Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Klasifikácia nekonzistentnosti generovaných dát s vysvetlením

Bakalárska práca

2026
Martin Hovorka

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Klasifikácia nekonzistentnosti generovaných dát s vysvetlením

Bakalárska práca

Študijný program:	Informatika
Študijný odbor:	Aplikovaná informatika
Školiace pracovisko:	Katedra informatiky
Školiteľ:	doc. RNDr. Martin Madaras, PhD.

Bratislava, 2026
Martin Hovorka

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Martin Hovorka
Študijný program: aplikovaná informatika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Inconsistency Classification of Generated Data with Explanations
Klasifikácia nekonzistentnosti generovaných dát s vysvetlením

Anotácia: Táto práca navrhuje prístup k detekcii nezrovnalostí v generovaných 2D alebo 3D dátach, klasifikácii nezrovnalostí a generovaniu vysvetlení týchto nezrovnalostí. Výskum skúma metódy detekcie umelo vygenerovaných alebo manipulovaných fotografií so zameraním na tri hlavné techniky: úběžníky, zrkadlové odrazy a tiene. Tieto metódy boli aplikované na sériu obrázkov s cieľom vyhodnotiť, či zobrazené objekty vykazujú realistickú perspektívu, konzistentné odrazy a fyzikálne presné tiene.

Cieľ: Cieľom práce je preskúmať najmodernejšie metódy generovania syntetických súborov údajov a nájsť otvorený súbor údajov s klasifikovanými nekonzistentnými obrázkami. Preskúmať využitie modelu SMPL a SMPLify na detekciu siluet ľudí. Následne definovať metódu na detekciu hrán a tieňových bodov v obraze. Nakoniec detegovať nekonzistentnosti v detekovaných tieňoch a úběžníkoch premietnutých hrán. Vytvoriť nástroj a definovať postup na vysvetlenie detekovaných nekonzistentností.

Literatúra: OpenFake: An Open Dataset and Platform Toward Large-Scale Deepfake Detection
https://arxiv.org/abs/2509.09495

Learnable SMPLify: A Neural Solution for Optimization-Free Human Pose Inverse Kinematics
https://arxiv.org/abs/2508.13562

Kľúčové slová:

Vedúci: doc. RNDr. Martin Madaras, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.

Dátum zadania: 24.09.2025

Dátum schválenia: 29.09.2025 doc. RNDr. Damas Gruska, PhD.
garant študijného programu

Čestné vyhlásenie: Čestne vyhlasujem, že celú bakalársku prácu na tému „Klasifikácia nekonzistentnosti generovaných dát s vysvetlením“, vrátane všetkých jej príloh a obrázkov, som vypracoval/vypracovala samostatne, a to s použitím literatúry uvedenej v priloženom zozname.

Pri príprave tejto práce boli tiež použité nástroje umelej inteligencie ChatGPT a Gemini za účelom prehľadávania internetu pre užitočné zdroje. Nástroje umelej inteligencie som použil/použila v súlade s príslušnými právnymi predpismi, akademickými právami a slobodami, etickými a morálnymi zásadami za súčasného dodržania akademickej integrity. Som si vedomý/vedomá, že plne zodpovedám za správnosť výsledného textu.

Poďakovanie: Veľmi venujem moje poďakovanie školitelovi, ktorý ma previedol cez rôzne úskalia, ktoré táto téma priniesla.

Abstrakt

V tejto práci si viacej priblížime, ako funguje detekovanie výtvoru umelej inteligencie najmä vo fotkách. Existujú metódy, ktoré za vhodných podmienok, vedia veľmi jednoducho odhaliť umelú inteligenciu. Odhalovanie je založené zväčša na tom, ako umelá inteligencia vie ľahko zlyhať na fyzických javoch. No dokazovanie toho, kde umelá inteligencia schybí, nie je úplne prosté. Musí si to prejsť mnohými fázami pozorného sledovania, skúšania a overovania. Za pomoci jednoduchých priamok, priesečníkov a vlastností geometrických útvarov, sa vieme dopracovať ku bezpochynému dôkazu, či fotka je produktom umelej inteligencie, alebo nie. Čiže často sa budem odvolávať na sedliacky rozum, a múdrych matematických predchodcov našej doby.

Kľúčové slová: umelá inteligencia, detekovanie, fyzika, geometria

Abstract

In this work, we will take a closer look at how the detection of artificial intelligence creations, especially in photos, works. There are methods that, under the right conditions, can very easily reveal the presence of AI. Detection is mostly based on how artificial intelligence tends to fail when it comes to physical phenomena. However, proving where AI makes mistakes is not entirely straightforward. It must go through many stages of careful observation, testing, and verification. With the help of simple lines, intersections, and properties of geometric shapes, we can arrive at undeniable evidence of whether a photo is a product of artificial intelligence or not. Thus, I will often refer to common sense and the wisdom of the mathematical forebears of our time.

Keywords: artificial intelligence, detection, physics, geometry

Obsah

Zoznam obrázkov

Úvod

Mnohí ľudia v tejto dobe, sú vystavený nesmiernemu riziku dezinformácii na internete. A preto je nesmierne dôloežité sa vyvarovať zbytočným predsudkom, ktoré vôbec nemusia byť založené na pravde. Tento klam je iba možný vďaka súčasnej technológii, ktorá dokáže vymyslenému klamstvu dať na prvý pohľad nádych pravdivosti. Na technológiu, ktorú najmä podotýkam je umelá inteligencia, ktorá je dobrý sluha, ale zlý pán.

Vďaka tomu ako presvedčivo vyzerajú spomenuté fotky a videá, je momentálne táto problematika celkom populárna. Dnes, keď si prezrieme internet, vieme nájsť mnoho efektívnych algoritmov, ktoré dokážu s nejakou pravdepodobnosťou odhaliť použitie umelej inteligencie. Vďaka týmto výskumom vieme povedať, že produkty umelej inteligencie sú zdanlivo klamlivé len na vonok.

Preto motiváciou pre túto prácu je zamedziť šírenie dezinformácií, produkovaných umelou inteligenciou. Súhlasím, že nám spomenuté generatívne modely vedia veľmi uľahčiť život. No keď sa pozrieme na AI ako na nepriateľa, tak vie slúžiť ako tá najsilnešia zbraň propagandy. Veď dostali sme sa do reality, kedy nevieme rozoznať niektoré videá, alebo fotky vyprodukované umelou inteligenciou.

A teda cieľom práce je analyzovať geometrické a fyzikálne vlastnosti obrazu, ktoré umelá inteligencia často nedokáže správne modelovať. Konkrétne sa zameriavam na metódy založené na perspektíve (úbežníky), zrkadlových odrazoch a tieňoch. Na základe týchto poznatkov následne navrhujem semiautomatizovaný algoritmus na detekciu nezrovnalostí v obraze.

Práca využíva prístup založený na analýze geometrie a fyziky, na rozdiel od čisto dátovo riadených metód, ako sú neurónové siete. Tento prístup umožňuje lepšie pochopiť dôvody, prečo je určitý obraz považovaný za autentický alebo manipulovaný.

Práca je členená nasledovne. V prvej kapitole sa venujem súčasnému stavu problematiky a teoretickým základom detekcie manipulovaných obrazov. A rozvinieme konkrétne metódy detekcie založené na analýze perspektívy, odrazov a tieňov spolu s praktickými ukážkami ich použitia. Tretia kapitola sa zameriava na návrh semiautomatizovaného algoritmu na detekciu zrkadlových odrazov. Vo štrvtej kapitole sa pozrieme na samotnú implementáciu algoritmu. V ďalšej časti práce sú prezentované výsledky testovania a ich diskusia. Záver práce sumarizuje dosiahnuté výsledky a predkladá možnosti

ďalšieho výskumu.

Kapitola 1

Súčasný stav riešenej problematiky

Modely na generovanie obrazu z textu, najmä moderné difúzne modely, sú na oko realistické, avšak stále vykazujú zásadné nedostatky pri presnom prenose textového zadania do výsledného obrazu. Častým problémom je nesprávne prepojenie vlastností s objektami, kde model napríklad priradí farbu nesprávnemu objektu alebo nedokáže správne zachytiť počet objektov či ich priestorové usporiadanie. Tieto chyby vyplývajú zo skutočnosti, že modely sa učia na základe štatistických vzťahov v dátach, bez skutočného pochopenia významu alebo fyzikálnych zákonitostí, čo vedie k nekonzistentným alebo nelogickým scénam.

Ďalším významným problémom je generovanie štruktúrovane presného obsahu, ako sú texty v obraze, diagramy alebo scény vyžadujúce presné priestorové vzťahy. Modely síce dokážu vytvoriť vizuálne presvedčivé výstupy, no často zlyhávajú pri zachovaní detailnej štruktúry, napríklad pri generovaní čitateľného textu alebo logicky konzistentných objektov. Tieto nedostatky poukazujú na obmedzenú schopnosť modelov vykonávať kompozičné uvažovanie a presnú kontrolu nad generovaným obsahom, čo predstavuje významnú výzvu pre ich spoľahlivé využitie v praxi.

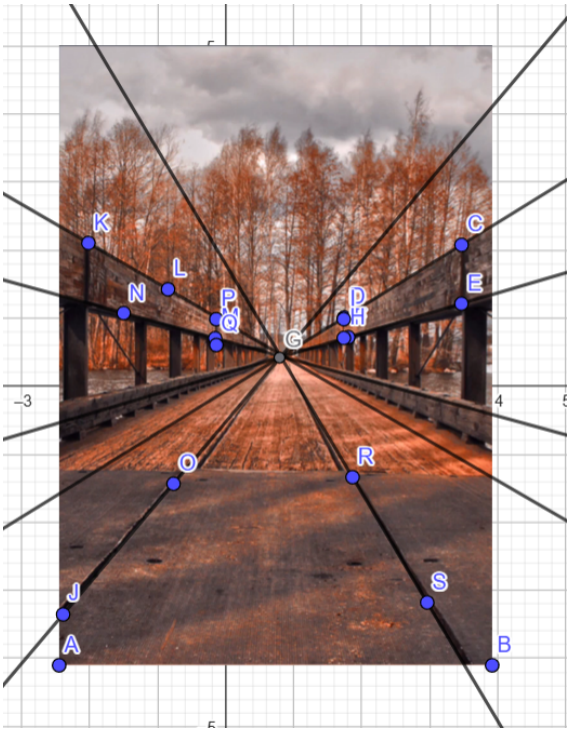
Najhľavnejšie zraniteľné oblasti genratívnej umelej inteligencie, ktorým sa budem v tejto kapitole venovať, sú perespektíva a úbežníky, tiene a zrkadlové odrazy. Venujem sa týmto oblastiam hlavne z dôvodu, že sú na fotkách najviac viditeľné a človeku najviac intuitívne. Moderný človek vie s pomerne veľkou pravdepodobnosťou odhadnúť tieto zraniteľne oblasti na fotke, lenže umelá inteligencia už tak pokročila za posledné roky, že sa sami ľudia dokážu ľahko nachytať. Preto je dobré najprv pochopiť týmto fyzikálnym javom od ich základov.

1.1 Perespektíva a úbežníky

Budem parafrázovať definíciu úbežníkov z práce [7] a [1]:

Určite ste už videli fotografiu železničných koľají, ktoré sa vzďaľujú od vás a pritom sa zdá, že medzera medzi nimi sa zužuje. V skutočnosti, v trojrozmernej scéne, je táto medzera samozrejme konštantná, no na fotografii pôsobí užšie vďaka základným vlastnostiam perspektívnej projekcie. Pri nej je veľkosť objektu premietnutého na snímač fotoaparátu (alebo na vašu sietnicu) nepriamo úmerná jeho vzdialenosti od kamery. Ak by koľajnice mali nekonečnú dĺžku, ich obraz by sa zbiehal do jediného bodu, tzv. vanishing pointu.

Obr. 1.1: Autentická fotka na ukážku úbežníka (bod G je úbežník).



Jednoducho povedané, je to miesto kde sa paralelné čiary v perspektívnom obraze zbiehajú. Je to základ skoro každej detekovacej metódy.

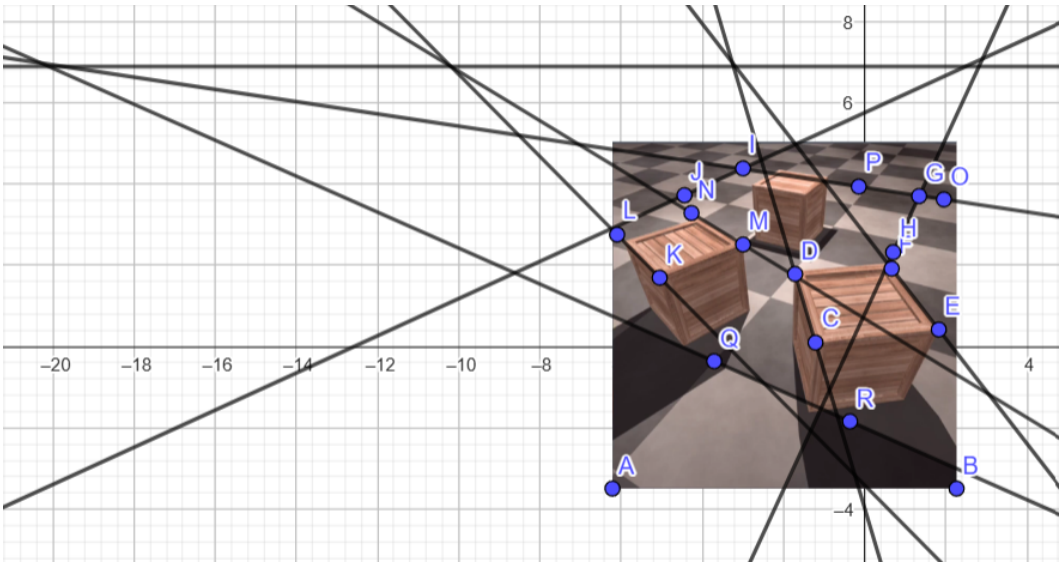
Ešte si musíme vystvetliť čo v perespektíve robí úbežnica. Úbežnica je priamka, v ktorej sa zbiehajú priesečníky priamok rovnobežných rovín (úbežníky).

Keď si načrtneme priamky, ktoré kopírujú hrany objektov, tak máme pár paralelných priamok. Geometrické objekty ako kocka a kváder sú naši veľký priatelia, lebo majú jasné paralelné hrany, ktoré sa dajú ľahko obtiahnuť priamkou.

Nasledovne treba odsledovať, či sa tieto páry paralelných priamok zbiehajú do jedného bodu (úbežníka). Keď toto zopakujeme viac krát na rôznych objektoch, tak vieme odsledovať, že tieto jednotlivé úbežnice nám vytvárajú peknú priamku. Toto je teda úbežnica, a je to jeden z dôkazov, že obrázok je reálny. Keby sa vanishing pointy nezbiehali do vanishing liny, tak to je značný indikátor, že fotka bola zmanipulovaná.

1.1.1 Postup

Ukážeme si to na príklade. Naším predmetom analýzy bude simulácia objektov 1.2, na ktorom si overíme jeho autenticitu.



Obr. 1.2: Fotka reprezentuje reálnu simuláciu objektov.

Ako objekty, ktoré určite budeme chcieť použiť na našu analýzu, sú v tomto obrázku 3 krabice. Keďže sa očividne jedná o kocky, nakreslíme priamky obťahujúc hrany krabíc. Na vytvorenie priamky si teda zoberieme dva body cez ktoré bude prechádzať, a to rohy krabíc. V našom obrázku, pre krabicu najbližšiu ku nám na pravo, sme načrtli body D a C, a jej paralelný pár bodov bude E a F (skrytý pod bodom H). Cez tieto páry bodov nám budú prechádzať priamky, ktoré sa stretnú vo úbežníku. Pre krabicu na ľavo, si načrtneme L a K, a jej paralelný pár bodov bude N a M. Cez tieto páry bodov nám budú prechádzať priamky, ktoré sa stretnú vo vanishing pointe.

Teraz to skúsme vyskúšať aj na iných objektoch, a to bude v našom prípade podľaha. Tu vieme s ľahkosťu kopírovať čiary, ktoré nám podľaha ponúka. Páry bodov budú H a G, a im prislúchajúci pár paralelných bodov bude J a I. Cez tieto páry bodov nám budú prechádzať priamky, ktoré sa stretnú vo úbežníku. To isté pre Q a R, a im prislúchajúci pár paralelných bodov bude P a O.

Teraz, keď máme hodný počet dát s ktorými môžeme pracovať, vieme cez jednotlivé úbežníky, nakresliť priamku. A teda si všimneme, že každý úbežník leží na úbežnici.

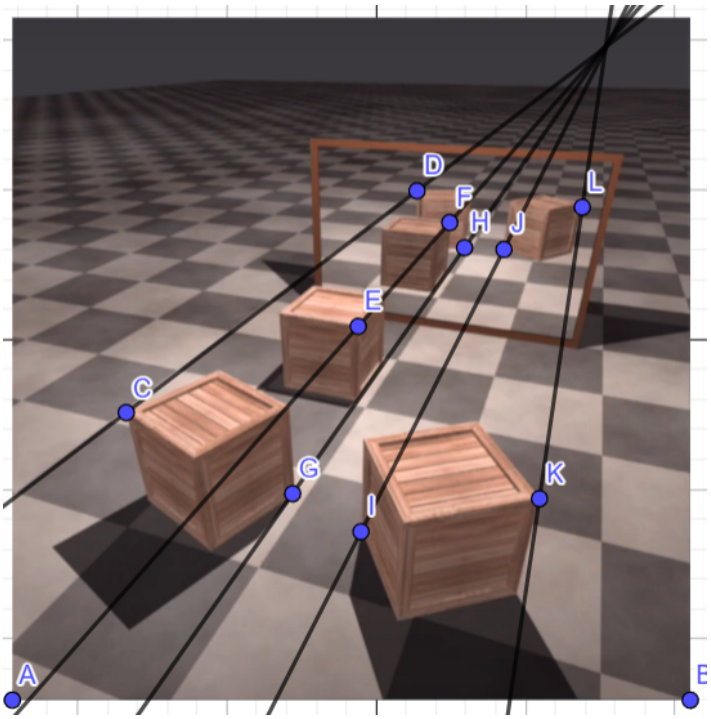
Dosiahli sme náš cieľ, že úbežníky ležia na úbežnici, a preto je daný obrázok autentický.

1.2 Zrkadlové odrazy

Skôr než sa pozrieme na to, ako odhaliť zmanipulované odrazy, je dôležité pochopiť, prečo ich vôbec dokážeme tak ľahko prehliadnuť. Croucher, Bertamini a Hecht vo svojej štúdii o tzv. „naívnej optike“ [2] zistili, že ľudská intuícia ohľadom správania zrkadiel je prekvapivo nepresná. Ich experimenty odhalili fenomén známy ako „predčasná chyba“.

Ľudia nesprávne odhadujú, z akých uhlov a pozícií by mal byť objekt v zrkadle viditeľný, a nedokážu správne aplikovať fyzikálny zákon odrazu do praxe. Náš mozog je z pohľadu perspektívy vysoko tolerantný voči geometricky nemožným odrazom. Práve táto naša slepota umožňuje falšovateľom fotografií vytvárať vizuálne uveriteľné, no fyzikálne nesprávne manipulácie.

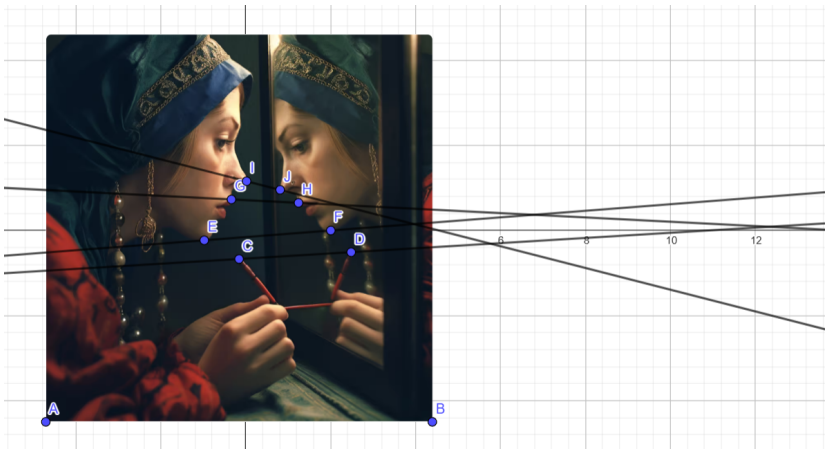
Čiže prácu ktorú tiež spomenieme, bude o zrkadlových odrazoch [8]. Keď do fotografie vložíme alebo upravíme odrazy objektov, výsledný odraz často porušuje základné geometrické pravidlá rovnakej roviny a perspektívy. Ak je odraz vložený neprirodzene, napríklad nesúhlasí s originálnymi úbežníkmi, takúto nesúladnosť možno posúdiť ako dôkaz manipulácie. Tento článok [14] sa venuje konkrétnym príkladom, na ktorých funguje ich vybudovaný detekovacii algoritmus. Postup algoritmu si vieme spraviť ručne na ukážke 1.3.



Obr. 1.3: Autentická simulácia na ukážku vanishing pointu v odraze zrkadla.

1.2.1 Postup

Ukážeme si to na príklade. Naším predmetom analýzy bude ukážka 1.4, na ktorom si overíme jeho autenticitu.



Obr. 1.4: Fotka generovaná umelou inteligenciou.

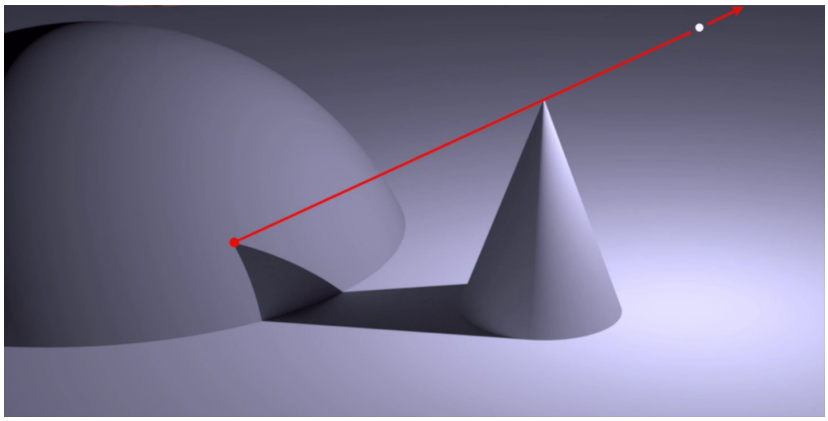
Objekt na obrázku ktorému sa budeme venovať, je žena držiaca prehnutú paličku. Teraz si nájdeme najvýraznejšie črty, ktoré budeme porovnávať s ich odrazom v zrkadle, a spravíme z nich body. Špička nosa je bod I, horná pera je bod G, spodná časť brady je bod E a vrch paličky je bod C. Teraz ku každému spomenutému bodu, nájdeme ich príslušný odraz v zrkadle. Ku bodu I prislúcha J, ku G prislúcha H, ku E prislúcha F a ku C prislúcha D. Cez každý pár ktorý som spomenul sa nakreslí priamka.

Najvýraznejšie črty sme vyčerpali, a môžeme prejsť na analýzu. Z geometrie je zjavné, že priamky nemajú spoločný bod v ktorom sa priamky zbiehajú (vanishing point). A teda vieme zhodnotiť, že fotka je vygenerovaná umelou inteligenciou.

1.3 Tiene

Hoci sa tiene zdajú byť priamočiare, spoliehať sa na ľudský zrak pri detekcii takýchto manipulácií je vysoko nepostačujúci. Štúdia od Nightingale a kol. [11] ukázala, že ľudia majú len veľmi obmedzenú schopnosť odhaliť zmanipulované tiene alebo odrazy voľným okom, a to aj v prípadoch, keď sú inštruovaní, aby hľadali chyby. Náš zrak totiž pri bežnom vnímaní často filtruje informácie o tieňoch ako takzvaný „šum“.

A teda v poslednej detekovacej metóde sa budeme odvolávať na zákonitosti tieňov [10]. V tomto článku autori ukazujú, ako sa dajú odhaliť upravené alebo sfalšované fotografie pomocou analýzy tieňov. V reálnom svete sa tiene správajú podľa presných fyzikálnych pravidiel. A teda ak vieme, kde je zdroj svetla, vieme predpovedať, ako budú tiene padať. A naopak ak vieme, kde tiene smerujú, vieme zistiť, kde by mal byť svetelný zdroj. Viacej sa vieme dočítať aj v tomto papiery [6].



Obr. 1.5: Autentická simulácia na ukážku polohy tieňa voči zdroju svetla.

1.3.1 Postup

Ukážeme si to na príklade. Analýzu vykonáme na fotke vygenerovanej umelou inteligenciou 1.6, a si overíme jeho autenticitu.

Obr. 1.6: Fotka generovaná umelou inteligenciou.

Objektom, ktorým sa budeme venovať na obrázku sú deti. Teraz hľadáme na deťoch najvýraznejšie črty, ktoré sú zároveň aj ľahko viditeľné v tieňoch detí. Poďme postupne skúmať od ľava do prava. Na druhom dieťati, je zakončenie pravej ruky dostatočne výrazný bod a jemu prisluchajúci bod v tieni je tiež jasne viditeľný. Čiže bod na ruke je I a na tieni je to bod J. Na treťom dieťati, budeme jednať o prste ľavej ruky. Bod

prstu je C a v tieni bod D. Ešte má druhú výraznú črtu a to ľavú stranu sukne, ktorá sa tiež nachádza v tieni. Bod na sukni je G a v tieni bod H. Posledného štvrtého dieťaťa najvýraznejšia črta sú končeky prstov ľavej ruky, ktorým prislúcha jasný tieň. Bod končekov prstov je E a bod na tieni F.

Teraz keď skonštruujeme priamky cez tieto páry bodov, môžeme odsledovať, že všetky priamky nemajú spoločný vanishing point. A teda sa jedná o fotku, vygenerovanú umelou inteligenciou.

Kapitola 2

Návrh detekovacieho algoritmu

2.1 Zvolený prístup detekcie

Zvolil som prístup detekcie založený na analýze zrkadlových odrazov v obraze. Tento prístup vychádza z predpokladu, že generatívne modely umelej inteligencie síce dokážu vytvárať vizuálne presvedčivé obrázky, avšak často nedokážu správne zachovať geometrickú konzistenciu scény. Zrkadlové odrazy predstavujú špecifický prípad, kde musí byť dodržaná presná geometrická transformácia medzi objektom a jeho odrazom.

Na prvý pohľad môžu byť takéto obrazy realistické, avšak pri detailnej analýze sa často objavujú nezrovnalosti, najmä v oblasti perspektívy a úbežníkov. V reálnych podmienkach platí, že priamky objektu a jeho odrazu by mali byť navzájom konzistentné a mali by sa zbiehať do rovnakých alebo zodpovedajúcich úbežníkov. Generatívne modely však túto vlastnosť nedokážu vždy správne reprodukovať, čo vytvára priestor pre detekciu manipulovaného obsahu.

2.2 Schéma analýzy obrazu

Navrhovaný algoritmus by mal pracovať s jedným vstupným obrazom, ktorý je potrebné pred samotnou analýzou vhodne predspracovať. Predspracovanie zahŕňa úpravu obrazu do formy vhodnej pre detekciu príznakov, napríklad normalizáciu, prípadne prevod do odtieňov sivej.

Kľúčovým krokom algoritmu je identifikácia dvoch hlavných oblastí obrazu: samotného objektu a jeho zrkadlového odrazu. V týchto oblastiach sa následne detekujú hlavné príznaky. Tieto body reprezentujú charakteristické časti obrazu, ktoré sú vhodné na ďalšiu analýzu.

Po detekcii bodov nasleduje ich párovanie medzi objektom a jeho odrazom. Na základe týchto korešpondencií je možné vytvoriť množinu priamok, ktoré reprezentujú geometrické vzťahy medzi jednotlivými časťami obrazu.

V ďalšom kroku by mal algoritmus vyhodnotiť konzistenciu získaných úbežníkov. Pokúsime sa naimplementovať postup z kapitoly 1.2.1. Naopak, pri umelo generovaných obrazoch sa často vyskytujú odchýlky, ktoré bude možné identifikovať.

2.3 Výstupy a interpretácia výsledkov

Výstupom navrhovaného algoritmu je vizuálne aj numerické vyhodnotenie analyzovaného obrazu. Jedným z hlavných cieľov je poskytnúť výsledok, ktorý umožní používateľovi pochopiť, na základe čoho bol obraz označený ako potenciálne manipulovaný.

Čiže jednotlivé kroky spracovania priamo v obraze, napríklad zobrazenie detekovaných bodov, priamok a odhadnutých úbežníkov. Na úkor ľahšieho pochopenia konkrétnym oblastiam, aby sa dalo ľahko upraviť parametre algoritmu do budúcna.

Ako hlavná časť algoritmu bude verdikt, že či je analyzovaný obraz umelo generovaný alebo manipulovaný. Tento verdikt je založený na miere geometrickej nekonzistencie zistených vzťahov a môže byť doplnený o kvantitatívne metriky.

Kapitola 3

Implementácia riešenia

V tejto kapitole sa implementuje algoritmus, ktorý je založený na detekcii zrkadlového odrazu v digitálnom obraze. Riešením bude žiaľ semiautomatizovaný algoritmus, keďže bude vyžadovaný zásah užívateľa pri spracovaný vstupu. V ďalších podkapitolách sa vysvetlí celý algortimus.

3.1 Hlavná myšlienka algoritmu

Hlavnou myšlienkou navrhovaného algoritmu je detekcia geometrických nekonzistencií medzi objektom a jeho zrkadlovým odrazom prostredníctvom analýzy zbiehavosti priamok. V ideálnom prípade by mal zrkadlový odraz predstavovať geometricky korektnú transformáciu reálneho objektu, čo znamená, že zodpovedajúce body by mali vytvárať konzistentné projekčné vzťahy a ich spojnica by mala smerovať k spoločnému úbežníku.

Prvým krokom je spracovanie vstupu. Počas implementácie algoritmu som si rýchlo uvedomil, že pri práci s nespracovaným vstupným obrazom a priamom hľadaní párov totožných bodov je síce možné, ale získame veľmi malý počet dát, čo nie je vhodné na ďalšie spracovanie. Hlavným problémom je, že algoritmus SuperGlue nie je zrkadlovo invariantný pri hľadaní korešpondujúcich príznakov. To znamená, že ak sa určitý príznak nachádza aj v zrkadlovom odraze, algoritmus ho nedokáže správne spárovať, pretože ho nepovažuje za ekvivalentný – aj napriek tomu, že ide len o jeho zrkadlovo otočenú verziu.

Ako alternatívny prístup bol testovaný algoritmus GLS-MIFT, ktorý zrkadlovú invariantnosť poskytuje. Tento algoritmus sa však ukázal ako nedostatočný v poskytnutí dostatočného počtu zhodujúcich sa párov, keďže GLS-MIFT potrebuje viac vstupov, a to my v našom príapde vieme poskytnúť iba jeden.

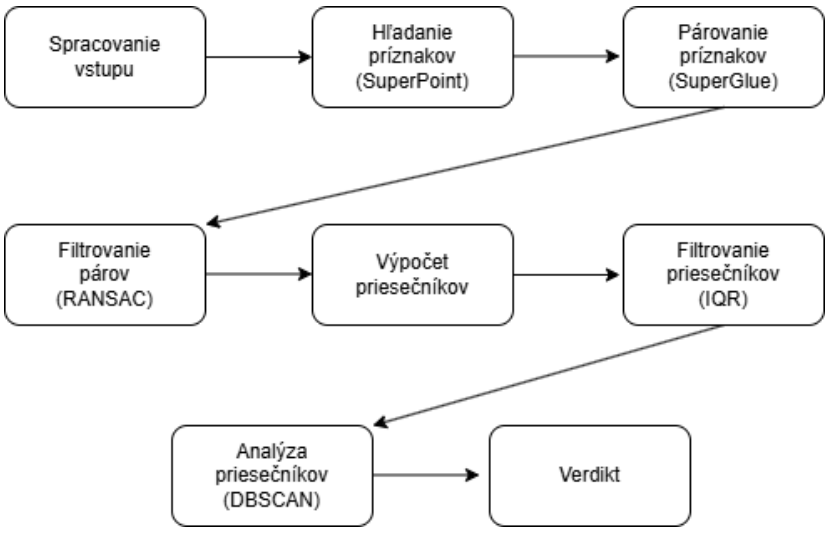
Na základe týchto pozorovaní bol zvolený prístup predspracovania vstupného obrazu rozdelením na dve časti, pričom jedna z nich je následne zrkadlovo otočená (ľubovoľne ktorá). Rozdelenie fotky si môžeme dovoliť, keďže predpokladáme, že objekt a

jeho zrkadlový odraz sa nachádzajú na rozdielnych častiach fotky. Na takto upravené časti je následne aplikovaný algoritmus SuperGlue, ktorý už nemá problém s párovaním príznakov, keďže zrkadlová invariantnosť je zabezpečená v predspracovaní. Týmto spôsobom je možné získať dostatočné množstvo konzistentných a relevantných dát, ktoré sú následne využité v ďalších krokoch algoritmu. Jediné negatívum je, že štruktúra detekovacieho algoritmu nebude úplne automatizovaná, keďže na rozdelenie obrazu je potrebný priamy zásah užívateľa.

Charakteristické body (príznaky) v oblasti objektu a jeho odrazu získame pomocou metódy SuperPoint. SuperGlue sa pozerá na tieto príznaky a pospája, čím vznikne množina korešpondencií medzi dvoma časťami obrazu. Na základe týchto korešpondencií sa vytvárajú priamky reprezentujúce geometrické vzťahy medzi bodmi.

Kľúčovým krokom je analýza priesečníkov týchto priamok. V prípade reálneho obrazu by sa mali tieto priesečníky koncentrovať v malej oblasti (úbežník), zatiaľ čo pri umelo generovaných obrazoch bývajú rozptýlené. Miera rozptylu týchto priesečníkov je následne použitá ako rozhodovací faktor pre určenie geometrickej konzistencie obrazu.

Na zvýšenie robustnosti algoritmu sa využívajú metódy ako RANSAC na odstránenie nesprávnych korešpondencií, IQR filtrovanie na elimináciu odľahlých priesečníkov a zhlukovacie algoritmy (napr. DBSCAN) na identifikáciu dominantného úbežníka. Výsledkom je kvantitatívne vyjadrenie kvality konvergencie priamok, ktoré slúži ako základ pre finálny verdikt.



Obr. 3.1: Diagram hlavnej myšlienky algoritmu.

3.2 Použité technológie a modely

3.2.1 GoogleColab

GoogleColab je cloudová služba, ktorá umožňuje spúšťanie Python kódu v prostredí Jupyter Notebook bez potreby lokálnej inštalácie. Colab poskytuje prístup k výpočtovým zdrojom vrátane GPU a umožňuje jednoduchú integráciu s knižnicami pre strojové učenie a počítačové videnie, ako sú PyTorch alebo OpenCV.

3.2.2 SuperPoint

SuperPoint [3] algoritmus predstavuje moderný prístup na detekciu významných bodov (keypoints) v obraze a zároveň na výpočet ich deskriptorov. Na rozdiel od tradičných metód, ako sú SIFT alebo ORB, využíva hlbokú neurónovú sieť, ktorá je trénovaná pomocou samo-učiaceho (self-supervised) prístupu. Model je schopný robustne identifikovať charakteristické body aj pri zmenách perspektívy, rotácie alebo osvetlenia. Vďaka tomu poskytuje spoľahlivé vstupné dáta pre ďalšie spracovanie, najmä pre analýzu geometrických vzťahov v obraze.

3.2.3 SuperGlue

SuperGlue [13] je algoritmus, kde ide o metódu založenú na neurónových sieťach, ktorá využíva mechanizmus pozornosti na hľadanie optimálnych korešpondencií medzi bodmi. Na rozdiel od klasických prístupov, ktoré porovnávajú deskriptory bodov nezávisle, SuperGlue berie do úvahy aj globálny kontext scény a vzťahy medzi bodmi. Vďaka tomu dosahuje výrazne presnejšie párovanie aj v zložitých alebo šumom zaťažených situáciách.

3.2.4 RANSAC

RANSAC [4] používa robustnú metódu, ktorá umožňuje odhadnúť matematický model (napr. priamku alebo transformáciu) aj v prípade, že vstupné dáta obsahujú veľké množstvo chybných alebo odľahlých hodnôt (outlierov). Algoritmus pracuje tak, že opakovane vyberá náhodné podmnožiny bodov, z ktorých odhaduje model, a následne vyhodnocuje, koľko bodov s týmto modelom súhlasí. Najlepší model je vybraný na základe maximálneho počtu súhlasných bodov

3.2.5 IQR

IQR filterovanie [9] je štatistická metóda používaná na detekciu a odstránenie šumu v obraze na základe identifikácie odľahlých hodnôt. Metóda vychádza z interkvartilového

rozpätia, ktoré je definované ako rozdiel medzi horným (Q3) a dolným (Q1) kvartilom dát. Pixely, ktorých hodnota leží mimo intervalu určeného interkvartilovým rozpätím, sú považované za odľahlé a sú nahradené odhadovanou hodnotou, typicky pomocou lokálneho priemerovania susedných pixelov.

3.2.6 K-NN

K-NN [12] je metóda strojového učenia, ktorá rozhoduje o triede nového objektu na základe podobnosti s už známymi dátami. Pri klasifikácii sa najprv vypočíta vzdialenosť medzi novým objektom a všetkými bodmi v trénovacej množine, následne sa vyberie počet najbližších susedov. Výsledná trieda sa určí podľa toho, ktorá trieda sa medzi týmito susedmi vyskytuje najčastejšie.

3.2.7 DBSCAN

DBSCAN [5] je algoritmus zhlukovania, ktorý identifikuje zhluky dát na základe hustoty bodov v priestore. Základnou myšlienkou je, že zhluk predstavuje oblasť s vysokou hustotou bodov, zatiaľ čo body nachádzajúce sa v oblastiach s nízkou hustotou sú považované za šum alebo odľahlé hodnoty. Algoritmus využíva dva hlavné parametre: polomer okolia a minimálny počet bodov, pričom body sú klasifikované ako jadrové, hraničné alebo šumové.

3.3 Štruktúra algoritmu

Navrhovaný algoritmus pozostáva z viacerých na seba nadväzujúcich krokov, ktoré postupne transformujú vstupný obraz na výsledný verdikt o jeho geometrickej konzistencii.

3.3.1 Predspracovanie

Na zabezpečenie kvality vstupných dát sú snímky konvertované do bezstratového formátu PNG, čím sa vytratia artefakty, ktoré by mohli negatívne ovplyvniť presnosť detekcie príznakov.

Potom sa prerobí rozlíšenie na 800×600 pixelov (Goldilocks pravidlo) so zachovaním pôvodného pomeru strán. Toto rozlíšenie predstavuje kompromis medzi dostatočným množstvom detailov pre extrakciu kľúčových bodov a obmedzením šumu, ktorý je výpočtovo naročný na vizuálne deskriptory.

Takto prerobený vstup je dôležitý pre spoľahlivejšie hľadanie korešpondujúcich párov pomocou modelu SuperGlue.

Obr. 3.2: Vytvorený graf pre vstup (vstup je autentická fotka odfotená mnou).

Keďže objekt a jeho zrkadlový odraz nemusí byť iba na ľavej a pravej časti fotky, ale aj na hornej a dpľnej časti fotky, tak algoritmus sa ešte pred vytvorením grafu spýta používateľa, či chce fotku rezať vertikálne, alebo horizontálne. Pri prípade horizontálneho rezania (objekt a zrkadlová časť sa nachádzajú na hornej a doľnej časti fotky), si otočíme vstupnú fotku o 90 stupňov, a až potom orežeme. Toto otáčanie je dôležité, lebo keď chceme neskôr jednu orezanú časť zrkadlovo otočiť, tak sa zachovajú všetky zrkadlové pravidlá. (MOZNO VIAC VYSVETLIT)

Na to aby sa neskôr použil SuperGlue, tak potrebujeme ešte jednu orezanú časť fotky zrkadlovo otočiť. Táto časť je teda triviálna, no veľmi dôležitá, keďže SuperGlue nie je zrkadlovo invariantný algoritmus. V nasledujúcej prílohe, sa vieme presvedčiť, že po zrkadlovom obrátení, sú orezané časti takmer identické.

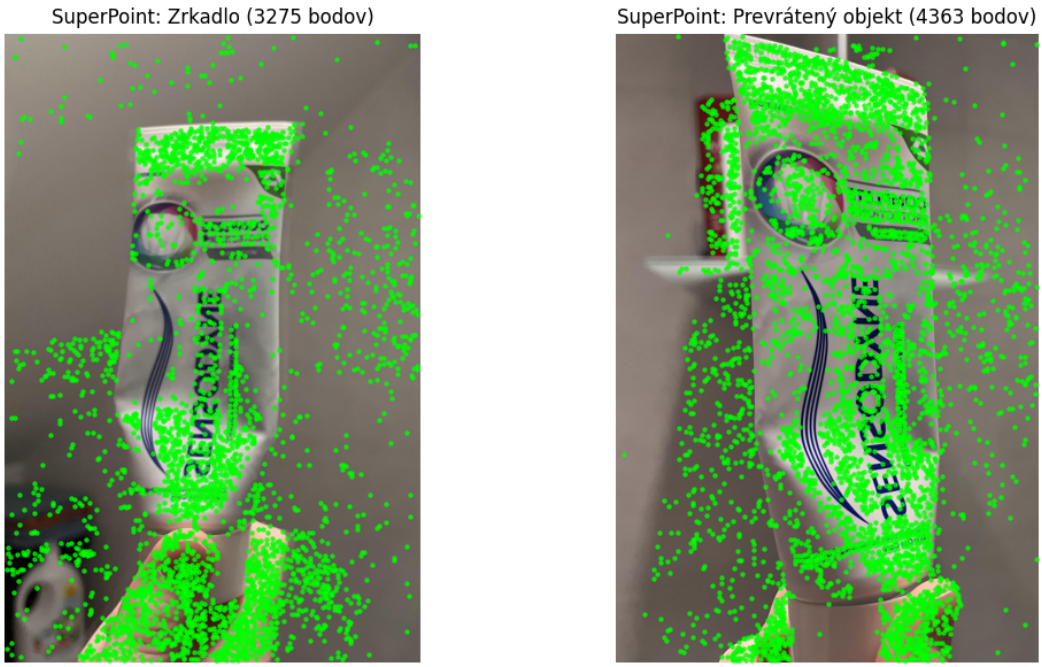
Obr. 3.3: Ukážka zkradlového pretočenia orezanej časti.

Posledným krokom predspracovania, je prehodenie fotiek na jednokanálové (sivé) obrázky. Dôvodom je, že algoritmus SuperPoint vie pracovať iba na jednokanálových obrázkoch, keďže hľadá kde sa intenzita svetla mení (vysoký kontrast alebo gradient). Na prevod používam OpenCV knižnicu v Pythone, ktorá má ľahkú implementáciu (TREBA FOTKA PREMENY FOTIEK NA SIVU + KOD?)

3.3.2 Hľadanie príznakov

Na hľadanie príznakov som najprv používal algoritmy SIFT a ORB, ktoré fungujú, ale prešiel som na SuperPoint, pretože SuperGlue bol doslova stvorený pre SuperPoint. SuperGlue bol navrhnutý po modeli SuperPoint, čiže jeho vstupné vrstvy boli explicitne naprogramované tak, aby prijímali presne ten formát dát (tzv. tenzor), aký SuperPoint produkuje. SuperGlue bol navyše priamo trénovaný na výstupoch zo SuperPointu. Počas tréningovej fázy mu výskumníci poskytli milióny deskriptorov vygenerovaných SuperPointom.

Tento krok nám poskytne dve množiny bodov (vytvorené z orezaných častí), ktoré sú vstupom pre ďalší krok.



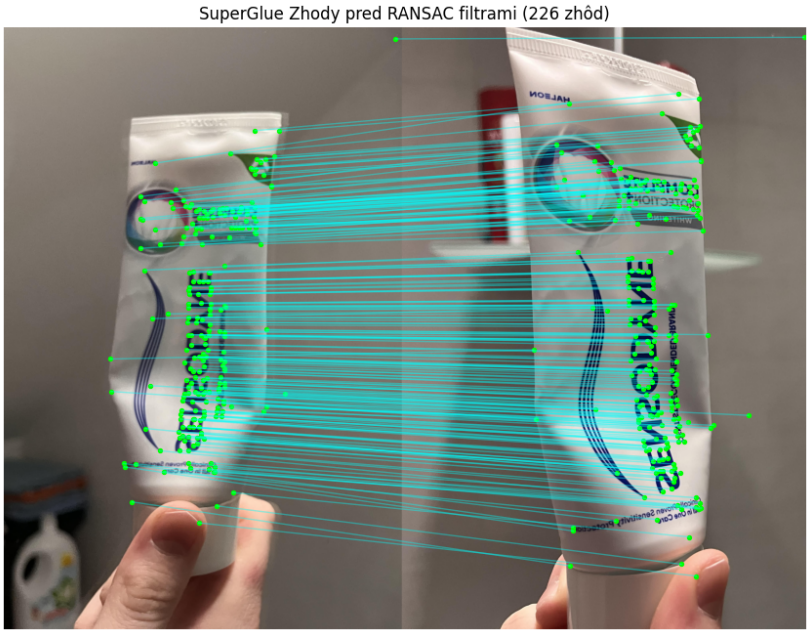
Obr. 3.4: Vizualizácia SuperPointu.

3.3.3 Hľadanie párov príznakov

Na rozdiel od klasických metód párovania, ktoré porovnávajú deskriptory bodov nezávisle, SuperGlue využíva neurónovú sieť. Tento prístup umožňuje zohľadniť nielen lokálne vlastnosti jednotlivých príznakov, ale aj globálny kontext scény a vzťahy medzi bodmi. Výsledkom je výrazne presnejšie a robustnejšie párovanie, najmä v prípadoch, kde sú dáta zaťažené šumom alebo obsahujú nejednoznačné štruktúry.

Vstupom pre algoritmus SuperGlue sú príznaky a ich deskriptory získané z oboch častí obrazu (objekt a jeho zrkadlovo upravený odraz). Výstupom je množina párov bodov, ktoré algoritmus považuje za vzájomne zodpovedajúce. Každému páru je zároveň priradená miera dôvery, ktorá vyjadruje pravdepodobnosť správnosti daného párovania.

Na obrázku 3.5 je zobrazený výsledok párovania pomocou algoritmu SuperGlue. Je možné si všimnúť, že aj napriek vysokej presnosti algoritmu sa medzi nájdenými pármi nachádzajú aj nesprávne korešpondencie. Tieto chyby môžu byť spôsobené napríklad opakujúcimi sa štruktúrami v obraze alebo nedokonalosťami generovaného obsahu. Z tohto dôvodu je potrebné tieto páry ďalej filtrovať.



Obr. 3.5: Vizualizácia SuperGlue.

3.3.4 Filtrovanie nájdených párov

Na odstránenie nesprávnych párov príznakov je využitý algoritmus RANSAC, ktorý predstavuje metódu na odhad geometrických modelov v prítomnosti odľahlých hodnôt.

Algoritmus RANSAC pracuje iteratívne. V každej iterácii náhodne vyberie malú podmnožinu párov bodov, z ktorých odhadne geometrický model (homografiu). Následne vyhodnotí, koľko z celkového množstva bodov je s týmto modelom konzistentných. Tieto body sú označené ako konzistentné body. Proces sa opakuje viackrát a ako výsledný model je vybraný ten, ktorý má najväčší počet konzistentných bodov.

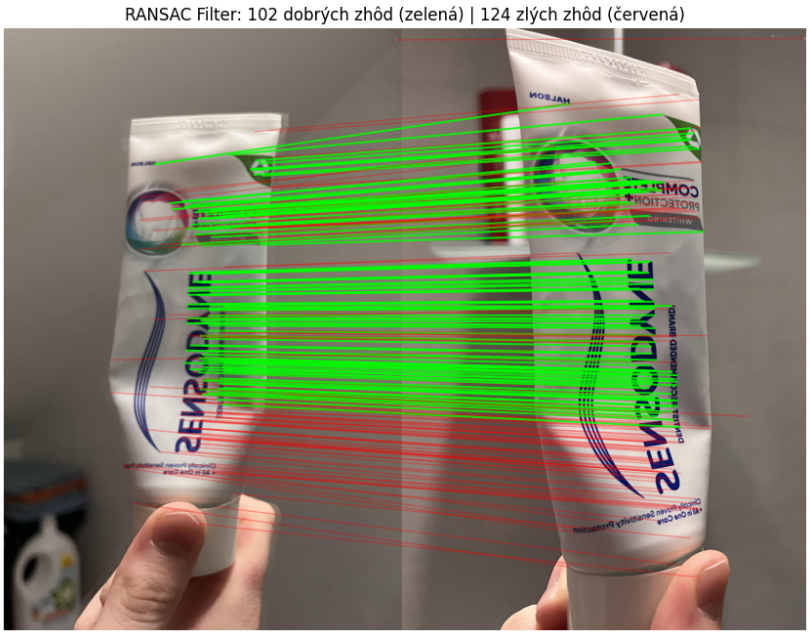
Dôležitým parametrom algoritmu RANSAC je prahová hodnota, ktorá určuje toleranciu chyby pri posudzovaní zhody bodov. Táto hodnota (v pixeloch) definuje maximálnu vzdialenosť medzi skutočnou pozíciou bodu a jeho predikovanou pozíciou podľa odhadovaného modelu. Inými slovami, prahová hodnota určuje, ako veľká odchýlka je ešte akceptovaná na to, aby bol pár bodov považovaný za správny.

Voľba tejto hodnoty predstavuje kompromis medzi presnosťou a robustnosťou. Pri príliš nízkej hodnote je algoritmus veľmi prísny a akceptuje iba takmer dokonalé zhody, čo môže viesť k tomu, že väčšina bodov bude zamietnutá, najmä v prípade šumu, skreslenia objektívu alebo nedokonalostí zrkadla. Naopak, pri príliš vysokej hodnote sa algoritmus stáva menej selektívnym a môže akceptovať aj nesprávne páry, čo negatívne ovplyvní výsledný geometrický model.

V implementácii je použitá vyššia prahová hodnota (10 pixelov), ktorá umožňuje zachovať väčší počet potenciálne správnych párov bodov. Vychádzam z predpokladu,

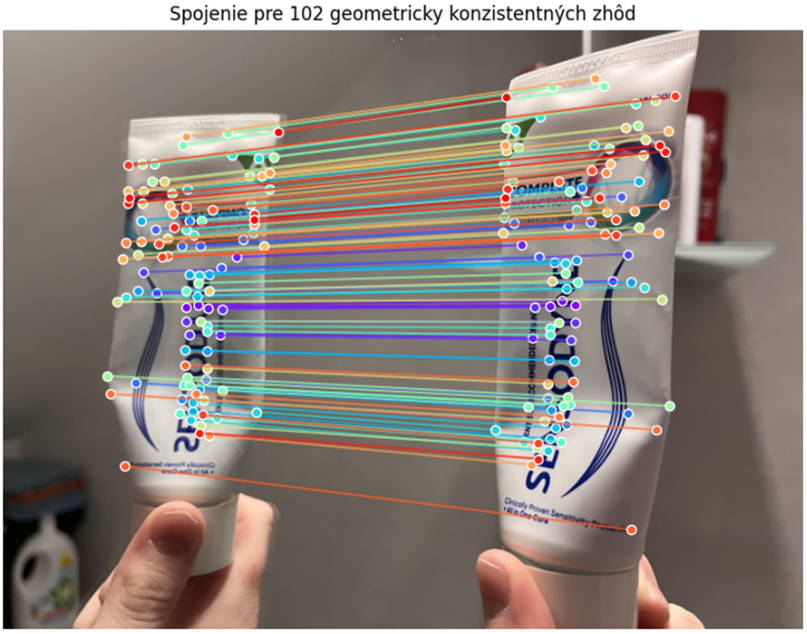
že následné kroky spracovania sú schopné dodatočne eliminovať nesprávne korešpondencie. Použitie vyššej prahovej hodnoty síce vedie k zvýšeniu výpočtovej náročnosti, avšak tento kompromis je nevyhnutný.

Na obrázku 3.6 je znázornený výsledok po aplikovaní algoritmu RANSAC, kde sú zachované iba konzistentné páry príznakov. Je možné vidieť, že väčšina nesprávnych korešpondencií bola úspešne odstránená, čo vedie k presnejšej analýze geometrických vlastností obrazu.



Obr. 3.6: Vizualizácia RANSAC filtrovania.

Ako posledný krok, pretočím naspäť pôvodne zrkadlovo pretočenú časť vstupu, aby som mohol v ďalších krokoch pracovať s pôvodným vstupom. Samozrejme, musím pri tom zachovať nájdené príznaky, čiže ich súradnice tiež pretočím zrkadlovo. (TREBA KOD?) Obrázok 3.7 predstavuje výsledok filtrovania.



Obr. 3.7: Vizualizácia hotového filtrovania.

3.3.5 Priesečníky nájdených párov

V tejto fáze algoritmu sa použijú nájdené páry z predošlého kroku. Hlavnou myšlienkou je nájsť úbežník priamok, vytvorených z párov totožných bodov.

Implementácia je nasledovná. Cez páry totožných bodov sa natiahnu priamky a nájdu ich priesečníky. Keďže priestor kde sa môžu priamky pretnúť je nekonečný 2D priestor (graf), tak môžeme predpokladať, že skoro každá priamka sa pretne s každou inou. To znamená, že množstvo priesečníkov rastie kvadraticky voči počtu párom.

3.3.6 Analýza potencionálnych úbežníkov

Za ideálnych podmienok je úbežník definovaný ako bod, v ktorom sa pretínajú všetky priamky. V reálnych dátach však, v dôsledku šumu a nepresností v detekcii príznakov, nedochádza k presnému prieniku všetkých priamok v jednom bode. Preto je v implementácii úbežník definovaný ako bod, do ktorého sa koncentruje najväčší počet priesečníkov priamok.

Pred samotnou identifikáciou úbežníka je nevyhnutné odstrániť šum z množiny priesečníkov. Šum v tomto prípade predstavujú odľahlé priesečníky, ktoré vznikajú napríklad pri takmer rovnobežných priamkach alebo nesprávne spárovaných príznakoch, a nachádzajú sa výrazne mimo hlavnej koncentrácie bodov.

Prvý krok filtrácie využíva štatistickú metódu interkvartilového rozpätia (IQR). V porovnaní so štandardnou odchýlkou je IQR robustnejšie voči extrémnym hodnotám, keďže nevyužíva aritmetický priemer, ktorý môže byť výrazne ovplyvnený vzdialenými

bodmi. Pre každú os priestoru (súradnice x a y) sa nezávisle vypočíta prvý kvartil (Q_1) a tretí kvartil (Q_3), pričom samotné interkvartilové rozpätie je definované ako:

$$\mathbf{IQR} = \mathrm{Q}_3 - \mathrm{Q}_1. \tag{3.1}$$

Na základe tejto hodnoty sa určí akceptovateľná oblasť pre platné priesečníky:

$$\mathrm{H}_{dolna} = \mathrm{Q}_1 - k \cdot \mathbf{IQR}, \tag{3.2}$$

$$\mathrm{H}_{horna} = \mathrm{Q}_3 - k \cdot \mathbf{IQR}. \tag{3.3}$$

V implementácii je použitá konštanta $k = 1.5$, čo zodpovedá štandardnému Tukeyho filtru pre detekciu odľahlých hodnôt. Táto voľba predstavuje vyvážený kompromis medzi odstránením extrémnych priesečníkov a zachovaním dostatočného množstva relevantných dát.

```
for i in range(3):
        if len(filtrovane) < 5:
            break

        x_suradnice = filtered[:, 0]
        y_suradnice = filtered[:, 1]

        q1_x, q3_x = np.percentile(x_suradnice, [25, 75])
        iqr_x = q3_x - q1_x
        h_dolna_x = q1_x - (1.5 * iqr_x)
        h_horna_x = q3_x + (1.5 * iqr_x)

        q1_y, q3_y = np.percentile(y_suradnice, [25, 75])
        iqr_y = q3_y - q1_y
        h_dolna_y = q1_y - (1.5 * iqr_y)
        h_horna_y = q3_y + (1.5 * iqr_y)

        validne_x = (x_suradnice >= h_dolna_x) & (x_suradnice <=
    h_horna_x)
        validne_y = (y_suradnice >= h_dolna_y) & (y_suradnice <=
    h_horna_y)
        maska = validne_x & validne_y

        filtrovane = filtrovane[maska]
```

Algoritmus 3.1: Výpočet priesečníkov a ich hodnoty.

Tento proces je aplikovaný iteratívne v troch krokoch. Iteratívny prístup je dôležitý, pretože extrémne vzdialené priesečníky v prvých iteráciách umelo zväčšujú hodnotu IQR. Po ich odstránení sa v ďalších krokoch hranice prirodzene zúžia a presnejšie ohraničia jadro dát.

Po odstránení hrubých odľahlých hodnôt môže množina priesečníkov stále obsahovať viacero lokálnych zhlukov, ktoré vznikajú v dôsledku komplexnej geometrie scény. Na identifikáciu hlavného bodu konvergencie je preto použitý algoritmus DBSCAN.

Výhodou algoritmu DBSCAN je, že nevyžaduje vopred určiť počet zhlukov a zároveň dokáže identifikovať zhluky ľubovoľného tvaru. Izolované body automaticky klasifikuje ako šum. Kľúčovým parametrom algoritmu je ϵ, ktorý určuje maximálnu vzdialenosť bodov v rámci jedného zhluku.

Hodnota parametra ϵ je určená pomocou analýzy k-najbližších susedov. Najprv sú dáta štandardizované (nulová stredná hodnota a jednotková variancia), aby sa zabránilo dominancii jednej osi. Následne sa pre každý bod vypočíta vzdialenosť ku jeho k-tému najbližšiemu susedovi ($k \leq 5$). Získané vzdialenosti sú zoradené a parameter ϵ je definovaný ako:

$$\epsilon = 1.5 \cdot med(\mathrm{D}_k)$$

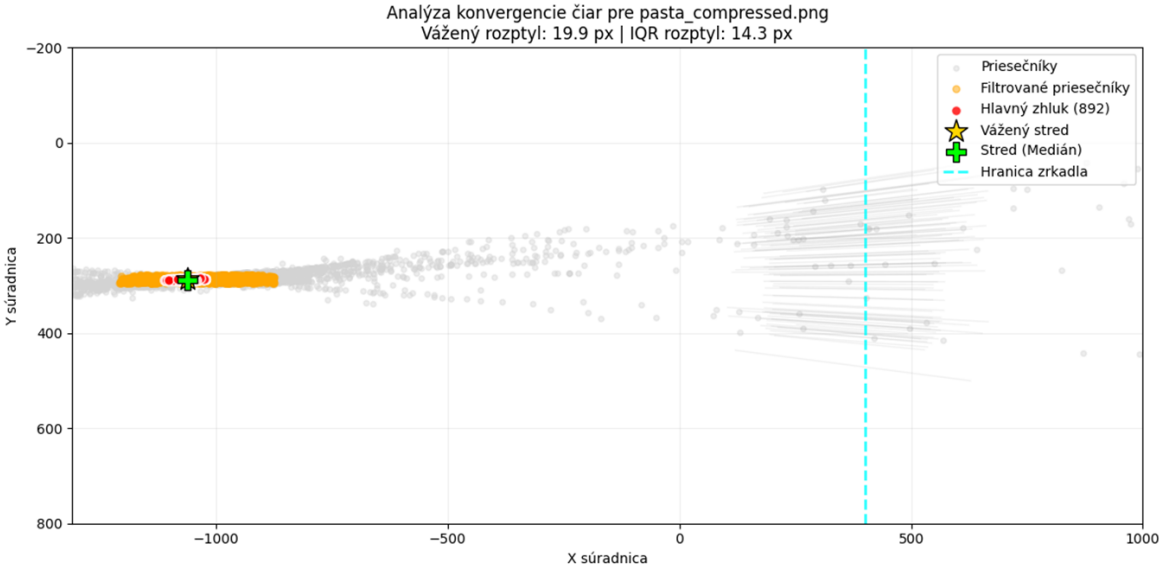
kde $med(\mathrm{D}_k)$ je medián vzdialeností ku k-tému susedovi. Konštanta 1.5 kompenzuje variabilitu hustoty a zabezpečuje spojenie bodov v hlavnom zhluku.

```
if len(filtered) >= 10:
        skalovane = StandardScaler().fit_transform(filtrovane)
        neigh = NearestNeighbors(n_neighbors=min(5, len(filtrovane)-1)
    ).fit(skalovane)
        vzdialenosti, _ = neigh.kneighbors(skalovane)
        eps_auto = np.median(np.sort(vzdialenosti[:,-1])) * 1.5

        skore = DBSCAN(eps=eps_auto, min_samples=min(3, len(filtrovane
    )//5)).fit_predict(skalovane)

        zhluky = {}
        rozlisne_skore = set(skore)
        for lbl in rozlisne_skore:
            if lbl != -1:
                body_v_zhluku = np.sum(skore == lbl)
                zhluky[lbl] = body_v_zhluku
        if zhluky:
            najlepsie_skore = max(zhluky, key=zhluky.get)
            hlavny_zhluk = filtrovane[skore == najlepsie_skore]
        else:
            hlavny_zhluk = filtrovane
    else:
        hlavny_zhluk = filtrovane
```

Algoritmus 3.2: Zhlukovanie a štatistika.

Po aplikácii algoritmu DBSCAN sú priesečníky rozdelené na jednotlivé zhluky. Následne je vybraný hlavný zhluk na základe jeho veľkosti (počtu bodov) a vzdialenosti

od predpokladanej hranice zrkadla. V poslednej fáze je z vybraného zhluku určený výsledný úbežník ako reprezentatívny bod konvergencie. Tento bod je vypočítaný pomocou kombinácie mediánu a váženého priemeru, čím sa zabezpečí robustnosť voči šumu. Zároveň je možné určiť aj jeho rozptyl, ktorý slúži na posudzovanie geometrickej konzistencie.



Obr. 3.8: Vizualizácia analýzy.

3.3.7 Výstup

Výstupom algoritmu je finálny verdikt, ktorý určuje, či je analyzovaný obraz autentický alebo umelo generovaný. Na rozhodnutie sa využívajú dáta váženého priemeru a interkvartilového rozpätia (IQR), ktoré boli získané v predchádzajúcich krokoch spracovania. Vážený priemer slúži na určenie polohy úbežníka, zatiaľ čo IQR poskytuje informáciu o rozptyle priesečníkov okolo tejto polohy. Tým, že obe hodnoty sú dôležité na vyhodnotenie verdiktu, som sa rozhodol z nich spraviť priemer. A podľa priemeru určujem, či je daný vstup autentický, alebo nie. Hodnoty, podľa ktorých algoritmus usudzuje, či je priemer dostatočný, sme získali vo fáze testovania v kapitole 4.

Ak je priemer menší ako 40, tak vieme zhodnotiť, že konvergencia je výborná a teda jedná sa o autentický vstup. Pre $40 < priemer < 130$ sa považuje vstup ako akceptovateľný. Pre $130 < priemer < 200$ považujeme vstup za nedôveryhodný no potencionálne správny. A pre priemer vyšší ako 200 vieme s určitosťou povedať, že fotka je umelo vygenerovaná.

Ak sa v procese našlo menej párov ako 5, tak vieme zhodnotiť, že fotka nie je autentická. A to práve preto, lebo ak SuperGlue nájde toľko málo zhôd, tak sa objekt pravdepodobne nezhoduje s jeho zkradlovým odrazom.

3.4 Vstupné dáta a obmedzenia

Vstupné dáta pre algoritmus sú digitálne fotografie obsahujúce objekt a jeho zrkadlový odraz. Pre správnu funkčnosť algoritmu je dôležité to, aby boli na vstupnom obraze viditeľné obe časti – samotný objekt aj jeho odraz v zrkadle. Tieto časti sa musia nachádzať buď v ľavej a pravej polovici obrazu, alebo v jeho hornej a dolnej časti. Ešte predpokladáme, že zrkadlá sú ploché (rovné). Kebyže nie sú ploché, tak sa analýza zjavne nedá spraviť.

V prípade, že podhodíme algoritmu vstup nespĺňajúci spomenuté ktitéria (napríklad ak objekt alebo jeho odraz nie sú dostatočne viditeľné, alebo sa nachádzajú v inom priestorovom usporiadaní), algoritmus síce technicky vykoná spracovanie, avšak nedokáže vykonať analýzu. V takom prípade dôjde k nesprávnemu vyhodnoteniu vstupu.

Dataset použitý v tejto práci pozostáva z vlastných fotografií, autentických fotografií z webovej platformy Roboflow, vygenerovaných obrázkov použitím nástrojov umelej inteligencie, ako sú Bing Image Creator a Gemini. Dataset momentálne pozostáva z 17 autentických fotiek a 29 umelo vygenerovaných fotiek.

Pri generovaní obrázkov použitím nástrojov umelej inteligencie, som používal text-to-image prompty takéhoto typu: "Vytvor mi fotografiu (X objektu) na pravej strane, s jeho zrkadlovým odrazom na ľavej strane fotografie, na ktorej je viditeľná väčšina častí objektu. Zachovajte realistickú kompozíciu."

Kapitola 4

Výsledky a diskusia

Algoritmus bol testovaný na datasete obsahujúcom 46 obrázkov, z toho 17 autentických a 29 umelo generovaných. Pri vyhodnocovaní geometrickej konzistencie sa ako hlavný ukazovateľ používal rozptyl priesečníkov (v pixeloch).

Pre autentické fotografie sa hodnoty rozptylu pohybovali v intervale 9.3 px až 187.0 px, pričom mediánová hodnota dosiahla 42.6 px. Naopak, pri umelo generovaných obrázkoch bol rozptyl výrazne vyšší, v rozsahu od 21.1 px až po 4694.4 px, s mediánom 493.0 px.

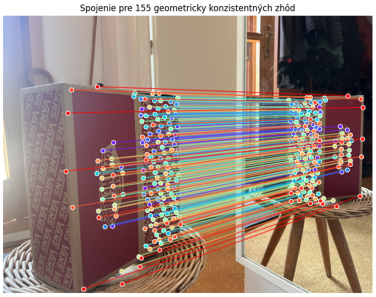
Celková presnosť algoritmu dosahuje približne 89.1 %, čo predstavuje pomer správne klasifikovaných obrázkov (41) k celkovému počtu testovaných vzoriek (46). Slabá konvergencia sa nepočíta ako správny výsledok, je to tzv. "šedá zóna".

Pri autentických fotografiách algoritmus správne identifikoval 94.1 % prípadov (16 zo 17), pričom nedošlo k žiadnemu prípadu, kedy by bola reálna fotografia označená ako umelo generovaná. To znamená, že algoritmus nevykazuje falošne pozitívne výsledky, čo je z pohľadu praktického nasadenia veľmi dôležitá vlastnosť.

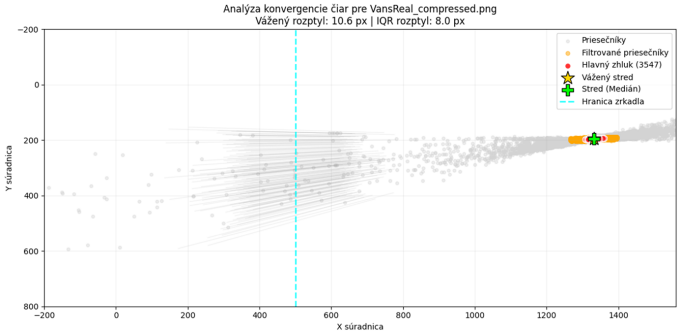
Pri umelo generovaných obrázkoch algoritmus správne detegoval 86.2 % prípadov (25 z 29) pri použití striktnej rozhodovacej hranice. V tomto hodnotení sú za správne považované iba tie obrázky, ktoré vykazujú jednoznačnú geometrickú nekonzistentnosť alebo zlyhajú už vo fáze párovania príznakov.

Ak by sme do úspešnej detekcie zahrnuli aj prípady so slabou konvergenciou (tzv. podozrivú kategóriu), celková úspešnosť detekcie by sa zvýšila približne na 93 %. Táto hodnota však zahŕňa aj menej jednoznačné prípady.

Ako ukážku správne fungujúceho algoritmu na autentickej fotografii (odfotenej mnou) sa uvádza príloha 4.1. Priemerný rozptyl pre tento prípad je 9.3 px. Priemer je teda menší ako 40 px, a teda algoritmus správne rozhodol, že sa jedná o autentickú fotku.



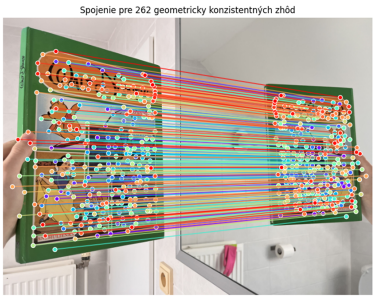
(a) nájdenie korešpondujúcich párov.



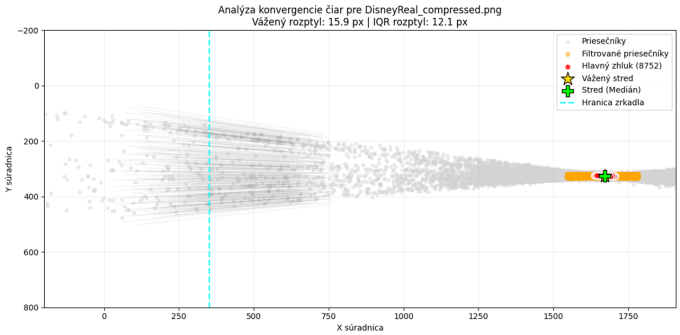
(b) analýza geometrie

Obr. 4.1: Vizualizácia výsledku pre autentický vstup.

Druhú ukážku správne fungujúceho algoritmu na autentickej fotografii (odfotenej mnou) sa uvádza príloha 4.2. Priemerný rozptyl pre tento prípad je 14 px. Priemer je teda menší ako 40 px, čo algorimus správne vyhodnotí, že sa jedná o autentickú fotku.



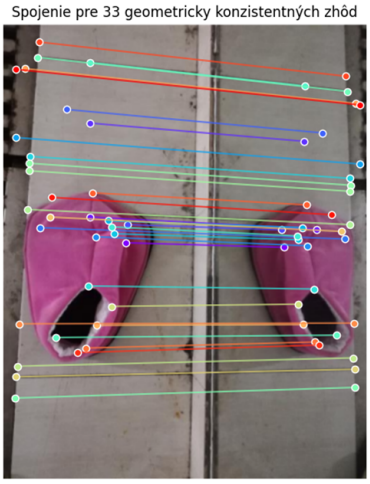
(a) nájdenie korešpondujúcich párov



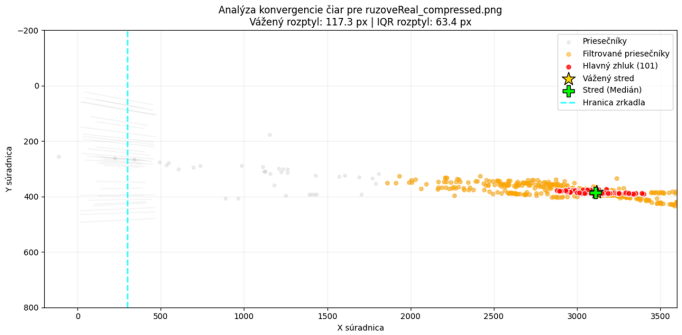
(b) analýza geometrie

Obr. 4.2: Vizualizácia výsledku pre autentický vstup

Tretiu ukážku správne fungujúceho algoritmu na autentickej fotografii (z datasetu Roboflow) sa uvádza príloha 4.3. Priemerný rozptyl pre tento prípad je 90.4 px. Priemer je teda väčší ako 40 px a menší ako 130 px, čo algoritmus akceptuje ako autentickú fotku.



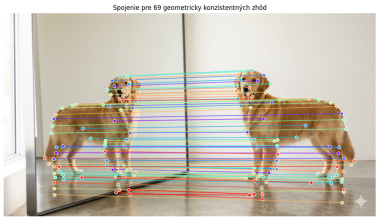
(a) nájdenie korešpondujúcich párov.



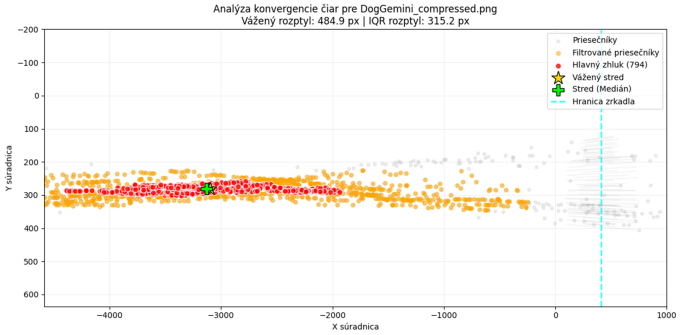
(b) analýza geometrie

Obr. 4.3: Vizualizácia výsledku pre autentický vstup.

Na obrázku 4.4 je znázornený výsledok pre umelo vygenerovaný obraz (Gemini), kde je možné pozorovať výraznú geometrickú nekonzistentnosť. Priemerný rozptyl v tomto príapde je 403.1 px. Priemer je teda väčší ako 200 px, a teda algoritmus správne rozhodol, že sa jedná o umelo vygenerovanú fotku.



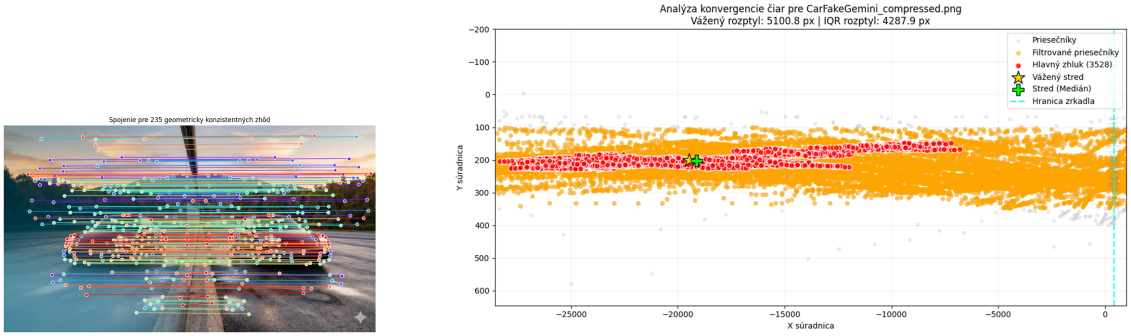
(a) nájdenie korešpondujúcich párov.



(b) analýza geometrie.

Obr. 4.4: Vizualizácia výsledku pre umelo vygenerovaný vstup.

Ako ďalšiu ukážku 4.5 máme znázornený výsledok pre umelo vygenerovaný obraz (Gemini), na ktorom sa vieme ľahko presvedčiť, že je geometricky nesprávny. Priemerný rozptyl v tomto príapde je 4694.35 px. Priemer je teda väčší ako 200 px, a čo znamená, že fotka je zmanipulovaná.



(a) nájdenie korešpondujúcich párov.

(b) analýza geometrie.

Obr. 4.5: Vizualizácia výsledku pre umelo vygenerovaný vstup.

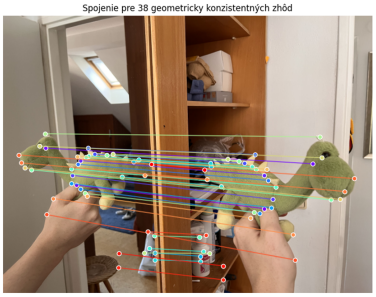
Zaujímavým zistením je, že pri niektorých AI obrázkoch algoritmus zlyhal už v počiatočnej fáze párovania príznakov (SuperGlue nenašiel dostatočný počet bodov). V kontexte detekcie to predstavuje pozitívny výsledok, keďže ide o prípady, kde umelo generovaný obraz nedodržuje základné vlastnosti zrkadlového odrazu (zachovať textúry, ...). Tento prípade si môžeme všimnúť na ukážke 4.6.

Obr. 4.6: Fotka vygenerovaná umelou unteligenciou.

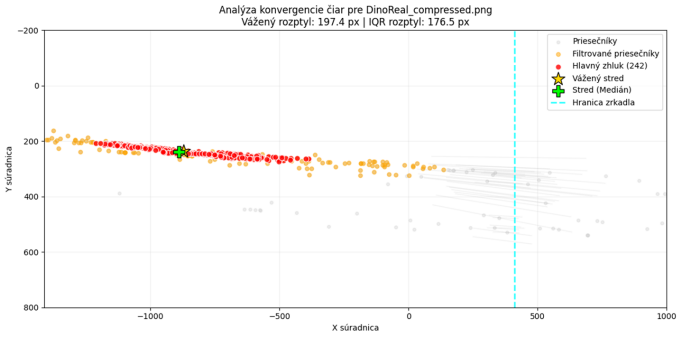
4.1 Anomálie

Napriek vysokej presnosti algoritmus nedosahuje 100 % spoľahlivosť. V tejto časti sú uvedené prípady, kde došlo k nesprávnemu vyhodnoteniu.

Na obrázku 4.7 je znázornený prípad autentickej fotografie, ktorá bola vyhodnotená ako podozrivá. Hlavným dôvodom bolo nesprávne párovanie príznakov v dolnej časti obrazu, čo negatívne ovplyvnilo výpočet geometrie a následnú analýzu.



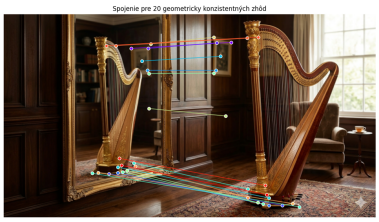
(a) nájdenie korešpondujúcich párov.



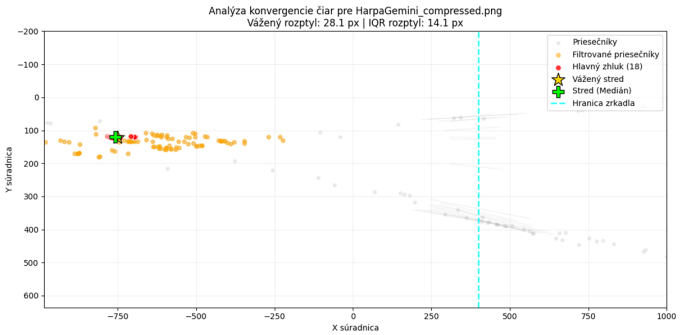
(b) analýza geometrie.

Obr. 4.7: Vizualizácia chybného vyhodnotenia pre autentický vstup.

Naopak, na obrázku 4.8 je príklad umelo generovaného obrazu, ktorý algoritmus nesprávne vyhodnotil ako autentický. Tento prípad predstavuje tzv. falošne negatívny výsledok. Dá sa na obrázku všimnúť, že objekt je s jeho zrkadlovým odrazom jasne odlišný v textúre, samotnej veľkosti a natočení. Algoritmus si ale vytypoval len tie body, ktoré zachovávali správny odraz, a keď iba s tými pracuje, tak sa samozrejme dopracuje ku verdiktu, že fotka je autentická.



(a) nájdenie korešpondujúcich párov.



(b) analýza geometrie.

Obr. 4.8: Vizualizácia chybného vyhodnotenia pre umelo generovaný vstup.

4.2 Diskusia

Ako sa vo výsledkoch algoritmu spomínalo, tak algoritmus má celkovú úspešnosť približne 89.1%. Je to očakávaný výsledok, keďže pracovanie so samotným obrazom ako vstup je veľmi náročná úloha pre počítač.

Algoritmus má mnoho parametrov, ktoré sa dajú ľadiť. Tieto parametre sú napríklad:

1. prahová hodnota pre detekciu príznakov (SuperPoint),
2. prahová hodnota pre párovanie príznakov (SuperGlue),
3. tolerancia chyby v algoritme RANSAC,
4. multiplikátor k interkvartilového rozpätia (IQR),
5. adaptívny parameter ϵ pre algoritmus DBSCAN,
6. minimálny počet bodov v zhluku (`min_samples`),
7. rozhodovacie prahové hodnoty pre finálny verdikt.

Pri testovaní algoritmu, sa najviac menili hodnoty tolerancie RANSAC v kombinácii s multiplikátorom ϵ pre DBSCAN. Práve preto, lebo priamo pracujú v analýze priesečníkov. Skúšali sa mnohé hodnoty pre tieto parametre, no tie s ktorými momentálne algoritmus pracuje, mali najväčšiu úspešnosť.

Ďalší parameter pri ktorom treba robiť veľký kompromis, je prahová hodnota pre SuperGlue. Čím menšia prahová hodnota, tým lepšia kvalita v totožných pároch. Ale na to aby algoritmus fungoval správne, nechceme iba najlepšie totožné páry, ale aj tie nie úplne správne. Dôvodom je, aby sme mohli fotky generované umelou inteligenciou ľahko nachytať. Skvelým príkladom je fotka ktorú sme už testovali 4.5. Na nej sa dá všimnúť, že príznaky v reálnej scéne nie sú úplne totožné s príznakmi v zrkadlovom odraze. SuperGlue ich aj tak spojí, takto dostaneme páry, ktoré si umelá inteligencia myslela že sú správne, čo sú skvelé dáta pre algoritmus na analýzu.

Napriek dosiahnutým výsledkom má algoritmus aj svoje obmedzenia. Experimenty ukázali, že ho je možné oklamať v prípadoch, keď iba malá časť obrazu vykazuje geometrickú konzistenciu. V takýchto situáciách algoritmus pracuje len s podmnožinou správne spárovaných bodov, zatiaľ čo nesprávne alebo nekonzistentné časti obrazu sú ignorované, pretože sa pre ne nepodarí nájsť korešpondencie. Typickým príkladom sú umelo generované obrázky ako 4.8, kde je zrkadlový odraz vizuálne nesprávny (napr. odlišná textúra alebo deformácia), avšak obsahuje oblasti, ktoré náhodne spĺňajú geometrické podmienky. Algoritmus tieto oblasti vyhodnotí ako konzistentné a na ich základe môže nesprávne určiť obraz ako autentický.

Taktiež je tu riziko, kedy SuperGlue zlyhá v párovaní bodov. Aj keď sa RANSAC implementuje práve kvôli tomuto problému, môže sa stať, že niektoré páry sa aj tak dostanú cez filtrovanie. Málo kedy sa to stane, ale je to stále veľke riziko, ktoré veľmi negatívne ovplyvní výsledok algoritmu.

Záver

Cieľom tejto práce bolo preskúmať možnosti detekcie umelo generovaných alebo manipulovaných obrazov na základe ich geometrických a fyzikálnych vlastností. Práca sa zameriavala najmä na analýzu fyzikálnych javoch perespektívy, tieňov a zkradlových odrazov. Pričom práca vychádzala z predpokladu, že generatívne modely umelej inteligencie často nedokážu presne zachovať fyzikálnu konzistenciu scény.

Hlavným prínosom práce je návrh a implementácia semiautomatizovaného algoritmu, ktorý analyzuje geometrickú konzistenciu medzi objektom a jeho zrkadlovým odrazom. Algoritmus využíva moderné metódy počítačového videnia, konkrétne modely SuperPoint a SuperGlue na detekciu a párovanie príznakov, spolu s robustnými štatistickými a geometrickými metódami (RANSAC, IQR, DBSCAN) na filtrovanie dát a identifikáciu hlavného úbežníka odrazu.

Experimentálne testovanie na datasete obsahujúcom 46 obrázkov (17 autentických a 29 umelo generovaných) preukázalo, že navrhovaný prístup je schopný efektívne rozlišovať medzi reálnymi a syntetickými obrazmi. Algoritmus dosiahol celkovú presnosť približne 89,1 % pri striktnej interpretácii výsledkov. Pri autentických fotografiách dosiahol úspešnosť 94,1 % bez výskytu falošne pozitívnych výsledkov, zatiaľ čo pri umelo generovaných obrázkoch bola úspešnosť detekcie 86,2 %. Pri zohľadnení aj hraničných (podozrivých) prípadov sa úspešnosť detekcie približuje k hodnote 93 %.

Výsledky zároveň ukázali, že geometrická analýza zrkadlových odrazov je efektívna detekčná metóda. V reálnych fotografiách dochádza k jasnej konvergencii priamok do úbežníka, zatiaľ čo pri AI generovaných obrazoch sú tieto priesečníky výrazne rozptýlené alebo nekonzistentné.

Napriek dosiahnutým výsledkom má navrhovaný prístup aj svoje obmedzenia. Algoritmus môže zlyhať v prípadoch, keď umelo generovaný obraz obsahuje lokálne konzistentné oblasti, na základe ktorých je nesprávne vyhodnotený ako autentický. Ďalším limitujúcim faktorom je závislosť od kvality vstupných dát a od správneho párovania príznakov, ktoré môže byť ovplyvnené šumom alebo komplexnosťou scény.

Do budúcna by bolo vhodné rozšíriť algoritmus o ďalšie detekčné metódy, napríklad analýzu tieňov alebo perspektívnych vzťahov, a vytvoriť tak všeobecnejší detekčný systém. Navyše aj plná automatizácia algoritmu, vrátane automatického rozdelenia obrazu.

Literatúra

[1] Paul Beardsley and David Murray. Camera calibration using vanishing points. In *Proceedings of the British Machine Vision Conference (BMVC)*. BMVA Press, 1992.

[2] Catherine J. Croucher, Marco Bertamini, and Heiko Hecht. Naïve optics: Understanding the geometry of mirror reflections. *Journal of Experimental Psychology: Human Perception and Performance*, 28(3):546–562, 2002.

[3] Andrew Rabinovich Daniel DeTone, Tomasz Malisiewicz. Superpoint: Self-supervised interest point detection and description. *Arxiv.org*, 26(2), 2018.

[4] Konstantinos G. Derpanis. Overview of the ransac algorithm. 2010.

[5] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. pages 226–231, 1996.

[6] H. Farid. A survey of image forgery detection. *IEEE Signal Processing Magazine*, 26(2), 2009.

[7] H. Farid. Photo forensics: Beyond the pixels. *American Scientist*, 105(5), 2017.

[8] A. Gallagher and H. Farid. Image forgery detection. *ACM Transactions on Graphics (TOG)*, 31(4), 2012.

[9] Firas Ajil Jassim. Image denoising using interquartile range filter with local averaging. *International Journal of Soft Computing and Engineering (IJSCE)*, 2013.

[10] E. Kee, J. F. O'Brien, and H. Farid. Exposing photo manipulation with inconsistent shadows. *ACM Transactions on Graphics (TOG)*, 32(3), 2013.

[11] Sophie J. Nightingale, Kimberley A. Wade, Hany Farid, and Derrick G. Watson. Can people detect errors in shadows and reflections? *Attention, Perception, & Psychophysics*, 81:2917–2943, 2019.

[12] Sarah Jane Delany Padraig Cunningham. k-nearest neighbour classifiers: 2nd edition (with python examples). *Arxiv.org*, 2020.

[13] Tomasz Malisiewicz Paul-Edouard Sarlin, Daniel DeTone. Superglue: Learning feature matching with graph neural networks. *Arxiv.org*, 2020.

[14] Amped Software. How to reveal ai-generated images by checking shadows and reflections in amped authenticate. *Forensic Focus*, 2023. Accessed: 2025-06-09.

Príloha A: obsah elektronickej prílohy

V tejto prílohe sa nachádzajú linky na všetky príklady detekovania autenticity fotiek. Používal som online stránku GeoGebra, ktorú som už v práci spomínal. Tu sú teda príklady vo forme linkov:

1. vanishing points
   (a) https://www.geogebra.org/classic/whdj3xc4
   (b) https://www.geogebra.org/classic/decxfugq
   (c) https://www.geogebra.org/classic/f5ezsarc
   (d) https://www.geogebra.org/classic/xbxr7qvr
2. zrkadlové odrazy
   (a) https://www.geogebra.org/classic/mequgrcu
   (b) https://www.geogebra.org/classic/kcppztak
3. tiene
   (a) https://www.geogebra.org/classic/f7usezw6
   (b) https://www.geogebra.org/classic/dskdmc9y

Príloha B: Používateľská príručka

Manuál GeoGebri je ľahko dostupný online a obsah je veľmi intuitívny.