Analiza robustnosti globokih nenadzorovanih metod za detekcijo vizualnih anomalij

Jakob Božič, Vitjan Zavrtanik, Danijel Skočaj

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani
E-pošta: {jakob.bozic, vitjan.zavrtanik, danijel.skocaj}@fri.uni-lj.si

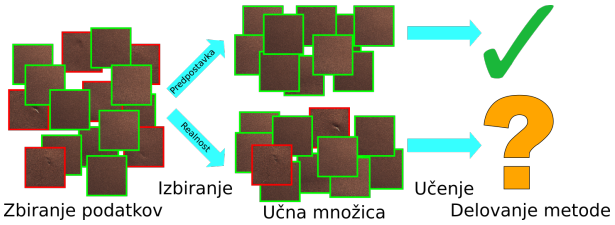
Analysis of robustness of deep unsupervised methods for visual anomaly detection

Unsupervised generative methods have recently attracted significant attention in the field of industrial visual anomaly detection, mainly owing to their ability to learn from non anomalous data withouth requiring anomalous samples and pixel-level labels, which are costly to obtain. An assumption that anomalous data are always correctly identified and consequently removed from the training set underlies all of the generative methods. In practice, however, correctly identifying every single anomalous image can either be very costly to do or it can not be done at all due to the nature of the problem. In this paper, we analyze how robust some of the recently proposed generative methods for anomaly detection are, by introducing anomalous data in the training process. Our analysis covers 3 methods and 4 datasets with 8 categories in total, and we conclude that while some of the methods are more robust than others, introducing a minor percentage of anomalous data in the training set does not significantly deteriorate the performance.

1 Uvod

Detekcija anomalij v slikah naslavlja problem identifikacije primerov, ki odstopajo od pričakovanega izgleda; na ta problem pogosto naletimo pri detekciji površinskih napak na izdelkih pri nadzoru kvalitete na proizvodnih linijah. V zadnjem obdobju je bilo razvitih veliko metod, ki temeljijo na paradigmi globokega učenja, primernih za detekcijo izstopajočih slik ter lokalizacijo anomalnih regij v slikah. Metode se po načinu delovanja v grobem delijo na dva dela: diskriminativne metode, ki modelirajo predstavitev tako normalnih kot anomalnih primerov ter nato razlikujejo med temi, in generativne metode, ki modelirajo le predstavitev normalnih primerov, detekcijo anomalij pa nato izvajajo z ocenjevanjem odstopanja predstavitve primera od pričakovane. Zbiranje množice anomalnih primerov, ki ustrezno predstavljajo porazdelitev možnih napak, je v praksi lahko težko izvedljivo, saj se anomalni primeri praviloma pojavljajo redko, posledično pa diskriminativne metode ne zmorejo dovolj dobro modelirati predstavitve anomalnih primerov. Dodaten problem za diskriminativne metode predstavlja tudi potreba po označevanju na nivoju slikovnih elementov, kar zahteva veliko vloženega dela, zato se namesto diskriminativnih metod vse bolj uveljavljajo generativne metode, ki teh dveh problemov nimajo.

Za učenje predstavitve normalnih primerov generativne metode učimo le na normalnih primerih, kar pomeni, da moramo pred učenjem vse učne primere podrobno pregledati, morebitne anomalne primere pa izločiti. To je lahko časovno zelo potratno, poleg tega pa je včasih za nekatere primere težko oz. skoraj nemogoče natančno določiti, ali so normalni ali anomalni. V dosedanji literaturi [4, 7, 3] se predpostavlja, da je proces izločanja anomalnih primerov iz učne množice izveden brez napak, kar pa je v praksi pogosto nemogoče zagotoviti. V tem prispevku zato analiziramo robustnost generativnih metod na prisotnost anomalnih primerov med učenjem (slika 1).



Slika 1: Proces zbiranja in ročnega izbiranja normalnih primerov za učenje generativnih metod. Navadno predpostavljamo, da proces poteka brez napak [7, 4, 3], kar pa je v realnih situacijah včasih nemogoče zagotoviti. Analizirali bomo, kaj se zgodi, če v učni množici ostanejo anomalni primeri.

2 Analizirane generativne metode

Generativne metode modelirajo predstavitev normalnih primerov. Med seboj jih razlikujemo po dveh ključnih lastnostih; načinu predstavitve in merjenju odstopanja od normalnosti. Razvitih je bilo veliko generativnih metod, analizirali bomo tri metode, ki so bile nedavno predlagane in dosegajo odlične rezultate.

RIAD RIAD (Reconstruction by Inpaiting for Anomaly Detection) [7] se predstavitve normalnih primerov uči preko globoke nevronske mreže, ki je naučena, da dopolnjuje manjkajoče regije v sliki (angl. inpainting). Med učenjem se na normalnih slikah posamezne regije maskira (vse slikovne elemente v regiji se nastavi na 0), od mreže pa pričakujemo, da bo iz konteksta sosednih regij znala rekonstruirati maskirano regijo. Predpostavljamo, da se bo mreža tako naučila rekonstruirati normalne regije, anomal-

nih pa ne, saj le-te niso prisotne med učnimi primeri. V fazi razpoznave sliko, za katero nas zanima, ali je anomalna ali ne, najprej razbijemo na več kopij, v vsaki kopiji je nekaj regij maskiranih, nekaj pa ne, vsaka regija pa je maskirana v natančno eni kopiji. Nato vse maskirane regije v vseh kopijah rekonstruiramo z naučeno mrežo, iz rekonstruiranih regij pa sestavimo rekonstrukcijo celotne slike. Oceno anomalnosti slike dobimo tako, da izmerimo odstopanje prvotne slike od rekonstruirane.

Gaussian AD Metoda, predlagana v [4], uporablja za predstavitev primerov prednaučeno nevronsko mrežo. Posamezne primere opiše z značilkami, ki jih proizvede mreža, ki je bila diskriminativno učena za razpoznavanje različnih kategorij slik na podatkovni zbirki ImageNet. Značilke, pridobljene na normalnih slikah, nato opiše z multivariantno Gaussovo porazdelitvijo in tako zgradi model normalnosti. V fazi razpoznave za sliko, za katero nas zanima, ali je anomalna ali ne, najprej pridobimo značilke iz prednaučene mreže, nato pa izračunamo odstopanje od normalnega modela s pomočjo Mahalanobisove razdalje. Metoda daje zelo dobre rezultate za razpoznavo, slabost metode pa je pomanjkanje vpogleda v njeno delovanje, saj deluje v prostoru značilk, posledično je tudi ne moremo uporabiti za lokalizacijo anomalnih regij.

PaDiM PaDiM (Patch Distribution Modeling Framework for Anomaly Detection and Localization) [3] je po delovanju podobna prejšnji metodi, saj se ravno tako nauči porazdelitve značilk iz prednaučene nevronske mreže. Metoda poleg detekcije omogoča tudi lokalizacijo anomalij, saj se nauči normalnosti za posamezne regije, ne le za celotno sliko. V fazi razpoznave za sliko, za katero nas zanima, ali je anomalna ali ne, najprej pridobimo značilke iz prednaučene mreže, nato pa za vse regije izračunamo odstopanje od normalnosti. Za vsako regijo tako dobimo stopnjo odstopanja, kar nam omogoča lokalizacijo, maksimalna stopnja za regijo pa predstavlja tudi končno oceno anomalnosti celotne slike.

3 Eksperimenti

Temeljito analizo robustnosti metod smo izvedli na 4 podatkovnih bazah, iz katerih smo izbrali skupno 8 kategorij. Za vse izbrane kategorije smo preverili robustnost metod, če je učni množici dodanih 1%, 5%, 10% ali 20% anomalnih primerov.

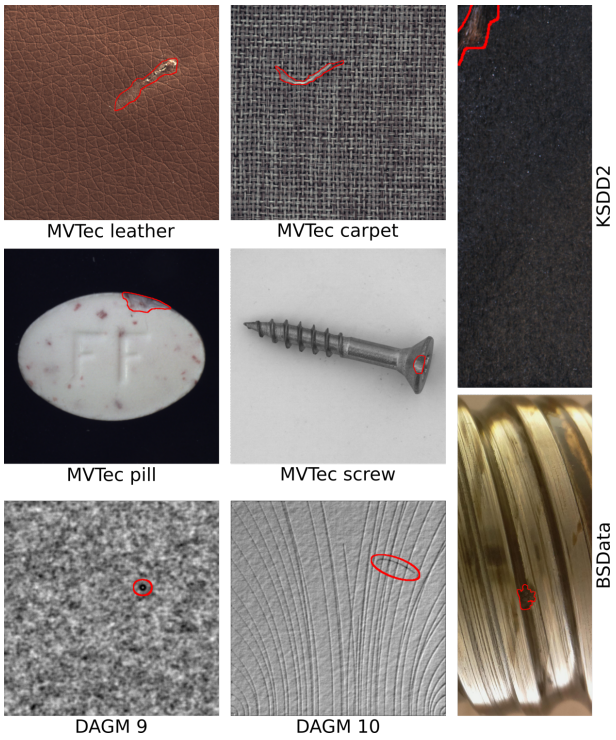
Za vsako kombinacijo metode, kategorije in deleža anomalnih primerov smo izvedli 5 ponovitev eksperimenta, s tem smo omejili vpliv naključnosti, ki lahko izvira ali iz naključnosti metode, ali iz izbire različnih primerov za učno množico. Za vsako izmed 5 ponovitev eksperimenta smo iz nabora vseh normalnih in anomalnih primerov v prvotni učni množici naključno izbrali podmnožici normalnih in anomalnih primerov ustrezne velikosti. Za podatkovno bazo MVTec smo število učnih normalnih primerov omejili na 150, saj kategorije ne vsebujejo dovolj primerov, da bi lahko zagotovili večjo heterogenost pri selekciji učnih podmnožic za ponovitve

Zbirka	Kategorija	U. N.	U. A. (1%, 5%, 10%, 20%)	T. N.	T. A.
MVTec AD	pill	150	2, 8, 15, 30	26	70
	screw	150	2, 8, 15, 30	41	59
	carpet	150	2, 8, 15, 30	28	44
	leather	150	2, 8, 15, 30	32	46
DAGM	9	250	3, 13, 25, 50	1000	150
	10	250	3, 13, 25, 50	1000	150
KolektorSDD2	/	250	3, 13, 25, 50	894	110
BSData	/	250	3, 13, 25, 50	210	75

Tabela 1: Povzetek podatkovnih zbirk, na katerih je bila izvedena analiza. Okrajšave U. N., U. A., T. N. in T. A. predstavljajo števila učnih normalnih, učnih anomalnih, testnih normalnih in testnih anomalnih slik.

eksperimenta. Dodatno smo zaradi velikega števila eksperimentov omejili število normalnih primerov v učnih množicah za preostale kategorije na 250. Kjer je bilo mogoče, smo ohranili prvotne testne množice, zato v nekaterih kategorijah vsebujejo veliko več primerov kot učne.

3.1 Podatkovne zbirke



Slika 2: Primeri slik iz podatkovnih zbirk, uporabljenih za analizo. Za vsako kategorijo je prikazana po ena anomalna slika, anomalne regije so obkrožene z rdečo barvo.

3.1.1 MVTec AD

Podatkovna zbirka MVTec AD [1] vsebuje 15 različnih kategorij, 4 kategorije predstavljajo teksture, preostalih 11 pa predmeti. Kategorije vsebujejo različna števila učnih in testnih slik, zaradi majhnega števila primerov v nekaterih kategorijah smo se omejili na 2 kategoriji tekstur in 2 kategoriji predmetov, ki vsebujejo največ slik, saj lahko tako simuliramo dodajanje anomalnih primerov v učno množico. Za izbrane kategorije smo novo testno množico sestavili iz vseh normalnih testnih primerov in iz

polovice anomalnih testnih primerov, preostali anomalni testni primeri pa so bili postopoma dodani v učno množico. Učno množico za vsako kategorijo smo sestavili iz 150 normalnih primerov in 2, 8, 15 ali 30 anomalnih primerov.

3.1.2 KolektorSDD2

Podatkovna zbirka KolektorSDD2 [2] vsebuje več kot 3000 slik površine izdelka iz industrijske aplikacije. Površina na slikah je praviloma uniforma, večina anomalnih regij pa zavzema le manjši del slike. V naših eksperimentih smo ohranili prvotno testno množico, ki vsebuje 894 normalnih in 110 anomalnih primerov, učno množico pa smo skrčili na 250 normalnih primerov in 3, 13, 25 ali 50 anomalnih primerov.

3.1.3 DAGM

Podatkovna zbirka DAGM [6] vsebuje 10 kategorij umetno generiranih površin in anomalij. Površine in anomalije se razlikujejo med kategorijami, v vseh pa se videz površin znotraj kategorije ne spreminja, anomalije pa zavzemajo le majhen del slike. Uporabili smo le zadnje dve kategoriji, 9. in 10., ohranili smo prvotni testni množici, učni množici pa smo skrčili na 250 normalnih primerov in 3, 13, 25 ali 50 anomalnih primerov.

3.1.4 BSData

Podatkovna zbirka BSData [5] vsebuje 1035 slik delov krogličnega navojnega vretena (angl. ball screw), od katerih je 394 slik anomalnih, preostale pa so normalne. V testno množico smo uvrstili 210 normalnih ter 75 anomalnih primerov, v učno pa 250 normalnih in 3, 13, 25 ali 50 anomalnih primerov.

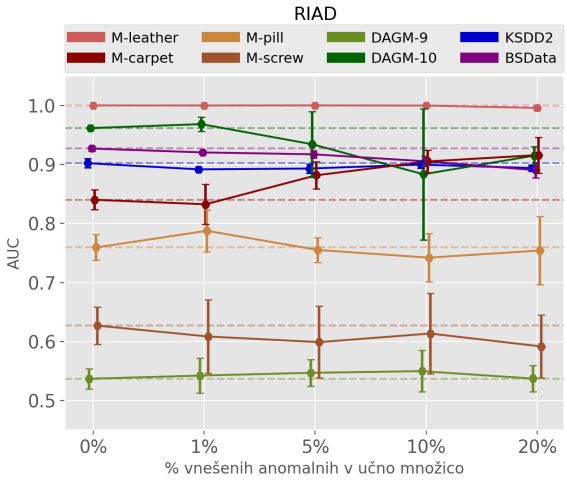
V tabeli 1 so povzeti ključni podatki o uporabljenih podatkovnih zbirkah, na sliki 2 pa lahko vidimo po eno anomalno sliko iz vsake kategorije. Opazimo lahko, da anomalne regije predstavljajo majhen del slik.

4 Rezultati

Za primerjavo robustnosti metod poročamo mero AUC, ki se za domeno detekcije anomalij najpogosteje pojavlja v literaturi. Kljub temu da se poleg mere AUC na nivoju slik včasih poroča tudi AUC na nivoju slikovnih elementov, se omejimo le na AUC na nivoju slik, saj je za domeno detekcije anomalij na industrijskih izdelkih le-ta pomembnejša. Za vse kombinacije metod in kategorij poročamo povprečje in standardni odklon 5 ponovitev eksperimenta.

RIAD Rezultati kažejo, da je metoda RIAD zelo robustna na prisotnost anomalnih primerov v učni množici. Tudi ko je v učno množico dodanih 20% anomalnih primerov, se AUC ne zniža občutneje, rezultati so skoraj povsod v intervalu napake. Povprečen AUC vseh kategorij pade iz začetnega 81.9% na 81.9%, 81.6%, 81.2% in 81.2%, ko dodamo 1%, 5%, 10% ali 20% anomalnih primerov v učno množico. Verjamemo, da gre tolikšno robustnost metode v največji meri pripisati njenemu delovanju, saj se uči rekonstrukcije posameznih regij slike, ki jih med učenjem maskiramo. Tudi v primeru anomalnih slik je zaradi narave problema večji del slike še vedno normalen, tako da se mreža tudi na anomalnih slikah v večini uči rekonstrukcije normalnih regij. Ker je mera AUC neodvisna od praga, ki loči normalne in anomalne primere, je končni rezultat odvisen le od razvrstitve testnih primerov, tako da izboljšana rekonstrukcija anomalnih regij zaradi dodajanja anomalnih primerov ne vpliva ključno na rezultat. Opazimo lahko tudi, da so intervali napake pri rezultatih metode večji kot pri ostalih dveh metodah, kar gre pripisati dodatnemu viru naključnosti, ki izhaja iz naključne inicializacije nevronske mreže in naključnega vrstnega reda slik med učenjem za vsako ponovitev eksperimenta. Ravno ta naključnost je najbrž vzrok za rahlo izboljšanje rezultata na podatkovni zbirki M-Carpet ob povišanju števila anomalnih slik, ki imajo lahko tudi regularizacijski efekt.

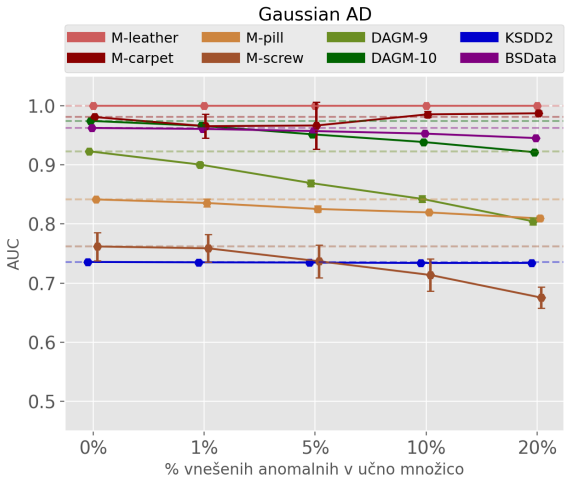
Na sliki 3 so prikazani rezultati eksperimentov za metodo RIAD. Prikazana so povprečja in standardni odkloni 5 ponovitev eksperimenta, zaradi boljše preglednosti so točke rahlo zamaknjene levo in desno. Skrajno leve točke predstavljajo AUC, ko učna množica ne vsebuje anomalnih primerov, nadaljnje pa AUC ob dodajanju naraščajočega števila anomalnih primerov v učno množico.



Slika 3: Rezultati eksperimentov za metodo RIAD. Prikazani so AUC za učno množico brez anomalnih primerov in za učne množice z naraščajočim deležem anomalnih primerov. Y os prikazuje AUC, X pa delež anomalnih primerov, ki je bil dodan glede na prvotno velikost učne množice. Črtkana črta za vsako kategorijo prikazuje AUC za učno množico brez anomalnih primerov.

Gaussian AD Metoda Gaussian AD kaže večjo občutljivost na prisotnost anomalnih primerov v učni množici. Pri skoraj vseh kategorijah opazimo jasen trend slabšanja rezultatov, ko se delež anomalnih primerov v učni množici viša, izjemi sta kategoriji *MVTec leather*, kjer metoda konstanto proizvede popoln rezultat, in *MVTec carpet*, kjer se rezultat malenkostno izboljša. Povprečen AUC vseh kategorij pade iz začetnega 89.7% na 89.0%, 88.0%, 87.3% in 86.0%, ko dodamo 1%, 5%, 10% ali 20% anomalnih primerov v učno množico. Opazimo lahko tudi, da so rezultati za iste kombinacije kategorij in deleža anomalnih primerov razmeroma podobni oz. da so intervali napake majhni, večje intervale napake opazimo predvsem pri kategoriji *MVTec screw*, ki vsebuje slike vijakov, rotirane za različne kote, ki so si vizualno zelo

različni, zato verjetno izbor slik igra večjo vlogo. Opazimo tudi, da so pri tej metodi intervali napake manjši, saj ne učimo nevronske mreže, temveč le ocenimo parametre za značilke, ki jih proizvede prednaučena nevronska mreža. Dodatno k zmanjšanju naključnosti pripomore tudi ocenjevanje anomalnosti na nivoju slik, kar pomeni, da se značilke iz prednaučene mreže za anomalne primere ne razlikujejo toliko od značilk normalnih primerov, saj je tudi na anomalnih slikah večji del slike še vedno normalen. Zaradi tega so parametri ocenjene distribucije manj občutljivi na posamezne anomalne slike in se ocenjena distribucija značilk ne premakne bistveneje. Iz istega razloga pa verjetno dobimo tudi izrazitejše poslabšanje ob vnosu večjega števila anomalnih primerov v učno množico, saj se distribucija značilk učne množice premakne bliže distribuciji značilk anomalne množice, posledično pa je težje določiti, kateri distribuciji pripada testni primer.



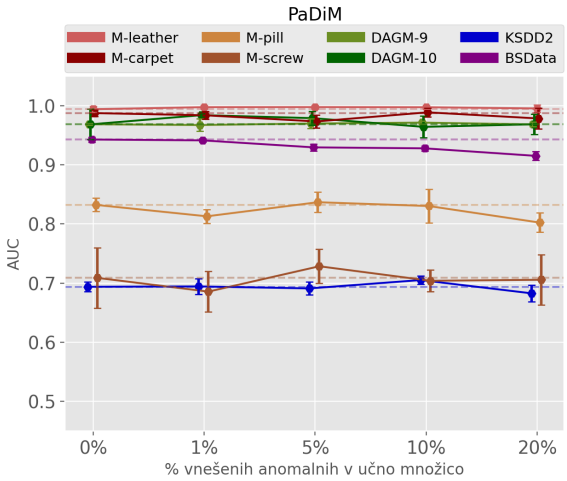
Slika 4: Rezultati eksperimentov za metodo Gaussian AD. Pomen elementov je razložen pod sliko 3.

PaDiM Tudi metoda PaDiM kaže visoko robustnost na prisotnost anomalnih primerov v učni množici. Povprečen AUC vseh kategorij se spremeni iz začetnega 88.7% na 88.3%, 88.8%, 88.6% in 87.7%, ko dodamo 1%, 5%, 10% ali 20% anomalnih primerov v učno množico. Podobno kot pri Gaussian AD lahko opazimo, da so rezultati metode dokaj ponovljivi, saj so z izjemo kategorije *MVTec screw* intervali napake zelo ozki. Delovanje metode PaDiM ima lastnosti tako metode RIAD kot Gaussian AD, saj podobno kot prva deluje na posameznih regijah slike, podobno kot druga pa za modeliranje posameznih regij uporablja značilke iz prednaučene mreže, za katere nato oceni parametre multivariantne Gaussove porazdelitve. Tako metoda prevzame dobre lastnosti obeh metod, robustnost prve in ponovljivost druge.

5 Zaključek

Analizirali smo, kako prisotnost anomalnih primerov v učni množici vpliva na delovanje treh nedavno predlaganih generativnih metod za detekcijo anomalij. Z obsežno eksperimentalno evaluacijo na 8 kategorijah iz 4 podatkovnih bazah smo pokazali, da so vse 3 metode robustne na prisotnost majhnega števila anomalnih primerov v učni množici, od teh pa sta 2 robustni tudi na večja števila anomalnih primerov. Zahteva, da učna množica pri učenju generativnih metod ne sme vsebovati anomalnih primerov, je torej lahko prekršena, dokler je število anomalnih primerov relativno majhno glede na število normalnih primerov, do katere mere pa je lahko prekršena, pa je odvisno tudi od posamezne metode in od zahtevnosti problema.

Zahvala: To delo je podprla Javna agencija za raziskovalno dejavnost Republike Slovenije (ARRS), projekt J2-9433 in program P2-0214.



Slika 5: Rezultati eksperimentov za metodo PaDiM. Pomen elementov je razložen pod sliko 3.

Literatura

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and C. Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9584–9592, 2019.

[2] Jakob Božič, Domen Tabernik, and Danijel Skočaj. Mixed supervision for surface-defect detection: from weakly to fully supervised learning. *Comput. Ind.*, 129:103459, 2021.

[3] Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: a patch distribution modeling framework for anomaly detection and localization. In *ICPR Workshops*, 2020.

[4] Oliver Rippel, Patrick Mertens, and D. Merhof. Modeling the distribution of normal data in pre-trained deep features for anomaly detection. *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 6726–6733, 2021.

[5] Tobias Schlagenhauf, Magnus Landwehr, and Juergen Fleischer. Industrial machine tool component surface defect dataset, 2021.

[6] Daniel Weimer, Bernd Scholz-Reiter, and Moshe Shpitalni. Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection. *CIRP Annals - Manufacturing Technology*, 65, 05 2016.

[7] Vitjan Zavrtanik, M. Kristan, and D. Skočaj. Reconstruction by inpainting for visual anomaly detection. *Pattern Recognit.*, 112:107706, 2021.