line režimu bylo vyhodnocení spuštěno s prázdnými modely. Výsledky vyhodnocení navržených metod v off-line i on-line režimu je na obrázku 9.

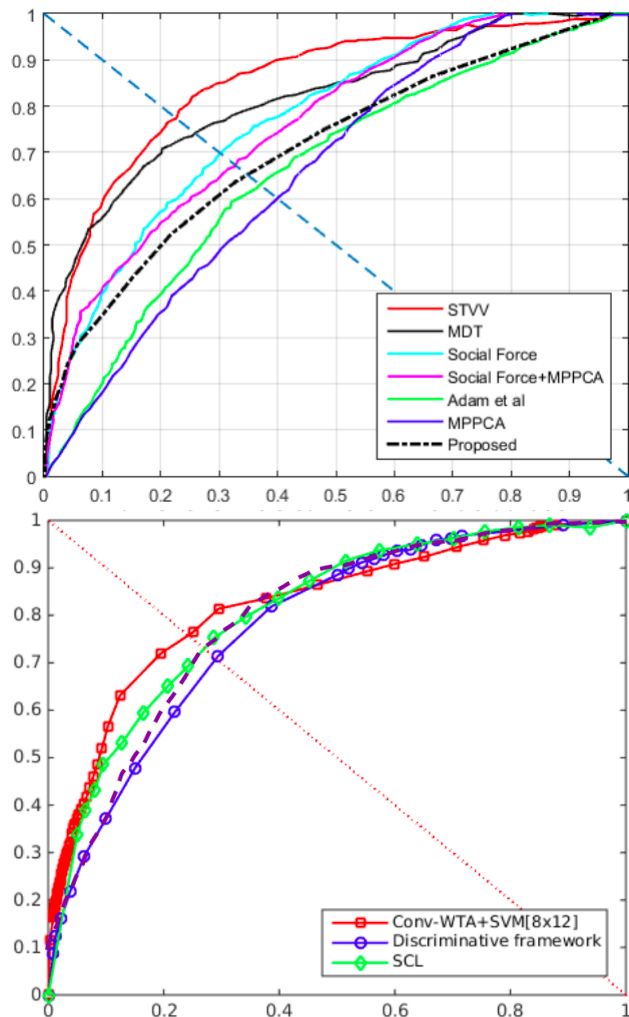
Zkoumaný detektor anomálie byl v kontextu uživatelské úlohy navržen pro scény velmi specifického charakteru. V rámci experimentů bylo provedeno jeho vyhodnocení i na datech odlišných od řešené úlohy. K tomuto účelu posloužila datová sada Avenue [6]. Výsledky vyhodnocení ukazují jeho schopnosti řešit úlohu detekce anomálie obecně. Výsledky vyhodnocení porovnávající úspěšnost on-line i off-line režimu jsou na obrázku 10.



Obrázek 10. Vyhodnocení navržených metod pro on-line i off-line režim na Avenue datasetu. (Převzato z [18])

Z výsledků vyhodnocení je patrné, že přidání plně propojené vrstvy neuronů pomohlo ve všech případech vylepšit úspěšnost metody. Architektura neuronové sítě OF + Ext + Seq pouze zbytečně zvyšuje výpočetní náročnost a nepřináší žádné zlepšení detekce. Poměrně dobrých výsledků dosahuje v porovnání s variantami založenými na konvolučním autoenkodéru i základní statistická metoda.

Nejstabilnějších výsledků bylo v rámci veškerých experimentů dosahováno variantou kombinující optický tok společně s vizuální informací. Vzhledem důležitosti vizuální informace v praktických úlohách jsem se rozhodl tuto metodu implementovat i do uživatelského nástroje. Výsledky vyhodnocení existujících přístupů na datových sadách UCSD a Avenue je na obrázku 11. Navrhovaná metoda dosahuje v porovnání s existujícími přístupy velmi dobrých výsledků a zároveň nabízí výhodu v podobě on-line analýzy.



Obrázek 11. Vyhodnocení existujících přístupů na datových sadách UCSD a Avenue. (Převzato z [18])

8. Závěr

Cílem této práce byl návrh a implementace nástroje, který umožňuje automatizovanou analýzu chování davu na videodatech z dronu a napomáhá lokalizovat případnou anomálii na mapovém podkladu. V rámci práce byla navržena a otestována spolehlivá metoda stabilizace videa specializovaná na zvláštnosti videodat z dronu.

Dále byl navržen vlastní detektor anomálie založený na extrakci příznaků optického toku a vizuální informace pomocí konvolučního autoenkodéru. Výhodou tohoto detektoru je jeho nezávislost na cílové scéně a možnost zpracování on-line dat.

Detektor byl vyhodnocen na třech neznámějších datových sadách zaměřených na detekci anomálie. Byla prokázána jeho srovnatelná úspěšnost s nejlepšími současnými přístupy. Současně se ukázalo, že jeho použití není omezeno jenom na davové scény snímané z paluby dronu.

Funkčnost nástroje byla integrována do intuitivní webové aplikace sloužící pro jeho konfiguraci a vizualizaci výsledků analýzy.

Poděkování

Závěrem bych chtěl poděkovat svému vedoucímu diplomové práce Ing. Vítězslavu Beranovi Ph.D. za dlouhodobou a velmi produktivní spolupráci, která výrazným způsobem prohloubila moje znalosti nejen v oblasti řešení úloh počítačového vidění.

Literatura

- [1] Teng Li, Huan Chang, Meng Wang, Bingbing Ni, Richang Hong, and Shuicheng Yan. Crowded scene analysis: A survey. *IEEE transactions on circuits and systems for video technology*, 25(3):367–386, 2014.
- [2] Mahdyar Ravanbakhsh, Moin Nabi, Enver Sangineto, Lucio Marcenaro, Carlo Regazzoni, and Nicu Sebe. Abnormal event detection in videos using generative adversarial nets. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 1577–1581. IEEE, 2017.
- [3] Dan Xu, Elisa Ricci, Yan Yan, Jingkuan Song, and Nicu Sebe. Learning deep representations of appearance and motion for anomalous event detection. *arXiv preprint arXiv:1510.01553*, 2015.
- [4] B Ravi Kiran, Dilip Mathew Thomas, and Ranjith Parakkal. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *Journal of Imaging*, 4(2):36, 2018.
- [5] Antoni Chan and Nuno Vasconcelos. Ucsd pedestrian dataset. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 30(5):909–926, 2008.
- [6] Cewu Lu, Jianping Shi, and Jiaya Jia. Abnormal event detection at 150 fps in matlab. In *Proceedings of the IEEE international conference on computer vision*, pages 2720–2727, 2013.
- [7] R Raghavendra, AD Bue, and M Cristani. Unusual crowd activity dataset of university of minnesota, 2006.
- [8] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. ” O’Reilly Media, Inc.”, 2008.
- [9] Konstantinos G Derpanis. The harris corner detector. *York University*, 2, 2004.
- [10] Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Nelson Education, 2014.
- [11] Douglas A Reynolds. Gaussian mixture models. *Encyclopedia of biometrics*, 741:659–663, 2009.
- [12] Bolei Zhou, Xiaogang Wang, and Xiaoou Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2871–2878. IEEE, 2012.
- [13] Wilko Guilluy, Laurent Oudre, and Azeddine Beghdadi. Video stabilization: challenges, methods and evaluation.
- [14] Md Alamgir Hossain and Eui-Nam Huh. A novel performance metric for the video stabilization method. In *International Conference on Ubiquitous Information Management and Communication*, pages 486–496. Springer, 2019.
- [15] Ioannis Mademlis, Vasileios Mygdalis, Nikos Nikolaidis, Maurizio Montagnuolo, Fulvio Negro, Alberto Messina, and Ioannis Pitas. High-level multiple-uav cinematography tools for covering outdoor events. *IEEE Transactions on Broadcasting*, 65(3):627–635, 2019.
- [16] Saad Ali and Mubarak Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6. IEEE, 2007.
- [17] S Bansod and A Nandedkar. Detection and localization of anomalies from videos based on optical flow magnitude and direction. In *International Conference on Communication and Signal Processing 2016 (ICCASP 2016)*, pages 454–462. Atlantis Press, 2016.
- [18] Hanh TM Tran and David Hogg. Anomaly detection using a convolutional winner-take-all autoencoder. In *Proceedings of the British Machine Vision Conference 2017*. British Machine Vision Association, 2017.