

Diskriminativna metoda za detekcijo 3D anomalij

Vitjan Zavrtanik¹, Matej Kristan¹, Danijel Skočaj¹

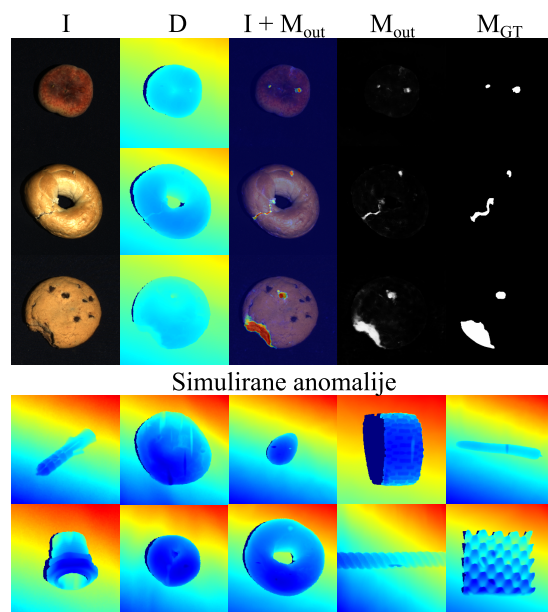
¹Univerza v Ljubljani, Fakulteta za Računalništvo in informatiko
E-pošta: vitjan.zavrtanik@fri.uni-lj.si

A discriminative method for 3D anomaly detection

Recent methods for surface anomaly detection are based on feature extraction using pretrained networks. On RGB anomaly detection datasets these methods achieve excellent results, but the results for 3D anomaly detection are worse due to the lack of pretrained networks suitable for this domain. There is also a lack of industrial depth image datasets that allow learning of networks that could be used in these methods. Discriminative anomaly detection methods do not require pretrained networks and learn with simulated anomalies. The process of simulating anomalies appropriate to the domain of industrial depth data is non-trivial and is necessary for training discriminative methods. We propose a novel 3D anomaly simulation process that is suitable for learning discriminative methods. We demonstrate the effectiveness of the process using DRÆM-3D, a strong discriminative method for 3D anomaly detection. The proposed approach achieves excellent results on the MVTec3D anomaly detection database, where DRÆM-3D outperforms all previous state-of-the-art methods on both the 3D anomaly detection problem and the 3D+RGB anomaly detection problem.

1 UVod

Problem detekcije anomalij je lokalizacija območij na sliki, ki odstopajo od normalnega videza predmeta. Pri standardni postavitvi problema so med učenjem na voljo le slike predmeta brez anomalij. Metode se učijo modela brezhibnega videza predmeta, odstopanja od tako naučenega modela pa kategorizirajo kot anomalije. Nedavne metode, ki dosegajo odlične rezultate [9, 11, 15], temeljijo na značilkah, pridobljenih s pomočjo mrež, ki so bile predhodno učene na podatkovnih množicah slik RGB [5]. Te informativne značilke omogočajo zelo dobre rezultate na podatkovnih množicah detekcije površinskih anomalij RGB, kot je MVTec AD [1]. V praksi ni mogoče zaznati vseh anomalij le na podlagi slik RGB, zato se nedavna dela [13, 2, 12] osredotočajo na novo raziskovalno vprašanje detekcije anomalij v 3D in RGB. V tej postavitvi problema sta za detekcijo anomalij na voljo RGB slika in 3D sken predmeta. Metoda M3DM [13], ki dosega odlične rezultate, uporablja predhodno naučeno mrežo za obdelavo oblakov točk, da iz 3D podatkov objekta pridobi informativne značilke. Nato se za detekcijo anoma-



Slika 1: Predlagana simulacija anomalij globine omogoča učenje učinkovitih modelov diskriminativne zaznave anomalij, s čimer se zaobide potreba po predhodno naučenih mrežah.

lij uporabi model, podoben modelu PatchCore [9]. Predstavitve, pridobljene s pomočjo mreže za obdelavo oblakov točk, so pomanjkljive, saj obstaja velika domenska razlika med podatki, uporabljenimi za predhodno učenje mreže, in industrijskimi globinskimi podatki. To vodi do suboptimalne učinkovitosti M3DM pri zaznavanju 3D anomalij [13]. V metodi AST [12] se za modeliranje normalnega videza predmeta tako v RGB kot v 3D domeni uporablja tokovni model. Slike RGB predstavljajo značilke pridobljene s predhodno naučeno mrežo, medtem ko se za predstavitev 3D podatkov uporabljajo surove vrednosti posameznih regij globinskih slik. Surove globinske vrednosti ne vsebujejo informacij, ki bi bile semantično pomenljive glede na večje območje predmeta, zato niso optimalna predstavitev globinskih informacij in povzročajo slabo učinkovitost AST pri zaznavanju 3D anomalij [12].

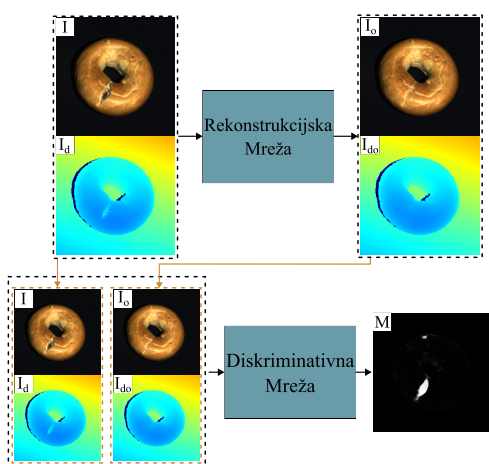
Diskriminativni modeli detekcije površinskih anomalij, kot je DRÆM [16], ne zahtevajo predhodno naučenih mrež, temveč zahtevajo dobro opredeljen proces simulacije anomalij. Predlagamo nov pristop k simulaciji ano-

malij, ki je primeren za detekcijo 3D površinskih anomalij, poleg tega pa predlagamo DRÆM-3D, prilagoditev metode DRÆM, ki anomalije zaznava tako v 3D+RGB kot samo v 3D (Slika 1). Predlagani DRÆM-3D doseže odlične rezultate na zahtevni podatkovni množici detekcije 3D površinskih anomalij MVTEc3D [2].

2 Metoda

DRÆM-3D uporablja diskriminativni model za detekcijo anomalij DRÆM [16] kot svoj osrednji model, pri čemer je edina prilagoditev števila kanalov, ki jih uporabljata rekonstrukcijska in diskriminativna mreža, tako da upošteva globinske informacije. Metoda DRÆM [16] in postopek učenja DRÆM-3D sta opisani v Poglavlju 2.1. Postopek simulacije 3D anomalij je opisan v Poglavlju 2.2.

2.1 DRÆM-3D



Slika 2: Arhitektura DRÆM-3D. Operacija združevanja po kanalih je prikazana s črnimi prekinjenimi črtami. Rekonstrukcijska mreža sprejme vhodno RGB (I) in globinsko (I_d) sliko ter vrne izhodni rekonstrukciji brez anomalij I_o in I_{od} . Diskriminativna mreža iz vhodnih slik I, I_d, I_o in I_{od} vrne masko segmentacije anomalij M .

DRÆM [16] je sestavljen iz rekonstrukcijske in diskriminativne mreže. Vhodna slika I je poslana skozi rekonstrukcijsko mrežo, ki obnovi videz slike I brez anomalij, če je I anomalna, in ohrani videz I , če je I brez anomalij. Rekonstrukcijska mreža implicitno zazna anomalne regije in jih rekonstruira z brezhlebnim videzom predmeta. Rezultat rekonstrukcijske mreže I_o se nato združi po dimenziji kanalov s sliko I in se vstavi v diskriminativno mrežo. Diskriminativna mreža vrne masko segmentacije anomalij M . Slike I in I_o se znatno razlikujejo na območjih, kjer je I anomalna, kar zagotavlja močan signal za diskriminativno mrežo in omogoča natančno segmentacijo anomalij.

DRÆM je učen samo s slikami brez anomalij, saj so med učenjem anomalije simulirane in vstavljene v slike brez anomalij. Simulirane anomalije so ustvarjene tako, da najprej vzamemo sliko brez anomalij I iz učne množice. Nato uporabimo generator Perlinovega šuma [8], da ustvarimo masko anomalij M . Iz nabora slik izven distribucije nato izberemo sliko I_{ood} . Nato konstruiramo

umetno anomalno sliko, ki vsebuje simulirane anomalije na območjih, določenih z masko M :

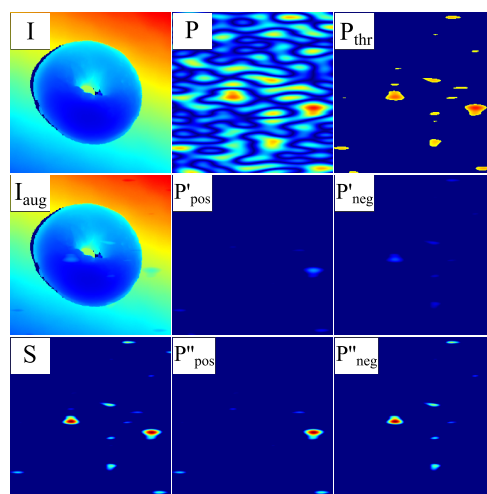
$$I_{aug} = \bar{M}I + M(\beta I + (1 - \beta)I_{ood}), \quad (1)$$

kjer je β parameter intenzitete anomalije, ki je vzorčen v intervalu $[0.1, 1.0]$, \bar{M} pa je inverz maske M .

DRÆM je tako učen na slikah brez anomalij I , slikah s simuliranimi anomalijami I_{aug} in pripadajočimi maskami anomalij M . Rekonstrukcijska mreža je naučena z uporabo funkcij izgube $L2$ in SSIM [14] za rekonstrukcijo slik, medtem ko je diskriminativna mreža učenca z uporabo fokalne funkcije izgube [7] za segmentacijo.

DRÆM-3D sledi postopku učenja DRÆM, vendar zahteva 4 vhodne kanale, 3 za RGB in 1 za globinsko sliko. Arhitektura DRÆM-3D je prikazana na Sliki 2 in tesno sledi arhitekturi DRÆM [16].

2.2 Proces simulacije 3D anomalij



Slika 3: Proces simulacije anomalij.

DRÆM uporablja strategijo kopiranja in lepljenja za simulacijo anomalij na podatkih RGB, vendar isti pristop ni primeren za globinske podatke zaradi nekaterih lastnosti industrijskih globinskih podatkov. Prvič, globina od točke najbližje senzorju do najbolj oddaljene točke se v veliko primerih povečuje postopoma, medtem ko v slikah RGB zaradi pogojev osvetlitve ali lastnosti predmeta pogosto pride do ostrih sprememb v vrednostih. Drugič, minimalne lokalne spremembe v globinskih vrednostih lahko vsebujejo pomembne informacije, kot je sprememba teksture, vendar pa se take minimalne spremembe v slikah RGB pogosto lahko zanemarijo. Preprosto lepljenje območij iz globinskih slik izven distribucije je zato slaba strategija za simulacijo globinskih anomalij.

Z namenom določitve enotnega postopka simulacije za vse razrede predmetov se slike globine I_d najprej normalizirajo v vrednosti med 0 in 1 na podlagi najmanjše in največje globinske vrednosti iz nabora slik za učenje. Nato se vrednosti skalirajo med 0.1 in 0.9, da se zagotovi prostor za morebitne anomalije, ki odstopajo višje ali nižje od najmanjše in največje vrednosti iz učne množice.

Podobno kot pri DRÆM, je začetna maska anomalij P najprej ustvarjena s pomočjo generatorja Perlinovega

Tabela 1: Rezultati detekcije anomalij na podatkovni množici MVTec3D za postavitev problema 3D, RGB in 3D+RGB. Rezultati so navedeni kot AUROC vrednosti na ravni slike. Rezultati ocenjenih metod so razvrščeni, pri čemer so prvo, drugo in tretjevrščeni označeni.

	Method	Bagel	Cable Gland	Carrot	Cookie	Dowel	Foam	Peach	Potato	Rope	Tire	Mean
3D	FPFH [6]	82.5	55.1	95.2	79.7	88.3 ^③	58.2	75.8	88.9 ^③	92.9	65.3 ^③	78.2
	3D-ST [3]	86.2	48.4	83.2	89.4	84.8	66.3	76.3	68.7	95.8 ^②	48.6	74.8
	AST _{3D} [12]	88.1 ^③	57.6 ^③	96.5 ^①	95.7 ^③	67.9	79.7 ^②	99.0 ^①	91.5 ^②	95.6 ^③	61.1	83.3 ^③
	M3DM _{3D} [13]	94.1 ^②	65.1 ^②	96.5 ^①	96.9 ^②	90.5 ^②	76.0 ^③	88.0 ^③	97.4 ^①	92.6	76.5 ^②	87.4 ^②
	DRÆM-3D	96.5 ^①	88.5 ^①	96.5 ^①	99.7 ^①	95.0 ^①	92.6 ^①	95.3 ^②	81.7	96.8 ^①	91.4 ^①	93.4 ^①
RGB	PatchCore [9]	87.6	88.0	79.1	68.2	91.2	70.1	69.5	61.8	84.1	70.2	77.0
	DifferNet [10]	85.9	70.3	64.3	43.5	79.7	79.0	78.7	64.3 ^③	71.5	59.0	69.6
	PADiM [4]	97.5 ^①	77.5	69.8	58.2	95.9	66.3	85.8 ^③	53.5	83.2	76.0	76.4
	CS-Flow [11]	94.1	93.0 ^①	82.7	79.5 ^②	99.0 ^①	88.6 ^③	73.1	47.1	98.6 ^②	74.5	83.0
	AST _{RGB} [12]	94.7 ^②	92.8 ^②	85.1 ^②	82.5 ^①	98.1 ^②	95.1 ^①	89.5 ^②	61.3	99.2 ^①	82.1 ^②	88.0 ^①
	M3DM _{RGB} [13]	94.4 ^③	91.8 ^③	89.6 ^①	74.9 ^③	95.9 ^③	76.7	91.9 ^①	64.8 ^②	93.8 ^③	76.7 ^③	85.0 ^②
DRÆM [16]	94.2	90.6	81.6 ^③	66.6	78.3	91.9 ^②	65.6	78.4 ^①	98.6 ^②	92.5 ^①	83.8 ^③	
3D+RGB	PatchCore+FPFH [6]	91.8	74.8	96.7 ^③	88.3	93.2	58.2	89.6	91.2 ^③	92.1	88.6 ^②	86.5
	AST [12]	98.3 ^③	87.3 ^③	97.6 ^①	97.1 ^②	93.2 ^③	88.5 ^③	97.4 ^①	98.1 ^①	100 ^①	79.7	93.7 ^③
	M3DM [13]	99.4 ^②	90.9 ^②	97.2 ^②	97.6 ^①	96.0 ^②	94.2 ^②	97.3 ^③	89.9	97.2 ^②	85.0 ^③	94.5 ^②
	DRÆM-3D	99.6 ^①	92.8 ^①	93.1	88.8 ^③	98.4 ^①	99.2 ^①	92.2 ^③	95.1 ^②	100 ^①	92.3 ^①	95.2 ^①

šuma [8], kjer so vrednosti pikslov na lokaciji $P_{xy} \in (-1, 1)$. Slika P je binarizirana na podlagi absolutne vrednosti s pragom $T = 0.5$, kar pripelje do maske anomalij M . DRÆM nato ohrani le M in zavrne P . Vendar P vsebuje lokalno zvezne regije, ki tesno sovpadajo s lastnostmi industrijskih globinskih podatkov in se lahko uporabijo za simuliranje 3D anomalij. Najprej so vrednosti P z absolutnimi vrednostmi manjšimi od T nastavljeni na 0 v P_{thr} . Absolutne vrednosti P_{thr} na območjih, ki ustrezajo pozitivnim vrednostim v M , so vedno večje od 0.5 zaradi postopka binarizacije. Ta območja so najprej razdeljena na pozitivne in negativne vrednosti, s čimer dobimo dve komponenti P_{pos} in P_{neg} , kjer je $P_{thr} = P_{pos} + P_{neg}$. Vrednosti obeh komponent so premaknjene tako, da so najmanjše absolutne vrednosti komponent enake 0, torej: $P'_{neg} = P_{neg} + T$ in $P'_{pos} = P_{pos} - T$. Nato se komponenti skalira na podlagi največjih absolutnih vrednosti: $P''_{pos} = P'_{pos}/\max(\text{abs}(P'_{pos}))$, $P''_{neg} = P'_{neg}/\max(\text{abs}(P'_{neg}))$. Končne vrednosti anomalij se nato seštejejo in dobimo $S = P''_{pos} + P''_{neg}$, ki se uporablja kot slika za vir globinskih anomalij. Koraki postopka simulacije anomalij so prikazani na Sliki 3.

Končna slika anomalije globine je nato sestavljena iz globinske slike brez anomalij I_d , generirane maske segmentacije M in slike za vir anomalij S :

$$I_{aug} = I_d \odot M + \overline{M} \odot (I_d + \lambda S), \quad (2)$$

kjer simbol \odot predstavlja elementno množenje (Hadamardov produkt), \overline{M} je obratna vrednost maske M , in λ je utežni parameter, vzorčen iz intervala (0.02, 0.1), ki določa intenziteto dodanih anomalij.

3 Rezultati

Podatkovna množica in vrednotenje DRÆM-3D je ocenjen na standardni podatkovni množici detekcije anomalij MVTec3D [2]. Podatkovna baza MVTec3D [2] vsebuje 4147 3D skenov, pridobljenih s pomočjo industrijskega 3D senzorja, ki hkrati beleži tudi RGB podatke. 894 skenov vsebuje anomalije, vidne bodisi v RGB, 3D ali 3D+RGB.

Vse metode so vrednotene na 3 postavitvah problema v podatkovni množici MVTec3D [2]. V postavitvi 3D je na voljo samo 3D informacija, v postavitvi RGB so na voljo samo slike RGB, v postavitvi 3D+RGB pa sta na voljo obe modalnosti. Nekatere anomalije so vidne le v 3D ali RGB, zato popolne vrednosti evalvacijskih metrik ne bi smele biti dosegljive v postavitvah samo 3D ali samo RGB.

Postopek učenja DRÆM-3D je učen za 30000 iteracij s paketno velikostjo 16 in stopnjo učenja 0.0002. Približno polovica slik v vsakem paketu vsebuje simulirane anomalije med učenjem.

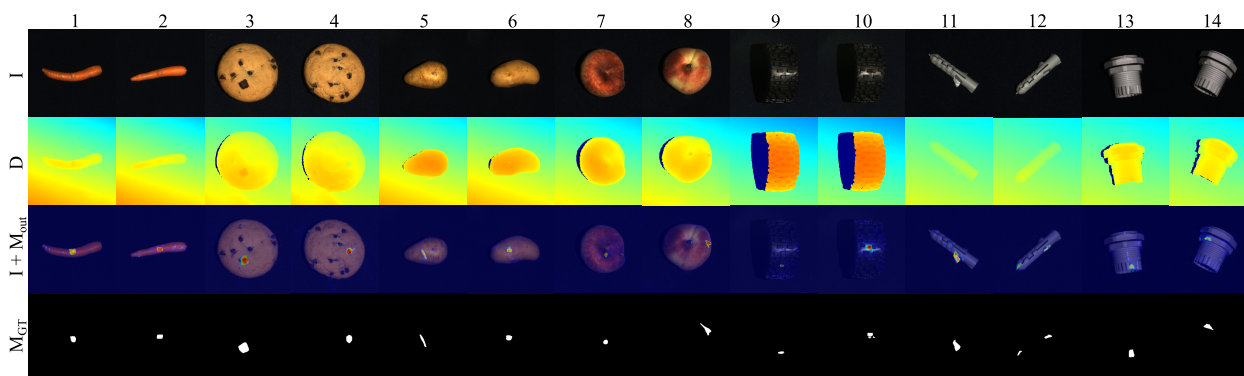
3.1 Rezultati detekcije anomalij

Tabela 1 vsebuje rezultate detekcije anomalij na ravni slike, ocenjene z metriko AUROC. V postavitvi 3D DRÆM-3D presega prejšnjo najboljšo metodo M3DM [13] za 6 odstotnih točk, kar kaže na močnejšo sposobnost izkoriščanja razpoložljivih 3D podatkov. V postavitvi 3D+RGB DRÆM-3D še vedno presega M3DM [13], vendar se razlika v zmogljivosti zmanjša na 0.8 odstotne točke. M3DM uporablja predhodno naučeno mrežo za ekstrakcijo značilik iz RGB slik in zelo dobro modelira RGB podatke, kar je verjetno razlog, da lahko zmanjša razliko med M3DM in DRÆM-3D v postavitvi 3D+RGB. V postavitvi RGB je DRÆM-3D zamenjan z DRÆM, saj se med učenjem uporabljajo samo anomalije simulirane v RGB. Novejše metode, kot so M3DM [13] in AST [12], presegajo DRÆM [16] v postavitvi RGB, verjetno zaradi uporabe močnih prednaučenih mrež pri gradnji modelov normalnega videza.

Slika 4 prikazuje kvalitativne rezultate DRÆM-3D na podatkovni množici MVTec3D [2]. Tudi pri razredih kjer je globinska informacija ključnega pomena za detekcijo anomalij (stolpci 1-8, Slika 4), DRÆM-3D natančno lokalizira anomalije. DRÆM-3D zazna tudi anomalije vidne v RGB (stolpci 9-14).

4 Zaključek

Predstavljena je bila nova metoda za simulacijo anomalij za industrijske globinske slike, ki omogoča učenje diskriminativnih metod detekcije 3D anomalij, kjer je nujen



Slika 4: Kvalitativni rezultati DRÆM-3D na podatkovni množici MVTEC3D. Prikazane so vhodna slika RGB I , globinska slika D , pravilna maska anomalij M_{GT} in izhodna maska vrisana na sliki RGB ($I+M_{out}$).

postopek simulacije anomalij za 3D domeno. Predlagana je bila metoda DRÆM-3D za detekcijo površinskih anomalij v 3D, ki dosega rezultate, ki presegajo doslej najboljšo metodo M3DM [13] za 6 odstotnih točk v postavitvi 3D in za 1 odstotno točko v postavitvi 3D+RGB na podatkovni množici detekcije anomalij MVTEC3D [2].

Predlagani postopek simulacije anomalij bi se lahko uporabil pri diskriminativnih pristopih detekcije anomalij, ki uporabljajo predhodno naučene mreže za ekstrakcijo značilk iz RGB slik, pri čemer bi se normalen videz RGB naučile iz značilk prednaučenih mrež, normalen videz 3D pa bi se lahko naučile s pomočjo simuliranih anomalij. To bi verjetno izboljšalo zmogljivost na postavitvi RGB+3D.

5 Zahvala

To delo je bilo delno financirano s strani projekta ARIS J2-3169 (MV4.0) in raziskovalnega programa Računalniški vid (P2-0214).

Literatura

- [1] Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: MV-Tec AD – A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 9592–9600 (2019)
- [2] Bergmann, P., Jin, X., Sattlegger, D., Steger, C.: The mvtec 3d-ad dataset for unsupervised 3d anomaly detection and localization. In: Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 5: VISAPP. pp. 202–213 (2022)
- [3] Bergmann, P., Sattlegger, D.: Anomaly detection in 3d point clouds using deep geometric descriptors. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 2613–2623 (2023)
- [4] Defard, T., Setkov, A., Loesch, A., Audigier, R.: Padim: A patch distribution modeling framework for anomaly detection and localization. In: International Conference on Pattern Recognition. pp. 475–489. Springer (2021)
- [5] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
- [6] Horwitz, E., Hoshen, Y.: An empirical investigation of 3d anomaly detection and segmentation. arXiv preprint arXiv:2203.05550 (2022)
- [7] Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
- [8] Perlin, K.: An image synthesizer. ACM Siggraph Computer Graphics **19**(3), 287–296 (1985)
- [9] Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., Gehler, P.: Towards total recall in industrial anomaly detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14318–14328 (2022)
- [10] Rudolph, M., Wandt, B., Rosenhahn, B.: Same same but different: Semi-supervised defect detection with normalizing flows. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 1907–1916 (2021)
- [11] Rudolph, M., Wehrbein, T., Rosenhahn, B., Wandt, B.: Fully convolutional cross-scale-flows for image-based defect detection. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 1088–1097 (2022)
- [12] Rudolph, M., Wehrbein, T., Rosenhahn, B., Wandt, B.: Asymmetric student-teacher networks for industrial anomaly detection. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 2592–2602 (2023)
- [13] Wang, Y., Peng, J., Zhang, J., Yi, R., Wang, Y., Wang, C.: Multimodal industrial anomaly detection via hybrid fusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8032–8041 (2023)
- [14] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
- [15] Yu, J., Zheng, Y., Wang, X., Li, W., Wu, Y., Zhao, R., Wu, L.: Fastflow: Unsupervised anomaly detection and localization via 2d normalizing flows. arXiv preprint arXiv:2111.07677 (2021)

- [16] Zavrtnik, V., Kristan, M., Skočaj, D.: Draem - a discriminatively trained reconstruction embedding for surface anomaly detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 8330–8339 (October 2021)