



Zadání diplomové práce

Název:	Aplikace pro marketingovou analýzu služby Instagram
Student:	Bc. Adam Formánek
Vedoucí:	Ing. Jakub Novák
Studijní program:	Informatika
Obor / specializace:	Znalostní inženýrství
Katedra:	Katedra aplikované matematiky
Platnost zadání:	do konce letního semestru 2023/2024

Pokyny pro vypracování

Cílem práce je vytvořit automatizované softwarové metody (algoritmy) pro vyhodnocení interakce uživatelů s aplikací sociální sítě Instagram na platformě iOS s využitím záznamů obrazovek. Vstupem je video (záznam obrazovky) a výstupem jsou strukturovaná textová data popisující chování uživatele v aplikaci. Prostředkem pro dosažení cíle je návrh algoritmů zpracování obrazu. Myšlenka celé úlohy je ze záznamu vytáhnout tzv. transakce. Transakce reprezentuje celou interakci uživatele s reklamním sdělením + Informace o reklamě.

Stávající metriky digitálních platforem jako jsou "líky", komentáře nebo sdílení jsou již neefektivní a uživatelé se v digitálním prostředí stále více a vědoměji snaží skrýt své chování viditelným metrikám s cílem předejít reklamnímu cílení a zmást algoritmy. Níže zmíněné reference ukazují možnosti použití alternativních metrik popisující chování v digitálním prostředí pomocí metrik typu „see“, „search“ a „save“.

Úkoly:

- 1) Seznamte se s aplikací Instagram a způsobem její prezentace reklamního sdělení uživateli.
- 2) Proveďte rešerši algoritmů využitelných pro zisk a popis interakce uživatele s aplikací.
- 3) Navrhněte formát strukturovaného textového výstupu.
- 4) Navrhněte a implementujte algoritmy zpracování obrazu (minimálně 10) pro zisk interakcí s uživatelem.
- 5) Otestujte výsledky navržených algoritmů na datech z aplikace Instagram a zhodnoťte



**FAKULTA
INFORMAČNÍCH
TECHNOLOGIÍ
ČVUT V PRAZE**

úspěšnost.

Reference:

- AUDY MARTÍNEK, Petra. Brand-lurking Phenomenon on Social Media. Praha, 2021.

Dizertační práce. Univerzita Karlova, Fakulta sociálních věd, Institut komunikačních studií a žurnalistiky. Vedoucí práce Supa, Markéta.

- Martínek, Petra. (2021). Mapping methods of research on consumer engagement with brands on social media: A literature review. Methodological Innovations. 14. 205979912098538. 10.1177/2059799120985384.





**FAKULTA
INFORMAČNÍCH
TECHNOLÓGIÍ
ČVUT V PRAZE**

Diplomová práce

Aplikace pro marketingovou analýzu služby Instagram

Bc. Adam Formánek

Katedra aplikované matematiky
Vedoucí práce: Ing. Jakub Novák

4. května 2023

Poděkování

Rád bych poděkoval Ing. Jakubu Novákovi za vzorné vedení této práce, přátelský přístup, rychlou zpětnou vazbu a užitečné rady a postřehy. Dále bych rád poděkoval své rodině za nesmírnou podporu.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů, zejména skutečnost, že České vysoké učení technické v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 citovaného zákona.

V Praze dne 4. května 2023

.....

České vysoké učení technické v Praze
Fakulta informačních technologií

© 2023 Adam Formánek. Všechna práva vyhrazena.

Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí a nad rámec oprávnění uvedených v Prohlášení na předchozí straně, je nezbytný souhlas autora.

Odkaz na tuto práci

Formánek, Adam. *Aplikace pro marketingovou analýzu služby Instagram*. Diplomová práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2023.

Abstrakt

Práce se zabývá automatickou detekcí reklamních sdělení a korepondujících interakcí uživatele ze záznamu obrazovky sociální sítě Instagram na platformě iOS. Řešení se v pořadí skládá z detekce lokality, nového příspěvku, reklamního sdělení, typu média a následně samotných interakcí. Na jednotlivé podproblémy jsou využity algoritmy počítačového vidění a zpracování obrazu jako template matching, Hough circles, SSIM index nebo segmentace prahováním. Algoritmy jsou implementovány v aplikaci, která na základě vstupního videa, barevného režimu operačního systému a modelu zařízení vytvoří strukturovaný textový soubor skládající se z tzv. transakcí. Transakce popisují reklamní sdělení a obsahují informace jako je lokalita detekce, délka trvání manipulace s reklamou, typ reklamy, typ média a v neposlední řadě interakce. Aplikace dokázala na testovacím datasetu detekovat reklamní sdělení v 92,16 % případů a korepondující interakce s úspěšností 93,53 %. Výsledky na datech z uživatelského testování mají úspěšnost detekce reklamního sdělení 78,57 %. Úspěšnost určení korepondujících interakcí je 81,25 %, avšak s velmi vysokou mírou falešných detekcí.

Klíčová slova počítačové vidění, Instagram, záznam obrazovky, mobilní zařízení, iOS, reklamní komunikace, detekce interakcí uživatele, Houghovy kruhy, template matching, SSIM

Abstract

This thesis aims to automate advertisement detection and corresponding interactions done by a user from a screen recording of the Instagram iOS application. The solution consists of detecting a location, a new post, an advertisement, a type of media, and the actual interactions. Computer vision techniques, such as template matching, Hough circles, SSIM index, or segmentation by thresholding, are used to resolve each subproblem. Algorithms are implemented in the form of an application, which receives video, system skin, and type of the device as input and outputs a structured text file. The output file consists of transactions, which are composed of the advertisement, additional information such as the location of detection, length of interactions with the advertisement or its type, and interactions. On the test dataset, the proposed application achieved to detect the advertisement in 92,16 % of cases and corresponding interactions with a success rate of 93,53 %. Results on the dataset from user testing show 78,57 % correctly detected advertisements and interactions with a success rate of 81,25 %, however with a high rate of false positives.

Keywords computer vision, Instagram, screen recording, mobile device, iOS, advertisement communication, detection of user interactions, Hough circles, template matching, SSIM

Obsah

Úvod	1
1 Marketingový výzkum	3
2 Aplikace Instagram a prezentace reklamních sdělení	5
2.1 Lokace	5
2.1.1 Story	5
2.1.2 Reels	6
2.1.3 Feed	6
2.1.4 Profil	7
2.1.5 Lokace po prokliku CTA tlačítka	8
2.1.6 Ostatní	9
2.2 Typy reklamních sdělení	11
2.2.1 Sponzorovaný	11
2.2.2 Placená partnerství	12
2.2.3 Organický	12
3 Definice interakcí	13
3.1 See	13
3.2 Search	15
3.3 Save	16
3.4 Social	17
3.5 Shop a Skip	19
4 Výzkum problematiky	21
4.1 Získání informací ze záznamu obrazovky	22
4.2 Metody rozpoznání interakcí	23
4.2.1 Pohybové interakce	23
4.2.2 Statické interakce	24

5	Teoretické zázemí	25
5.1	Terminologie komprese videa	25
5.1.1	Parametry komprese	25
5.1.1.1	Bit rate	25
5.1.1.2	Frame rate	25
5.1.2	Codec	26
5.2	Zpracování obrazu	26
5.2.1	Segmentace	27
5.2.2	Background subtraction	27
5.2.3	Filtrace	28
5.2.4	Hranová detekce	28
5.2.5	Binární morfologické operace	29
5.2.6	Template matching	30
5.2.7	Kontury	30
5.2.8	Hough circles	30
5.2.9	SSIM index	31
5.3	OCR	32
6	Návrh algoritmů	33
6.1	Vstupní data	33
6.2	Předzpracování dat	36
6.3	Inicializační část	37
6.3.1	Metadata videa	37
6.3.2	Škála	37
6.3.3	Určení škály pomocí TM	38
6.4	Návrh řešení algoritmů	41
6.4.1	Určení lokality	41
6.4.2	Detekce pohybu a nového příspěvku	44
6.4.3	Detekce reklamy	46
6.4.4	Rozpoznání typu média	48
6.4.5	Rozpoznání interakcí	49
7	Transakce a sktrukturovaný výstup	59
7.1	Transaction	59
7.2	Action	60
7.3	Výstup	61
8	Implementace	63
8.1	Převod VFR na CFR	64
8.2	Rozhraní	64
8.3	Inicializační část a předzpracování	65
8.4	Detekce lokality	65
8.5	Detekce pohybu a nového příspěvku	66
8.6	Detekce reklamního sdělení	66

8.7	Uzavření a otevření transakce	67
8.8	Detekce typu média	67
8.9	Detekce interakcí	67
8.10	Uzavření transakce	70
8.11	Evaluace	70
9	Výsledky	73
9.1	Vlastní dataset	73
9.2	Reálná data	77
10	Diskuze	81
	Závěr	83
	Literatura	85
A	Seznam Použitých zkratk	95
B	Obsah přiloženého archivu	97

Seznam obrázků

1.1	Diagram chování zákazníka od povědomí, po zájem o koupi [1].	4
2.1	Ukázka lokace feed a typů příspěvků.	7
2.2	Ukázka lokací story (a), reels (b) a profesionální profil (c).	8
2.3	Ukázka odkazovaných stránek při prokliku CTA tlačítka.	9
2.4	Klasické obrazovky. Komentáře na lokaci feed (a), feed příspěvky na profilu (b) a obrazovka pro hledání profilů (c).	9
2.5	Modální okna. Rozšířený popis na story (a), možnosti na reels a feed (b, c), komentáře na reels (d), uložení příspěvku (e) a sdílení příspěvku (f).	10
2.6	Sponsored reklamní sdělení na různých lokacích.	11
2.7	Ukázka příspěvku placené partnerství (a) a označeného reklamního sdělení na konci popisku (b).	12
3.1	Progress bar na reels a story.	14
3.2	See-Comment na feed (a) a See-Pause na story (b).	15
3.3	BrowseProfile interakce.	16
3.4	Screenshot na iOS 16.	17
3.5	ShareIG na lokaci reels.	17
3.6	Like a SavePostIG u příspěvku na feed.	18
3.7	Follow interakce na profilu a reels.	18
3.8	Zveřejnění komentáře.	19
4.1	Segmentace a mapování kolonek z [2].	22
5.1	Ukázka MOG background subtractoru [3].	28
5.2	Ukázka morfologických operací eroze (b) a dilatace (c) na binární masce (a) [4].	29
5.3	Převod a reprezentace bodů v akumulátoru [5].	31
5.4	Porovnání MSE a SSIM indexu [6].	32

6.1	Ukázka sekvence navštívení profilu, iPhone 11 Pro, světlý režim. Uživatel prohlíží feed (a), klikne na profil (b) a posune se dolů (c).	34
6.2	Ukázka různých barevných režimů a zařízení.	35
6.3	Porovnání horních systémových prvků na lokaci story.	37
6.4	Detekce ikon na lokaci reels pomocí TM.	42
6.5	Detekce lokace modálního okna shareIG. Celá obrazovka bez oříznutí (a), oříznuté modální okno se zvýrazněným indikátorem (b) a oříznuté specifické UI prvky (c).	43
6.6	Navigační menu scrollovacích obrazovek s fixním navigačním menu.	43
6.7	Navigační menu na feed před posunutím dolů (a) a po posunutí dolů (b).	44
6.8	Přechodová animace story.	45
6.9	Přechodová animace reels.	45
6.10	Typy hlaviček sponzorovaný a placené partnerství.	46
6.11	Typy profilových obrázků. Klasický (a), aktivní story, účet zveřejnil nějaké story, které si může uživatel stisknutím na obrázek prohlédnout (b) a kolaborace mezi dvěma profily (c).	47
6.12	Celá hlavička příspěvku s ikonou more na feed.	47
6.13	CTA tlačítka reklamních sdělení. Feed (c) obsahuje i tlačítko určující typ média.	48
6.14	See–ExpandText interakce. Uživatel si právě zobrazil reels reklamní sdělení (a), po chvíli se automaticky objeví CTA tlačítko (b) a nakonec rozkliknul popis (c).	50
6.15	Původní progress bar a jeho maska.	51
6.16	Ukázka Save–SavePostIG pro reels scénáře jedna (a) a alertu scénáře dva (b).	51
6.17	Celý proces uložení příspěvku scénáře tři. Uživatel si prohlíží příspěvek (a), pomocí ikony more zobrazí reelsOptions (b), klikne na ikonu save, čímž se zobrazí saveIG (c). Alternativní scénář je vyobrazen na (d), kdy uživatel kliknul na ikonu save_fill a tím zrušil uložení — interakce není detekována.	52
6.18	Ukázka See–BrowseImages (a), lišty pod příspěvkem s ikonami a indikátorem (b) a segmentace indikátoru z obrázku 6.18a (c).	53
6.19	Průběh interakce See–ExpandText na feed.	53
6.20	Zoom (a) a WatchWhole (b) interakce pro feed.	54
6.21	Ukázka zpětné vazby u Social–Follow na feed. Uživatel si zobrazil příspěvek od uživatele, kterého nesleduje (a) a následně kliknul na sledovat (b).	55
6.22	Follow interakce na profilu prvním (a) a druhým (b) způsobem.	56
6.23	Ukázka interakce ShareIG. Vybrání příjemce (a), odeslání (b) a následné automatické zavření modálního okna (c).	57
7.1	Ukázky strukturovaného výstupu. Detekované akce pouze podle časové známky (a) a sjednocené kontinuální akce (b).	62

8.1	Vývojový diagram zpracování snímků aplikací.	63
8.2	Zjednodušený diagram třídy AppModel.	64
8.3	Segmentace a detekce slov z hlavičky feed příspěvku.	67
8.4	Výstup aplikace (a) a korespondující referenční řešení (b).	71
9.1	Confusion matrix detekce reklamních sdělení (transakcí).	74
9.2	Výsledky detekce atributů.	74
9.3	Počty TP, FN a FP detekcí interakcí.	75
9.4	Výsledky detekce transakcí pro reálná data.	78
9.5	Úspěšnost detekce interakcí pro reálná data.	79

Úvod

Sociální sítě jsou velmi populární fenomén dnešní doby. Díky rozsáhlému počtu uživatelů mají obrovský dosah na naše vnímání a myšlení, čehož využívají jak zakladatelé těchto sítí, tak firmy, které zde mohou propagovat své produkty a služby.

Uživatelé sociálních sítí mohou na příspěvky reagovat několika možnými způsoby, jakými jsou například „like“, komentář, sdílení, či proklik. Popsané reakce jsou následně využívány jako ukazatel zájmu konzumentů o reklamní sdělení nebo pro vytvoření personalizovaného obsahu. Ukazuje se, že nynější metody pro cílení reklamních sdělení nejsou příliš účinné, protože čím dál větší procento uživatelů jsou si daných metrik vědomi, a snaží se tak svoji aktivitu a chování na sociálních sítích skrýt. Následkem je zkreslení personalizovaného obsahu a výsledná marketingová neefektivita. Nový způsob analýzy interakcí s reklamním sdělením bere v potaz i tzv. „neviditelné“ uživatele s cílem zvýšit účinnost marketingové komunikace.

Hlavním cílem práce je vytvořit aplikaci, která z obrazových dat dokáže identifikovat jak viditelné, tak neviditelné interakce. Vstupem je video záznam obrazovky aplikace Instagram na platformě iOS a výstupem strukturovaný textový soubor, který popisuje detekovaná reklamní sdělení, dodatečné informace o reklamě a interakce uživatele.

Marketingový výzkum

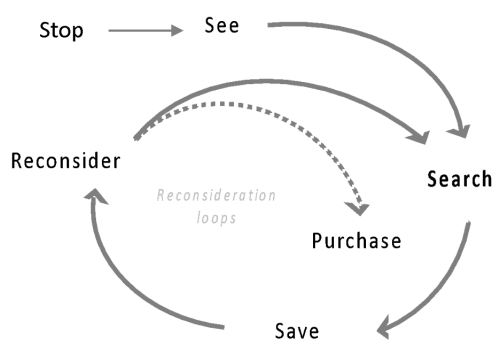
V posledních letech se digitální komunikace řadí mezi hlavní způsoby marketingové komunikace firem, které cílí jak na koncové zákazníky, tak i velkoobchodní partnery. O tomto tvrzení svědčí i fakt, že již nyní se na digitální komunikaci vynaloží přes 60 % výdajů spojených s reklamou a je očekáváno, že trend dále poroste [7].

V současné době jsou hlavními metrikami zájmu uživatele o reklamní sdělení dosah, prokliky, konverze a viditelné interakce, jako například „like“, okomentování nebo sdílení příspěvku (tzv. „viditelný engagement“). Popsané metriky jsou brány jako klíčový ukazatel popisující chování uživatele v digitálním prostředí, což má ale značné nevýhody [1].

Ukazuje se, že způsobem „viditelného engagementu“ na reklamní sdělení nereaguje většina uživatelů, konkrétně až 90 % [8, 9]. Následkem je, že analýzy, které jsou na zmíněných metrikách založeny, zkreslují reálné chování cílených uživatelů, což vede k vysoké míře neefektivity. Odhaduje se, že 80 % digitální komunikace je cíleno na irelevantní cílovou skupinu, téměř 50 % lidí začne značku ignorovat a až 38 % zákazníků od inzerované značky úplně odejde kvůli časté a nerelevantní digitální komunikaci [10, 11, 12].

Uživatelé, kteří se snaží svoji aktivitu před stávajícími metrikami schovat, jsou tzv. „neviditelní účastníci“. Neviditelný účastník je konzument obsahu v online prostředí, který s reklamním sdělením interaguje způsobem, aby viditelnými metrikami nebyl „zpozorován“. Následkem takového chování je ztráta podstatného množství dat, které popisují vztah inzerované značky (či produktu) a konzumenta reklamy [1].

Nový přístup, založený na konceptech „see-think-do-care“ a tzv. consumer decision journey, zohledňuje kromě měřitelných metrik i „neviditelné chování“ [13, 14, 1, 15]. Opírá se o tři základní skupiny interakce „see“, „search“ a „save“, které v pořadí odpovídají třem marketingovým cílům povědomí (awareness), zvažování (consideration) a zájem (purchase intent), viz obrázek 1.1.



Obrázek 1.1: Diagram chování zákazníka od povědomí, po zájem o koupi [1].

Aplikace Instagram a prezentace reklamních sdělení

Instagram je bezplatná sociální síť dostupná ve formě webové, iOS, či Android aplikace, která umožňuje snadné sdílení fotek a videí mezi přáteli nebo komunitou [16]. Počet uživatelů Instagramu celosvětově přesáhl jednu miliardu již v roce 2020 a stále roste, čímž se řadí mezi prvních pět nejpoužívanějších sociálních sítí [17, 18].

Díky dosahu a vlivu platformy, a sociálních sítí obecně, se hojně využívá pro digitální marketingovou komunikaci prostřednictvím reklamních sdělení. Reklamní sdělení lze definovat jako příspěvek, který propaguje jakýkoliv typ produktu (oblečení, elektronika, služba, apod.).

2.1 Lokace

Jelikož se Instagram řadí mezi přední sociální sítě vůbec, přirozeně nabízí uživatelům spoustu lokací a funkcionalit. Pro účely práce jsou popsány pouze lokace, které mají spojitost s marketingovou komunikací a interakcemi s nimi.

2.1.1 Story

Prvním formátem, který Instagram nabízí jsou stories, viz obrázek 2.2a. Zpravidla se jedná o krátká videa, nebo fotografie, které jsou zobrazeny pouze pár sekund a po uplynuté době se automaticky přejde na nové story, pokud ho uživatel sám nezastaví podržením obrazovky. Doba zobrazení a počet příspěvků od jednoho uživatele určuje posuvná lišta v horní části obrazovky. Obsah vydrží na sociální síti pouze po omezenou dobu, kdy se na něj uživatel může podívat, pak se automaticky vymaže.

Uživatel má možnost příspěvek „olajkovat“ a sdílet. Příspěvky zde neobsahují přímo sekci s popiskem, ten se nachází buď přímo ve fotografii či videu nebo po prokliku v modálním okně.

2.1.2 Reels

Dalším formátem prezentace příspěvků jsou reels, viz obrázek 2.2b. Jedná se o vertikálně posuvné příspěvky, které jsou zobrazeny přes celou obrazovku, tedy vždy je zobrazen pouze jeden, obdobně jako u stories. Reels příspěvek má vždy pouze jedno video nebo fotografii, které není časově omezené. Pokud se jedná o video, přehrává se ve smyčce.

Možností reakcí jsou, kromě již zmíněných u story, okomentování, sledování profilu uživatele, který příspěvek sdílel a uložení na platformě. Opět je zde možnost pozastavit si příspěvek podržením obrazovky nebo rozšířit celý popis.

2.1.3 Feed

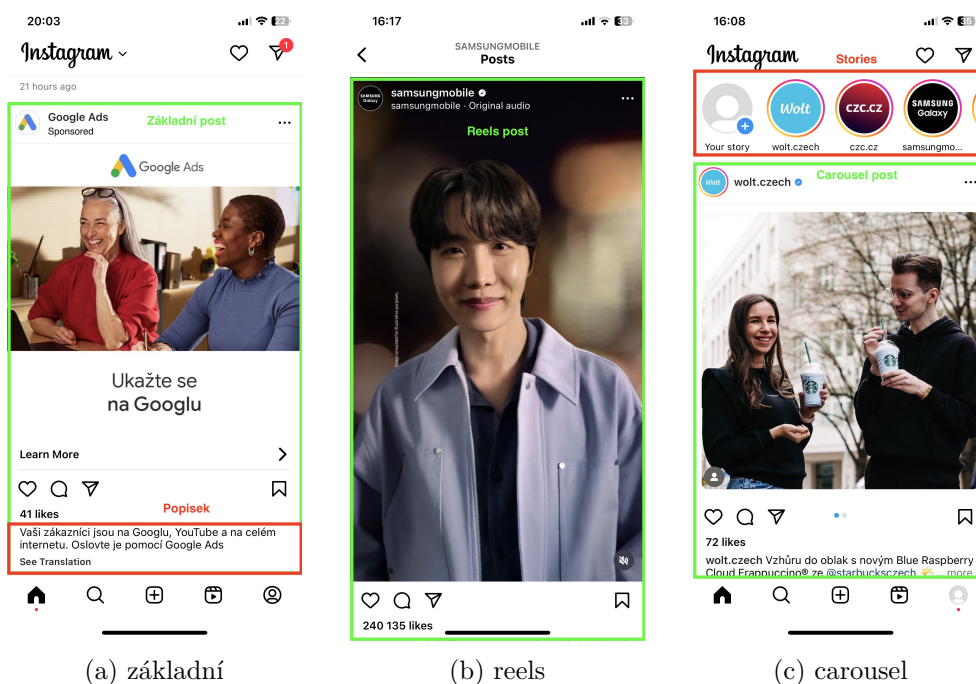
Feed lze přirovnat k vertikálně posuvné „nástěnce“, na které se vyskytují příspěvky, které jsou persistentní, stejně jako u reels. Oproti předchozím lokacím zde však není pravidlo, že je na obrazovce pouze jeden příspěvek. Ty navíc nejsou uniformní a lze je roztřídit do následujících kategorií, viz obrázek 2.1:

- základní – pouze fotka či video
- reels – reels příspěvek zobrazený na feedu
- carousel – série fotek a videí
- kolekce – více fotek v jednom příspěvku

Kromě zmíněných akcí u reels, jsou možnosti uživatele rozšířeny o přiblížení fotografie či videa nebo horizontální posun obsahu, pokud je příspěvek typu carousel. Oproti stories nebo reels nelze příspěvek pozastavit.

V aplikaci je hned několik obrazovek, které mají stejné charakteristiky jako hlavní feed obrazovka, ale slouží k jinému účelu:

- hlavní stránka – všechny příspěvky sledovaných i nových uživatelů doporučených platformou
- pouze sledovaní uživatelé
- oblíbení uživatelé – oblíbení a sledovaní
- feed na profilu uživatele – příspěvky daného instagramového profilu
- explore – příspěvky pouze uživatelů doporučených platformou



Obrázek 2.1: Ukázka lokace feed a typů příspěvků.

2.1.4 Profil

Instagramový profil reprezentuje uživatele nebo firmu (značku) a uvádí obecné informace jako počet sledujících a příspěvků společně s profilovým obrázkem. Profil spadá do jedné z následujících dvou kategorií [19]:

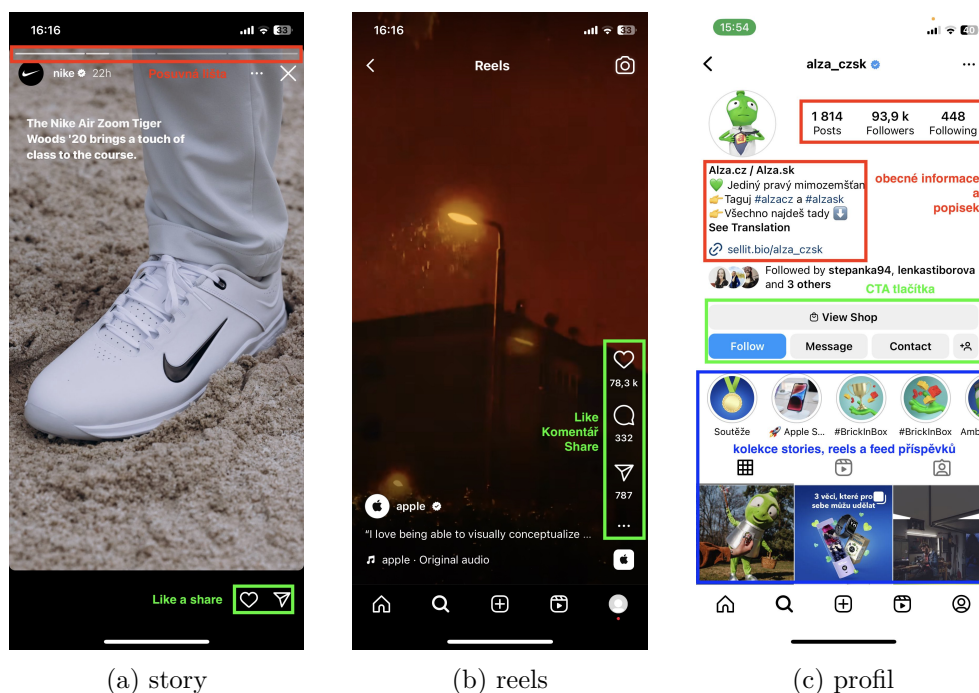
- osobní
- profesionální – influencer či podnikatelský

Osobní profil nemusí obsahovat vůbec žádné další informace. Navíc je možné ho nastavit jako soukromý, což znamená, že ostatní uživatelé, kteří účet nesledují, uvidí pouze obecné informace.

Profesionální profil je vždy veřejný a zpravidla obsahuje navíc vlastní popis s dodatečnými informacemi a odkazy na stránky nebo profily na jiných sociálních sítích. Volitelně také „Call To Action“ (dále CTA) tlačítko, které je popsáno v podsekcí 2.1.5.

Na všech profilech je poté seznam všech story, reels nebo feed příspěvků, pokud daný uživatel nějaký zveřejnil a samozřejmě tlačítko pro sledování uživatele. Ukázku lze vidět na obrázku 2.2c.

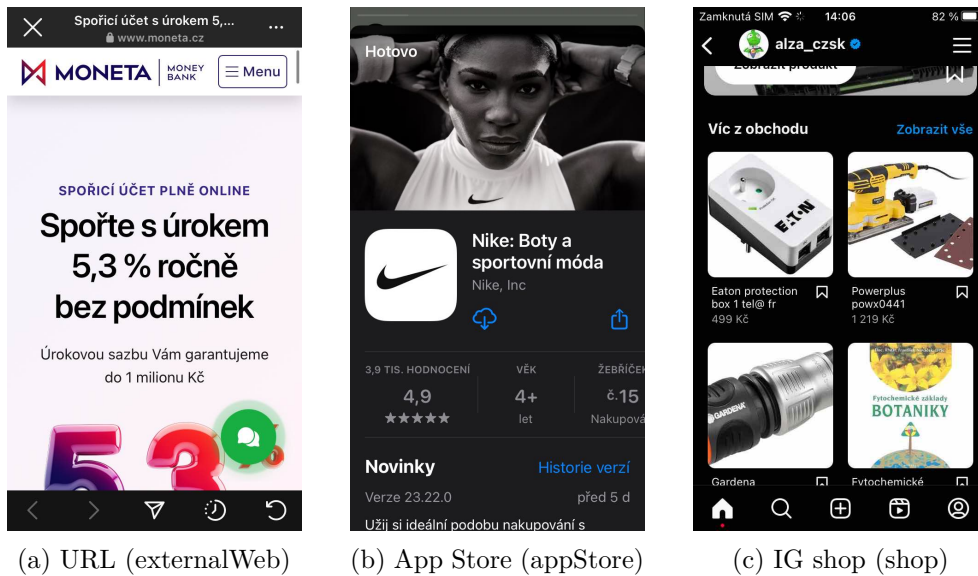
2. APLIKACE INSTAGRAM A PREZENTACE REKLAMNÍCH SDĚLENÍ



Obrázek 2.2: Ukázka lokací story (a), reels (b) a profesionální profil (c).

2.1.5 Lokace po prokliku CTA tlačítka

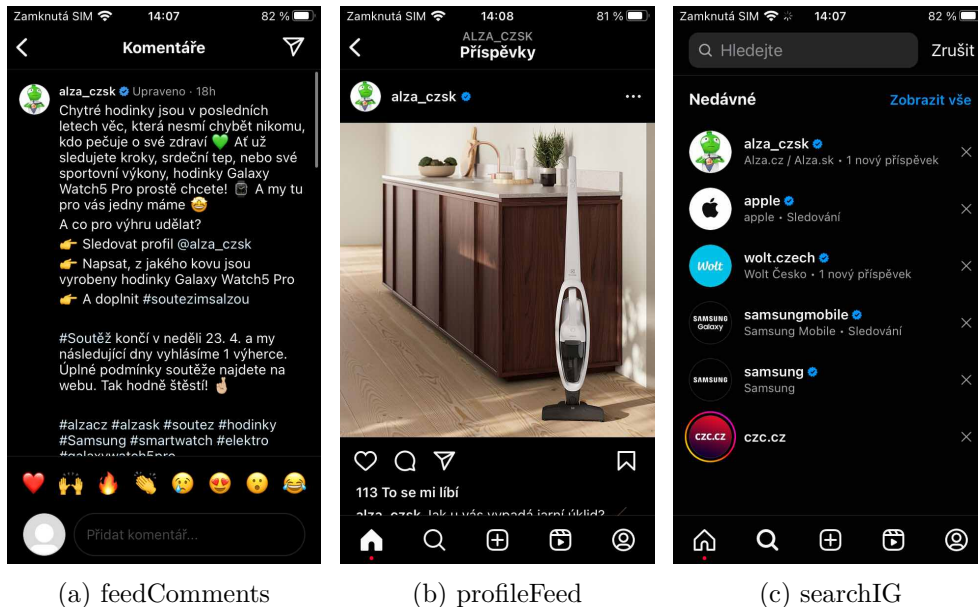
CTA je speciální tlačítko, které uživatele odkáže na nastavený URL link, Instagramový profil, Shop, nebo deep link do aplikace třetí strany (typicky App Store, rezervace nebo objednávka jídla) [20, 21]. Jedná se o funkcionalitu, která je dostupná pouze pro profesionální typy účtů. Na obrázku 2.3 jsou vidět nejčastější příklady externích lokalit po prokliku.



Obrázek 2.3: Ukázka odkazovaných stránek při prokliku CTA tlačítka.

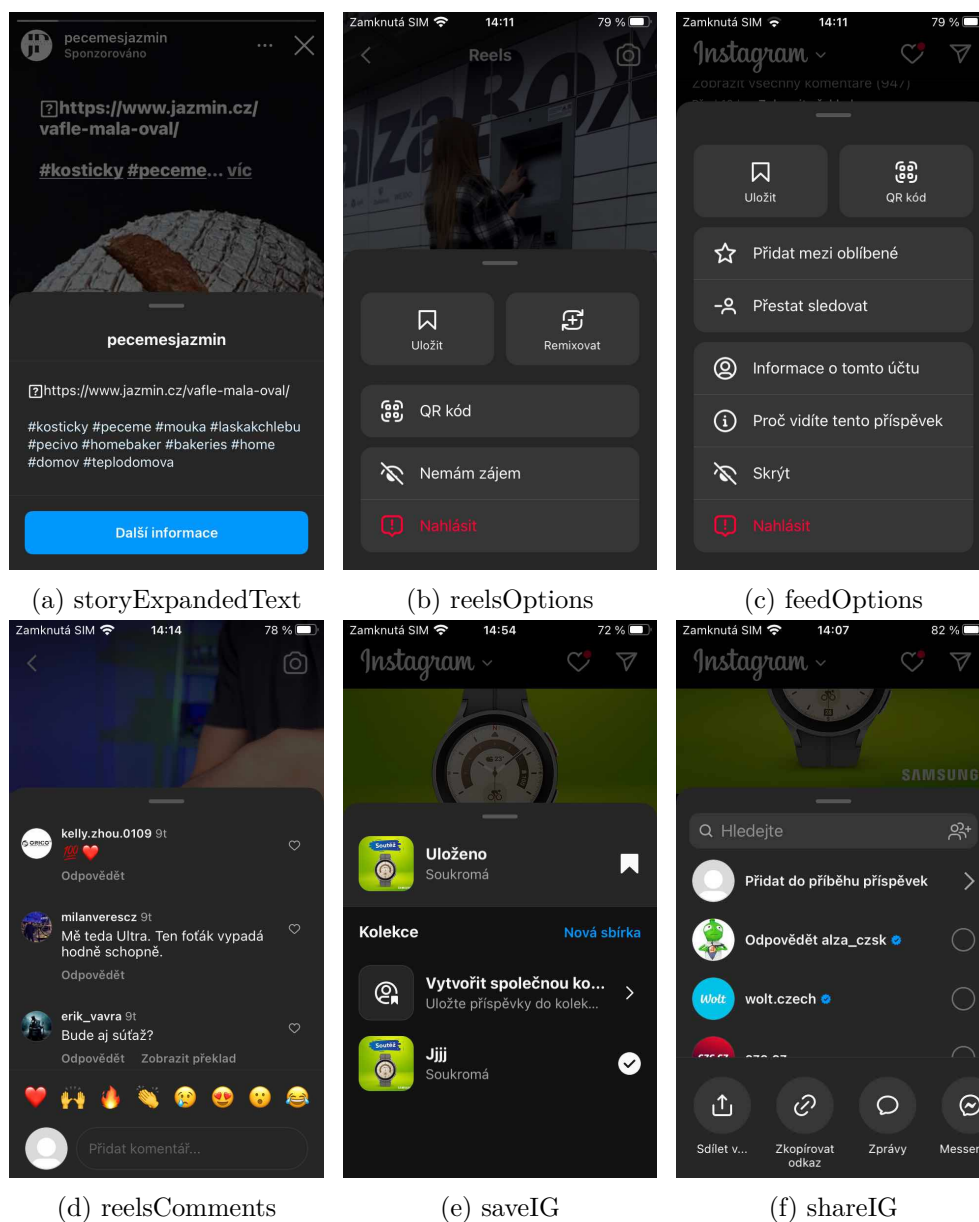
2.1.6 Ostatní

Zde jsou vyobrazeny vedlejší lokality, které ještě nebyly zmíněny.



Obrázek 2.4: Klasické obrazovky. Komentáře na lokaci feed (a), feed příspěvky na profilu (b) a obrazovka pro hledání profilů (c).

2. APLIKACE INSTAGRAM A PREZENTACE REKLAMNÍCH SDĚLENÍ



Obrázek 2.5: Modální okna. Rozšířený popis na story (a), možnosti na reels a feed (b, c), komentáře na reels (d), uložení příspěvku (e) a sdílení příspěvku (f).

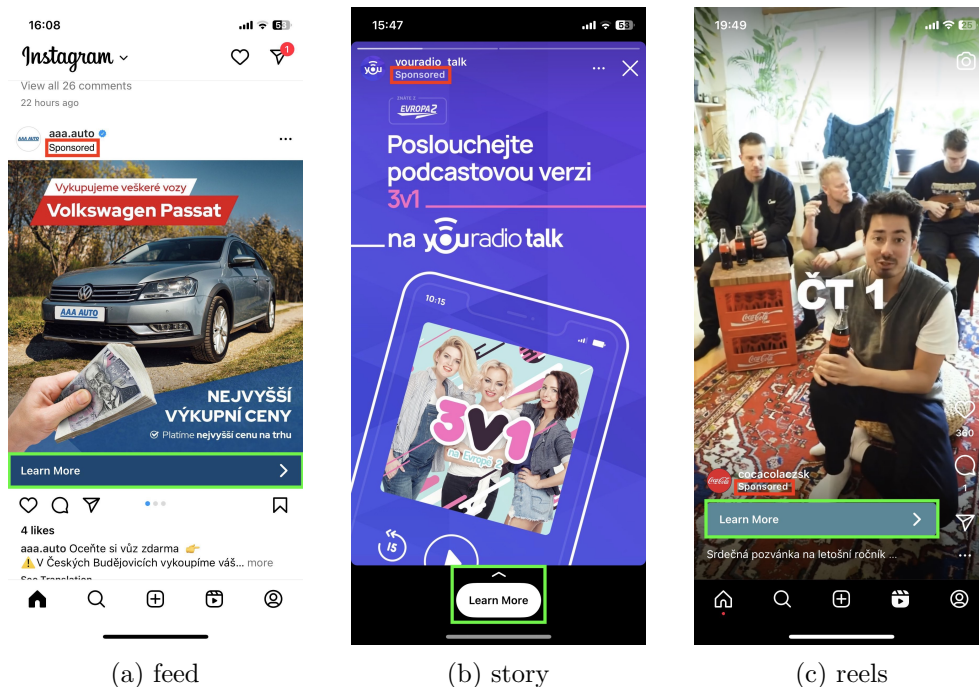
2.2 Typy reklamních sdělení

Marketingový příspěvek má velmi podobné vlastnosti jako příspěvek standardní. Jedná se buď o fotografii nebo video, uživatel je schopný na příspěvek reagovat komentářem, „likem“ či sdílením. Může se vyskytovat pouze na lokacích jako příspěvek, tedy story, reels nebo feed. Obecně je lze rozdělit do tří následujících kategorií.

2.2.1 Sponzorovaný

Jedná se o placenou službu zprostředkovanou přímo platformou Instagram, popřípadě platformou Facebook, dostupnou všem uživatelům s profesionálním účtem. Pomocí nástrojů jako Instagram Boost nebo Facebook Add Manager lze vytvořit příspěvek, který se na základě navolených parametrů od inzerenta reklamního sdělení a peněžní alokace na propagaci, zobrazí potenciálním zákazníkům, které sama platforma určí [22, 23, 24, 25].

Všechny příspěvky, které jsou propagovány tímto způsobem, jsou označené popiskem „Sponzorováno“ (resp. „Sponsored“), viz 2.6, a zároveň mohou obsahovat CTA tlačítko [26].



Obrázek 2.6: Sponsored reklamní sdělení na různých lokacích.

2. APLIKACE INSTAGRAM A PREZENTACE REKLAMNÍCH SDĚLENÍ

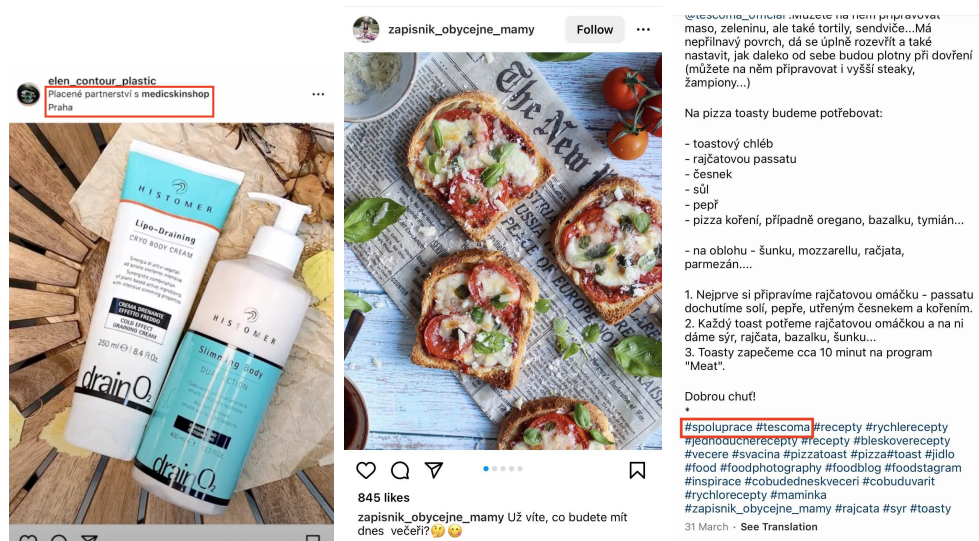
2.2.2 Placená partnerství

Mezi formu digitální marketingové komunikace spadá i spolupráce s různými internetovými influencery nebo celebritami. Instagram pro tento případ nabízí funkci placeného partnerství, která by se podle pravidel platformy měla při placené spolupráci inzerenta a propagátora využít [27, 28]. Příspěvky jsou poté označeny popisem „Placené partnerství“ (resp. „Paid partnership“), viz 2.7a.

2.2.3 Organický

Do kategorie spadají všechny příspěvky, které nejsou explicitně označeny jako reklamní sdělení. Je časté, že influenceři zmiňují spolupráci s danou značkou nebo firmou pouze v popisku pomocí „hashtagů“ a většinou na konci. Typickými označeními jsou například „#spoluprace“ nebo „#ad“ a uživatelské jméno účtu společnosti, která reklamní sdělení financuje. Pokud je popis dlouhý, uživatel ho musí rozkliknout, aby se zobrazil celý, viz obrázek 2.7b.

Dále do kategorie patří příspěvky firem samotných, které uživatel již sleduje. Ačkoliv je v drtivé většině jejich účel propagace vlastních produktů nebo služeb, nejsou označeny ani jednou ze zmíněných značek, jelikož se nejedná o propagaci zajištěnou sociální sítí samotnou. Výjimkou může být přítomnost CTA tlačítek, které napovídají, že se jedná reklamní sdělení.



(a) Placené partnerství

(b) Organický příspěvek

Obrázek 2.7: Ukázka příspěvku placené partnerství (a) a označeného reklamního sdělení na konci popisku (b).

Definice interakcí

Interakce uživatele s reklamním sdělením lze podle charakteristik kategorizovat do následujících reakčních skupin [29].

3.1 See

Veškeré interakce reakční skupiny see odpovídají marketingovému cíli „awareness“, viz kapitola 1. Uživatel s reklamním sdělením manipuluje pouhým pozorováním, což splňují následující interakce.

Stop

Uživatel se zastavil na reklamním sdělení.

BrowseImages

Uživatel prohlíží více obrázků či videí (horizontální posouvání) u jednoho příspěvku, což se může stát jen na lokaci feed u příspěvku typu carousel, viz obrázek 2.1c.

ExpandText

Interakce je možná tehdy, když příspěvek obsahuje popisek. ExpandText poté odpovídá rozkliknutí a zvětšení popisku.

WatchWhole

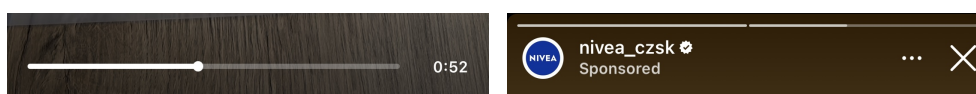
Značí přehrání celého videa a definice se liší v závislosti na lokaci.

Na feed je podmíněna typem příspěvku. Pokud se jedná o příspěvek typu reels, viz obrázek 2.1b, zobrazí se tlačítko, které nabízí přehrání videa znovu. Ostatní typy příspěvků přehrávají obsah ve smyčce bez žádného explicitního značení.

3. DEFINICE INTERAKCÍ

Lokace reels může mít dva typy v závislosti na délce videa. Krátké video se přehrává ve smyčce bez jakéhokoliv značení, dlouhé má ve spodní části obrazovky posuvnou lištu (progress bar), viz obrázek 3.1a, značíci jeho délku a umístění ve videu. WatchWhole poté znamená, že je progress bar celý vyplněný.

Nakonec u lokace story je interakce podmíněna horní posuvnou lištou, která signalizuje počet příspěvků a dobu zobrazení, viz obrázek 3.1b. Pokud je na story pouze jeden příspěvek, význam WatchWhole je stejný jako u dlouhého videa na reels. V případě, že je na story pod jedním účtem více než jeden příspěvek, interakce je validní, pokud si uživatel zobrazí poslední z nich.



(a) reels

(b) story

Obrázek 3.1: Progress bar na reels a story.

Zoom

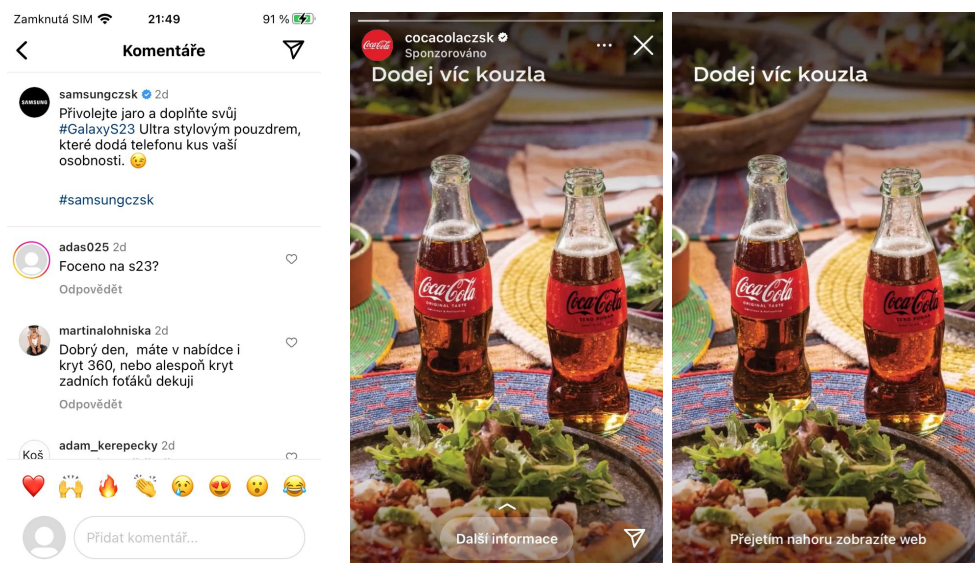
Uživatel si přiblížil fotku či video na příspěvku, což je možné pouze u příspěvku na feed.

Comment

Uživatel navštívil obrazovku s komentáři u příspěvku, viz obrázek 3.2a. Interakce je možná pouze pro lokace feed a reels.

Pause

Uživatel pozastavil obsah na reels nebo story. Při pozastavení jsou schovány veškeré ikony a popisky u příspěvku a je zobrazen pouze samotný obsah, viz obrázek 3.2b.



(a) See–Comment

(b) Pause

Obrázek 3.2: See–Comment na feed (a) a See–Pause na story (b).

3.2 Search

Search reakční skupina odpovídá marketingovému cíli „consideration“, viz kapitola 1 a spadají pod ní následující interakce.

Search

Uživatel vyhledal instagramový profil přímo v aplikaci Instagram. Tedy jedná se pouze o vyhledávání v rámci platformy.

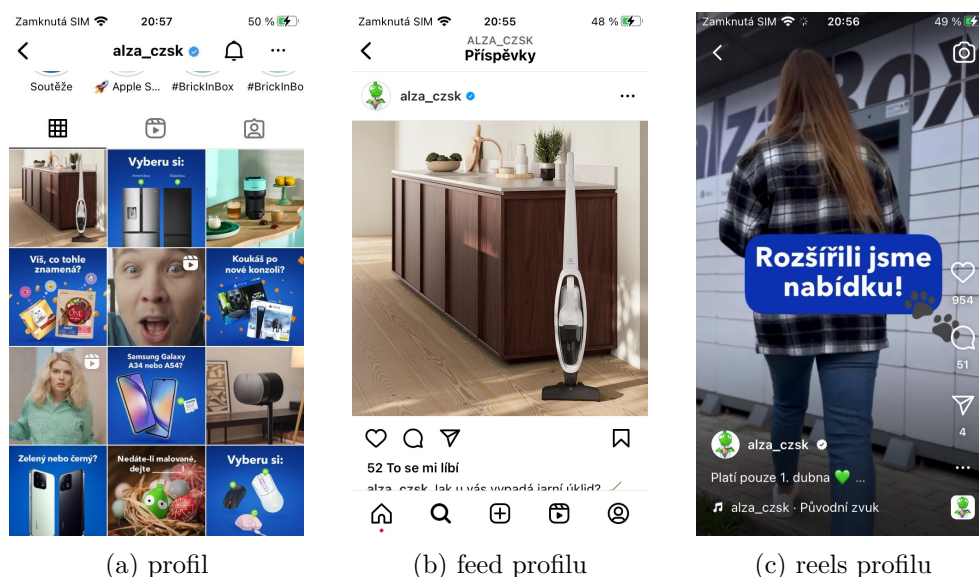
BrowseProfile

Uživatel si prohlíží profil, což platí i na korespondující příspěvky daného profilu. Tedy pokud je uživatel na lokaci profil a přejde na profileFeed, která obsahuje pouze příspěvky daného profilu, pořád se jedná o stejnou interakci. To samé platí pro korespondující story a reels, ukázkou lze vidět na obrázku 3.3.

Redirect

Uživatel kliknul na link v popisku příspěvku, či na profilu nebo na CTA tlačítko. To způsobí, že je přesměrován na nastavený URL link a může se ocitnout na nějaké z externích obrazovek, viz podsekcce 2.1.5. Redirect interakce se nevztahuje na lokaci shop, která je definována speciální reakční skupinou.

3. DEFINICE INTERAKCÍ



Obrázek 3.3: BrowseProfile interakce.

3.3 Save

Poslední skupina, která koresponduje s marketingovým cílem „purchase intent“, viz kapitola 1, obsahuje následující.

Screenshot

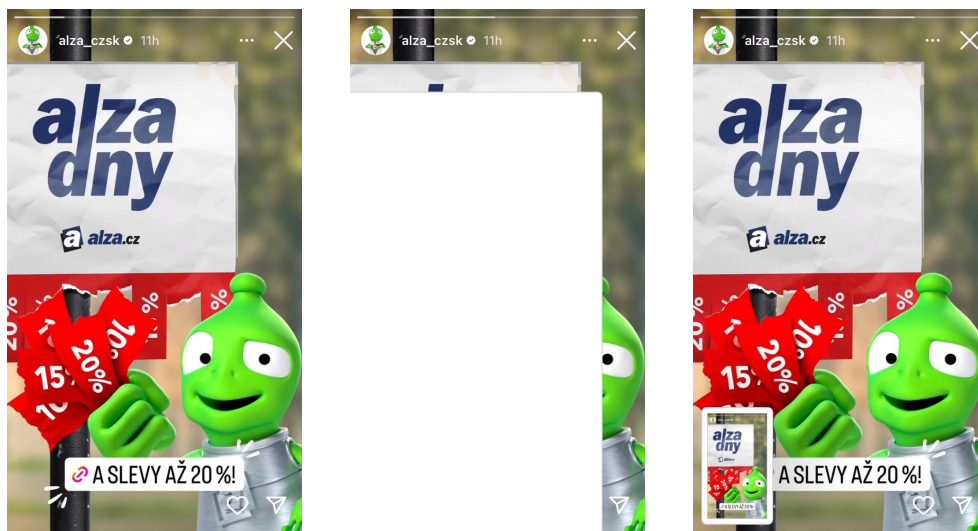
Uživatel vytvořil snímek obrazovky pomocí standardních systémových kontrolních prvků. Ukázka pro operační systém iOS je vidět na obrázku 3.4

SavePostIG

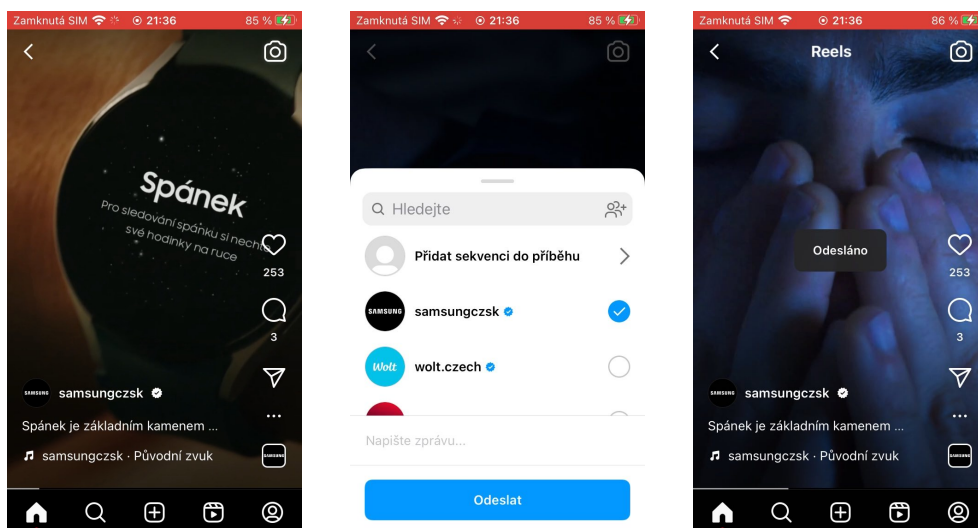
Instagram nabízí možnost uložení příspěvku do knihovny přímo v aplikaci, viz obrázek 3.6, ale pouze v případě, že se nejedná o story nebo sponzorovaný typ. Funkcionalitu lze provést buď pomocí save tlačítka u příspěvku, nebo přes tlačítko more, které zobrazí modální okno (reelsOptions nebo feedOptions), kde se vyskytují další možnosti, včetně uložení. Pokud má uživatel uložené příspěvky roztříděné do kolekcí, zobrazí se poté navíc modální okno saveIG, které dává možnost výběru kolekce, kam příspěvek uložit.

ShareIG

Interakce ShareIG odpovídá sdílení příspěvku v rámci platformy Instagram. Jedná se o vícekový scénář, kdy uživatel nejdříve klikne na tlačítko sdílení, poté vybere účty, kterým chce příspěvek sdílet a nakonec svůj výběr potvrdí tlačítkem odeslat. Celá interakce je zachycena na obrázku 3.5.



Obrázek 3.4: Screenshot na iOS 16.



Obrázek 3.5: ShareIG na lokaci reels.

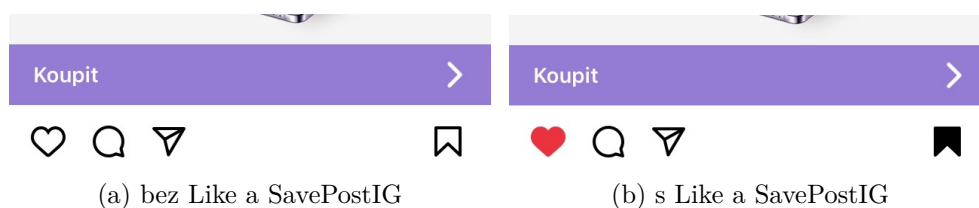
3.4 Social

Skupina Social koresponduje s viditelnými akcemi, kterými uživatel interaguje s příspěvkem.

Like

Kladné ohodnocení příspěvku, které se projevuje zčervenáním srdíčka, viz obrázek 3.6.

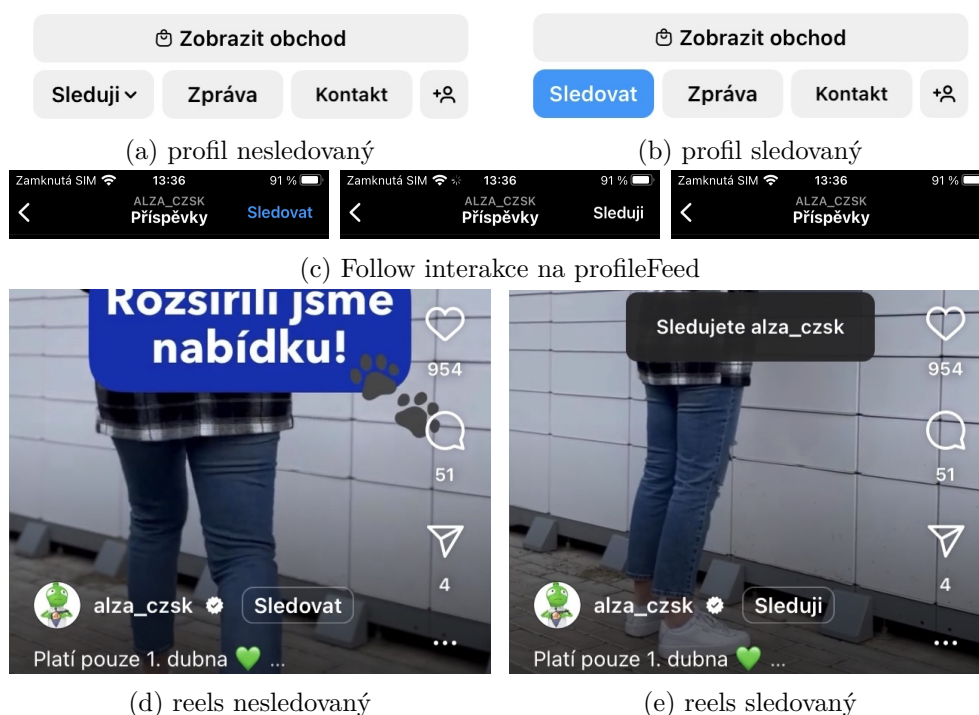
3. DEFINICE INTERAKCÍ



Obrázek 3.6: Like a SavePostIG u příspěvku na feed.

Follow

Pokud se uživateli příspěvku daného profilu líbí, může využít možnosti sledování (neboli Follow). Interakci lze učinit buď přímo na profilu, při prohlížení obsahu profilu, nebo u příspěvku na reels kliknutím na tlačítko, viz obrázek 3.7. Follow lze učinit i na feed, opět pomocí tlačítka, avšak zpětná vazba je pouhá změna popisku tlačítka bez upozornění, jako v případě reels.

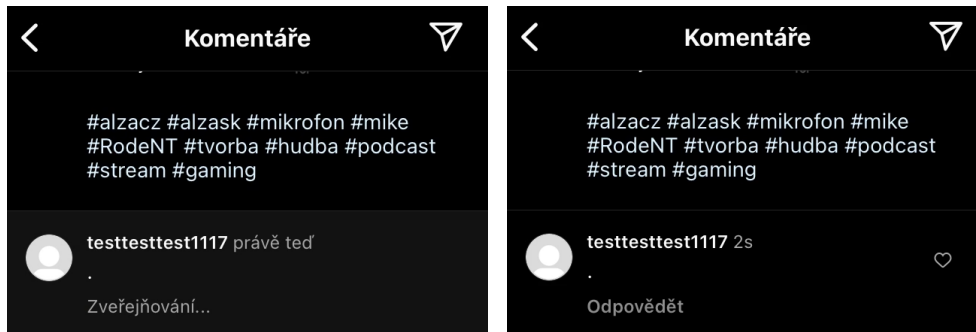


Obrázek 3.7: Follow interakce na profilu a reels.

Comment

Uživatel okomentuje příspěvek tím, že nejdříve přejde do sekce s komentáři, napíše nějakou zprávu a odešle. Nově přidaný komentář je na chvíli zvýrazněn

změnou pozadí, viz obrázek 3.8. Interakce se vztahuje pouze na příspěvky lokace feed nebo reels.



(a) právě odeslaný

(b) zveřejněný

Obrázek 3.8: Zveřejnění komentáře.

3.5 Shop a Skip

Poslední dvě reakční skupiny jsou speciální případy, kde každá obsahuje pouze jednu stejnojmennou interakci.

Shop

Uživatel navštívil lokaci shop, viz obrázek 2.3c.

Skip

Skip odpovídá stavu, kdy uživatel narazil na reklamní sdělení, ale neprovedl nad ním žádnou interakci popsanou výše. Výjimkou je Stop, která je implicitně vykonána pokaždé, kdy uživatel vidí reklamní sdělení.

Výzkum problematiky

Hlavním zdrojem zisku informací aktivit uživatelů jsou v dnešní době samotní poskytovatelé aplikací a služeb. Analýzy na základě získaných informací jsou dále poskytovány ve formě nástrojů, které umožňují snadné použití digitální inzerce na daných platformách, například Facebook ads nebo Google ads [24, 30].

Pro extrakci dat ve webovém prohlížeči lze využít například HTML kód, jako v případě [31], kde je popsán systém, který na základě navštívených webových stránek dokáže uložit a zpětně zobrazit přesnou posloupnost historie. Ačkoliv se již nejedná o aktuální problém, článek ukazuje na možnou formu zisku informací.

O něco pokročilejší přístup se objevuje v [32], kde je nad webový prohlížeč implementována HTTP proxy, která modifikuje zdrojový HTML kód přidáním JavaScript kódu ještě před zobrazením stránky uživateli. Nově implementované části poté slouží pro získání informací jako je vstup z klávesnice nebo pohyb myši.

V [9] je popsána analýza chování uživatelů na sociálních sítích ze sady HTTP relací, které byly získány pomocí externí webové stránky. Na základě datasetu bylo poté možné analyzovat jednotlivé interakce uživatelů s příspěvky a navštívené lokace.

Všechny přístupy popsány výše vyžadují externí aplikaci, která běží nad zařízením uživatele nebo přístup k datům, které patří společnostem vlastníci platformu. Apple má velmi striktní pravidla a běh aplikací na pozadí omezuje jen na pár základních funkcionalit, které jsou časově omezeny [33]. Navíc není možné naprogramovat nastavbu nad existující aplikaci, jelikož každá je operačním systémem vnímána jako samostatná entita. Z toho plyne, že zisk informací musí probíhat čistě z nahrávek.

4.1 Zisk informací ze záznamu obrazovky

Jednou z prací, která se zabývá problematikou zisku informací z nahrávek je [2]. Autoři popisují vývoj nástroje scvRipper, který za pomoci algoritmů počítačového vidění dokáže zaznamenat lokace a informace na obrazovce softwarového vývojáře v čase. Nejdříve jsou využitím metody image differencing z obrazovky vyfiltrovány screenshoty, které obsahují téměř totožný obsah. Následně jsou pomocí Cannyho detektoru, morfologických operací eroze a dilatace a shlukovacího algoritmu DBSCAN [34] nalezeny hranice jednotlivých kolonek a oken na obrazovce. Pro detekci druhu oken je použita kombinace FAST a SURF algoritmu, která nalezne klíčové body a následně referenční šablonu v obrázku [35, 36]. Jelikož na obrazovce může být více oken, je potřeba jednotlivé kolonky a okna sloučit pomocí normalizovaného min-max cut algoritmu [37]. Informace z vysegmentovaných oken jsou získány pomocí OCR.



Obrázek 4.1: Segmentace a mapování kolonek z [2].

Aplikace Waken [38] využívá diferencování obrázků ze záznamu obrazovky pro identifikaci kurzoru, jeho pohybu a ikon v aplikaci, bez nutnosti jejich předchozí znalosti.

[39] se zabývá segmentací a určením lokalit v mobilní aplikaci. Mobilní zařízení je součástí aparatury pro uživatelské testování, která navíc obsahuje kameru a zařízení pro detekci pohybu očí. Zpracovávané video je tedy z pohledu třetí osoby. Předzpracování zahrnuje manuální omezení počtu obrázků, segmentaci displeje a prstu použitím Otsuova prahování a detekci klávesnice pomocí SIFT [40]. Lokace jsou následně klasifikovány použitím SIFT a SSIM [41, 42]. Algoritmus SIFT se objevuje i v [43]. V kombinaci s OCR a indexováním představuje nástroj Sikuli Script, který dokáže v sadě obrázků nalézt

ikonku. Rozšiřující studie [44] popisuje nasazení Sikuli Scriptu v aplikaci Sikuli Test pro zjednodušené GUI testování definované pomocí obrázků. Definované obrázky od testerů se při samotném testování ve snímku hledají metodou template matching.

V [45] je popsána implementace převodu nahrávek android aplikací na zreprodukovatelné scénáře. Detekce se opírá o touch indikátor, což je systémová funkcionální umožňující grafické znázornění doteku uživatele na obrazovce. Kombinací dvou implementací konvolučních neuronových sítí (R-CNN a opacity CNN) se z obrázku rozpozná pozice a průhlednost indikátoru. Výstup detekce se následně vyfiltruje a přeformátuje do reprodukovatelných scénářů, které lze spustit přímo na zařízení. Podobnému principu se věnuje [46] s účelem automatizovaného vytvoření UI testů.

4.2 Metody rozpoznání interakcí

Interakce definované v kapitole 3 mají dané vizuální vlastnosti, podle kterých se pozná, zda proběhly či nikoliv. Ačkoliv jsou interakce vesměs rozdílné (někdy i pro specifickou lokalitu), lze uvažovat následující základní rozdělení.

4.2.1 Pohybové interakce

Zde spadají interakce, které v definici obsahují nějaký pohyb, tedy Expand-Text, BrowseImages a některé případy WatchWhole, které zahrnují progress bar.

Mezi obvyklé metody sledování pohybu objektů patří background subtraction, která staví na inicializaci pozadí ze sekvence obrázků nebo referenci a porovnávání dalších snímků s pozadím [47]. Tím vznikne vysegmentovaná část, která reprezentuje pouze pohybující se objekt. Metoda nachází využití například pro detekci pohybu aut nebo osob při použití statické kamery [48, 49, 50, 51]. Nevýhodou je citlivost na dynamické změny ve snímku [52]. Bylo pro ni navrženo spoustu optimalizací a variant, například pro větší odolnost vůči stínům, proměnlivému pozadí nebo pro zvýšení rychlosti, ačkoliv všechny cílí na výpočet v reálném čase [53].

Dalším používaným přístupem je optical flow, která sleduje pohyb stejných bodů mezi snímky, z kterých lze určit směr a vzdálenost. Ačkoliv lze metodu různými předpoklady či úpravami použít i s pohyblivou kamerou [54], je velmi výpočetně náročná a náchylná na šum [47]. Použití optical flow lze najít například pro detekci pohybu očí nebo dopravy [55, 56].

Frame differencing detekuje pohyb objektu odečtením nynějšího snímku od předchozího, což už naznačuje, že metoda není vhodná, pokud jsou snímky velmi proměnlivé a kromě objektu se hýbe i pozadí [47, 54]. Temporal differencing má snahu tuto nevýhodu vyřešit použitím více než jednoho snímku do minulosti. Ačkoliv je metoda odolnější oproti dynamickým změnám v obrázku,

výsledný objekt obecně není ideálně vysegmentován (maska například obsahuje díry) [52].

4.2.2 Statické interakce

Like a Save interakce jsou definovány speciální ikonou. Pro nalezení ikony lze využít algoritmy jako SURF, SIFT nebo template matching, které se při zpracování nahrávky obrazovky hojně využívají [2, 39, 43, 44], viz sekce 4.1.

Interakce jako je Follow, ShareIG a Screenshot (popřípadě i Like) lze identifikovat na základě barev a tvarů UI prvků. Standardně využívané metody segmentace jsou například adaptivní a manuální prahování, shlukování nebo hranová detekce [57, 58]. Hranovou detekci Cannyho detektorem nebo Houghovu transformaci, lze využít pro určení jednoduchých tvarů [59, 60]. Využití Cannyho detektoru pro detekci tvarů v mobilní aplikaci je popsána v [61].

Zbývající interakce jako například Pause a WatchWhole jsou definovány podobností snímků na obrazovce. Možný přístup je použití SSIM indexu nebo image differencingu jako v [2, 39], kde slouží pro filtraci neměnných sekvencí obrázků ve videu. Místo SSIM lze použít i jiné metriky popisující rozdílnost snímků, například FSIM, MSE nebo PSNR [42].

Teoretické zázemí

K detailnějšímu pochopení problematiky je potřeba si přiblížit základní pojmy komprese videa, vybrané metody zpracování obrazu a technologii OCR, čemuž se věnuje následující kapitola.

5.1 Terminologie komprese videa

Video může mít spoustu kompresních formátů v závislosti na požadavcích, které jsou na něj kladeny.

5.1.1 Parametry komprese

Každé video, se kterým běžný uživatel přijde do styku, je v nějaké formě zkomprimováno. Nehledě na sofistikovanosti a účelu kompresního formátu, každý z nich pracuje s následujícími parametry.

5.1.1.1 Bit rate

Bit rate lze chápat jako hodnotu v bitech, která určuje, do jakého detailu může software popsat snímek ve videu. Čím větší bit rate, tím lepší je kvalita obrázku, ale také větší paměťová náročnost. Rozlišují se dva základní typy [62]:

- Constant bit rate (CBR) – na každý snímek je použito stejné množství bitů
- Variable bit rate (VBR) – množství bitů se v průběhu času mění

5.1.1.2 Frame rate

Určuje frekvenci snímků, které jsou ve videu sekvenčně zobrazeny za sebou [62]. Častěji se vyskytuje ve formě frames per second (FPS), který určuje počet

po sobě jdoucích snímků, popisujících jednu sekundu videa. Opět se rozlišují dva základní typy:

- Constant frame rate (CFR) – každá sekunda videa má stejný počet snímků
- Variable frame rate (VFR) – počet snímků videa za sekundu se v průběhu může měnit

Hlavní výhodou VFR je možnost vynechání totožných snímků v sekvenci, čímž se zmenší paměťová náročnost.

5.1.2 Codec

Stěžejní částí komprese videa je codec, jehož název vznikl spojením slov coder a decoder. Jedná se většinou o kus softwaru, ale ne nutně, který zajišťuje kompresi a dekompresi videa. Každý codec má své výhody i nevýhody a jeho volba a použití se vždy liší od požadavků, které jsou kladeny (kvalita obrazu, velikost komprimovaného videa, rychlost komprese a dekomprese, ...) [62].

Nejrozšířenějším standardem je Advanced Video Coding (AVC), který je podporován na drtivě většině platform (Xbox, PlayStation, Blu-ray, ...), operačních systémů (Windows i macOS) a mobilních zařízeních. Pro identifikaci různých konfigurací používá profily a úrovně, které ovlivňují vlastnosti zakódovaného a dekodovaného videa [62].

Zde jsou důležité pouze kombinace úrovně L3.x a L4.x a novějšího profilu High, který oproti předchůdcům Baseline, Extended a Main podporuje více optimalizací a větší škálu videí. Úroveň definuje hranici výpočetní síly a paměti potřebnou pro implementaci pomocí parametrů jako bitrate a frame rate, viz tabulka 5.1. Čím menší je daná hranice, tím méně kvalitní, ale zároveň méně paměťově náročné, bude výsledné video [63].

Tabulka 5.1: Porovnání různých úrovní v AVC [63].

Úroveň	Typické rozlišení	Typický frame rate	Maximální bit rate
3.1	1280 × 720p	30	14 Mbps
3.2	1280 × 720p	60	20 Mbps
4	720p nebo 1080i	60p/30i	20 Mbps
4.1	720p nebo 1080i	60p/30i	50 Mbps
4.2	1920 × 1080p	60p	50 Mbps

5.2 Zpracování obrazu

Metody zpracování obrazu se opírají o vlastnosti snímku, na základě kterých lze získat užitečné informace, potlačit nechtěné nebo snímek opravit, či vylepšit.

5.2.1 Segmentace

Jedna ze stěžejních oblastí je segmentace, jelikož jejím úkolem je nalézt oblast zájmu v obrázku. Mezi nejjednodušší patří prahování, které na základě stanovené hranice přiřadí pixelu hodnotu 1, pokud je jeho jasová hodnota větší, jinak 0. Výsledkem je binární obrázek, kde pixely s nenulovou hodnotou určují vysegmentovanou část. V případě definování dvou hranic jsou vysegmentovány ty pixely, které spadají do hranicemi určeného intervalu. Prahování se běžně používá jak pro šedotónové, tak barevné snímky. Barevný obrázek je vhodné nejdříve převést ze standardního RGB barevného prostoru na HSV, který pixel reprezentuje vektorem, jehož hodnoty značí barvu, míru hustoty a světlosti. Výhodou HSV je jednodušší reprezentace, jelikož barva je zastoupena pouze jednou hodnotou [64].

Sofistikovanější je Otsuova metoda, která dokáže na základě světelných hodnot pixelů šedotónového obrázku odvodit hranici automaticky. Metoda předpokládá, že v obrázku jsou dvě třídy pixelů, které popisují popředí (světlá oblast) a pozadí (tmavá oblast). Minimalizací rozptylu mezi pixely stejné třídy a maximalizací rozptylu mezi pixely rozdílných tříd lze docílit nalezení optimální hranice, která nejlépe rozděluje popředí od pozadí [5].

5.2.2 Background subtraction

Background subtraction je hojně využívaná technika pro segmentaci popředí od pozadí z video záznamu statické kamery. Metoda nejdříve provede inicializaci modelu pozadí. Rozdíl mezi aktuálním snímkem a předpokládaným modelem se získá porovnáním každého pixelu aktuálního snímku s předpokládanou barevnou mapou pozadí. V případě, že rozdíl mezi barvami je větší než stanovená hranice, pixel je považován za popředí. Model pozadí musí být průběžně aktualizován, aby se adaptoval změnám v obraze. Různé přístupy se poté liší ve volbě modelu inicializace pozadí a algoritmu určení hranice pro rozdělení pixelů na popředí a pozadí [47, 53].

Jedním z používaných přístupů je Mixture of Gaussians (MOG), který předpokládá, že pravděpodobnost výskytu určité hodnoty pixelu v čase t je popsána K průměry Gaussovských distribucí, kde K je parametr určený uživatelem (běžně mezi 3 a 5). Čím větší a kompaktnější je distribuce, tím více je pravděpodobné, že pixel reprezentuje pozadí. Finální klasifikace poté závisí na manuálně stanovené hranici T [53].



(a) původní obrázek (b) maska pohyblivých objektů

Obrázek 5.1: Ukázka MOG background subtractoru [3].

5.2.3 Filtrace

Účelem filtrace v počítačovém vidění je zejména potlačení šumu nebo očištění snímku od nežádoucích vlastností. Filtr je definován maticí (jádro) s lichými rozměry, kterou lze vnímat jako posuvné okno nad originálním snímek. Všechny hodnoty filtru $H(i, j)$ jsou pronásobeny s korespondujícími hodnotami v originálním snímku $I(u + i, v + j)$ a poté sečteny. Proces lze vyjádřit vztahem pro lineární konvoluci [65]:

$$I'(u, v) \leftarrow \sum_{(i,j) \in R_H} I(u + i, v + j) \cdot H(i, j) \quad (5.1)$$

Populární je Gaussův filtr, který pro vytvoření jádra používá Gausovu funkci. Je velmi efektivní pro eliminaci šumu a lze ho popsat následujícím vzorcem [65]:

$$G_\sigma(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (5.2)$$

5.2.4 Hranová detekce

Hrany v obraze jsou reprezentovány velkým rozdílem hodnot sousedních pixelů a mohou být užitečné pro identifikaci objektů nebo segmentaci ve snímku. Oblasti největší difference lze zjistit pomocí gradientu. Jelikož má ale snímek diskrétní hodnoty, používá se lineární konvoluce v kombinaci s filtry, které výpočet derivací aproximují. Příkladem aproximace první derivace ve vertikálním směru může být filtr H_{Px} Prewittové nebo horizontálním směru Sobelův operátor H_{Sy} [66].

$$H_{Px} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad H_{Sy} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (5.3)$$

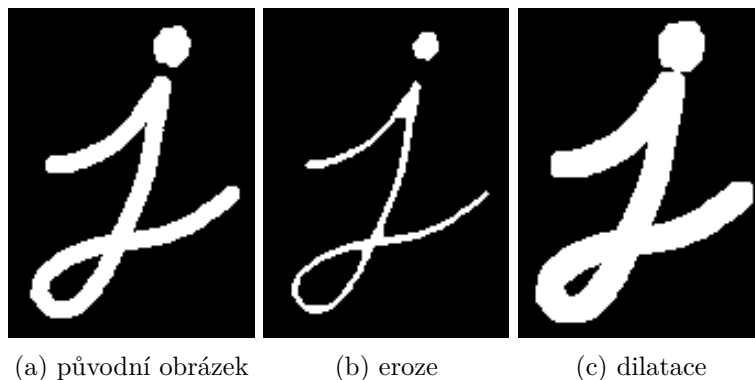
Popsané filtry jsou náchylné na falešné detekce způsobené šumem. Komplexnější a populární přístup je Cannyho detektor, jehož detekce hran sestává

ze čtyř kroků. Obraz je nejdříve vyhlazen Gaussovým filtrem. Následně je aplikován gradientní filtr pro výpočet aproximací derivací použitím Sobelova operátoru. Třetím krokem je nonmaximum suppression, který potlačí body, které nejsou největší v nějakém okolí, což pomáhá ztenčit hrany a zachovat pouze nejvíce významné. Posledním krokem je dvojité prahování. Body ležící nad oběma hodnotami prahů jsou považovány jako silné hrany. Pixely, které leží uvnitř intervalu definovaného hranicemi, se klasifikují jako hrana, pokud v blízkém okolí leží bod, který byl určen jako silná hrana [59, 67].

5.2.5 Binární morfologické operace

Výsledkem segmentace obrazu je binární maska, kterou je většinou potřeba analyzovat a dále upravovat. Binární morfologické operace fungují na principu lineární konvoluce nad binární maskou, při kterém jsou aplikovány logické operace namísto aritmetických. Jádro se nazývá strukturní element a jeho střed kotva. Mezi základní patří operace eroze (logický AND) a dilatace (logický OR). Jak lze vidět na obrázku 5.2, operace eroze binární objekt zúžuje, jelikož jsou zachovány pouze ty pixely pod kotvou, jejichž okolí odpovídá tvaru strukturního elementu. Operace dilatace na druhé straně binární masku zvětšuje. Pokud se pod kotvou vyskytne nenulová hodnota, zkopíruje se celý strukturní element do výsledného obrázku na dané místo [68, 67].

Kombinací dvou základních operací lze vytvořit operace další. Operace otevření provede nejdříve erozi a poté dilataci, čímž se lze zbavit například vypouklých částí masky nebo šumu bez změny velikosti masky. U operace zavření je nejdříve aplikována dilatace a následně eroze, což lze využít například pro zaplnění chybějících míst v masce [68].



(a) původní obrázek

(b) eroze

(c) dilatace

Obrázek 5.2: Ukázka morfologických operací eroze (b) a dilatace (c) na binární masce (a) [4].

5.2.6 Template matching

Template matching (dále TM) je metoda pro hledání lokace referenční šablony v obrázku za použití 2D korelace (cross korelace). Cross korelaci r_{xy} mezi snímkem $x(m, n)$ a $y(m, n)$, kde m značí výšku a n šířku snímků, lze vyjádřit vztahem [67]:

$$r_{xy}(m, n) = \sum_k \sum_l x(k, l)y(k - m, l - n) \quad (5.4)$$

Ze vzorce vyplývá, že největší hodnoty budou ve snímku na pozici, kde jsou pixely nejsvětější, nehledě na šablonu. Řešením je použití normalizované cross korelace rn_{xy} [67].

$$rn_{xy}(m, n) = \frac{\sum_k \sum_l (x(k, l) - \bar{x}_l)(y(k - m, l - n) - \bar{y})}{\sqrt{\sum_k \sum_l (x(k, l) - \bar{x}_l)^2 \sum_k \sum_l (y(k - m, l - n) - \bar{y})^2}} \quad (5.5)$$

Referenční šablona je posouvána po obrázku a postupně jsou počítány normalizované cross korelace mezi šablonou a korespondujícími výřezy v obrázku. Vypočítané korelace jsou následně uloženy do korelační mapy, kde souřadnice pixelu s maximální hodnotou značí nejvíce podobnou oblast referenční šabloně [69].

Pro správné fungování je potřeba, aby objekt v šabloně velikostně odpovídal hledanému objektu ve snímku, jelikož korelace neřeší afinní transformace, ani škálování.

5.2.7 Kontury

Kontury lze chápat jako křivku spojující všechny body, které ohraničují nějaký objekt ve snímku. Jsou užitečným nástrojem pro detekci, rozpoznání objektů nebo určení tvarových charakteristik jako kulatost, vypouklost nebo poměr stran [70]. Pomocí kontur lze vypočítat i prostorové a centrální momenty [71].

5.2.8 Hough circles

Algoritmus Hough circles se používá pro detekci kruhů ve snímku a je založen na principu Houghovy transformace (dále HT). Předpokladem pro HT je popsání kontury objektu parametrickou rovnicí, v případě kruhu:

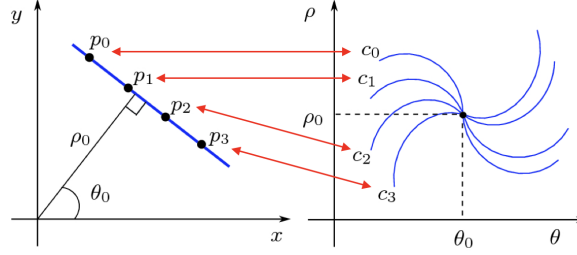
$$(x - a)^2 + (y - b)^2 = r^2 \quad (5.6)$$

Rovnice popisující tvar objektu se převede do prostoru parametrů, neboli akumulátoru — jednotlivé osy jsou reprezentovány příslušnými parametry (u kruhu se jedná o 3D prostor, jelikož má tři parametry a , b a r). Vstupem HT je binární obrázek z hranové detekce. Jednotlivé body z masky se převedou do akumulátoru, kde jsou reprezentovány křivkou. Body v akumulátoru, které

jsou protínány největším počtem křivek, formují lokální maxima a určují nalezené parametry popisující kruhy v původní souřadnicové soustavě [5].

Příklad převodu bodů do akumulátoru lze vidět na obrázku 5.3, kde je využito parametrického vyjádření přímky vzorcem [5]:

$$x \cdot \cos \delta + y \cdot \sin \delta = r \quad (5.7)$$



Obrázek 5.3: Převod a reprezentace bodů v akumulátoru [5].

5.2.9 SSIM index

Structural Similarity (SSIM) index je motivován faktem, že lidské oko je přizpůsobeno k extrahování strukturálních informací z vizuálních scén [72].

Pro porovnání dvou snímků se proto využívá tří podobnostních metrik. První dvě složky jsou jasová $l(\mathbf{x}, \mathbf{y})$ a kontrastní $c(\mathbf{x}, \mathbf{y})$ a jsou definovány vztahy:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (5.8)$$

kde μ_x a μ_y jsou průměrné jasové hodnoty obrázků x a y , σ_x a σ_y jsou směrodatné odchylky pixelových hodnot. C_1 a C_2 jsou konstanty pro vyvážení nestability, pokud je jmenovatel blízko nule. Třetí složka je strukturální $s(\mathbf{x}, \mathbf{y})$, která měří strukturální podobnost mezi dvěma snímky vztahem:

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \quad (5.9)$$

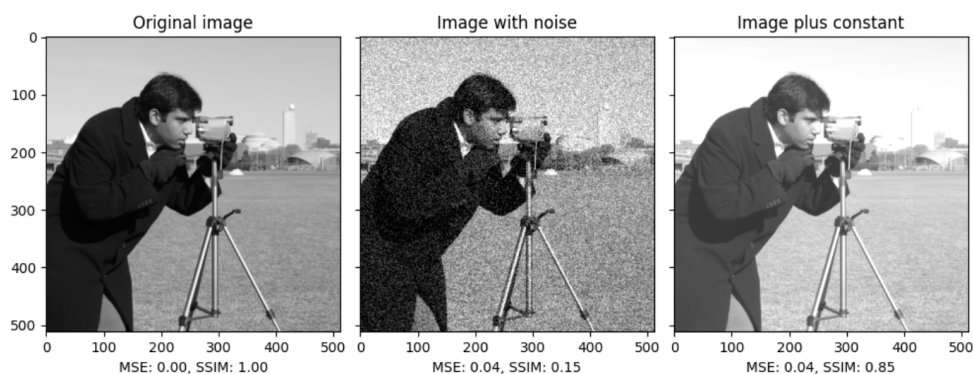
kde C_3 je opět malá konstanta a σ_{xy} lze estimovat vztahem:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y). \quad (5.10)$$

SSIM index se následně vypočítá pronásobením všech složek, které jsou váženy koeficienty $\alpha > 0, \beta > 0$ a $\gamma > 0$ určující důležitost jednotlivých složek [41]:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma \quad (5.11)$$

Hodnota indexu se pohybuje v intervalu mezi 0 a 1, přičemž čím větší je index, tím více jsou si snímky podobné. Příklad porovnání MSE a SSIM indexu lze vidět na obrázku 5.4.



Obrázek 5.4: Porovnání MSE a SSIM indexu [6].

5.3 OCR

Optical Character Recognition (OCR) je software, který převádí text z obrázu do podoby čitelné počítačem. Text se může vyskytovat v jakékoliv podobě fontu či velikosti, od nápisů na značkách či plakátech, až po psaný text rukou.

Ačkoliv je modelů, které se specializují na určitý problém, celá řada, OCR se obecně skládá z několika částí. První a kritickou částí je preprocessing zahrnující převod obrázku na vhodný formát, odstranění šumu a různé druhy filtrace. Následně se provádí segmentace stránky či částí textu a úpravy masky morfologickými operacemi. Ve čtvrtém kroku je každému písmenu přidělen tzv. feature vektor, který jej charakterizuje. Extrahované vektory jsou následně použity pro klasifikaci. Mezi používané klasifikátory spadají Support Vector Machines (SVM), K-Nearest Neighbor (KNN) nebo pravděpodobnostní neuronové sítě (PNN). Posledním krokem je post processing, který slouží například pro odladění výstupu podle definovaných požadavků [73].

Příklady OCR frameworků jsou Google Drive OCR, Tesseract nebo Omni-Page [73].

Návrh algoritmů

Teoretická část práce se zabývá návrhem algoritmů pro získání interakcí uživatelů s reklamními sděleními z obrazových dat. Záznamy z obrazovek je nutno analyzovat a předzpracovat. Na základě vstupů je dále provedena inicializace služeb k získání potřebných informací pro následné zpracování videa.

6.1 Vstupní data

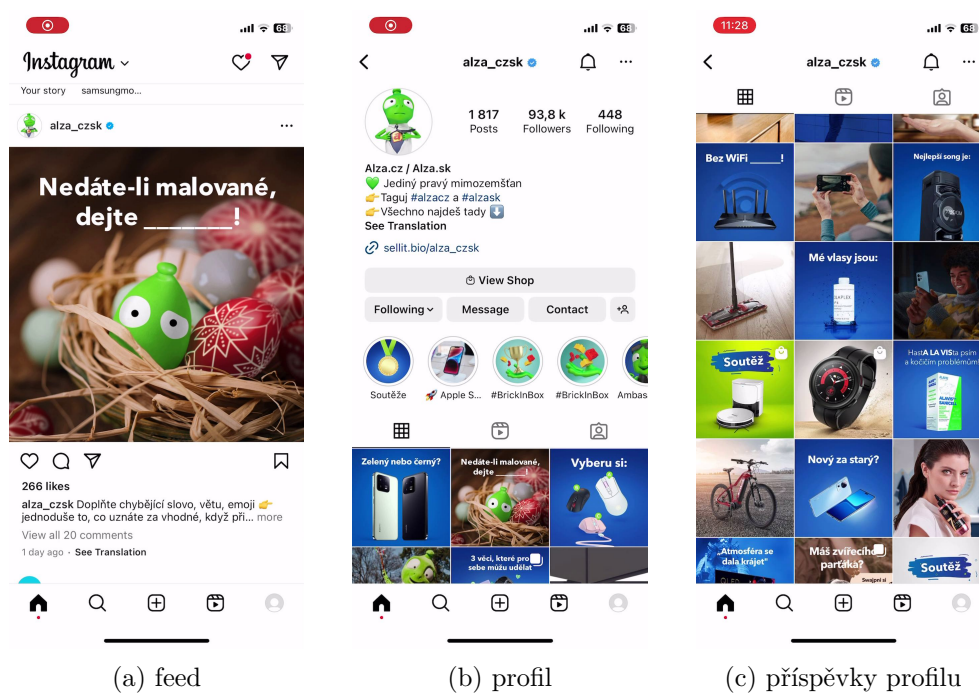
Sběr dat probíhá prostřednictvím panelu dobrovolných účastníků, kteří zaznamenávají svou aktivitu pomocí systémové Record funkcionality [74]. Rozlišení nahrávky, které Apple nezveřejňuje, se liší podle zařízení a je omezeno na 1920×1920 px [75], což se potvrdilo i z vlastních nahrávek, které byly pořízeny na několika zařízeních s různými rozlišeními (11 Pro, SE 1st gen, SE 2nd gen, 14 Pro, 11 a 12 mini). Tedy záznamy z modelů, které mají nativní rozlišení větší než 1920 px, jsou systémem naškálovány na přesně tuto hranici při zachování poměru stran. Nahrávky ze zařízení s menším rozlišením, než je stanovená hranice, zůstávají ve své nativní formě.

Paměťová náročnost záznamů se liší na základě zvoleného zařízení a délky, ale orientačně se pohybují v řádech desítek až stovek MB za minutu. Záznam má konstantní bit rate (dále CBR), variabilní frame rate (dále VFR) a codec AVC. Počet snímků za sekundu (dále FPS) se opět liší podle zařízení a obsahu nahrávky, ačkoliv spodní a horní hranice je nastavena na 30 a 60 FPS.

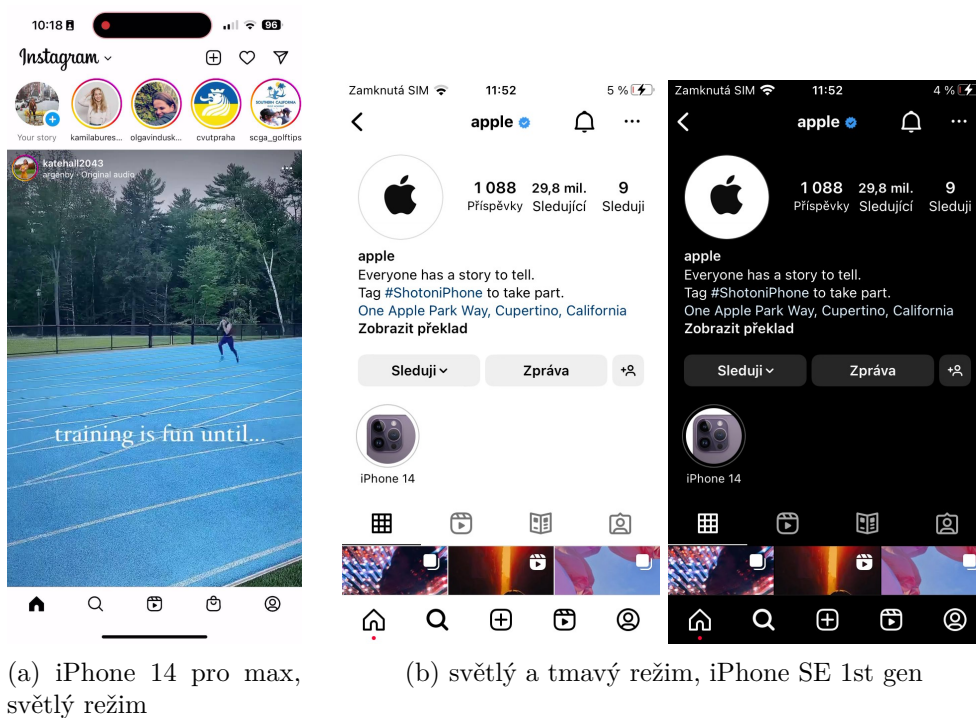
Nynější verze aplikace Instagram je možné nainstalovat na zařízení, které mají verzi operačního systému iOS 13.4 nebo vyšší [76]. Tím se lze omezit pouze na zařízení, které podporují alespoň iOS 13, viz tabulka 6.1.

Aplikace na vstupu kromě videa, viz obrázek 6.1, přijímá také model zařízení a systémový barevný režim, který může být buď světlý nebo tmavý, viz obrázek 6.2b. Barevné režimy mění vzhled aplikací v celém operačním systému. Není nutné, aby aplikace třetích stran funkcionality více barevných režimů podporovaly, ale v dnešní době se již jedná o standard, který je implementován téměř všude a Instagram není výjimkou.

6. NÁVRH ALGORITMŮ



Obrázek 6.1: Ukázka sekvence navštívení profilu, iPhone 11 Pro, světlý režim. Uživatel prohlíží feed (a), klikne na profil (b) a posune se dolů (c).



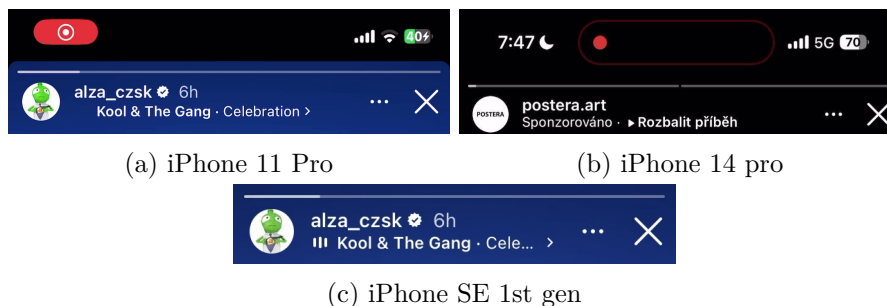
Obrázek 6.2: Ukázka různých barevných režimů a zařízení.

6.2 Předzpracování dat

Jak již bylo zmíněno v sekci 6.1, vstupní video má VFR. Z pohledu komprese je to sice vlastnost žádoucí, ale pro zpracování přináší komplikace při načítání jednotlivých snímků a korespondujících časových značek. Navíc například interakce Pause ze sekce 3.1, potřebuje duplicitní snímky, jelikož ji charakterizují. Video je tedy nejprve potřeba překonvertovat na konstantní frame rate (dále CFR), který zajistí, že ve videu se budou vyskytovat i duplicitní snímky.

Z hlediska optimalizace je žádoucí omezit počet zpracovávaných FPS. Byl proveden experiment manuálního testování aplikace Instagram za účelem zjištění délky trvání animací (zobrazení nového story, reels, profilu, komentářů, ...). Při standardně rychlém internetovém připojení (alespoň 20 Mbps) animace trvají většinou zhruba 0,5 sekundy. Pokud je součástí obrazovky i asynchronní načtení dat, může animace trvat déle. Existují ale i krajní případy jako zavření story, kdy animace trvá zhruba 0,3 sekundy. Vyjde-li se z délky trvání nejkratší animace, hranici je možné stanovit nejméně na 3 FPS. Jelikož ale existují interakce, které potřebují dostatečný počet snímků pro rozpoznání (Pause nebo WatchWhole), takto nízký počet FPS by mohl způsobit ztrátu dat. Hranice je tedy stanovena na 10 FPS jako kompromis mezi optimalizací a zachováním rozumného množství informací. Výjimkou je lokace feed s hranicí posunutou na dvojnásobek, protože je potřeba detekovat vertikální pohyb uživatele, viz podsekce 6.4.2, a pokud je rychlost příliš velká, 10 FPS není dostačující.

Poslední krok předzpracování, je oříznutí systémových prvků. Apple design svých zařízení modernizuje a mění, čemuž se nevyhnul ani displej. Od jednoduchého čtvercového designu, jako je vidět na obrázku ze zařízení iPhone SE 1st gen 6.2b, se vyvinul k displeji s rozšířenou horní systémovou lištou a novou spodní částí pro systémové ovládání (tzv. Home indicator), viz obrázek ze zařízení iPhone 11 Pro 6.1a. Nejnovější inovace přišla s nástupem iPhone 14 Pro a Dynamic Island funkcionalitou, která zvětšuje výšku systémové horní lišty, viz obrázek 6.2a [77]. Rozměry pro oříznutí systémových prvků se tedy liší v závislosti na zařízení, navíc pro každý snímek separátně, jelikož modely, které nemají Home indicator, zobrazují story přes horní systémovou lištu, viz obrázek 6.3. Pokud by se tedy ořezávalo bez ohledu na lokalitu, u vybraných modelů by nebylo možné zpracovat interakci WatchWhole, protože by došlo k oříznutí progress baru, který ji definuje. Seznam zařízení, které mají Home indicator (tedy i rozšířenou horní systémovou lištu) nebo Dynamic island lze vidět v tabulce 6.1.



Obrázek 6.3: Porovnání horních systémových prvků na lokaci story.

6.3 Inicializační část

Prvním krokem při spuštění aplikace je inicializační část, která na základě vstupního videa a parametrů provede kroky, které jsou uplatněny u algoritmů pro získání interakcí.

6.3.1 Metadata videa

Mimo obecné informace jako FPS nebo délka videa je nutné získat i čas vytvoření záznamu, jelikož detekované interakce musí být označeny časovou značkou. Z indexu snímku, který se zpracovává, a získaných metadat lze poté snadno dopočítat přesný čas, kdy se interakce udála.

Dalším důležitým údajem je formát profilu AVC codeců. Ačkoliv všechny modely využívají High profil, úroveň jednotlivých profilů se liší. Hlavní rozdíl je u modelu SE 1st gen, který díky nejnižšímu výkonu a nejmenšímu rozlišení používá úroveň 3.2, kdežto ostatní modely mají alespoň úroveň 4.x. Následkem je, že barvy výsledného videa jsou od ostatních modelů rozdílné.

6.3.2 Škála

Posledním krokem inicializační části je výpočet škály, která umožňuje převádět vzdálenosti naměřené na referenčním zařízení na hodnoty platné na jiném. Pomocí škály nelze počítat vzdálenosti UI prvků, které nemají fixní rozměry. Příkladem je progress bar u story, viz obrázek 6.3, jelikož závisí na šířce obrazovky. Výpočet hodnot je tedy platný pro fixní vzdálenosti a rozměry, mezi které spadají následující případy:

- rozměry ikon
- vzdálenosti od kraje obrazovky¹

¹To se týká pouze případu, kdy je UI prvek zarovnaný ke hraně obrazovky. Pokud je vycentrovaný, jeho umístění už závisí na velikosti displeje.

Mobilní zařízení mají na své rozměry poměrně vysoké a různorodé rozlišení. Aby byl zajištěn konzistentní vzhled, Apple používá pro vykreslení grafických prvků tzv. logický souřadnicový systém, díky kterému lze uživatelské rozhraní naimplementovat pouze jednou, jelikož se při zobrazení pronásobí měřítkem (scale factor), viz tabulka 6.1, který obsah namapuje na souřadnicový systém displeje zařízení [78].

Se znalostí měřítka je možné převádět vzdálenosti naměřené z referenčních hodnot na vzdálenosti jiného zařízení. Jak bylo zmíněno v sekci 6.1, rozlišení vstupního videa je omezeno na 1920 px, což je potřeba ve výpočtu zohlednit. Výsledný vzorec výpočtu škály pro převod vzdálenosti ze screenshotu zařízení A na vzdálenost ve videu zařízení B je tedy:

$$s = \frac{sf_A}{sf_B} \cdot \frac{hs_B}{hv_B}, \quad (6.1)$$

kde sf_B a sf_A je scale faktor zařízení B a A, hs_B je výška screenshotu zařízení B a hv_B je výška videa zařízení B². Výslednou škálou pak stačí daný rozměr podělit.

Pro příklad lze uvést převod velikosti ikony symbolizující profil uživatele na feed o šířce 96 px, která byla naměřena ze screenshotu ze zařízení iPhone 11 Pro. Cílený model je SE 1st gen. Z tabulky 6.1 lze vyčíst, že korespondující hodnoty scale factor jsou 3 a 2, a výška screenshotu SE 1st gen je 1136 px, což znamená, že výška videa je totožná. Po dosazení:

$$s = \frac{3}{2} \cdot \frac{1136}{1136} = 1,5.$$

Podělením rozměru škálou vyjde šířka ikony 64 px.

Do uvedeného vztahu je potřeba dosazovat rozlišení screenshotu a ne samotného zařízení, kvůli výjimkám jako je například model 12 mini nebo 8 plus, které mají fyzicky menší rozlišení než je samotný screenshot. Hypotézu potvrdily [79, 80] a také testování nad modelem 12 mini, který byl při vývoji k dispozici.

6.3.3 Určení škály pomocí TM

V sekci 6.1 je uvedeno, že vstupem je i název zařízení, pomocí kterého lze škálu vypočítat snadno z tabulkových hodnot. Pokud se parametr ve vstupu nevyskytuje, lze ji pomocí metody TM a brute force zjistit. Nejdříve se načtou referenční obrázky ikon a z tabulky se vytvoří veškeré možné hodnoty škály. Následně je vyzkoušena detekce ikon na prvním snímku videa pomocí dané škály. Metoda vrací pozici ikony, na které byla největší hodnota korelace, kterou lze vnímat jako míru jistoty. Škála, pomocí které se docílí největšího průměru korelací u všech vyzkoušených ikon je považována jako správná.

²Do výpočtu se uvádí výška, jelikož neexistuje iPhone, který by měl větší šířku, než výšku. To znamená, že jediný rozměr, který může být případně oříznut na 1920 px je výška.

Pokud je ze snímku zřejmý i barevný režim zařízení, také nemusí být uveden na vstupu, jelikož ikony je potřeba vyzkoušet pro obě varianty.

Ačkoliv se jedná o funkční způsob zjištění škály, popřípadě i barevného režimu, je doporučeno zadat parametry ručně, jelikož by měly být známé při sběru. Zjištění těchto parametrů lze navíc automatizovat pomocí modifikace aplikace, která se nyní používá pro ukládání videí na serverové úložiště.

Tabulka 6.1: Seznam modelů, které podporují iOS 13 nebo novější a jejich vlastnosti [81, 79, 80]

iPhone	Rozlišení screenshot (px)	Scale factor	Home indicator	Dynamic Island
14 Plus	1284 × 2778	3	Ano	Ne
14 Pro Max	1290 × 2796	3	Ano	Ano
14 Pro	1179 × 2556	3	Ano	Ano
14	1170 × 2532	3	Ano	Ne
SE 3rd gen	750 × 1334	2	Ne	Ne
13	1170 × 2532	3	Ano	Ne
13 mini	1125 × 2436	3	Ano	Ne
13 Pro Max	1284 × 2778	3	Ano	Ne
13 Pro	1170 × 2532	3	Ano	Ne
12	1170 × 2532	3	Ano	Ne
12 mini	1125 × 2436	3	Ano	Ne
12 Pro Max	1284 × 2778	3	Ano	Ne
12 Pro	1170 × 2532	3	Ano	Ne
SE 2nd gen	750 × 1334	2	Ne	Ne
11 Pro Max	1242 × 2688	3	Ano	Ne
11 Pro	1125 × 2436	3	Ano	Ne
11	828 × 1792	2	Ano	Ne
XR	828 × 1792	2	Ano	Ne
XS Max	1242 × 2688	3	Ano	Ne
XS	1125 × 2436	3	Ano	Ne
X	1125 × 2436	3	Ano	Ne
8 Plus	1242 × 2208	3	Ne	Ne
8	750 × 1334	3	Ne	Ne
7 Plus	1242 × 2208	3	Ne	Ne
7	750 × 1334	2	Ne	Ne
SE 1st gen	640 × 1136	2	Ne	Ne
6s Plus	1080 × 1920	3	Ne	Ne
6s	750 × 1334	2	Ne	Ne

6.4 Návrh řešení algoritmů

Jak plyne z kapitoly 3, některé interakce jsou společné (Screenshot, Shop, ShareIG . . .) a některé pro danou lokaci specifické (WatchWhole, SavePostIG, . . .). To znamená, že na základě lokality lze určit množinu možných interakcí. Pro jejich rozpoznání je potřeba vyřešit následující podproblémy:

1. určení lokality
2. detekce pohybu a nového příspěvku
3. detekce reklamy
4. rozpoznání interakcí

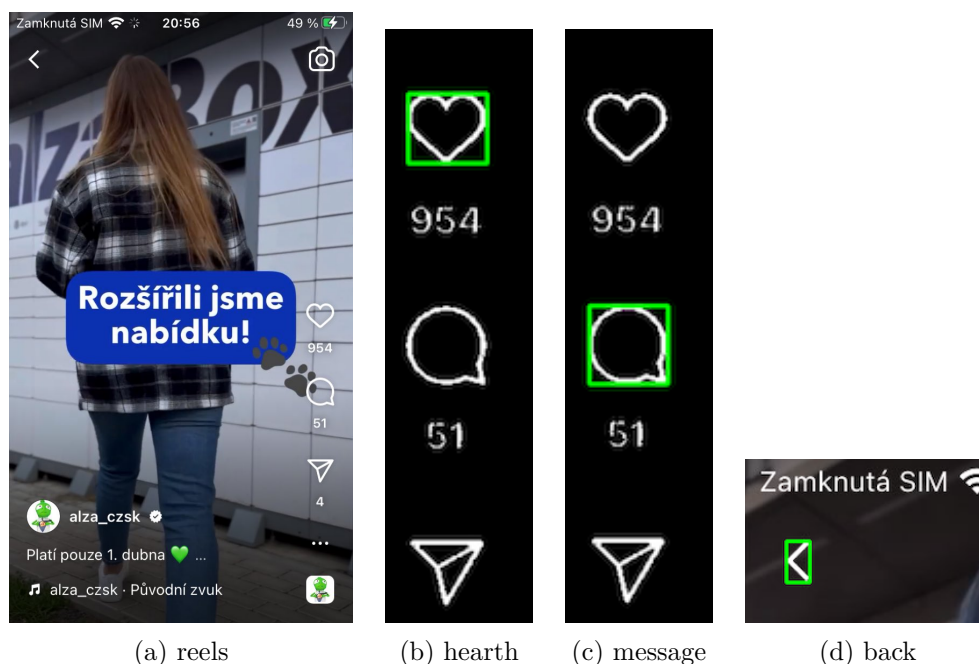
Kroky detekce nového příspěvku a reklamy jsou specifické pro feed, story a reels, jelikož pouze zde může uživatel narazit na příspěvky s reklamním sdělením. Tedy detekuje se nový příspěvek na obrazovce a vyhodnotí se, zda se jedná o reklamní sdělení. Pokud ano, lze začít rozpoznávat interakce.

Ačkoliv na profilu není prováděná detekce reklamy, uživatel má možnost přejít z jednoho profilu na jiný, popřípadě rovnou z příspěvku na profil, který nepatří inzerentovi. Z toho důvodu je i zde nutné kontrolovat pohyb a validitu.

6.4.1 Určení lokality

Lokalita uživatele je důležitá informace, pomocí které lze přímo identifikovat některé interakce. Ačkoliv jich Instagram nabízí celou škálu díky velkému počtu funkcionalit, v rámci práce se uvažuje pouze s takovými, které mají spojitost s reklamními sděleními a manipulaci s nimi.

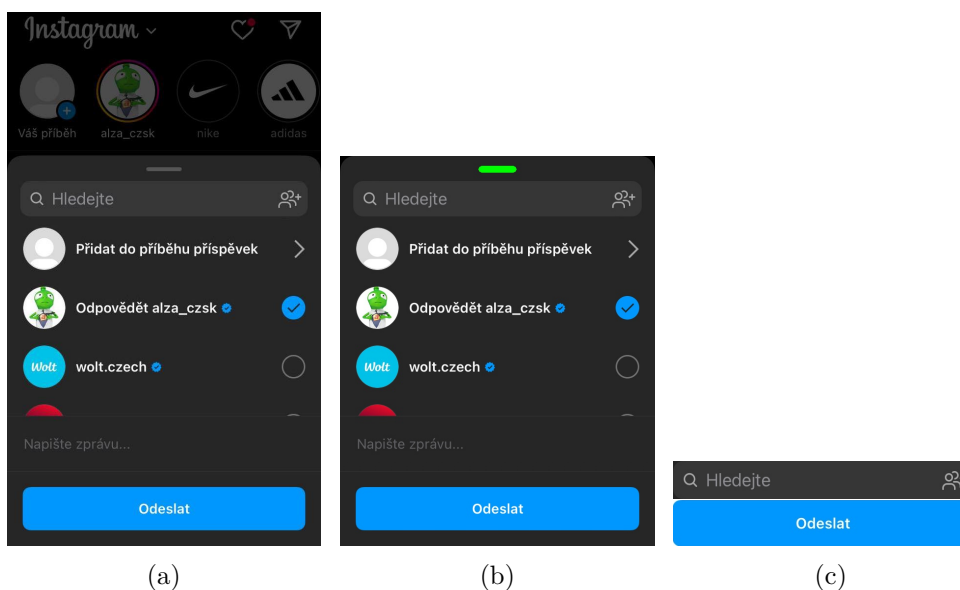
Každá obrazovka obsahuje ikony a UI prvky (tlačítka, profilový obrázek), které ji pomocí svého tvaru a umístění definují. Hlavní princip detekce lokality se opírá právě o přítomnost ikon, jejichž rozpoznání zajišťuje metoda TM, která má oproti SIFT a SURF výhodu ve své rychlosti. Nevýhodou je fakt, že ikony musí být stejné velikosti jako reference, podle které se hledají, což lze vyřešit převodem rozměrů pomocí škály. Algoritmus tedy vezme veškeré referenční ikony definující lokalitu, naškáluje a pokusí se je v konkrétní části snímku najít. Ikona je nalezena, pokud korelace mezi referencí a danou oblastí v obrázku přesahuje manuálně určený práh (většinou 0,8). Pokud snímek obsahuje dostatečné množství nálezů, lokalita je rozpoznána. Z popisu plyne, že musí existovat databáze referenčních obrázků ikon příslušející jednotlivým lokacím, případě barevným režimům. Součástí mapování pomocí TM je také předzpracování (výřez, segmentace, . . .), jako například u ikon na obrázku 6.4.



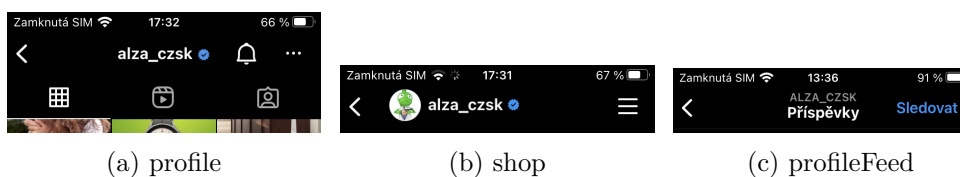
Obrázek 6.4: Detekce ikon na lokaci reels pomocí TM.

Lokality reprezentované modálním oknem je potřeba nejdříve očistit o pozadí, jelikož je průhledné a může tedy vést ke špatné detekci, viz obrázek 6.5a. Modální okno v rámci Instagramu je možné určit malým indikátorem v horní části, který lze pomocí velikosti, tvaru a prahování rozpoznat. Detekce poté probíhá pomocí TM nebo segmentací jiných UI prvků uvnitř modálního okna, jak je vidět na příkladu shareIG na obrázku 6.5. Obdobný postup je aplikován i na ostatní obrazovky s modálním oknem.

Mezi problematické lze zařadit ty, které umožňují uživateli vertikální posun, tedy shop, profil, profileFeed a feed. U prvních třech lokací je možné se opřít o grafické elementy v navigačním menu, viz obrázek 6.6. Profil je určen pomocí TM a na profileFeed lze využít kombinaci TM a segmentace prahováním, jelikož je to jediná obrazovka, kde jsou dva řádky textu pod sebou. Případně lze použít i OCR, které je ale časově náročnější a není odolné vůči lokalizaci (zvolenému jazyku). Pokud by algoritmus počítal s českými překlady, u anglické verze by nefungoval. Shop je vyřešen kombinací TM a algoritmem pro detekci kruhů Hough circles (dále HC). Pokud je v navigačním menu kruh o daném průměru a v určité vzdálenosti od levého okraje, lze tvrdit, že se jedná právě o lokaci shop.



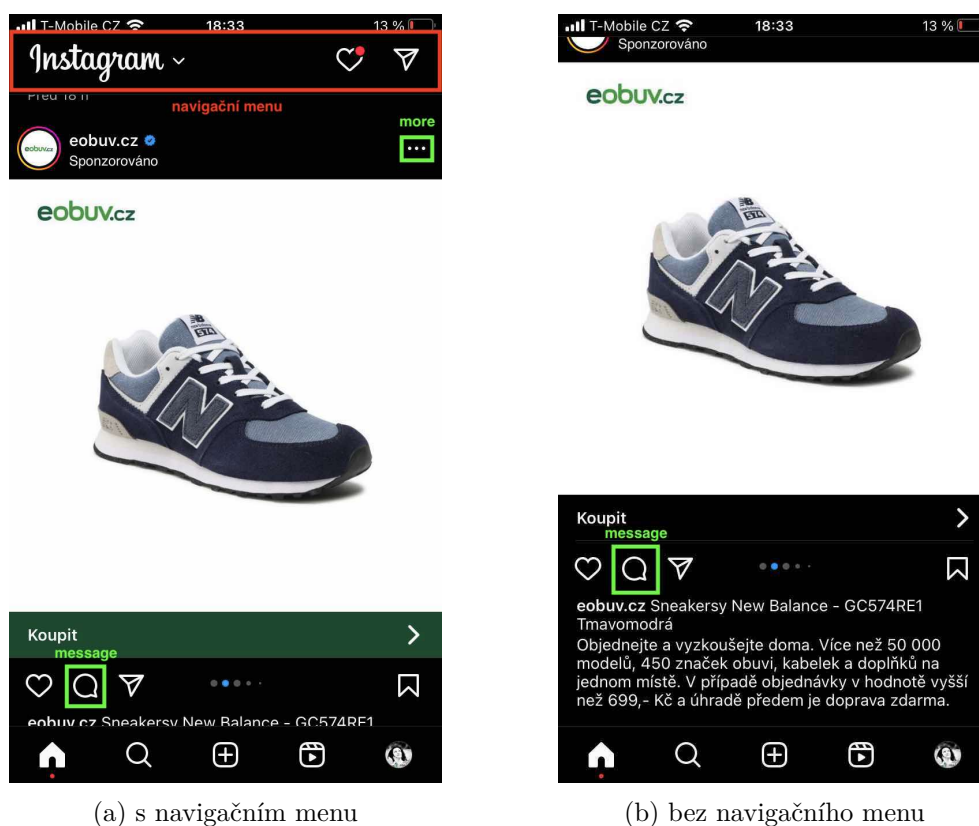
Obrázek 6.5: Detekce lokace modálního okna shareIG. Celá obrazovka bez oříznutí (a), oříznuté modální okno se zvýrazněným indikátorem (b) a oříznuté specifické UI prvky (c).



Obrázek 6.6: Navigační menu scrollovacích obrazovek s fixním navigačním menu.

Nejvíce problematickou lokací je feed, jelikož při posunu směrem dolů se horní lišta schová, viz obrázek 6.7. Tím lze, alespoň částečně, vyloučit určení lokality na základě navigačního menu. Možné řešení je využití TM na horní lištu. Pokud metoda selže, použije se TM na obsah uvnitř obrazovky na libovolnou ikonu pod příspěvkem (třeba message) a na ikonu more, které jsou v určité vzdálenosti od okrajů, jak je vidět na obrázku 6.7. Zmíněná kombinace ohraničuje právě jeden příspěvek a ten nemusí být na obrazovce celý, tedy nelze využít pouze ikony z oblasti pod, respektive nad ním. Komplikaci přináší fakt, že reels i story mají ikonu more ve stejné vzdálenosti od pravého kraje jako feed, což lze vyřešit segmentací pozadí feedu a kontrolou, zda má maska šířku totožnou s displejem a zabírá alespoň nějaké procento obrazovky (10 %). Pro ujištění se ještě zkontroluje, zda se nejedná o lokaci reels ani story. Po splnění všech podmínek a nálezů alespoň jedné ikony je rozpoznán feed.

6. NÁVRH ALGORITMŮ

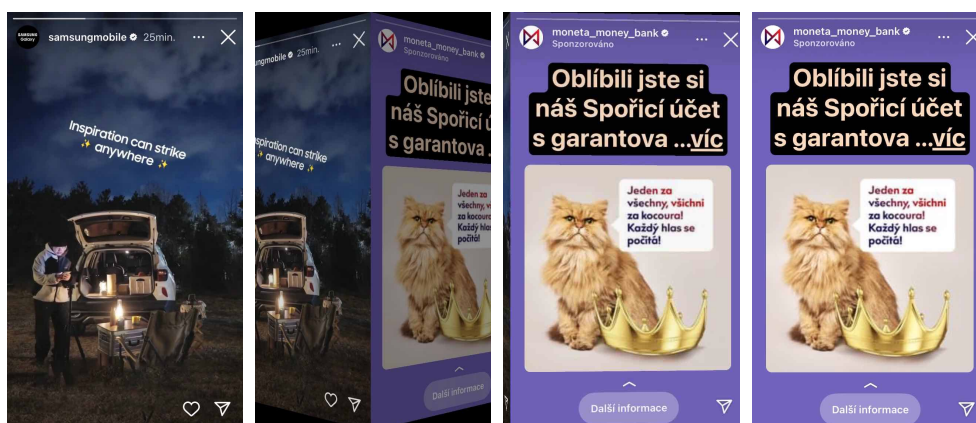


Obrázek 6.7: Navigační menu na feed před posunutím dolů (a) a po posunutí dolů (b).

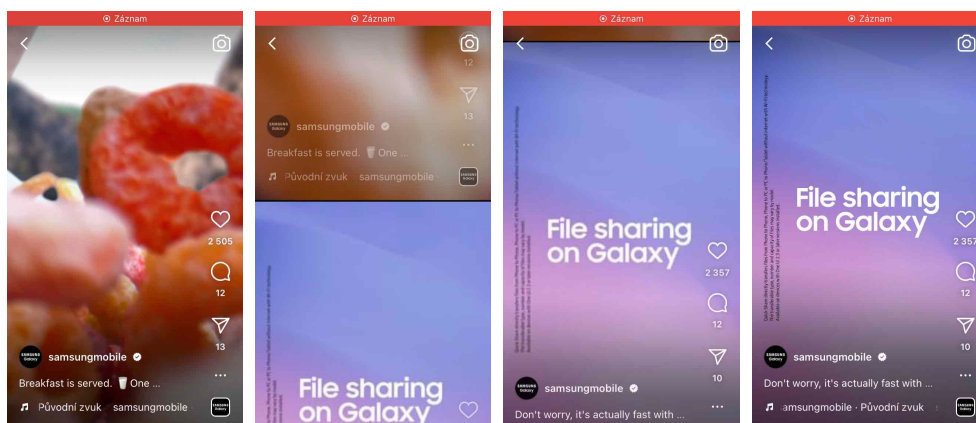
6.4.2 Detekce pohybu a nového příspěvku

Lokace story a reels prezentují obsah přes celou obrazovku, tedy uživatel vidí vždy pouze jeden příspěvek. Animace zobrazení nového příspěvku trvá mezi 0,3 a 0,5 sekundami. Jelikož je předešlým krokem detekce lokality a zpracovává se minimálně 10 FPS, je jisté, že v určitý moment bude snímek obsahovat přechod mezi dvěma příspěvky, což znamená, že lokalita nebude rozpoznána. Ve chvíli, kdy se pohyb na obrazovce ustálí a bude vidět pouze jeden příspěvek, lokalita bude opět odpovídat story nebo reels, což znamená, že byl zobrazen nový příspěvek.

Komplikací je konec přechodové animace, který je velmi jemný a rozdíl ustáleného a pohybující se příspěvku je minimální, což zapříčiní, že lokalita je detekována správně, ale obrazovka je stále v pohybu, ačkoliv nepatrném. Pro započítání dalších kroků je žádoucí, aby byl pohyb zcela eliminován. Toho lze docílit sledováním pohybu ikon, jelikož se hýbou společně s obrazovkou. Jakmile se ustálí, lze prohlásit, že byl zobrazen nový příspěvek. Řešení ustálení animace je aplikováno i při zobrazení obrazovky profilu.



Obrázek 6.8: Přechodová animace story.



Obrázek 6.9: Přechodová animace reels.

Feed umožňuje volný vertikální pohyb, což znamená, že ve snímku je možné vidět více než jeden příspěvek. Zároveň se nelze opřít o vlastnost, že při přechodu na nový příspěvek není detekována lokalita — je nutné sledovat pohyb, jeho směr a vzdálenost. Jelikož se obrazovka hýbe celá, z metod z podsekcce 4.2.1 se nabízí pouze optical flow. Metoda funguje dobře, pokud je rozdíl mezi snímky mírný, avšak při rychlejším pohybu dochází ke ztrátě referenčních bodů, což vede ke špatným výsledkům.

Vhodným řešením problému je manuální dopočítávání směru i vzdáleností na základě pozic ikon jako referenčních bodů. Metoda je odolnější vůči větším rozdílům mezi snímky a výpočetně méně náročná v porovnání s optical flow. Algoritmus se spléhá na to, že mezi dvěma snímky bude vždy alespoň jedna stejná ikona, což je zaručeno ve většině případů při správném určení typu ikon. Vhodné jsou message a more, jelikož ohraničují příspěvek, viz podsekcce 6.4.1 a obrázek 6.7. Předpoklad nefunguje ve dvou krajních scénářích:

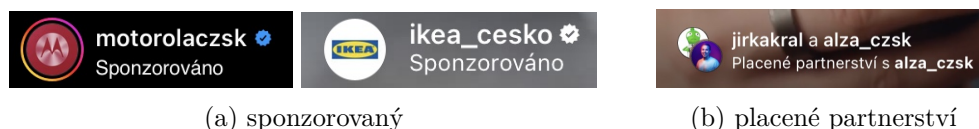
1. obsah příspěvku zabírá celou obrazovku
2. uživatel pohybuje obsahem tak, že nelze určit, zda se ve dvou snímcích jedná o stejnou ikonu

První scénář nastává pouze na zařízeních s malým displejem a u příspěvků typu reels. Jelikož referenční ikony leží v těsné blízkosti obsahu příspěvku, nedetekovaná vzdálenost posunutí bude velmi malá a tedy zanedbatelná. Druhý scénář nastane tehdy, když uživatel s obsahem hýbe tak rychle, že není možné okem rozpoznat jednotlivé příspěvky, čili jedná se opět o zanedbatelný případ.

Informace o pozici ikon ohraničujících příspěvek, směru a vzdálenosti pohybu jsou k detekci dostačující. Pokud uživatel posune obsah nějakým směrem o tři čtvrtiny výšky posuvné oblasti, což je hrubý odhad průměru výšky jednoho příspěvku, je vhodné zkusit detekovat reklamu. Hranice je důležitá, protože uživatel může libovolně pohybovat s obsahem oběma směry, ale jedním se v součtu posune minimálně, což znamená, že je pořád na stejném příspěvku.

6.4.3 Detekce reklamy

Jak je popsáno v sekci 2.2, existují tři typy reklamních sdělení. První dva jsou sponzorovaný a placené partnerství. Oba typy mají označení na pozici hlavičky pod uživatelským jménem, viz obrázek 6.10. Stačí tedy v obrazu alokovat hlavičku a za použití OCR zjistit uživatelské jméno a zda je příspěvek reklamní sdělení.



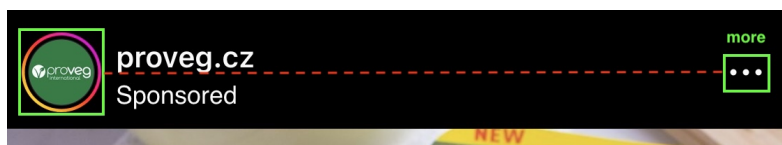
Obrázek 6.10: Typy hlaviček sponzorovaný a placené partnerství.

Pro alokaci lze využít předpokladu, že každá hlavička obsahuje profilový obrázek, který je kulatý, má fixní průměr, fixní vzdálenost od levého okraje a je ohraničený nepatrnou šedou hranicí. Nabízí se použití algoritmu HC a následnou filtraci detekovaných kruhů podle zmíněných vlastností. Někdy je vhodné předzpracování snímku v podobě hranové detekce, například pokud má profilový obrázek stejnou barvu jako pozadí. Příspěvek může mít několik druhů profilových obrázků, kde každý má jiný průměr a vzdálenost od levého okraje, viz obrázek 6.11, proto je nutné vyzkoušet detekci pro všechny možné varianty.



Obrázek 6.11: Typy profilových obrázků. Klasický (a), aktivní story, účet zveřejnil nějaké story, které si může uživatel stisknutím na obrázek prohlédnout (b) a kolaborace mezi dvěma profily (c).

U lokace feed je možné, že bude docházet ke špatným detekcím, protože na rozdíl od story nebo reels není profilový obrázek zobrazený ve fixní oblasti obrazovky. Jelikož v obsahu příspěvku může být zobrazeno cokoliv, nálezy jsou kromě vyfiltrování podle vlastností následně zkontrolovány, zda je na protější straně ikona more, viz obrázek 6.12. Z pozice profilového obrázku se následně získá pouze výseč, která obsahuje uživatelské jméno inzerenta a potenciální označení reklamy.

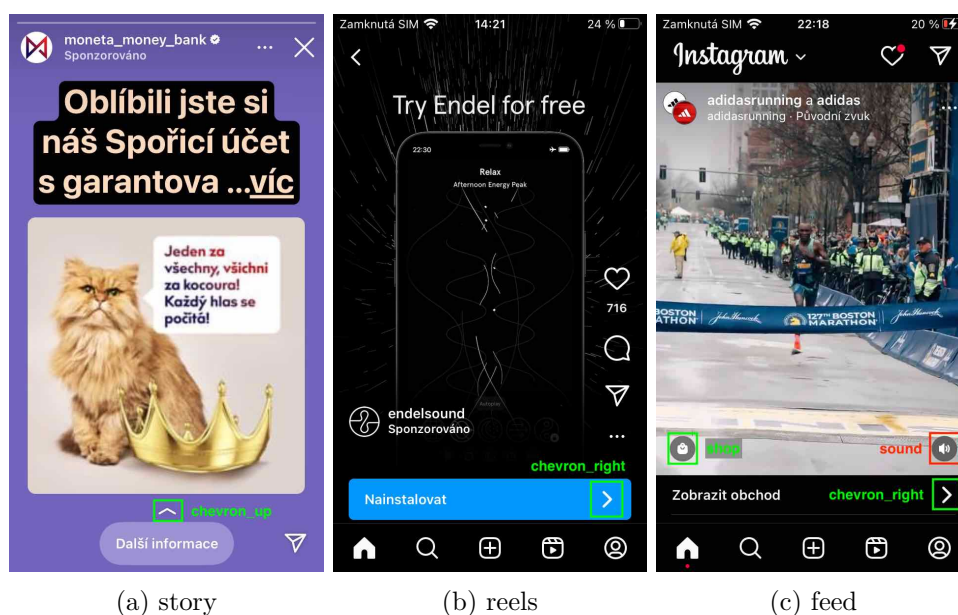


Obrázek 6.12: Celá hlavička příspěvku s ikonou more na feed.

Třetím typem reklamního sdělení je organic, který je v práci řešený pouze pro případ, že příspěvek obsahuje CTA tlačítka. Jelikož jsou tlačítka reprezentována ikonami (chevron_up, chevron_right a shop), viz obrázek 6.13, detekci lze zajistit metodou TM. Z obrázku 6.13 je možné pozorovat, že metodu je možné využít i jako sekundární pro určení reklamy, pokud byl špatně rozpoznán text pomocí OCR.

Důvodem omezení je nekonzistence značení a fakt, že reklama nemusí být označena vůbec. Pokud označena je, tak v popisku, který uživatel musí ve většině případů rozkliknout a tím může implicitně dojít ke ztrátě části interakcí, jelikož jejich rozpoznávání začíná až po detekci reklamního sdělení, ke které by došlo až při zobrazení označení.

6. NÁVRH ALGORITMŮ



Obrázek 6.13: CTA tlačítka reklamních sdělení. Feed (c) obsahuje i tlačítko určující typ média.

Ačkoliv se nejedná o detekci reklamy, pro detekci interakcí na profilu je důležité rozpoznat, zda koresponduje s inzerentem původního příspěvku. Při každém zobrazení profilu je tedy nutné provést kontrolu uživatelského jména, které se vyskytuje v navigačním menu, viz obrázek 6.6a. Uživatelské jméno stačí podle pozic ikon oříznout, vysegmentovat prahováním a detekovat pomocí OCR.

6.4.4 Rozpoznání typu média

Médium může být buď fotografie, video nebo kombinace obého, pokud se jedná o carousel příspěvek na feed. Story nemá žádný ukazatel, který by explicitně určoval, o jaký typ jde, rozpoznání je tedy nutné opřít o podobnost série snímků z videa. Po detekování reklamního sdělení se uloží n snímků celé obrazovky. Podle [82] je doba soustředění uživatele mezi 2 až 3 sekundami. [83] tvrdí, že uživatel mobilního zařízení průměrně stráví 1,7 sekundy u jednoho příspěvku. Z toho plyne, že n musí být menší, než tyto hranice, a proto je počet stanoven na 10, což odpovídá jedné sekundě pozorování příspěvku. Uložené snímky je potřeba následně porovnat. Pokud průměr podobností sousedních snímků přesáhne th_1 a zároveň podobnost prvního a posledního je větší než th_2 , jedná se o fotografii. První hranice slouží jako pojistka, pokud by video obsahovalo smyčku, která by se opakovala každou sekundu, což je nepravděpodobné, ale možné. Druhá hranice slouží pro samotnou detekci, jelikož fotografie musí být stejná jak ze začátku, tak po jedné sekundě. Hranice

by měly být určeny přísně, ale zároveň tak, aby například pohyb progress baru ve story neovlivnil detekci, čemuž experimentálně odpovídá číslo 0,98 (pro th_1 i th_2). Na výpočet podobnosti se hodí použít SSIM index, který bere v potaz celkovou texturu dvou snímků.

Ačkoliv reels může obsahovat progress bar, který indikuje, že typ média je video, není jisté, že bude zobrazen, pokud je video příliš krátké. Jelikož je příspěvek na reels zobrazen přes celou obrazovku, způsob určení typu média je totožný se story.

Poslední lokací je feed, kde z povahy prezentace příspěvků nelze SSIM index využít, protože uživatel může kdykoliv hýbat s obsahem. Z obrázku 6.13c je vidět, že pokud je médium video, obsahuje příspěvek ikonu pro ztlumení zvuku, což lze vyřešit metodou TM.

6.4.5 Rozpoznání interakcí

Popis algoritmů je rozdělen do podsekce podle typu obrazovky. Aby bylo jasné, o kterou reakční skupinu a interakci se jedná, je použita anotace „reakční skupina–interakce“.

Story

See–Stop je detekována tehdy, když je příspěvek na obrazovce rozpoznán jako reklamní sdělení a dokud uživatel nepřejde na jiný příspěvek nebo jinou lokalitu.

Tím, že je popis u story prezentován jako modální okno, interakce See–ExpandText je detekována, pokud je lokalita storyExpandText.

Složitější interakce je See–WatchWhole, jelikož vyžaduje sledování progress baru, který mimo indikace času reprezentuje i počet příspěvků. Komplikací je, že progress bar je průhledný, nelze tedy použít segmentační metody založené na barvě, jelikož se mění v závislosti na obsahu. Vhodnější je využití hranové detekce v kombinaci s morfologickými operacemi, pomocí které lze určit i počet příspěvků. Podle kontur lze oříznout pouze samotný progress bar a využitím background subtractoru pozorovat pohyb. V závislosti na pozici pohybující se linie a počtu příspěvků je interakce detekována.

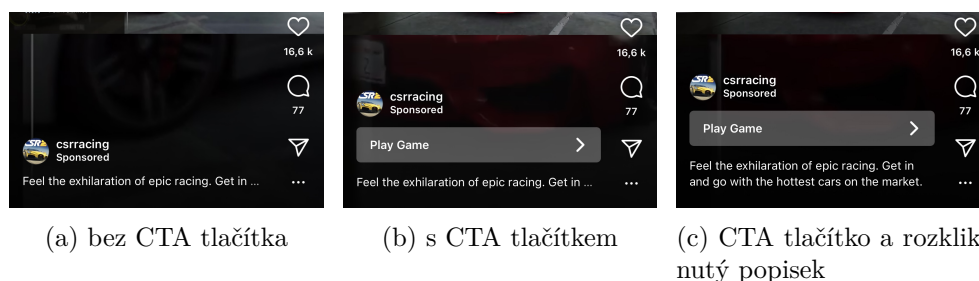
Další interakce je See–Pause. Z její definice je jasné, že při pozastavení nebude rozpoznána lokalita. Této vlastnosti lze využít a spustit algoritmus pro detekci pouze pokud předchozí lokalita byla story. Samotné rozpoznání je možné opřít o SSIM index a určení přísné hranice, kterou musí dva za sebou jdoucí snímky přesáhnout (zde 0,99).

Poslední je Social–Like. Ikonu srdíčka (hearth), která slouží k vykonání interakce lze detekovat vyříznutím spodní části obrazovky a metodou TM.

Reels

Kroky pro rozpoznání interakcí See–Stop a See–Pause jsou totožné s reels. Za zmínku stojí Social–Like, která je myšlenkově také stejná, ale barva ikony heart je závislá na obsahu. Před použitím TM pro detekci je proto aplikována segmentace prahováním.

Algoritmus detekující See–ExpandText se opírá o fakt, že profilový obrázek se po rozkliknutí popisku posune směrem nahorů, ale ostatní ikony zůstanou na místě, viz obrázek 6.14. Po detekování reklamního sdělení je potřeba uložit souřadnice profilového obrázku a kontrolovat, zda se přemístil. Komplikací je CTA tlačítko pokud je médium video, protože se objeví až po pár sekundách od zobrazení příspěvku, což způsobí, že se profilový obrázek posune. Součástí algoritmu je tedy i kontrola objevení tlačítka využitím TM.



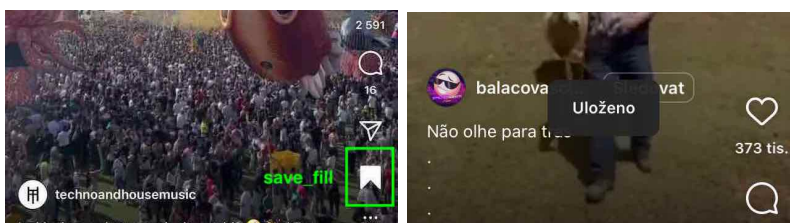
Obrázek 6.14: See–ExpandText interakce. Uživatel si právě zobrazil reels reklamní sdělení (a), po chvíli se automaticky objeví CTA tlačítko (b) a nakonec rozkliknul popisek (c).

Interakce See–WatchWhole má dva scénáře, viz kapitola 3, a nelze ji detekovat na fotografii. Prvním scénářem je krátké video, které nemá žádné explicitní označení opětovného přehrání, jelikož je zobrazeno ve smyčce. Nezbyvá tedy nic jiného, než si uložit první snímek a postupně kontrolovat pomocí SSIM indexu, kdy nějaký další překoná stanovenou hranici (0,98), která musí být opět přísná, ale zároveň musí zohlednit možnost, že uživatel dal příspěvku například like. Druhý scénář je progress bar ve spodní části videa. Ačkoliv se jedná o pohyb a nabízí se použití například background subtractoru, linie značící přehranou část videa není průhledná. To znamená, že pro detekci lze prahováním linii vysegmentovat a podle délky zkontrolovat, zda je vyplněna většina (alespoň 90 %). Procento vyplněnosti je možné snadno zjistit oříznutím progress baru podle referenčních hodnot vzdáleností od okraje, jelikož jsou fixní, narozdíl od samotné délky linie. Ukázku lze vidět na obrázku 6.15.



Obrázek 6.15: Původní progress bar a jeho maska.

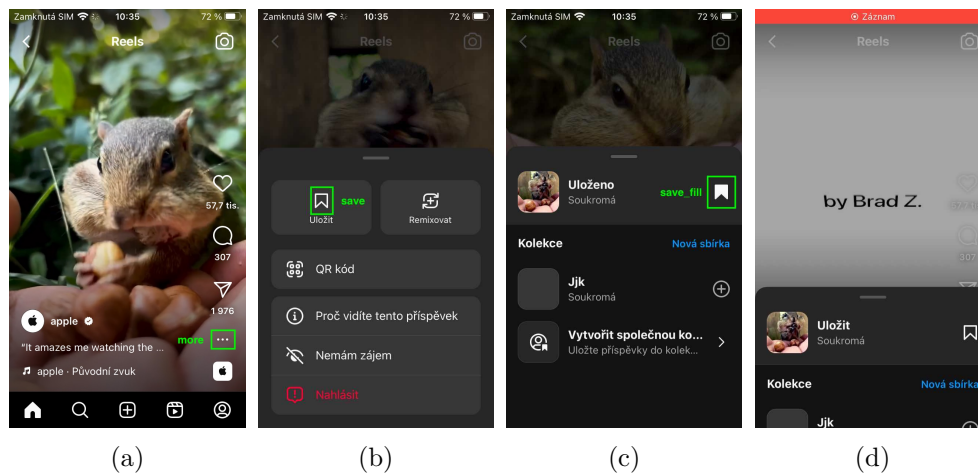
Další interakcí je Save–SavePostIG, která má tři scénáře. Na obrazovce může být přímo ikona pro uložení příspěvku (save). Pokud na ni uživatel klikne, vyplní se, viz obrázek 6.16a, což lze detekovat pomocí TM. Druhý a třetí scénář předpokládá, že se ikona save na obrazovce nevyskytuje. Začátek obou je vždy stejný, uživatel klikne na ikonu more a zobrazí se modální okno reelsOptions, kde následně klikne na save ikonu. Jednou možností je zobrazení alertu uprostřed obrazovky značící, že příspěvek byl uložen, viz obrázek 6.16b. Předpokladem pro zobrazení alertu je jasně daná posloupnost lokací, které si uživatel musel zobrazit: reels, reelsOptions, reels. Ve chvíli, kdy je posloupnost dodržena, lze po fixně stanovenou dobu (cca 2 sekundy), po kterou je alert zobrazen, určit zda se jedná o uložení nebo uživatel příspěvek z uložených odebral. Jelikož je alert textový, je použit OCR na výřezu obrazovky (alert je vždy uprostřed). Další možností je zobrazení saveIG obrazovky po prokliku z reelsOptions. Modální okno obsahuje v horní části ikonu save, která je buď vyplněná nebo ne. Pokud je při zavření modálního okna vyplněná, příspěvek je uložen. Celý postup je vidět na obrázku 6.17.



(a) vyplněná ikona save (save_fill) (b) alert po uložení

Obrázek 6.16: Ukázka Save–SavePostIG pro reels scénáře jedna (a) a alertu scénáře dva (b).

6. NÁVRH ALGORITMŮ



Obrázek 6.17: Celý proces uložení příspěvku scénáře tři. Uživatel si prohlíží příspěvek (a), pomocí ikony more zobrazí reelsOptions (b), klikne na ikonu save, čímž se zobrazí saveIG (c). Alternativní scénář je vyobrazen na (d), kdy uživatel kliknul na ikonu save_fill a tím zrušil uložení — interakce není detekována.

Social-Follow je definovaná alertem a změnou textu u uživatelského jména. Pro detekci je vhodný alert, jelikož se objeví hned po provedení interakce. Pozadí alertu je průhledné, nelze proto použít segmentaci prahováním ani hranovou detekci, jelikož v obsahu může být cokoliv a navíc médium video, které se přehrává i přes zobrazení alertu. Řešení se tedy opírá o pozici alertu, který je vždy uprostřed obrazovky a faktu, že text v alertu průhledný není. Kontinuálně se kontroluje, zda se na správné pozici vyskytuje kontura, která připomíná rozměry text. Detekce poté probíhá použitím OCR.

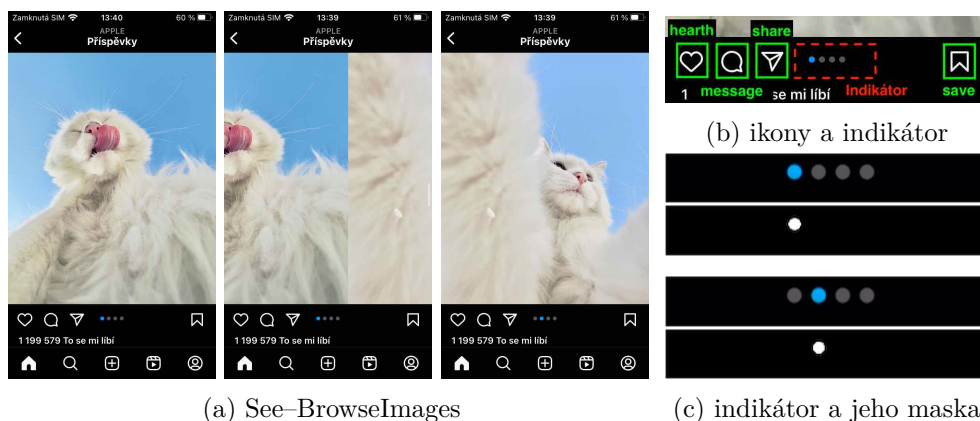
See-Comment je detekována na základě lokality, tedy detekce proběhne ve chvíli, kdy uživatel přešel z reels na reelsComments. Návaznou interakci Social-Comment lze detekovat pomocí segmentace prahováním, jelikož pozadí není průhledné a komentář je po zveřejnění zvýrazněn.

Feed

Detekce interakcí See-Stop a Social-Like je opět totožná s předchozími lokalitami. Stejně kroky mají i See-Comment a Social-Comment akorát s rozdílem, že obrazovka komentářů je reprezentována feedComments.

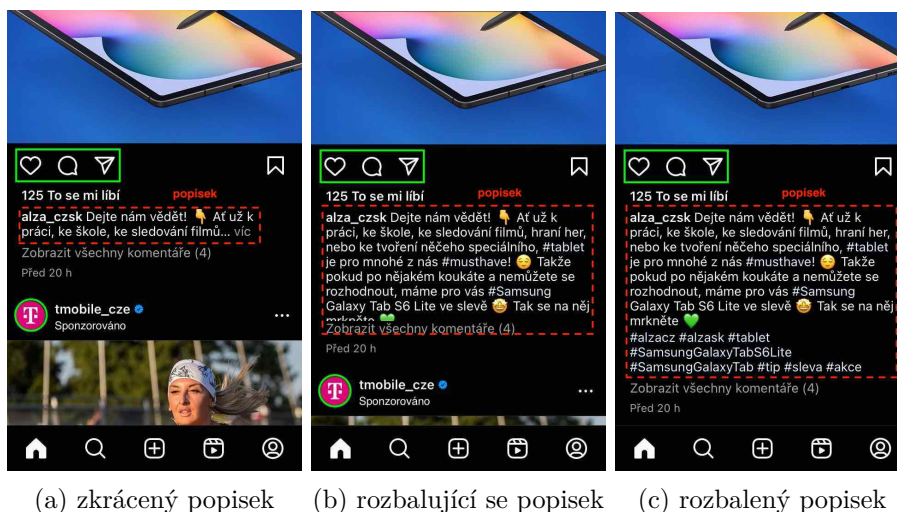
Interakci See-BrowseImages, jak bylo zmíněno v kapitole 3, lze detekovat pouze u příspěvku typu carousel, který je specifický indikátorem pozice uživatele v sérii fotografií/videí. Jak je vidět na obrázku 6.18b, díky TM lze zjistit souřadnice ikony sdílení (share) a save pod příspěvkem, které indikátor ohraničují. Jelikož je aktuální pozice zvýrazněna, je možné použít segmen-

taci prahováním na oříznutou část. Pro detekci poté stačí porovnávat, jak se předchozí pozice liší od stávající, viz obrázek 6.18c.



Obrázek 6.18: Ukázka See-BrowseImages (a), lišty pod příspěvkem s ikonami a indikátorem (b) a segmentace indikátoru z obrázku 6.18a (c).

Interakce See-Expand se projevuje posunutím následujícího příspěvku, ale zachováním pozice ikon stávajícího, viz příklad na obrázku 6.19. Pro detekci je potřeba si pamatovat a průběžně aktualizovat pozice ikon a následujícího profilového obrázku (popřípadě ikony more). Jakmile se začne pozice profilového obrázku měnit, ale ikony zůstanou na místě, proběhla interakce ExpandText. Předpokladem pro algoritmus je, že hlavička následujícího příspěvku je viditelná, což nemusí být vždy zaručeno.

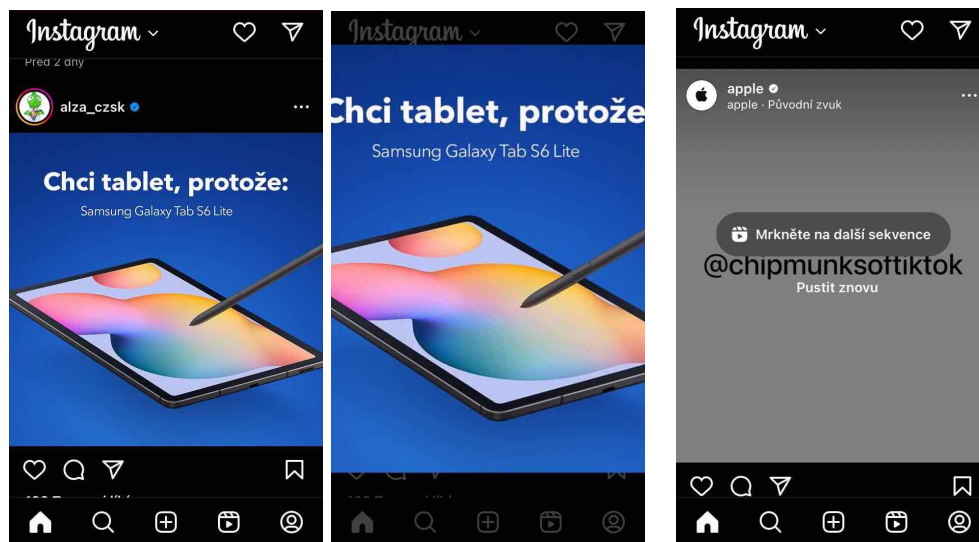


Obrázek 6.19: Průběh interakce See-ExpandText na feed.

Další interakcí je See-Zoom. Uživatel si může přiblížit jakýkoliv příspěvek na feed. Když tak učiní, pozadí s větším přiblížením postupně tmavne 6.20a.

Algoritmus pro detekci se může opřít o lokaci, která je v okamžik přiblížení neznámá a o hrany, které rozlišují pozadí od obsahu. Jakmile uživatel začne přibližovat obsah, detekují se hrany (například Cannyho detektorem) a zaznamená se jejich pozice. Jakmile se od sebe začnou vzdalovat nebo některá (či obě) zmizí, jedná se o interakci See–Zoom. Bohužel se může stát (v tmavém barevném režimu), že obsah příspěvku má stejnou barvu jako pozadí a v tu chvíli se nelze na detekci hran spolehnout. Další problém může nastat, pokud kvůli velikosti obsahu nebude možné hrany vůbec detekovat, což nastane, pokud je příspěvek (typicky typu reels) zobrazen přes celou obrazovku.

Interakce See–WatchWhole je opět možné rozpoznat pouze u videa a je závislá na typu příspěvku. Pokud se jedná o typ reels, na konci videa je zobrazeno tlačítko uprostřed obsahu, které značí, že video bylo přehráno, viz obrázek 6.20b a detekci lze vyřešit metodou TM. Ostatní typy příspěvků se přehrávají ve smyčce, což znamená, že je potřeba využít stejného přístupu jako u lokalit story nebo reels. Problémem je, že uživatel může kdykoliv obsah posunout a znemožní tak detekci, jelikož i když se ořízne pouze obsah příspěvku, rozměry pro porovnání nemusí být totožné. Pokud si během sledování chce uživatel přečíst popis, tak musí obsah posunout dolů a tím zmizí horní část. Možným řešením je oříznutí referenčního i stávajícího obrázku obsahu s ohledem na vzdálenost a směr pohybu. Nevýhodou je, že podobnost snímků se neměří z celého obsahu a dochází ke ztrátě informací — může dojít ke špatné detekci.



(a) ukázka See–Zoom interakce

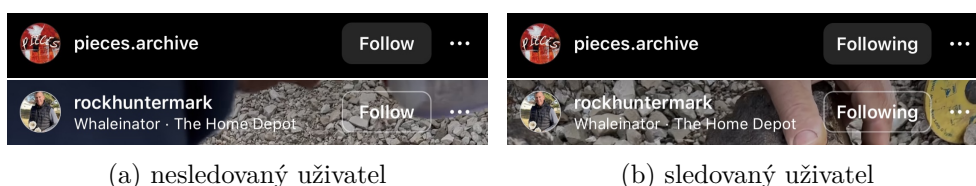
(b) See–WatchWhole typ reels

Obrázek 6.20: Zoom (a) a WatchWhole (b) interakce pro feed.

Ačkoliv se save ikona na obrazovce vyskytuje vždy, Save–SavePostIG může

proběhnout dvěma způsoby. Pokud uživatel nemá žádné uložené kolekce, interakci definuje pouze stisknutí ikony save, která se vyplní a detekuje pomocí TM. V druhém případě je po stisknutí prezentována obrazovka saveIG. Vyhodnocení poté probíhá stejně jako v případě reels, viz obrázek 6.17c.

Poslední interakcí, kterou lze vykonat přímo na feed je Social-Follow. Jediná zpětná vazba, kterou uživatel dostane po stisknutí tlačítka follow je změna jeho textu, viz obrázek 6.21. Nezbývá tedy nic jiného, než každý snímek text ve výřezu tlačítka kontrolovat pomocí OCR.



(a) nesledovaný uživatel

(b) sledovaný uživatel

Obrázek 6.21: Ukázka zpětné vazby u Social-Follow na feed. Uživatel si zobrazil příspěvek od uživatele, kterého nesleduje (a) a následně kliknul na sledovat (b).

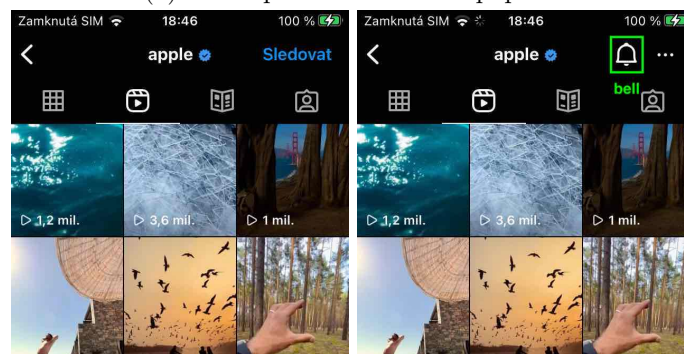
Profile

První interakce Search-BrowseProfile je detekována pokaždé, když je uživatel na lokaci profilu, profileFeed nebo korespondujících reels a story daného profilu. K rozpoznání je zapotřebí pouze typ obrazovky kromě reels, jelikož je možné z profilu přejít do reels, které profil nezveřejnil. Pro rozlišení je tedy nutné zkontrolovat, jestli inzerent příspěvku koresponduje s navštíveným profilem.

Druhou a poslední je Social-Follow, kterou lze vykonat tlačítkem v popisku profilu, tlačítkem v navigačním menu, pokud si uživatel prohlíží zed' příspěvků nebo na lokaci profileFeed, opět tlačítkem v navigačním menu, viz obrázky 6.22 a 3.7c. První dva způsoby lze vyřešit pomocí TM, ačkoliv jsou odlišné. Instagram totiž po provedení interakce zobrazí ikonu zvonečku (bell) v navigačním menu. Na třetí případ lze aplikovat segmentaci prahováním, jelikož po sledování se změní barva a text tlačítka, které po pár sekundách zmizí.



(a) follow pomocí tlačítka v popisku



(b) follow při prohlížení profilu

Obrázek 6.22: Follow interakce na profilu prvním (a) a druhým (b) způsobem.

Ostatní

Zde spadají interakce, jejichž detekce je totožná neohledně na lokalitu detekované reklamy a je možné je provést buď na jedné „společné“ nebo kdekoliv.

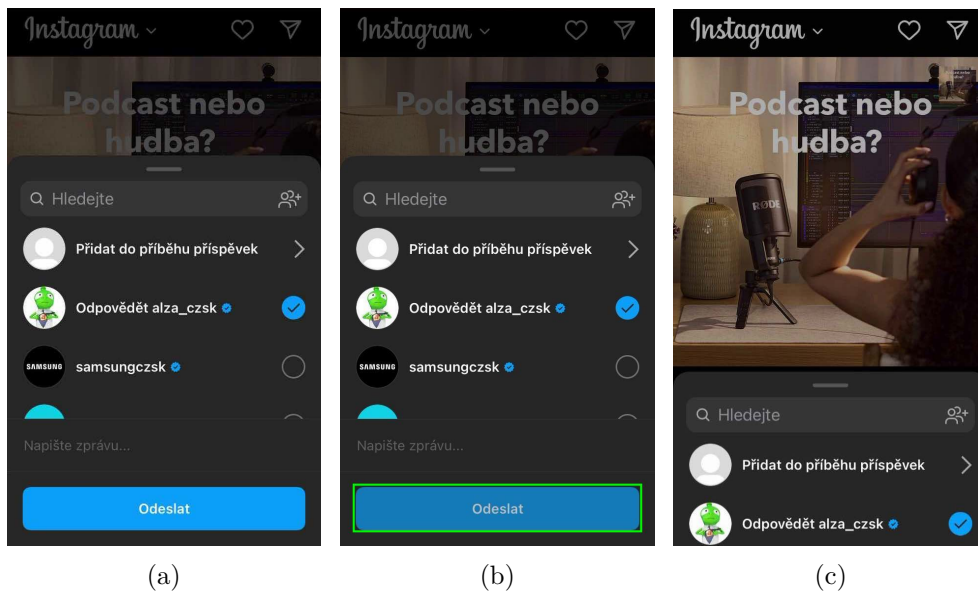
Search–Search je možné detekovat pouze na obrazovce searchIG a tehdy, kdy uživatel již narazil na reklamní sdělení. Určujícím prvkem interakce je text, který uživatel napíše do vyhledávacího pole. Nejdříve se tedy vysegmentuje vyhledávací pole podle barvy, pozice a tvaru. Následně stačí kontrolovat zadaný text, zda odpovídá uživatelskému jménu naposledy detekovaného reklamního sdělení.

Interakci Search–Redirect lze provést kdekoliv, kde se vyskytuje odkaz na externí webovou stránku nebo aplikaci. Detekce proběhne tehdy, když je lokalita externalWeb nebo appStore.

Jelikož je Save–Screenshot systémová funkce, není závislá na lokalitě a je umožněno ji provést kdekoliv. Jakmile uživatel udělá screenshot, celý snímek má chvíli uniformní odstín bílé a lokalita je neznámá. Algoritmus je tedy založen na počtu pixelů v obraze dané hodnoty a lokalitě.

Interakci Save–ShareIG je možné vykonat kdekoliv, kde se vyskytuje ikona share, například viz obrázek 6.18b. Součástí vykonání je vždy lokace shareIG,

kde uživatel musí vybrat příjemce a následně odeslat zmáčnutím tlačítka, viz obrázek 6.23. Zpětná vazba Instagramu, že byl nějaký obsah sdílen, se podle lokality, ze které byla interakce započata, liší (alert, specifická animace nebo žádná), proto je vhodné detekci provést přímo na lokalitě shareIG. Tlačítko po potvrzení výběru a odeslání jako zpětnou vazbu na malou chvíli změní barvu, následně se modální okno shareIG zavře. Detekce interakce je postavena na nalezení tlačítka pomocí rozměru a tvaru a následnou segmentací prahováním zmáčknutého stavu.



Obrázek 6.23: Ukázka interakce ShareIG. Vybrání příjemce (a), odeslání (b) a následné automatické zavření modálního okna (c).

Speciální interakce Shop–Shop, kterou je možné vykonat ze story, reels, feed nebo profile, je detekována na základě lokality shop.

Poslední speciální případ je Skip–Skip, kdy uživatel na příspěvku strávil méně než 1,5 sekundy. Stanovená hranice opět vyplývá z [82, 83].

Transakce a sktrukturovaný výstup

Rozpoznané interakce uživatele s reklamním sdělením je potřeba strukturovat. Výstup zpracování jednoho záznamu videa je reprezentován v JSON formátu, jehož parametry jsou popsány v tabulce 7.1. Nedefinované datové typy jsou popsány v následujících sekcích.

Tabulka 7.1: Atributy strukturovaného výstupu.

Atribut	Datový typ	Popis
amount	integer	Počet nalezených transakcí.
transactions	[Transaction]	Pole nalezených transakcí.

7.1 Transaction

Transakce reprezentuje reklamní sdělení, dodatečné informace o něm a veškeré interakce, které uživatel s reklamním sdělením učinil. V jednu chvíli se uvažuje pouze jedna transakce. Jakmile uživatel narazí na nové reklamní sdělení, které zveřejnil jiný inzerent, stávající transakce se ukončí a započne nová. Podrobný popis a datové typy jsou vidět v tabulce 7.2.

Atribut advertiser je reprezentováno polem, namísto řetězcem, jelikož platforma umožňuje zveřejnit příspěvek v „kolaboraci“ dvou účtů. Může se tedy stát, že uživatelské jméno bude uvedeno více než jedno. Také je vhodné zmínit, že délka transakce je odvozena od časových známek a délek jednotlivých interakcí, viz sekce 7.2.

Tabulka 7.2: Atributy objektu Transaction.

Atribut	Datový typ	Popis
duration	float	Doba trvání jedné transakce.
advertiser	[string]	Uživatelská jména inzerentů.
advertisementType	string	Typ reklamního sdělení: <i>sponsored</i> , <i>paid_promotion</i> , <i>organic</i> nebo <i>added</i> .
mediaType	string	Typ média reklamního sdělení: <i>photo</i> nebo <i>video</i> .
location	string	Lokace započetí transakce: <i>Story</i> , <i>Reels</i> nebo <i>Feed</i> .
actions	[Action]	Pole detekovaných interakcí.

7.2 Action

Akce reprezentuje interakci a reakční skupinu. Parametry jsou popsány v tabulce 7.3. Možné kombinace byly diskutovány v kapitole 3 a stručné shrnutí je popsáno v tabulce 7.4.

Tabulka 7.3: Atributy objektu Action.

Atribut	Datový typ	Popis
reaction	Reaction	Typ reakční skupiny.
type	ActionType	Typ interakce.
time	string	Časová známka uskutečnění interakce relativní k videu ve formátu "%Y-%m-%dT%H:%M:%S.%f".
duration	optional(float)	Doba trvání interakce.

Tabulka 7.4: Možné kombinace reakčních skupin a interakcí.

Reaction	ActionType	Kontinuální akce
see	stop	ano
	browseImages	ne
	expandText	ano
	watchWhole	ne
	zoom	ano
	comment	ano
	pause	ano
search	search	ne
	browseProfile	ano
	redirect	ano
save	screenshot	ne
	savePostIG	ne
	shareIG	ne
social	like	ne
	follow	ne
	comment	ne
shop	shop	ano
skip	skip	ne

7.3 Výstup

Detekované interakce se zaznamenávají pro každý snímek zvlášť. Jedna interakce je tedy objekt **Action**, který je označený časovou značkou reprezentující čas nálezu relativní k videu. Důsledkem je, že pokud si uživatel například prohlíží reklamní sdělení 3 sekundy, pole interakcí v transakci bude obsahovat 30 objektů **Action** z reakční skupiny see a typem interakce stop. Ačkoliv se jedná o validní výstup, může být poměrně obtížně čitelný.

Výstup lze převést na čitelnější variantu spojením kontinuálních akcí, viz tabulka 7.4, a přiřazením délky trvání odvozené z časových značek. Pokud je ale mezi dvěma kontinuálními akcemi větší časový rozestup, než je manuálně stanovená hranice (1,5 sekundy), nebudou spojeny. Příklady obou formátů výstupu je prezentován na obrázku 7.1.

Výjimkou jsou kontinuální akce Search–BrowseProfile a Search–Redirect, které se spojí nehledě na hranici rozestupu. Jedinou podmínkou je, že mezi nimi nesmí být další kontinuální akce, což u lokace externalWeb a appStore nemůže nastat. Důvodem jsou animace při přechodech a fakt, že při scrollování na obrazovce externalWeb se schovává horní i spodní lišta, která se využívá pro detekci lokace, což zapříčiní, že lokalita není detekována.

7. TRANSAKCE A SKTRUKTUROVANÝ VÝSTUP

```
{
  "amount": 1,
  "transactions": [
    {
      "location": "Story",
      "advertiser": [
        "aviko_gastro"
      ],
      "duration": 12.89679,
      "mediaType": "video",
      "advertisementType": "sponsored",
      "actions": [
        {
          "type": "stop",
          "reaction": "see",
          "time": "2023-04-05T19:53:59.056819"
        },
        ...
        {
          "type": "stop",
          "reaction": "see",
          "time": "2023-04-05T19:54:00.541342"
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": "2023-04-05T19:54:00.541342"
        },
        ...
        {
          "type": "stop",
          "reaction": "see",
          "time": "2023-04-05T19:54:11.953609"
        }
      ]
    }
  ]
}
```

(a) nespojené

```
{
  "amount": 1,
  "transactions": [
    {
      "location": "Story",
      "advertiser": [
        "aviko_gastro"
      ],
      "duration": 12.89679,
      "mediaType": "video",
      "advertisementType": "sponsored",
      "actions": [
        {
          "type": "stop",
          "reaction": "see",
          "time": "2023-04-05T19:53:59.056819",
          "duration": 12.89679
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": "2023-04-05T19:54:00.541342",
          "duration": 0.463913
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": "2023-04-05T19:54:04.902127",
          "duration": 0.742261
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": "2023-04-05T19:54:09.819608",
          "duration": 1.113392
        }
      ]
    }
  ]
}
```

(b) spojené

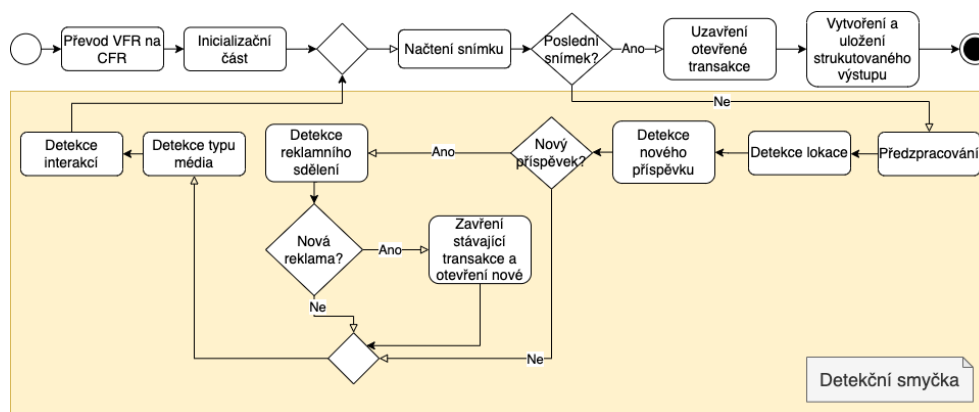
Obrázek 7.1: Ukázky strukturovaného výstupu. Detekované akce pouze podle časové známky (a) a sjednocené kontinuální akce (b).

Implementace

Kapitola popisuje detaily implementací jednotlivých částí aplikace. Obecná struktura zpracování je vidět na obrázku 8.1.

Pro prototypování byl použit jazyk Python v kombinaci s nástrojem Jupyter Notebook. Aplikace je napsána v jazyce Python. Mezi nejvýznamější použité knihovny patří Numpy³, OpenCV⁴, improutils⁵ a scikit-image⁶.

OCR je zajištěn knihovnou MMOCR⁷, která nabízí implementované modely pro detekci i čtení textu. Framework poskytuje i předtrénované modely na známých datasetech jako například icdar, synthtext nebo cute80, které jsou použity i v implementaci. Konkrétně model dbnet pro detekci textu a modely aster a master pro jeho rozpoznání [84].



Obrázek 8.1: Vývojový diagram zpracování snímků aplikací.

³<https://numpy.org>

⁴<https://opencv.org>

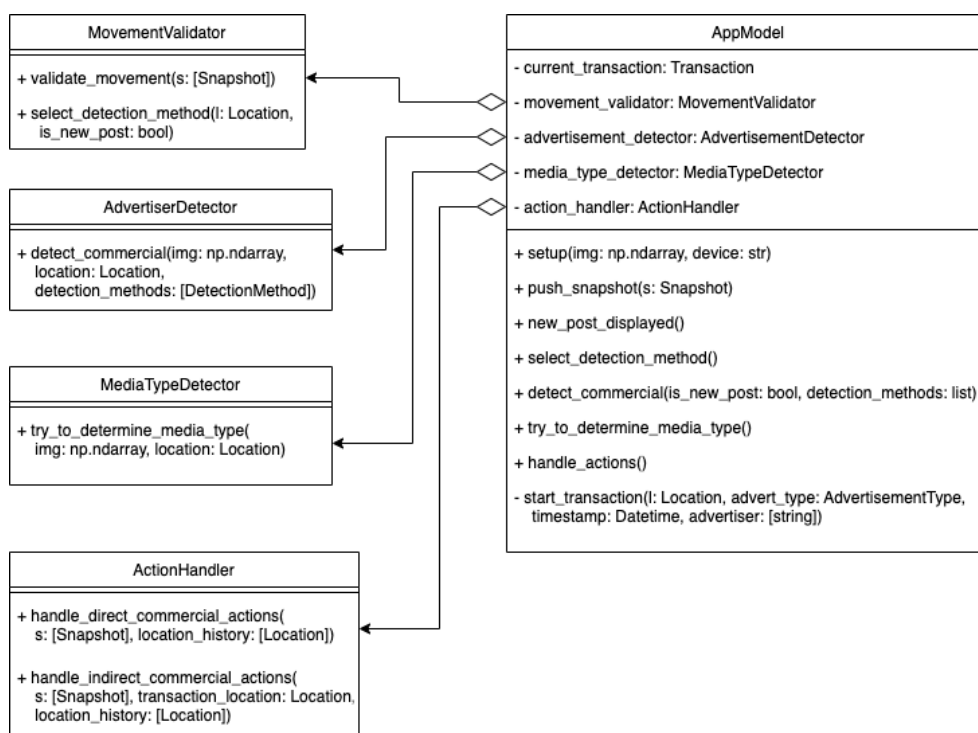
⁵https://github.com/ImprolabFIT/improutils_package

⁶<https://scikit-image.org>

⁷<https://github.com/open-mmlab/mocr>

Vývoj probíhal na vlastnoručně vytvořených záznamech ze zařízení iPhone 11 Pro a iPhone SE 1st gen.

Hlavní smyčka aplikace je naimplementována v souboru `app.py` a obsahuje třídy `LocationValidator` a `AppModel`, která slouží jako rozcestník pro ostatní třídy řešící jednotlivé podproblémy, drží aktuálně probíhající transakci a informace potřebné pro algoritmy jako škála, barevný režim nebo typ zařízení. Zjednodušený diagram třídy `AppModel` lze vidět na obrázku 8.2.



Obrázek 8.2: Zjednodušený diagram třídy `AppModel`.

8.1 Převod VFR na CFR

Prvním krokem je převedení videa do formátu CFR, který zajišťuje skript `cfr_converter.py`. Implementace využívá multiplatformní knihovny `ffmpeg`⁸, která dokáže na základě definovaných parametrů a filtrů upravovat vlastnosti fotografií, videí nebo audia.

8.2 Rozhraní

Jedná se o konzolovou aplikaci, která na vstupu přijímá:

⁸<https://ffmpeg.org>

- video záznam
- model zařízení (device) – "iPhone + <model>".
- barevný režim (skin) – "Light" nebo "Dark".

Dalším možným parametrem je pole uživatelských jmen oddělených čárkou (advertisers). Parametr slouží pro manuální přidání účtů, jejichž příspěvky budou automaticky detekovány jako reklama a jejich typ je ve výstupu reprezentován značkou *added*, viz tabulka 7.2.

Aplikace je vyvinuta pro český a anglický překlad Instagramu a díky co největšímu omezení používání OCR není potřeba lokalizaci v parametru specifikovat.

8.3 Inicializační část a předzpracování

Pro načtení FPS, trvání videa, datumu vytvoření videa a formátu profilu komprese je využito knihoven MediaInfo⁹, OpenCV a ffmpeg.

Následně se provede inicializace třídy `LocationValidator` a `AppModel`, která provádí výpočet škály metodou `setup`. Pro výpočet je k dispozici csv soubor obsahující veškeré informace o zařízeních, viz tabulka 6.1. Na základě zařízení, rozměrů videa a referenčních hodnot z tabulky se pomocí vzorce 6.1 vypočítá škála. Pokud na vstupu není zařízení k dispozici, je využit algoritmus popsáný v podsekcí 6.3.3, který je implementován pouze pro lokality feed, story a reels. Pokud jsou lokality reels nebo story, je nutné zadefinovat barevný režim. Vypočtená škála je následně použita pro úpravu rozměrů všech referenčních obrázků, které slouží pro detekci lokace a interakcí. Obrázky jsou uloženy ve složce `Templates`, kde jsou rozděleny do dalších složek podle typu obrazovky a následně barevného režimu.

Nakonec je na základě zařízení každý obrázek očištěn o systémové prvky a uložen do pole, jako objekt `Snapshot`, které reprezentuje paměť (zde 30 snímků do minulosti, což je ekvivalent 3 sekund). `Snapshot` obsahuje i časovou známku a lokaci snímku. Pokud se jedná o model, který nemá Home indicator, ani Dynamic Island, k oříznutí nedojde a je potřeba tento krok učinit až před detekcí jednotlivých interakcí v závislosti na lokalitě.

8.4 Detekce lokality

Rozpoznání lokality zajišťuje třída `LocationValidator`, která se snaží postupně detekovat jednotlivé nadefinované typy, dokud není nalezena správná. Pokud snímek neodpovídá ani jedné lokalitě, vrátí se neznámá. Jakmile je ale typ obrazovky určen správně, následující snímek už se kontroluje pouze pro stávající lokaci dokud nepřestane být validní, což vede k optimalizaci rychlosti.

⁹<https://mediaarea.net/en/MediaInfo>

Další optimalizací je omezení volání vyhodnocení lokality na základě podobnosti dvou snímků. Pokud si uživatel prohlíží fotografii, není důvod kontrolovat, zda se lokalita změnila. Validátor se tedy zavolá pouze v případě, když je SSIM index mezi dvěma snímky menší než nějaká hranice, která je experimentálně stanovena na hodnotu 0,95.

8.5 Detekce pohybu a nového příspěvku

Detekci nového příspěvku, a tedy i pohybu na vybraných lokalitách, zajišťuje třída `MovementValidator`. Metoda pro detekci pohybu má efekt pouze tehdy, když je lokalita story, reels, feed nebo profile.

Součástí třídy je i metoda, která vrací vhodné typy detekčních algoritmů reklamních sdělení. Určení vhodných metod pro detekci reklamy je důležité zejména pro lokaci feed, kde si uživatel může zobrazit příspěvek, který sice zabírá většinu obrazovky a je tedy validně klasifikován jako nový zobrazený, ale není mu vidět hlavička nebo CTA tlačítka.

8.6 Detekce reklamního sdělení

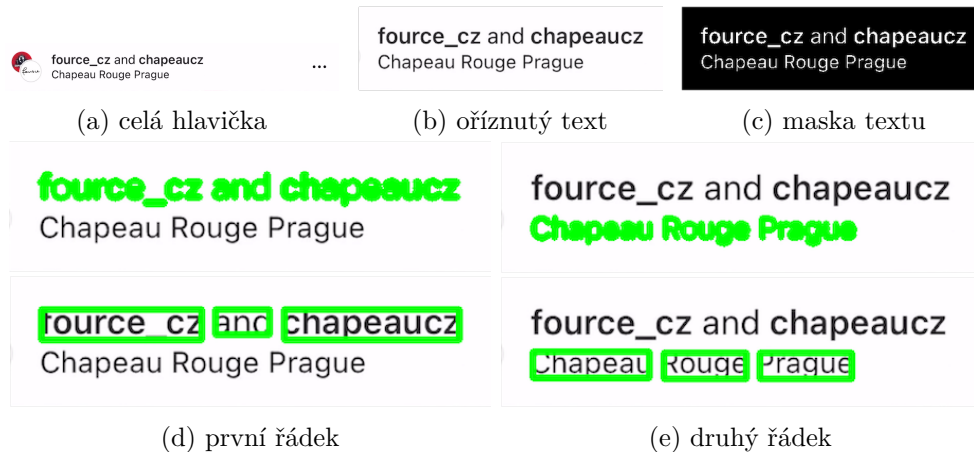
Jakmile je detekován nový příspěvek nebo bylo na základě pohybu vyhodnoceno aplikovat novou detekční metodu, zavolá se funkce pro rozpoznání reklamy ze třídy `AdvertisementDetector`. Funkce na vstupu přijímá obrázek, lokalitu a pole detekčních metod, které mají být aplikovány. Implementace uvažuje následující:

1. z hlavičky příspěvku
2. podle CTA tlačítek

Součástí je i získání uživatelského jména inzerenta, které se v hlavičce vyskytuje. Pokud je detekční metoda podle CTA tlačítek a lokalita je feed, získání uživatelského jména probíhá z popisku pod obsahem. Popisek nemusí být součástí příspěvku, což znamená, že pro detekci uživatelského jména musí uživatel následně hlavičku zobrazit. Story a reels mají hlavičku viditelnou vždy.

Pro detekci textu je využito tradičních metod zpracování obrazu, jelikož se vyskytuje vždy vedle profilového obrázku nebo ikon, jejichž lokace lze zjistit pomocí TM. Dalším důvodem je znak podtržítka, který se ve jménech vyskytuje poměrně často, i vícekrát za sebou, což činí detekčnímu modelu značné problémy. Text se tedy vysegmentuje prahováním, maska se upraví morfologickými operacemi a na základě nalezených kontur a jejich vzdáleností se písmena rozdělí do izolovaných slov a řádků, viz obrázek 8.3. Ty jsou následně jednotlivě rozpoznány pomocí OCR (model master).

Výjimkou je story, jelikož text pod uživatelským jménem je průhledný, což znamená, že segmentace prahováním nemusí stačit. Metoda tedy detekuje uživatelské jméno společně s popisem pod ním popsáním postupem. Pokud není popis nalezen, je použit OCR jak pro detekci, tak pro rozpoznání textu (dbnet a aster).



Obrázek 8.3: Segmentace a detekce slov z hlavičky feed příspěvku.

8.7 Uzavření a otevření transakce

Jakmile je stávající příspěvek vyhodnocen jako reklamní sdělení, aplikace otevře novou transakci se všemi doposud získanými informacemi (uživatelské jméno inzerenta, typ reklamního sdělení a lokalita). Pokud je nějaká otevřena, ukončí se, k čemuž dojde pokud uživatelské jméno nové transakce neodpovídá již probíhající. V opačném případě je zachována původní.

8.8 Detekce typu média

Rozpoznání typu média zajišťuje třída `MediaTypeDetector`, která na základě lokality použije vhodný algoritmus. Jakmile je médium určeno, přiřadí se otevřené transakci a detekce se ukončí, dokud není otevřena transakce nová.

8.9 Detekce interakcí

Implementované interakce, kterých je ve výsledku 26, jsou zobrazeny v tabulce 8.1.

Tabulka 8.1: Tabulka naimplementovaných interakcí podle lokalit.

Lokace	Interakce	Implementováno
story	See-Stop	ano
	See-ExpandText	ano
	See-WatchWhole	ne
	See-Pause	ano
	Social-Like	ano
reels	See-Stop	ano
	See-ExpandText	ano
	See-WatchWhole	ano
	See-Pause	ano
	See-Comment	ano
	Save-SavePostIG	ano
	Social-Like	ano
	Social-Follow	ano
Social-Comment	ano	
feed	See-Stop	ano
	See-BrowseImages	ano
	See-ExpandText	ne
	See-Zoom	ne
	See-WatchWhole (reels příspěvek)	ano
	See-WatchWhole (ostatní příspěvky)	ne
	See-Comment	ano
	Save-SavePostIG	ano
	Social-Like	ano
	Social-Follow	ne
Social-Comment	ano	
profile	See-BrowseProfile	ano
	Social-Follow	ano
obecné	Search-Search	ano
	Search-Redirect	ano
	Save-Screenshot	ano
	Save-ShareIG	ano
	Shop-Shop	ano
	Skip-Skip	ano

O detekci interakcí se stará třída `ActionHandler`, která obsahuje metody pro rozpoznání dvou typů interakcí:

- přímé – k validaci je nutné, aby lokalita odpovídala té, kde bylo detekováno reklamní sdělení
- nepřímé – lokalita neodpovídá té, kde bylo nalezeno reklamní sdělení

V obou případech musí být reklamní sdělení viditelné, což je do doby, než si uživatel zobrazí nový příspěvek, který není reklama, nebo si nezačne prohlížet profil, který neodpovídá stávajícímu inzerentovi. Rozdělení implementovaných interakcí podle typu na základě lokality je prezentováno v tabulce 8.2.

Tabulka 8.2: Přímé a nepřímé implementované interakce podle lokality.

Lokace	Interakce	Typ
story	See–Stop See–ExpandText See–Pause Social–Like	přímá nepřímá nepřímá/nepřímá přímá
reels	See–Stop See–ExpandText See–WatchWhole See–Pause See–Comment Save–SavePostIG Social–Like Social–Follow Social–Comment	přímá přímá přímá nepřímá nepřímá přímá/nepřímá přímá přímá nepřímá
feed	See–Stop See–BrowseImages See–WatchWhole (reels příspěvek) See–Comment Save–SavePostIG Social–Like Social–Comment	přímá přímá přímá nepřímá přímá/nepřímá přímá nepřímá
profile	See–BrowseProfile Social–Follow	nepřímá nepřímá
obecné	Search–Search Search–Redirect Save–Screenshot Save–ShareIG Shop–Shop Skip–Skip	nepřímá nepřímá nepřímá nepřímá nepřímá přímá

Interakce See–SavePostIG spadá do obou kategorií, jelikož její realizace je umožněna vícero cestami, které jsou naimplementovány separátně. See–Pause na lokaci story spadá také do obou kategorií, což je způsobeno optimalizační podmínkou detekce lokality, která požaduje, aby byly dva za sebou jdoucí snímky dostatečně odlišné, pro spuštění detekce lokality nové. Jelikož příspěvek na story nemá tolik prvků, které se při See–Pause schovají, nemusí být podmínka splněna, proto jsou implementovány oba typy.

Po splnění předpokladu viditelnosti reklamy je každý snímek zpracován příslušnou metodou ze třídy. Obě vrací pole detekovaných interakcí, které jsou následně ukládány do transakce.

8.10 Uzavření transakce

Při uzavření transakce se projdou veškeré detekované interakce a na základě časové známky první a poslední je určena celková délka manipulace uživatele s reklamním sdělením. Pokud je délka menší, než stanovená hranice pro Skip–Skip a zároveň v poli detekovaných interakcí je pouze See–Stop, jsou nahrazeny jednou Skip–Skip.

8.11 Evaluace

Pro evaluaci úspěšnosti je využit formát se sjednocenými kontinuálními interakcemi. Ke každému výstupu aplikace je potřeba vytvořit referenční řešení, což je JSON soubor s upravenou strukturou pro pohodlnější definici. Namísto časových značek u každé interakce se definuje čas jejího výskytu od začátku videa v sekundách. Aby bylo možné dopočítat reálné přesné časy interakcí, je nutné do reference přidat časovou známku vzniku videa. Příklad referenčního řešení je vidět na obrázku 8.4.

I přes zjednodušení je obtížné manuálně získávat přesné časy akcí z videa. Porovnání tedy počítá s tolerancí jedné sekundy a pro definici reference tak stačí odhad. Tolerance je uplatněna i pro délku trvání transakce. Výjimkou je údaj času vytvoření videa, jelikož pokud bude údaj posunutý například o půl sekundy, posunuté budou i dopočítané časy detekovaných interakcí.

Evaluaci zajišťuje skript `test_evaluation.py`. Po načtení hodnot řešení a výstupu jednoho záznamu je na základě časové známky porovnán výskyt transakcí. U správně nalezených proběhne kontrola veškerých obecných údajů jako lokace, uživatelské jméno inzerenta, délka trvání, typ média a typ reklamního sdělení. Naposled jsou zkontrolovány interakce u každé transakce podle časové známky, reakční skupiny a samotné interakce. Tímto způsobem skript porovná veškeré výstupy a reference. Výsledkem je text popisující statistiky úspěšnosti detekce transakcí, interakcí, popřípadě detaily vyhodnocení jednotlivých záznamů.

```
{
  "amount": 1,
  "transactions": [
    {
      "location": "Reels",
      "advertiser": [
        "alza"
      ],
      "duration": 13.14482,
      "mediaType": "video",
      "advertisementType": "sponsored",
      "actions": [
        {
          "type": "stop",
          "reaction": "see",
          "time": "2023-04-05T19:53:59.056819",
          "duration": 13.14482
        },
        {
          "type": "screenshot",
          "reaction": "save",
          "time": "2023-04-05T19:54:00.541342",
          "duration": null
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": "2023-04-05T19:54:04.902127",
          "duration": 0.742261
        },
        {
          "type": "like",
          "reaction": "social",
          "time": "2023-04-05T19:54:08.020000",
          "duration": null
        }
      ]
    }
  ]
}
```

(a) výstup

```
{
  "amount": 1,
  "vid_creation_date": "2023-04-05T19:53:55.014156",
  "transactions": [
    {
      "location": "Reels",
      "advertiser": [
        "alza"
      ],
      "duration": 13,
      "mediaType": "video",
      "advertisementType": "sponsored",
      "actions": [
        {
          "type": "stop",
          "reaction": "see",
          "time": 4,
          "duration": 13
        },
        {
          "type": "screenshot",
          "reaction": "save",
          "time": 5,
          "duration": null
        },
        {
          "type": "pause",
          "reaction": "see",
          "time": 10,
          "duration": 0.75
        },
        {
          "type": "like",
          "reaction": "social",
          "time": 13,
          "duration": null
        }
      ]
    }
  ]
}
```

(b) referenční řešení

Obrázek 8.4: Výstup aplikace (a) a korespondující referenční řešení (b).

Výsledky

Výsledky byly zhodnoceny na dvou datových sadách. První obsahuje vlastnoručně pořízené nahrávky s lokacemi feed, story a reels. Důraz je kladen především na hustotu jednotlivých interakcí v každém záznamu. Druhý dataset obsahuje čtyři videa z reálného sběru, kde je hustota interakcí velmi nízká. Zpracování videí probíhalo na zařízení MacBook Pro (13-inch, M1, 2020).

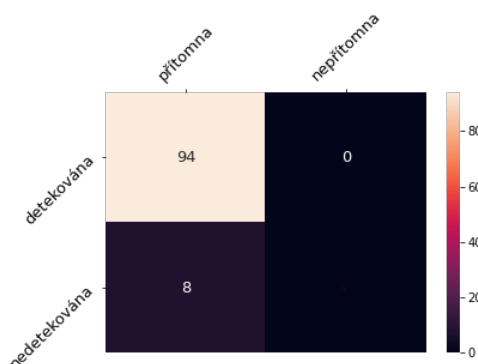
9.1 Vlastní dataset

Celková délka datasetu je 21 minut, obsahuje 102 transakcí (včetně uměle vytvořených pomocí parametru *advertisers*) a záznamy z obou barevných režimů na zařízení iPhone 11 Pro. Samotné zpracování bez převodu na CFR formát, inicializační části a vytváření strukturovaného výstupu, zabralo 113 minut, což je zhruba pětikrát více než délka datasetu. Rychlost zpracování se odvíjí od lokality, jelikož feed se zpracovává déle kvůli většímu počtu FPS. Dalším faktorem je počet příspěvků, na které uživatel narazí, protože použití OCR je časově náročné.

Aplikace dokázala detekovat 94 transakcí (TP), falešně detekovala 0 (FP) a nedetekovala 8 (FN), viz obrázek 9.1, což odpovídá úspěšnosti 92,16 %. Důvodem 8 nedetekovaných je rychlé projíždění příspěvků, především na story a reels, jelikož pro zapnutí detekce se nejdříve musí obsah přestat hýbat. Možnou opravou je zvýšení zpracovávaných FPS nebo snížení tolerance pohybu příspěvku výměnou za přesnost detekce.

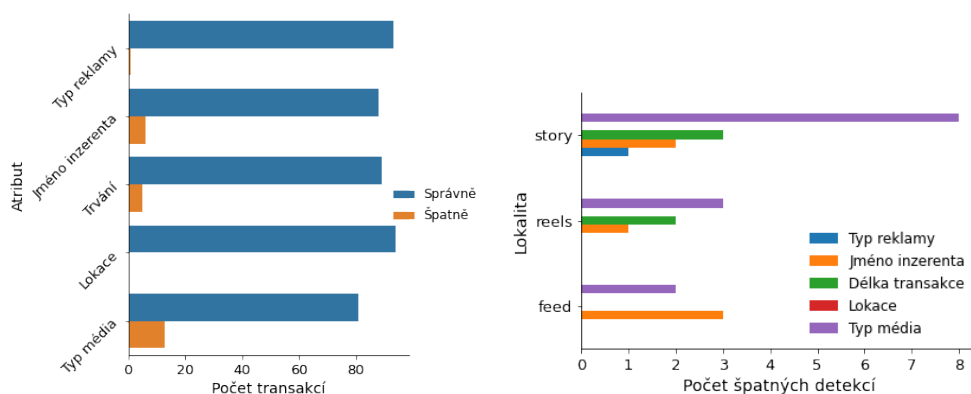
Případ správného nedetekování (TN) není v testech zohledněn, jelikož standardní příspěvky se ve strukturovaném výstupu nevyskytují. Obecně je v datasetu majorita reklamních sdělení typu *sponsored*, jelikož ale u všech typů (kromě *organic*) detekce probíhá na základě informací v hlavičce, jde především o správné určení uživatelského jména nebo textu pod ním. Pro typ *sponsored* je nutno provést oba kroky.

9. VÝSLEDKY



Obrázek 9.1: Confusion matrix detekce reklamních sdělení (transakcí).

U správně detekovaných transakcí jsou následně porovnány příslušející atributy. Počet transakcí na story je 33, na reels 20 a 41 na feed. Z obrázku 9.2a je vidět, že typ reklamního sdělení, délka transakce a lokace je určena převážně správně. Hůře dopadla detekce inzerenta, ale ve většině případů se jedná pouze o jeden nesprávně rozpoznávaný znak, což je způsobeno použitým modelem OCR. Určení typu média transakce dopadlo nejhůře. Z obrázku 9.2b plyne, že problém je především u lokací reels a story, kde je k určení použit SSIM index z prvních deseti snímků. Pokud se ve videu ze začátku nic nemění, jeví se jako fotografie a tím dojde i ke špatné detekci.

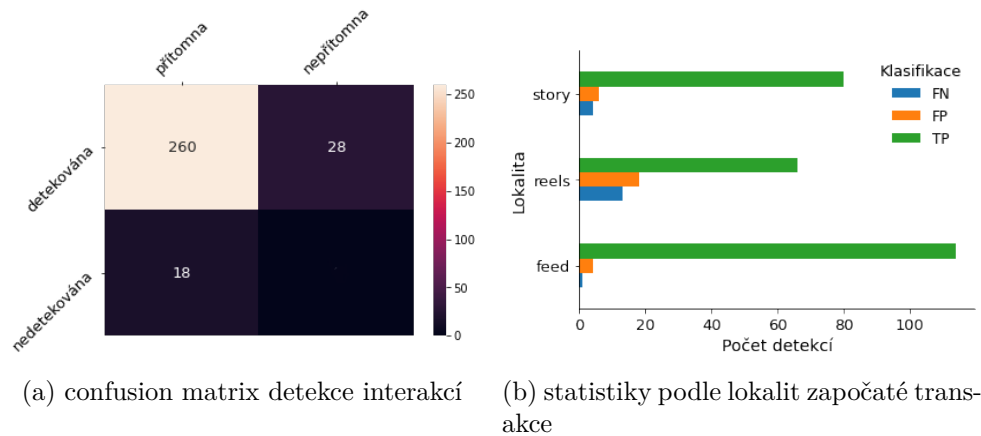


(a) počet správných a špatných detekcí (b) špatné detekce atributů podle lokalit atributů pro všechny lokace

Obrázek 9.2: Výsledky detekce atributů.

Data pro vyhodnocení interakcí se vztahují pouze na správně detekované transakce. Výsledky jsou rozděleny podle lokality, která značí, kde daná transakce započala. To znamená, že jsou zahrnuty i obecné interakce jako například Search-BrowseProfile, kterou lze vykonat na lokalitě jiné. Statistiky detekcí

všech interakcí souhrnně a podle lokalit lze vidět na obrázku 9.3, kde z počtu TP a FN plyne úspěšnost detekce 93,53 %. Ze správně detekovaných interakcí, které jsou kontinuální, se podařilo správně určit délku trvání ve 137 případech ze 149, což odpovídá úspěšnosti 91,94 %.



Obrázek 9.3: Počty TP, FN a FP detekcí interakcí.

Detailní údaje vyhodnocení pro story lze vidět v tabulce 9.1. Nejvíce problematická je interakce Search–Redirect. Po stisknutí CTA tlačítka, které odkazuje na webovou stránku, se zobrazí prázdná a částečně průhledná stránka s lokací feed na pozadí, což znamená, že se určí feed. Jelikož pro novou detekci je nutné, aby byly dva po sobě jdoucí snímky dostatečně odlišné, k nové detekci dojde až po načtení webu nebo vůbec, protože rozdíl snímků pořád nepřekonal stanovenou hranici. Důsledkem je většinou špatný čas určení interakce a zařazení do FP pro daný čas a FN vůči referenčnímu řešení.

Tabulka 9.1: Statistiky detekcí interakcí pro story.

Interakce	TP	FN	FP
See–Stop	36	0	1
See–ExpandText	2	0	0
See–Pause	10	0	0
Search–BrowseProfile	5	1	2
Search–Redirect	3	2	1
Search–Search	1	0	0
Save–Screenshot	4	0	0
Save–ShareIG	2	0	0
Social–Follow	3	0	2
Social–Like	4	0	0
Shop–Shop	1	0	0
Skip–Skip	9	1	0

Výsledky reels, viz tabulka 9.2, jsou nejhorší z trojice lokalit. U See–ExpandText dochází ke špatné detekci kvůli velké dynamice celého popisku, což lze napravit odladěním dalších krajních případů. Implementace See–Pause spoléhá na to, že lokalita je v moment interakce neznámá, což je vždy pravda. Bohužel to platí například pro animace při kliknutí na odkaz nebo zobrazení modálního okna. Algoritmus se tyto případy snaží vyřešit na základě podobnosti mezi dvěma snímky. Pokud je rozdíl příliš velký předpokládá se, že se nejedná o See–Pause. Reels má ale většinou na obrazovce velké množství prvků, které při schování zapříčiní, že se možnost interakce zamítne a není detekována. Nejhuře dopadla interakce See–WatchWhole, konkrétně případ bez progress bar, jelikož není nikdy jisté, jaké charakteristiky video má a neexistuje konkrétní ukazatel nového přehrání. Pokud je tedy první část videa neměnná nebo se opakují některé části ze začátku v průběhu, dojde k detekci. Video se přehrává i po rozbalení popisku či modálního okna, což znemožní interakci detekovat, jelikož není vidět.

Tabulka 9.2: Statistiky detekcí interakcí pro reels.

Interakce	TP	FN	FP
See–Comment	3	1	1
See–ExpandText	7	0	8
See–Pause	1	8	1
See–Stop	31	0	1
See–WatchWhole	3	2	7
Search–BrowseProfile	1	0	0
Search–Redirect	2	0	0
Save–SavePostIG	3	0	0
Save–Screenshot	1	1	0
Save–ShareIG	2	0	0
Social–Comment	1	1	0
Social–Follow	3	0	0
Social–Like	7	0	0
Skip–Skip	1	0	0

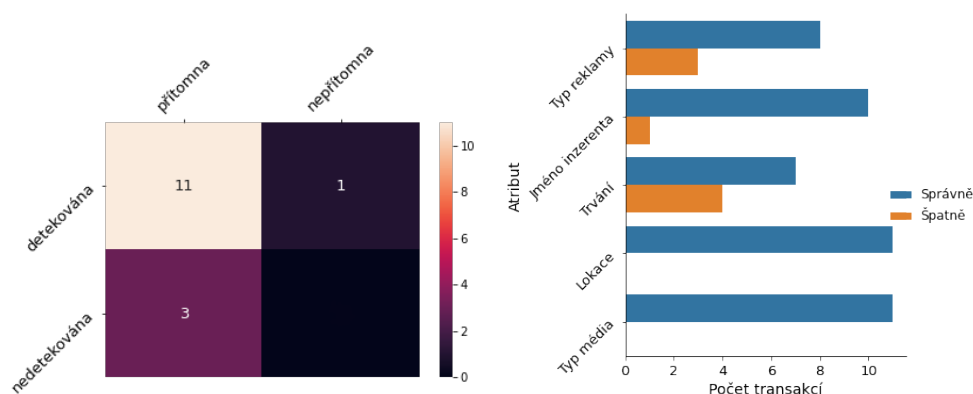
Poslední lokalitou je feed, viz tabulka 9.3, která dopadla nejlépe. Jedinou problematickou interakcí je Social–Follow (na profilu). Pokud jde uživatel z příspěvku na profil a začne ho sledovat, interakce je správně detekována. Pokud následně navštíví například profileFeed a po chvíli se vrátí zpět na profil, Social–Follow je na základě ikony, která je zobrazena vždy, když je profil sledován, rozpoznána podruhé. Nejedná se tedy o problém s algoritmem rozpoznání, ale o ošetření krajních případů pohybu uživatele. Stejný problém se vyskytuje u všech interakcí, které jsou detekovány pomocí TM.

Tabulka 9.3: Statistiky detekcí interakcí pro feed.

Interakce	TP	FN	FP
See-BrowseImages	12	0	0
See-Comment	4	0	0
See-Stop	32	0	0
See-WatchWhole	2	0	0
Search-BrowseProfile	6	0	0
Search-Redirect	4	0	0
Search-Search	2	0	0
Save-SavePostIG	5	0	0
Save-Screenshot	2	0	0
Save-ShareIG	4	0	0
Social-Comment	4	0	0
Social-Follow	1	1	2
Social-Like	10	0	2
Shop-Shop	1	0	0
Skip-Skip	25	0	0

9.2 Reálná data

Dataset je celkově dlouhý 20 minut a zpracování trvalo 3,65krát déle, tedy 73 minut. Rychlost je větší, než v případě přechozího datasetu, kvůli nízkému počtu transakcí, kterých je celkově 12. Nahrávky jsou ve světlém barevném režimu a podle odvozené škály pocházejí ze zařízení iPhone 14 Plus, 13 Pro Max nebo 12 Pro Max. Většina transakcí je uměle přidána parametrem `advertisers`, aby mohla být otestována správnost detekce. Z obrázků 9.4a a 9.4b plyne, že úspěšnost detekce transakcí je srovnatelná s vlastním datasetem, což bohužel nelze tvrdit o určení atributů. Největší problém činí určení typu reklamy a trvání transakce, která je odvozená od detekovaných interakcí. Důvodem je hlavně neodladěnost různých krajních scénářů, než samotné algoritmy pro detekci. Z hodnot TP a FN je úspěšnost detekce transakce 78,57 %.

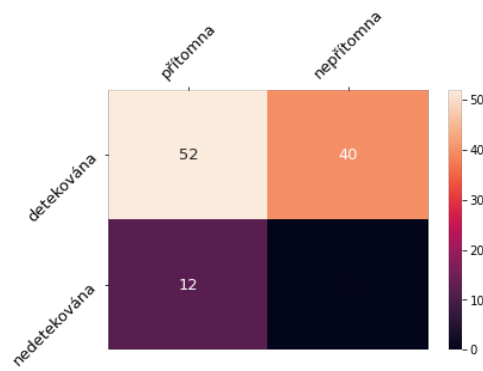


(a) počet správných a špatných detekcí (b) špatné detekce atributů podle lokalit atributů pro všechny lokace

Obrázek 9.4: Výsledky detekce transakcí pro reálná data.

Ze statistiky úspěšnosti určení interakcí ze správně detekovaných transakcí (celkově 64), viz obrázek 9.5, lze vidět značné zhoršení oproti vlastnímu datasetu. Ačkoliv z hodnot TP a FN plyne celková úspěšnost 81,25 %, počet FP je poměrně velký. Korektní určení délky detekovaných kontinuálních interakcí se povedlo v 78,57 % případech, konkrétně u 22 z 28.

Důvodem výsledků je opět neodladění krajních případů. Příkladem může být různorodost animací při prokliku na profileFeed z profilu. Příspěvek se v závislosti na umístění zobrazí vždy s jinou animací. To způsobí, že je detekována lokalita feed dříve než profileFeed, jelikož horní lišta, která profileFeed definuje, ještě nemusí být na svém místě. Kvůli optimalizaci podobnosti snímků je poté rozpoznání lokality zavoláno až po posunutí obrazovky nebo větší změně v příspěvku, pokud se jedná o video. Následkem je, že interakce Search–BrowseProfile je detekována dvakrát a zároveň neodpovídá délka. Dalším problémem je fakt, že Search–BrowseProfile musí být detekována, i když je uživatel na korespondujících příspěvcích profilu, které jsou reprezentovány lokalitami profileFeed, reels nebo story. V případě, že uživatel přišel na profil z reels nebo story, nemusí být z historie lokalit zřejmé, zda se uživatel vrátil na originální příspěvek nebo je stále na profilu. Rozpoznání podle inzerenta nepomůže, protože je téměř vždy totožný. Aplikace poté není schopna správně určit, zda se jedná o příspěvek na profilu nebo originální, což způsobí nesprávnou detekci jak interakcí, tak délky transakce, čímž se zvedne počet FP. Stejně jako v případě vlastního datasetu, je i zde problém s falešnou detekcí interakcí založených na TM kvůli neošetřeným případům pohybu.



Obrázek 9.5: Úspěšnost detekce interakcí pro reálná data.

Diskuze

Algoritmy navržené na dílčí podproblémy týkající se detekce reklamy, zobrazení nového příspěvku či lokality fungují izolovaně a na jednoduché scénáře slibně. To samé lze říci o interakcích založených na TM, segmentaci prahováním a detekci lokality. Hůře dopadly detekce založené na podobnosti snímků pomocí SSIM indexu, jelikož nejsou známy charakteristiky obsahu příspěvků.

Správnost rozpoznání transakce a korespondujících interakcí spolu úzce souvisí. Příkladem je profil, kde není povolena detekce nových transakcí. Díky velkému počtu obrazovek a alternativních scénářů, které má uživatel možnost na profilu provést, lze narazit na neošetřený případ, což může zapříčinit nesprávnost délky trvání interakcí nebo opakované detekce. Hlavní práce tedy spočívá v odlazování jednotlivých krajních případů a scénářů.

Rychlost zpracování lze vylepšit spuštěním aplikace na výkonějším zařízení nebo na grafické kartě, která je vhodnější pro OCR modely a algoritmy zpracování obrazu.

Operační systém iOS se pravidelně každý rok s příchodem nové verze mění, což může narušit systémovou interakci Save-Screenshot. Aplikace nebere v potaz systémové hlášky, push notifikace, ani podpůrné funkcionality poskytované operačním systémem jako AssistiveTouch, systémový zoom nebo Reachability. Zároveň je předpokládáno, že velikost systémového textu má výchozí hodnotu.

Díky velkému počtu uživatelů přichází Instagram s častými aktualizacemi. Podle záznamů z AppStore [76] se nová verze zveřejňuje každý týden. Vývoj nad takto živou aplikací je tudíž velmi náročný a vyžaduje častou údržbu a kontrolu jednotlivých interakcí. Během implementace proběhly změny rozhraní obrazovek feed, reels, shareIG a přibyla saveIG, čímž bylo potřeba implementovat nové algoritmy nebo upravovat stávající. Instagram je proto podporován do verze 275.1, která byla vydána 21.03.2023.

Bohužel se ukazuje, že nové funkcionality prochází nejdříve fází testování, tedy nemusí být přístupné všem uživatelům [85, 86]. Ke změně uživatelského rozhraní může tedy dojít i bez nutnosti aktualizace, pokud má uživatel nain-

stalovanou verzi, ve které je funkcionalita implementována a je skrytá. Projevení změny je následně závislé na Instagramu a ne na nové verzi. Důsledkem jsou možné lišící se UI prvky a kroky interakcí pro jednotlivé účty, což se potvrdilo z pozorování dlouhodobého a nově vytvořeného účtu, kde se lišila horní lišta obrazovky feed. Dostupné funkcionality jsou také ovlivněny zemí, kde uživatel Instagram používá. Příkladem je messenger funkcionalita [87], která není v Česku dostupná, čímž nelze vyloužit i absenci jiných. UI změny se projevují také mezi různými zařízeními, typicky na základě displeje (zařízení bez Home indicatoru vs. s Home indicatorem vs. s Dynamic Island), jelikož mají jiné systémové ovládací prvky a velikost. Ačkoliv změny nemusí být drastické, je žádoucí pro další rozvoj využít co nejvíce verzí a zařízení.

Závěr

Cíl práce je zaměřen na automatizaci detekce reklamních sdělení a definovaných interakcí uživatelů ze záznamů obrazovky mobilní aplikace Instagram na platformě iOS. Řešení zahrnuje návrh a implementaci rozpoznání lokality, nového příspěvku, reklamního sdělení, typu média a nakonec jednotlivých interakcí. Dílčí podproblémy jsou implementovány v aplikaci, která na vstupu přijímá video, barevný režim a model zařízení. Výstupem jsou dva strukturované textové soubory obsahující získané informace ve formě transakcí, které popisují reklamní sdělení, dodatečné informace a interakce uživatele.

Výsledná aplikace na datové sadě zaměřené zejména na hustotu interakcí dokázala detekovat transakci v 94 případech ze 102. U korektně detekovaných transakcí bylo určeno správně celkově 260 interakcí z 278. Na datové sadě reálných nahrávek uživatelů byla detekována transakce v 11 ze 14 případů a navazující interakce v 52 z 64. Bohužel kvůli velkému množství možností pohybu uživatele a počtu krajních případů je výstupem i 40 případů falešně detekovaných interakcí.

Práce staví na výzkumu chování a manipulace uživatelů s reklamními sděleními na sociálních sítích. Strukturovaný výstup bude následně použit pro analýzu chování uživatelů ve spolupráci se společností Tracey.ai, s. r. o.

Literatura

- [1] Martínek, P. A.: Brand-lurking Phenomenon on Social Media [online]. 2021, [cit. 2023-04-06]. Dostupné z: <https://dspace.cuni.cz/bitstream/handle/20.500.11956/152558/140096000.pdf?sequence=1&isAllowed=y>
- [2] Bao, L.; Li, J.; Xing, Z.; aj.: Reverse engineering time-series interaction data from screen-captured videos. In *2015 IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, 2015, s. 399–408, doi:10.1109/SANER.2015.7081850. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/7081850>
- [3] How to Use Background Subtraction Methods [online]. OpenCV docs. [cit. 2023-04-30]. Dostupné z: https://docs.opencv.org/4.x/d1/dc5/tutorial_background_subtraction.html
- [4] Morphological Transformations [online]. OpenCV docs. [cit. 2023-04-30]. Dostupné z: https://docs.opencv.org/master/d9/d61/tutorial_py_morphological_ops.html
- [5] M. Jiřina, L. B., J. Novák: BI-SVZ 9. přednáška. *Segmentace obrazu – Houghova transformace a plošné segmentace (BI-SVZ)* [online]. 2023. Poslední změna 2023-09-23 [cit. 2023-04-30]. Dostupné z: <https://courses.fit.cvut.cz/BI-SVZ/lectures/files/bi-svz-09-segmentace-obrazu-plosne.pdf>
- [6] Structural similarity index [online]. © scikit-image development team. [cit. 2023-04-30]. Dostupné z: https://scikit-image.org/docs/stable/auto_examples/transform/plot_ssim.html
- [7] Digital advertising spending in Europe from 2006 to 2022 [online]. Statista. [cit. 2023-04-06]. Dostupné z: <https://www.statista.com/statistics/307005/europe-online-ad-spend/>

- [8] Tsai, W.-H. S.; Men, L. R.: Consumer engagement with brands on social network sites: A cross-cultural comparison of China and the USA [online]. *Journal of Marketing Communications*, ročník 23, č. 1, 2017: s. 2–21, [cit. 2023-04-06].
- [9] Benevenuto, F.; Rodrigues, T.; Cha, M.; aj.: Characterizing User Behavior in Online Social Networks [online]. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*, IMC '09, New York, NY, USA: Association for Computing Machinery, 2009, ISBN 9781605587714, str. 49–62, doi:10.1145/1644893.1644900, [cit. 2023-04-06]. Dostupné z: <https://doi.org/10.1145/1644893.1644900>
- [10] 80 Trustmary. [cit. 2023-04-06]. Dostupné z: <https://trustmary.com/marketing/80-of-content-marketing-is-targeted-at-a-wrong-target-audience-avoid-this-common-mistake/>
- [11] Briggs, F.: Almost Half Of Consumers Will Reject Brands Sending Irrelevant Or Too Many Ads, Study Shows [online]. *Forbes*. Poslední změna 2016-04-18 [cit. 2023-04-06]. Dostupné z: <https://www.forbes.com/sites/fionabriggs/2016/04/18/almost-half-of-consumers-will-reject-brands-sending-irrelevant-or-too-many-ads-study-shows/?sh=316707f6769b>
- [12] Gartner Survey Shows Brands Risk Losing 38 Percent of Customers Because of Poor Marketing Personalization Efforts [online]. © 2023 Garter, Inc. Poslední změna 2019-03-11 [cit. 2023-04-06]. Dostupné z: <https://www.gartner.com/en/newsroom/press-releases/2019-03-11-gartner-survey-shows-brands-risk-losing-38-percent-of>
- [13] Kaushik, A.: See, Think, Do, Care Winning Combo: Content +Marketing +Measurement! [online]. [cit. 2023-04-06]. Dostupné z: <https://www.kaushik.net/avinash/see-think-do-care-win-content-marketing-measurement/>
- [14] Vollrath, M. D.; Villegas, S. G.: Avoiding digital marketing analytics myopia: revisiting the customer decision journey as a strategic marketing framework [online]. *Journal of Marketing Analytics*, 2021: s. 1–8. Dostupné z: <https://link.springer.com/article/10.1057/s41270-020-00098-0>
- [15] Audy Martínek, P.: Mapping methods of research on consumer engagement with brands on social media: A literature review [online]. *Methodological innovations*, ročník 14, č. 1, 2021: str. 2059799120985384, [cit. 2023-04-06]. Dostupné z: <https://journals.sagepub.com/doi/full/10.1177/2059799120985384>

-
- [16] About instagram [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: <https://www.statista.com/statistics/183585/instagram-number-of-global-users/>
- [17] Chaffey, D.: Global social media statistics research summary 2023 [online]. © Smart Insights (Marketing Intelligence) Ltd. [cit. 2023-04-07]. Dostupné z: <https://www.gartner.com/en/newsroom/press-releases/2019-03-11-gartner-survey-shows-brands-risk-losing-38-percent-of>
- [18] Number of Instagram users worldwide from 2020 to 2025 [online]. Statista. [cit. 2023-04-07]. Dostupné z: <https://www.statista.com/statistics/183585/instagram-number-of-global-users/>
- [19] Macready, H.: What is the Difference Between the 3 Instagram Profile Types [online]. © 2023 Business 2 Community. Poslední změna 2022-12-14 [cit. 2023-04-08]. Dostupné z: <https://www.business2community.com/instagram/what-is-the-difference-between-the-3-instagram-profile-types-02300390>
- [20] Add Call-To-Action. [online]. © 2023 Meta. [cit. 2023-04-07]. Dostupné z: <https://developers.facebook.com/docs/instagram/ads-api/guides/call-to-action/>
- [21] About Available Partners and Countries for Food Orders on Instagram [online]. © 2023 Meta. [cit. 2023-04-07]. Dostupné z: https://www.facebook.com/help/instagram/897533104016091?helpref=faq_content
- [22] Explore what's possible with Instagram ads. [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: <https://business.instagram.com/advertising>
- [23] About Instagram ads [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: https://help.instagram.com/537518769659039/?helpref=related_articles
- [24] Ads Manager. Create and manage ads on desktop and mobile. [online]. © 2023 Meta. [cit. 2023-04-07]. Dostupné z: <https://business.facebook.com/business/tools/ads-manager>
- [25] How Instagram decides which ads to show you. [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: https://help.instagram.com/173081309564229/?helpref=related_articles
- [26] About Instagram ads [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: https://help.instagram.com/526914357496585/?helpref=related_articles

- [27] How to use the paid partnership label to tag organic branded content on Instagram [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: <https://help.instagram.com/1109894795810258>
- [28] Branded Content Policies [online]. © 2023 Instagram from Meta. [cit. 2023-04-07]. Dostupné z: https://help.instagram.com/1695974997209192?helpref=faq_content
- [29] Martínek, P. A.; Caliandro, A.; Denegri-Knott, J.: Digital practices tracing: studying consumer lurking in digital environments [online]. *Journal of Marketing Management*, ročník 39, č. 3-4, 2023: s. 244–274, doi:10.1080/0267257X.2022.2105385, <https://doi.org/10.1080/0267257X.2022.2105385>. Dostupné z: <https://doi.org/10.1080/0267257X.2022.2105385>
- [30] Grow your business with Google Ads [online]. © 2023 Google. [cit. 2023-04-13]. Dostupné z: https://ads.google.com/intl/cs_cz/home/
- [31] Anupam, V.; Freire, J.; Kumar, B.; aj.: Automating Web navigation with the WebVCR [online]. *Computer Networks*, ročník 33, č. 1, 2000: s. 503–517, ISSN 1389-1286, doi:[https://doi.org/10.1016/S1389-1286\(00\)00073-6](https://doi.org/10.1016/S1389-1286(00)00073-6). Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1389128600000736>
- [32] Atterer, R.; Wnuk, M.; Schmidt, A.: Knowing the User's Every Move: User Activity Tracking for Website Usability Evaluation and Implicit Interaction [online]. In *Proceedings of the 15th International Conference on World Wide Web, WWW '06*, New York, NY, USA: Association for Computing Machinery, 2006, ISBN 1595933239, str. 203–212, doi:10.1145/1135777.1135811. Dostupné z: <https://doi.org/10.1145/1135777.1135811>
- [33] Framework – Background Tasks [online] © 2023 Apple Inc. [cit. 2023-04-12]. Dostupné z: <https://developer.apple.com/documentation/backgroundtasks>
- [34] Ram, A.; Jalal, S.; Jalal, A. S.; aj.: A density based algorithm for discovering density varied clusters in large spatial databases [online]. *International Journal of Computer Applications*, ročník 3, č. 6, 2010: s. 1–4. Dostupné z: https://www.researchgate.net/profile/Anant-Ram/publication/282334683_A_Density_based_Algorithm_for_Discovering_Density_Varied_Clusters_in_Large_Spatial_Databases/links/560d049d08aed543358d5e32/A-Density-based-Algorithm-for-Discovering-Density-Varied-Clusters-in-Large-Spatial-Databases.pdf

-
- [35] Rosten, E.; Drummond, T.: Machine learning for high-speed corner detection [online]. In *Computer Vision—ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7–13, 2006. Proceedings, Part I 9*, Springer, 2006, s. 430–443. Dostupné z: https://link.springer.com/chapter/10.1007/11744023_34%C2%A0
- [36] Bay, H.; Ess, A.; Tuytelaars, T.; aj.: Speeded-up robust features (SURF). *Computer vision and image understanding*, ročník 110, č. 3, 2008: s. 346–359. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1077314207001555>
- [37] Shi, J.; Malik, J.: Normalized cuts and image segmentation [online]. *IEEE Transactions on pattern analysis and machine intelligence*, ročník 22, č. 8, 2000: s. 888–905. Dostupné z: https://repository.upenn.edu/cgi/viewcontent.cgi?article=1101&context=cis_papers
- [38] Banovic, N.; Grossman, T.; Matejka, J.; aj.: Waken: Reverse Engineering Usage Information and Interface Structure from Software Videos [online]. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, New York, NY, USA: Association for Computing Machinery, 2012, ISBN 9781450315807, str. 83–92, doi:10.1145/2380116.2380129. Dostupné z: <https://doi.org/10.1145/2380116.2380129>
- [39] Simko, J.; Vrba, J.: Screen recording segmentation to scenes for eye-tracking analysis [online]. *Multimedia Tools and Applications*, ročník 78, 2019: s. 2401–2425. Dostupné z: <https://link.springer.com/article/10.1007/s11042-018-6369-7>
- [40] Lowe, D. G.: Distinctive image features from scale-invariant keypoints [online]. *International journal of computer vision*, ročník 60, 2004: s. 91–110. Dostupné z: <https://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94>
- [41] Wang, Z.; Bovik, A.; Sheikh, H.; aj.: Image quality assessment: from error visibility to structural similarity [online]. *IEEE Transactions on Image Processing*, ročník 13, č. 4, 2004: s. 600–612, doi: 10.1109/TIP.2003.819861. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/1284395>
- [42] Sara, U.; Akter, M.; Uddin, M. S.: Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study [online]. *Journal of Computer and Communications*, ročník 7, č. 3, 2019: s. 8–18. Dostupné z: <https://www.scirp.org/journal/paperinformation.aspx?paperid=90911>

- [43] Yeh, T.; Chang, T.-H.; Miller, R. C.: Sikuli: Using GUI Screenshots for Search and Automation [online]. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, New York, NY, USA: Association for Computing Machinery, 2009, ISBN 9781605587455, str. 183–192, doi:10.1145/1622176.1622213. Dostupné z: <https://doi.org/10.1145/1622176.1622213>
- [44] Chang, T.-H.; Yeh, T.; Miller, R. C.: GUI Testing Using Computer Vision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, New York, NY, USA: Association for Computing Machinery, 2010, ISBN 9781605589299, str. 1535–1544, doi:10.1145/1753326.1753555. Dostupné z: <https://doi.org/10.1145/1753326.1753555>
- [45] Bernal-Cárdenas, C.; Cooper, N.; Moran, K.; aj.: Translating Video Recordings of Mobile App Usages into Replayable Scenarios [online]. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, ICSE '20, New York, NY, USA: Association for Computing Machinery, 2020, ISBN 9781450371216, str. 309–321, doi:10.1145/3377811.3380328. Dostupné z: <https://doi.org/10.1145/3377811.3380328>
- [46] Hu, G.; Zhu, L.; Yang, J.: AppFlow: Using Machine Learning to Synthesize Robust, Reusable UI Tests [online]. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2018, New York, NY, USA: Association for Computing Machinery, 2018, ISBN 9781450355735, str. 269–282, doi:10.1145/3236024.3236055. Dostupné z: <https://doi.org/10.1145/3236024.3236055>
- [47] Kulchandani, J. S.; Dangarwala, K. J.: Moving object detection: Review of recent research trends [online]. In *2015 International Conference on Pervasive Computing (ICPC)*, 2015, s. 1–5, doi:10.1109/PERVASIVE.2015.7087138. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/7087138>
- [48] Mandellos, N. A.; Keramitsoglou, I.; Kiranoudis, C. T.: A background subtraction algorithm for detecting and tracking vehicles [online]. *Expert Systems with Applications*, ročník 38, č. 3, 2011: s. 1619–1631, ISSN 0957-4174, doi:<https://doi.org/10.1016/j.eswa.2010.07.083>. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0957417410007116>
- [49] Sanda, S. H.: Highly robust and efficient vehicle detection tracking and classification model for intelligent transportation system [online]. *Expert Systems with Applications*, 2017. Dostupné z:

https://shodhganga.inflibnet.ac.in/bitstream/10603/257943/14/14_annexure.pdf

- [50] McKenna, S. J.; Jabri, S.; Duric, Z.; aj.: Tracking Groups of People [online]. *Computer Vision and Image Understanding*, ročník 80, č. 1, 2000: s. 42–56, ISSN 1077-3142, doi:<https://doi.org/10.1006/cviu.2000.0870>.
- [51] Haritaoglu, I.; Harwood, D.; Davis, L.: W/sup 4/: real-time surveillance of people and their activities [online]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, ročník 22, č. 8, 2000: s. 809–830, doi:10.1109/34.868683. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/868683>
- [52] Hu, W.; Tan, T.; Wang, L.; aj.: A survey on visual surveillance of object motion and behaviors [online]. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, ročník 34, č. 3, 2004: s. 334–352. Dostupné z: <http://www.cs.cmu.edu/~.dgovinda/pdf/recog/01310448.pdf>
- [53] Piccardi, M.: Background subtraction techniques: a review [online]. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, ročník 4, 2004, s. 3099–3104 vol.4, doi:10.1109/ICSMC.2004.1400815. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/1400815>
- [54] Yuan, D.; Yu, Y.: Moving object detection under moving camera by using normal flows [online]. In *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*, 2014, s. 2517–2522, doi:10.1109/ROBIO.2014.7090719.
- [55] Nguyen, B. L.; Chahir, Y.; Molina, M.; aj.: Eye Gaze Tracking with Free Head Movements Using a Single Camera [online]. In *Proceedings of the 1st Symposium on Information and Communication Technology, SoICT '10*, New York, NY, USA: Association for Computing Machinery, 2010, ISBN 9781450301053, str. 108–113, doi:10.1145/1852611.1852632. Dostupné z: <https://doi.org/10.1145/1852611.1852632>
- [56] Aslani, S.; Mahdavi-Nasab, H.: Optical flow based moving object detection and tracking for traffic surveillance [online]. *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, ročník 7, č. 9, 2013: s. 1252–1256. Dostupné z: https://research.iaun.ac.ir/pd/mahdavinasab/pdfs/PaperM_8653.pdf
- [57] Cheng, H.; Jiang, X.; Sun, Y.; aj.: Color image segmentation: advances and prospects [online]. *Pattern Recognition*, ročník 34, č. 12, 2001: s. 2259–2281, ISSN 0031-3203, doi:<https://doi.org/10.1016/>

- S0031-3203(00)00149-7. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0031320300001497>
- [58] Pal, N. R.; Pal, S. K.: A review on image segmentation techniques [online]. *Pattern Recognition*, ročník 26, č. 9, 1993: s. 1277–1294, ISSN 0031-3203, doi:[https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J). Dostupné z: <https://www.sciencedirect.com/science/article/pii/003132039390135J>
- [59] Minichino, J.; Howse, J.: *Learning OpenCV 3 Computer Vision with Python*. Birmingham: Packt Publishing Ltd, druhé vydání, 2015, ISBN 978-1-78528-384-0.
- [60] Chung, B. W.: *Pro Processing for Images and Computer Vision with OpenCV: Solutions for Media Artists and Creative Coders*. Hong Kong: Apress, první vydání, 2017, ISBN 978-1-4842-2774-9.
- [61] Nguyen, T. A.; Csallner, C.: Reverse Engineering Mobile Application User Interfaces with REMAUI (T) [online]. In *2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2015, s. 248–259, doi:10.1109/ASE.2015.32. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/7372013>
- [62] Beach, A.; Owen, A.: *Video compression handbook [online]*. Peachpit Press, 2018. Dostupné z: <https://learning.oreilly.com/library/view/video-compression-handbook/9780134846736/>
- [63] Sullivan, G. J.; Topiwala, P. N.; Luthra, A.: The H. 264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions [online]. *Applications of Digital Image Processing XXVII*, ročník 5558, 2004: s. 454–474. Dostupné z: <https://last.hit.bme.hu/download/vidtech/k%C3%B6nyvek/H264.pdf>
- [64] M. Jiřina, L. B., J. Novák: BI-SVZ 4. přednáška. *Obraz jako matice – Strojové vidění a zpracování obrazu (BI-SVZ)* [online]. 2023. Poslední změna 2023-09-23 [cit. 2023-04-30]. Dostupné z: <https://courses.fit.cvut.cz/BI-SVZ/lectures/files/bi-svz-04-obraz-jako-matrice.pdf>
- [65] M. Jiřina, L. B., J. Novák: BI-SVZ 7. přednáška. *Filtrace v prostorové a frekvenční oblasti (BI-SVZ)* [online]. 2023. Poslední změna 2023-09-23 [cit. 2023-04-30]. Dostupné z: <https://courses.fit.cvut.cz/BI-SVZ/lectures/files/bi-svz-07-filtrace-v-prostorove-a-frekvencni-oblasti.pdf>
- [66] M. Jiřina, L. B., J. Novák: BI-SVZ 8. přednáška. *Segmentace obrazu – hranové (BI-SVZ)* [online]. 2023. Poslední změna 2023-09-23 [cit. 2023-04-30]. Dostupné z: <https://courses.fit.cvut.cz/BI-SVZ/lectures/files/bi-svz-08-segmentace-obrazu-hranove.pdf>

- [67] Sundararajan, D.: *Digital Image Processing*. Singapore: Springer Nature Singapore Pte Ltd., první vydání, 2017, ISBN 978-981-10-6112-7, doi: <https://doi.org/10.1007/978-981-10-6113-4>.
- [68] M. Jiřina, L. B., J. Novák: BI-SVZ 10. přednáška. *Morfologie a tvarové charakteristiky (BI-SVZ)* [online]. 2023. Poslední změna 2023-09-23 [cit. 2023-04-30]. Dostupné z: <https://courses.fit.cvut.cz/BI-SVZ/lectures/files/bi-svz-10-morfologie-a-tvarove-charakteristiky.pdf>
- [69] Template Matching [online]. OpenCV docs. [cit. 2023-04-30]. Dostupné z: https://docs.opencv.org/4.x/d4/dc6/tutorial_py_template_matching.html
- [70] Contours : Getting Started [online]. OpenCV docs. [cit. 2023-04-30]. Dostupné z: https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html
- [71] Structural Analysis and Shape Descriptors [online]. OpenCV docs. [cit. 2023-04-30]. Dostupné z: https://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html
- [72] Wang, Z.; Bovik, A. C.: Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures [online]. *IEEE Signal Processing Magazine*, ročník 26, č. 1, 2009: s. 98–117, doi:10.1109/MSP.2008.930649. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/4775883>
- [73] Mittal, R.; Garg, A.: Text extraction using OCR: A Systematic Review [online]. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, s. 357–362, doi:10.1109/ICIRCA48905.2020.9183326. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/9183326>
- [74] Take a screenshot or screen recording on iPhone [online] iPhone User Guide: © 2023 Apple Inc. [cit. 2023-04-08]. Dostupné z: <https://support.apple.com/en-gb/guide/iphone/iphc872c0115/ios>
- [75] Bouchard, A.: ReplayKit Max removes the screen resolution cap from iOS' native screen recording feature [online] © 2023 iDownloadBlog.com. Poslední změna 2021-11-29 [cit. 2023-04-08]. Dostupné z: <https://www.idownloadblog.com/2021/11/30/replaykit-max/>
- [76] App Store Preview: Instagram [online] © 2023 Apple Inc. [cit. 2023-04-08]. Dostupné z: <https://apps.apple.com/us/app/instagram/id389801252>
- [77] iPhone 14 Pro – Technical Specifications [online] Apple support: © 2023 Apple Inc. [cit. 2023-04-11]. Dostupné z: https://support.apple.com/kb/SP875?locale=en_GB

- [78] UIKit documentation – Instance Property scale [online] © 2023 Apple Inc. [cit. 2023-04-12]. Dostupné z: <https://developer.apple.com/documentation/uikit/uiscreeen/1617836-scale>
- [79] Belinski, E.: iOS Ref – Resolution by iOS device [online]. [cit. 2023-04-12]. Dostupné z: <https://iosref.com/res>
- [80] sauce2011: Complete iPhone screenshot resolutions [online]. Poslední změna 2022-11-07 [cit. 2023-04-12]. Dostupné z: https://www.reddit.com/r/iphone/comments/yof8mj/complete_iphone_screenshot_resolutions/
- [81] iOS Resolutions – Display properties of every iPhone, iPad and iPod touch Apple ever made [online]. Poslední změna 2023-01-25 [cit. 2023-04-12]. Dostupné z: <https://www.ios-resolution.com>
- [82] Kies, S. C.; aj.: Social media impact on attention span [online]. *Journal of Management & Engineering Integration*, ročník 11, č. 1, 2018: s. 20–27, [cit. 2023-04-18]. Dostupné z: https://niplav.github.io/doc/psychology/attention_span/social_media_impact_on_attention_span_carstens_et_al_2018.pdf
- [83] Capturing Attention in Feed: The Science Behind Effective Video Creative [online]. Poslední změna 2016-04-20 [cit. 2023-04-18]. Dostupné z: <https://www.facebook.com/business/news/insights/capturing-attention-feed-video-creative>
- [84] MMOCR documentation: Overview [online] © OpenMMLab 2020-2030. [cit. 2023-04-22]. Dostupné z: <https://mmocr.readthedocs.io/en/dev-1.x/modelzoo.html>
- [85] Tech, H.: Instagram is testing three layouts so don't be surprised if your friend sees a different one [online] © 2023 Tech HindustanTimes. Poslední změna 2022-08-20 [cit. 2023-04-28]. Dostupné z: <https://tech.hindustantimes.com/tech/news/instagram-is-testing-three-layouts-so-don-t-be-surprised-if-your-friend-sees-a-different-one-71599747712979.html>
- [86] Mosseri, A.: One Launch + Four Tests [online] In: *Twitter*. Poslední změna 2022-08-20 [cit. 2023-04-28]. Dostupné z: <https://twitter.com/mosseri/status/1603083799662690304>
- [87] Update to Messenger features on Instagram [online]. © 2023 Instagram from Meta. [cit. 2023-04-28]. Dostupné z: https://help.instagram.com/654906392080948/?cms_platform=iphone-app&helpref=platform_switcher

Seznam Použitých zkratek

- AVC** Advanced Video Coding
- CTA** Call To Action
- CFR** Constant frame rate
- FPS** Frames per second
- HSV** Hue Saturation Value
- HTML** HyperText Markup Language
- HTTP** HyperText Transfer Protocol
- Mbps** Megabit per second
- MB** Megabyte
- px** Pixel
- RGB** Red Green Blue
- UI** User Interface
- VFR** Variable frame rate

Obsah přiloženého archivu

	readme.md	stručný popis a nastavení aplikace
	app	zdrojový kód aplikace
	thesis	zdrojový kód práce ve formátu \LaTeX
	DP_Formánek_Adam_2023.pdf	text práce ve formátu PDF

