Baza slik za videonadzor prometa ob omejitvah vgrajene zasebnosti

Jakob Kreft, Marija Ivanovska, Janez Perš

Univerza v Ljubljani-Fakulteta za elektrotehniko, Tržaška cesta 25, 1000 Ljubljana
Tk2693@student.uni-lj.si, marija.ivanovska@fe.uni-lj.si,janez.pers@fe.uni-lj.si

Privacy-By-Design Constrained Traffic Surveillance Dataset

We developed a robust database to facilitate the adaptation of existing neural networks to operate on blurred images, catering to privacy-aware traffic surveillance systems. Using a pair of cameras (one focused, one defocused to achieve various levels of blurring), we collected 4077 images, comprised of sharp images and the images containing two levels of blur. Every blurred image has its sharp, pixel-aligned counterpart. Using the DETR algorithm and manual annotations, we provide reliable labels for object detection. Our analysis demonstrates that blurred images hold considerably less information than sharp ones, showcasing our effective blurring approach. This work paves the way for future research in computer vision based traffic surveillance, particularly in neural network adaptation for blurred images, offering developers a valuable tool for creating efficient, privacy-focused neural networks. The code used to create this dataset is publicly available.

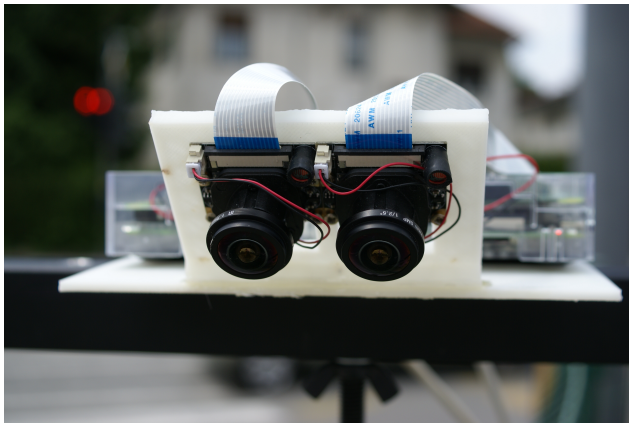
1 Uvod

V dobi digitalne povezanosti postajajo nadzorni sistemi vseprisotni del urbanih infrastruktur po vsem svetu. Ti sistemi, še posebej kamere za nadzor prometa, igrajo ključno vlogo pri zagotavljanju javne varnosti in učinkovitem upravljanju prometa. Čeprav so ti sistemi neprecenljivi za zgoraj omenjene namene, neizogibno sprožajo kritična vprašanja glede zasebnosti.

Zasebnost je temeljna človekova pravica, ki jo moramo spoštovati tudi v digitalnem svetu. Posebno pri nadzoru prometa, kjer se zbirajo obsežne količine podatkov, je nujno najti ravnotežje med zagotavljanjem varnosti in spoštovanjem zasebnosti posameznikov. Ta vprašanja so še toliko bolj pereča, saj tehnologija napreduje in se uporaba nadzornih sistemov širi.

1.1 Vgrajena zasebnost

Vgrajena zasebnost (ang. Privacy by Design) je koncept, ki spodbuja vključevanje zasebnosti in varstva podatkov v začetne faze načrtovanja in razvoja produktov, storitev ali sistemov, namesto da se te zadeve obravnavajo šele po razvoju. To vključuje načrtovanje in implementacijo tehničnih in organizacijskih ukrepov v skladu z načeli varstva podatkov. Glavni cilj koncepta je zagotoviti, da so vse zasebnostne zahteve vključene v oblikovanje in delovanje sistema od samega začetka, s čimer se zmanjšuje tveganje za kršitve zasebnosti in se izboljšuje zaupanje uporabnikov.



Slika 1: Za zajem realistične baze smo potrebovali dve kameri, postavljeni tesno skupaj, z dvema objektivoma z vidnim kotom 175 stopinj. Kameri sta bili napajani in krmiljeni preko povezave Ethernet PoE (angl. Power over Ethernet) in dveh razvojnih kompletov Raspberry Pi 4.

1.2 Vpliv na zasnovo strojne opreme

Vgrajena zasebnost neposredno vpliva na zasnovo strojne opreme z usmeritvijo k minimizaciji podatkov. Ta koncept poudarja, da se podatki, ki niso nujno potrebni za doseganje zastavljenega cilja, ne bi smeli niti zajeti. Koncept ne dovoljuje doseganja zasebnosti s poznejšo obdelavo podatkov, zato morajo biti omejitve že v strojni opremi.

To je bistveno ne le za ohranjanje zasebnosti posameznikov, ampak tudi za izpolnjevanje zakonodajnih zahtev in standardov za varstvo podatkov. Poleg tega pa to krepi zaupanje uporabnikov in javnosti v tehnologijo, ki zmanjšuje tveganje za zlorabo podatkov, takšne videonadzorne tehnologije pa je tako lažje tudi certificirati za uporabo v javnosti.

V tej raziskavi se osredotočamo na pomembno vprašanje: kako izboljšati vgrajeno zasebnost v sistemih za videonadzor prometa? Naša motivacija izhaja iz želje po izboljšanju zasebnosti posameznikov brez ogrožanja funkcionalnosti nadzornih sistemov. Zato smo se v tej

Slika 2: Primer zajetega in poravnanega para ostre in zamegljene slike. Na desni je prikazana zamegljena slika s prvo (najmanj intenzivno) stopnjo zameglitve.

raziskavi osredotočili na pripravo baze slik za razvoj in prilagajanje obstoječih nevronskih mrež za delovanje na zamegljenih slikah, kar bi lahko zagotovilo spoštovanje zasebnosti, hkrati pa ohranilo funkcionalnost nadzornih kamer.

2 Sorodna dela

Na področju obdelave slik so konvolucijske nevronske mreže (CNN) znatno napredovale na področju odstranjevanja zamegljenosti slik [1, 2, 3]. Zato smo se odločili za izdelavo baze z najmanj dvema stopnjama neostrosti slik. Obe stopnji sta bili bolj zamegljeni, kot primeri iz študije Koh in sod. [4].

Izognili smo se umetnemu glajenju oziroma zamegljevanju slik, kot je zameglitev z Gaussovim konvolucijskim filtrom [5]. Namesto tega smo se osredotočili na resnično zameglitev, ki bolje predstavlja kompleksnost in raznolikost zamegljenosti v resničnem svetu [6, 7].

Naravna slika se lahko zamegli iz različnih razlogov, vključno z izgubo fokusa in optičnimi aberacijami [8, 9]. Vendar pa je gibanje, bodisi kamere ali predmetov v prizoru, najpogostejši razlog za zamegljenost naravnih slik [10, 11, 12]. Gibanje povzroča zamegljenost, ki ovira vizualne naloge, kot so prepoznavanje predmetov ali besedila [4, 13, 2].

2.1 Raziskave na področju vgrajene zasebnosti

Obstajajo alternativne metode nadzora prometa, kot je VANET (Vehicular Ad-hoc Network) [15, 16, 17]. Kljub potencialu VANET, smo izbrali kamere, saj so neodvisne od tehnološke opreme vozil in omogočajo univerzalno obdelavo podatkov, ob upoštevanju zasebnosti. Indukcijske zanke, ki zaznavajo vozila preko spremembe elektromagnetnega polja, so pogosta alternativa, vendar imajo omejitve [18, 19]. Naše delo se osredotoča na zamegljene slike za ohranitev zasebnosti in pridobitev podrobnih informacij.

Zhu in sod. [20] so razvili cenovno ugoden sistem za sledenje parkirnih mest, ki ohranja zasebnost z uporabo zameglitvenega filtra na kameri. Ugotovili so, da sistem ostaja natančen kljub degradaciji slike. Sugianta in sod. [21] so razvili sistem za ugotavljanje upoštevanja varne razdalje na letališčih, pri čemer je zasebnost zagotovljena z uporabo prepoznavanja ljudi neposredno na kamerah.

V raziskavah Hinojosa in sod. [22] ter Xia in sod. [23], so bile uporabljene specialne kamere ki zameglijo oziroma popačijo sliko. Cena je lahko velika ovira za uporabo v videonadzornih sistemih. Srivastav in sod. [24] zmanjšajo sliko na 64x48 slikovnih elementov za ohranjanje zasebnosti, kar omogoča zaznavo drže ljudi. Kljub temu obstaja tveganje glede zasebnosti, saj kamera sama še vedno zajema jasno sliko.

2.2 Obstoječe baze slik

Han et al. [25] so razvili bazo slik obrazov z različnimi stopnjami zamegljenosti za izboljšanje algoritmov ocenjevanja zamegljenosti.

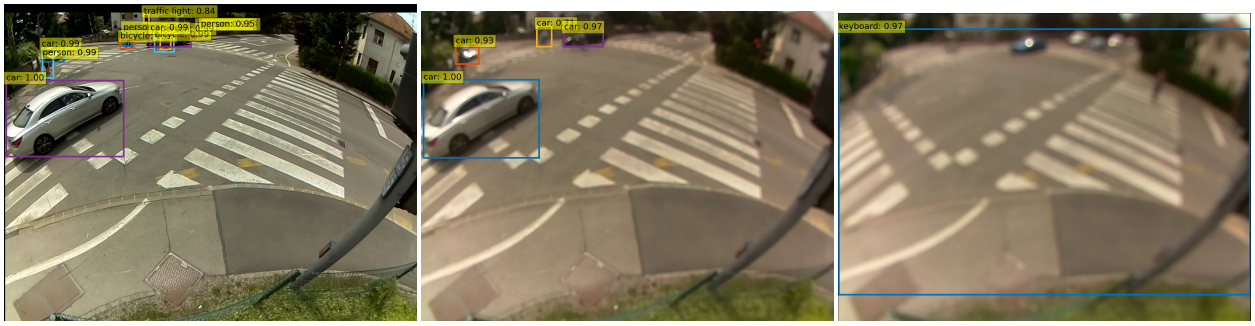
Obstale obstoječe baze, kot so ImageNet, COCO, PASCAL VOC, BDD100K, Visual Genome ne vsebujejo sistematično zamegljenih slik [26], obenem pa nismo zasledili podobnih sistematično urejenih baz fizično razostrenih slik za eksperimente z vgrajeno zasebnostjo.

3 Metodologija

Bazo slik smo zbrali in preizkusili njeno zahtevnost v naslednjih korakih:

1. **Zajem podatkov**: Za zajem prometa v realnem času smo uporabili dva modula Raspberry Pi 4, ki sta bila opremljena s kamerami Electreeks [27] z nastavljivima objektivoma M12. Prva kamera je imela objektiv nastavljen tako, da je zajemala ostre slike prometa, druga kamera pa je bila zamegljena z namerno nastavitvijo objektiva izven fokusa. S tem smo omogočili zajem zamegljenih slik in referenčnih ostrih slik istega prizora, kar je služilo kot podlaga za izdelavo naše baze podatkov. Postavitev je prikazana na Sliki 1.

2. **Obdelava in poravnava slik**: Zbrane slike smo poravnali s pomočjo značilnih točk, pri čemer smo kot transformacijski model uporabili homografijo,



Slika 3: Primer delovanja detekcijskega algoritma DETR [14] . Po vrsti od leve proti desni: ostra slika, zamegljena slika s prvo stopnjo defokusiranja ter slika z drugo stopnjo defokusiranja. Zaradi načina zajema slik imata prvi dve sliki enako vsebino, tretja pa je bila zajeta v drugem časovnem trenutku. Rezultati kažejo, da bo treba algoritme za detekcijo na defokusiranih slikah ciljno učiti na takšnih podatkih.

značilne točke pa so bile tako na zamegljenih kot na ostrih pridobljene avtomatsko preko algoritma SIFT[1]. Na ta način smo odprli možnost prenosa ročnih ali avtomatskih oznak (očrtanih pravokotnikov) iz ostrih na zamegljene slike. Primer zajetih in poravnanih slik je prikazan na Sliki 2.

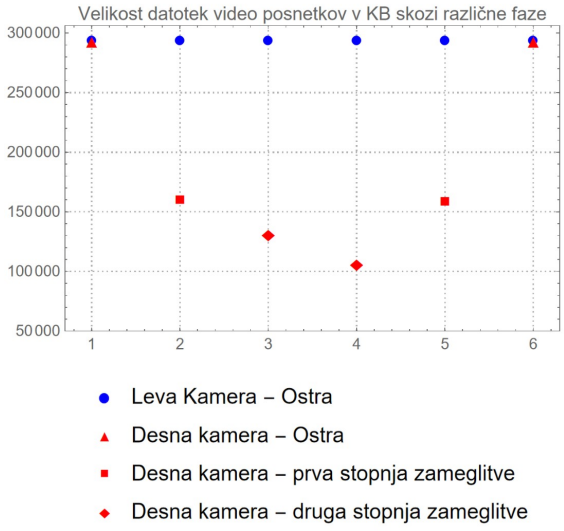
3. **Označevanje**: Z algoritmom DETR [14] smo pridobili očrtane pravokotnike (angl. bounding boxes) za vsako zaznano vozilo na sliki. Ker smo v prejšnjem koraku poravnali slike, so ti očrtani pravokotni veljali tudi za zamegljene slike. Da bi zagotovili natančnost očrtanih pravokotnikov, smo razvili dodatno kodo v jeziku Python z uporabniškim vmesnikom, ki je omogočala ročne korekcije očrtanih pravokotnikov vozil ter dodajanje ali odstranjevanje le-teh, kjer je bilo to potrebno.

4. **Pretvorba podatkov v format COCO**: V zadnji fazi smo vse informacije iz očrtanih pravokotnikov združili v format COCO. Ta korak je bil potreben za zagotovitev kompatibilnosti naših podatkov z obstoječimi orodji za strojno učenje in algoritmi za detekcijo slik.

4 Eksperimenti in rezultati

Prvotni cilj tega dela je bil izvesti primerjavo med delovanjem ene od obstoječih metod detekcije objektov na ostrih slikah na eni strani, in bolj ali manj zamegljenimi slikami na drugi strani. Med delom se je izkazalo, da sposobnost detekcije izbranega detektorja DETR [14] že pri majhni stopnji zamegljenosti radikalno pade, zato kvantitativnih rezultatov ne navajamo. Na Sliki 3 je primerjava med detekcijami na ostri sliki (kjer je algoritem DETR zagrešil samo eno napako) ter na naslednjih dveh stopnjah zameglitve.

Ocenili smo tudi padec količine informacije glede na stopnjo zamegljenosti slik, in sicer preko dolžine datotek, ki jih je strojno podprta kompresija videa po standardu H.264, vgrajena v platformo Raspberry Pi 4 generirala v fiksnem času zajema ob fiksnih nastavitvah, brez premikanja kamere. Na Sliki 4 vidimo, da z defokusiranjem objektiva dejansko informacijska vsebina posnetka močno pade, in da je padec sorazmeren stopnji defokusiranja, ki povzroči zameglitev.



Slika 4: Ocena upada količine informacije glede na stopnjo defokusiranja. Graf prikazuje velikost datotek petminutnih posnetkov, kompresiranih po standardu H.264, ki so bili posneti z različnimi stopnjami zameglitve. V vseh fazah je leva kamera snemala ostre slike. Nasprotno pa je desna kamera med fazami snemala z različnimi stopnjami zameglitve. V prvi fazi je snemala jasno, v drugi fazi je bila uporabljena prva stopnja zameglitve, v tretji fazi druga stopnja zameglitve. Po zamenjavi lokacije je v četrti fazi ponovno snemala z drugo stopnjo zameglitve, v peti fazi s prvo stopnjo zameglitve, v šesti fazi pa je snemala ostro.

5 Zaključek

Zajeli, uredili, označili in preliminarno preizkusili smo bazo parov ostrih in razostrenih slik, ki bo služila prilagoditvi algoritmov računalniškega vida na delovanje ob omejitvah vgrajene zasebnosti. V našem primeru smo se osredotočili na videonadzor prometa in zagotavljanje zasebnosti s pomočjo navadnih cenovno dostopnih objektivov, ki jih fiksiramo tako, da ne zagotavljajo fokusirane

[1] Algoritem SIFT je presenetljivo dobro deloval tudi na zamegljenih slikah.

slike v opazovanem območju. Vsa koda, ki je bila uporabljena pri tem delu je prosto dostopna na `https://github.com/jakobkreft/diplomsko_delo/`. Pomembna prednost našega dela pred preprostim megljenjem slik s pomočjo Gaussovega filtra je v tem, da naša baza popolnoma replicira fizikalne pojave, ki nastanejo ob dejanskem defokusiranju objektiva (razostritev se zgodi pred zajemom slike, vzorčenjem in kvantizacijo slikovnih elementov). Takšno realistično bazo lahko poleg učenja metod detekcije objektov uporabimo tudi za testiranje algoritmov odprave zamegljenosti in s tem preverimo, ali res zagotavlja predviden nivo zasebnosti.

Soavtorji so bili sofinancirani iz naslednjih virov ARIS: projekti J2-2506 in J2-2501 (A) ter raziskovalni programi P2-0095 in P2-0250 (B).

Literatura

[1] K. Zhang, W. Ren, Y. Zhang, W.-S. Lai, B. Stenger, M.-H. Yang, and H. Li, "Deep image deblurring: A survey," vol. 130, pp. 2103–2130, 06 2022.

[2] M. Hradis, J. Kotera, P. Zemcik, and F. Sroubek, "Convolutional neural networks for direct text deblurring," *Convolutional Neural Networks for Direct Text Deblurring*, 01 2015.

[3] K. Zhang, W. Ren, W. Luo, W. Lai, B. Stenger, M. Yang, and H. Li, "Deep image deblurring: A survey," *CoRR*, vol. abs/2201.10700, 2022. [Online]. Available: https://arxiv.org/abs/2201.10700

[4] J. Koh, J. Lee, and S. Yoon, "Single-image deblurring with neural networks: A comparative survey," *Computer Vision and Image Understanding*, vol. 203, p. 103134, 2021.

[5] E. S. Gedraite and M. Hadad, "Investigation on the effect of a gaussian blur in image filtering and segmentation," in *Proceedings ELMAR-2011*, 2011, pp. 393–396.

[6] Y.-Q. Liu, X. Du, H.-L. Shen, and S.-J. Chen, "Estimating generalized gaussian blur kernels for out-of-focus image deblurring," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 3, pp. 829–843, 2021.

[7] A. Ciancio, A. Costa, E. da Silva, A. Said, R. Samadani, and P. Obrador, "No-reference blur assessment of digital pictures based on multifeature classifiers," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 20, pp. 64–75, 01 2011.

[8] R. L. Lagendijk and J. Biemond, "Chapter 14 - basic methods for image restoration and identification," in *The Essential Guide to Image Processing*, A. Bovik, Ed. Boston: Academic Press, 2009, pp. 323–348.

[9] R. Wang and D. Tao, "Recent progress in image deblurring," 2014, arXiv.

[10] Y. Yitzhaky and N. Kopeika, "Identification of blur parameters from motion blurred images," *Graphical Models and Image Processing*, vol. 59, no. 5, pp. 310–320, 1997.

[11] S. K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 6, pp. 689–698, 2004.

[12] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *Acm transactions on graphics (tog)*, vol. 27, no. 3, pp. 1–10, 2008.

[13] X. Qi, L. Zhang, and C. Tan, "Motion deblurring for optical character recognition," in *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, 2005, pp. 389–393 Vol. 1.

[14] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Computer Vision–ECCV 2020: Proceedings, Part I 16*. Springer, 2020, pp. 213–229.

[15] K. A. Hafeez, L. Zhao, B. Ma, and J. W. Mark, "Performance analysis and enhancement of the dsrc for vanet's safety applications," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3069–3083, 2013.

[16] M. R. Ghori, K. Z. Zamli, N. Quosthoni, M. Hisyam, and M. Montaser, "Vehicular ad-hoc network (vanet): Review," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, 2018, pp. 1–6.

[17] C. Zhang, L. Zhu, C. Xu, X. Du, and M. Guizani, "A privacy-preserving traffic monitoring scheme via vehicular crowdsourcing," *Sensors*, vol. 19, no. 6, p. 1274, 2019.

[18] D. D. Romero, A. Aprabuwono, Taufik, and A. Hasniaty, "A review of sensing techniques for real-time traffic surveillance," *Journal of Applied Sciences*, vol. 11, no. 1, pp. 192–198, 2011.

[19] "Chapter 2, traffic detector handbook: Third edition—volume i - fhwa-hrt-06-108," dostopno: 5.7.2023. [Online]. Available: https://www.fhwa.dot.gov/publications/research/operations/its/06108/02.cfm

[20] H. Zhu, S. Fan, X. Wang, and S. C.-K. Chau, "Privacy-preserving camera-based monitoring and tracking system for parking spaces," in *Proceedings of the 7th ACM BuildSys conference*. New York, NY, USA: Association for Computing Machinery, 2020, p. 346–347.

[21] N. Sugianto, D. Tjondronegoro, R. Stockdale, and E. I. Yuwono, "Privacy-preserving ai-enabled video surveillance for social distancing: Responsible design and deployment for public spaces," *Information Technology & People*, no. ahead-of-print, 2021.

[22] C. Hinojosa, J. C. Niebles, and H. Arguello, "Learning privacy-preserving optics for human pose estimation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 2573–2582.

[23] Y. Xia, Y. Tang, Y. Hu, and G. Hoffman, "Privacy-preserving pose estimation for human-robot interaction," *arXiv preprint arXiv:2011.07387*, 2020.

[24] V. Srivastav, A. Gangi, and N. Padoy, "Human pose estimation on privacy-preserving low-resolution depth images," in *Proceedings of MICCAI 2019: 22nd International Conference*. Springer, 2019, pp. 583–591.

[25] Q. Han, M. Zhang, C. Song, Z. Wang, and X. Niu, "A face image database for evaluating out-of-focus blur," in *Proceedings of the 2008 Eighth International Conference on Intelligent Systems Design and Applications - Volume 02*, ser. ISDA '08. USA: IEEE Computer Society, 2008, p. 277–282.

[26] A. Salari, A. Djavadifar, X. Liu, and H. Najjaran, "Object recognition datasets and challenges: A review," *Neurocomputing*, vol. 495, pp. 129–152, 2022.

[27] "Raspberry pi kamera v2 nightvision full HD fisheye — electreeks," Electreeks, 11 2020. [Online]. Available: https://electreeks.de/shop/raspberry-pi-kamera-nachtsicht-mit-infrarot-sperrfilter-175-grad-weitwinkel-objektiv/