



UKAEA-RACE-PR(21)04

R Skilton, Y Gao

Combining object detection with generative adversarial networks for in-component anomaly detection

Enquiries about copyright and reproduction should in the first instance be addressed to the UKAEA Publications Officer, Culham Science Centre, Building K1/O/83 Abingdon, Oxfordshire, OX14 3DB, UK. The United Kingdom Atomic Energy Authority is the copyright holder.

The contents of this document and all other UKAEA Preprints, Reports and Conference Papers are available to view online free at scientific-publications.ukaea.uk/

Combining object detection with generative adversarial networks for in-component anomaly detection

R Skilton, Y Gao

Combining Object Detection with Generative Adversarial Networks for in-component Anomaly Detection

Robert Skilton¹ and Yang Gao²

Abstract—Present inspection techniques in place at the Joint European Torus (JET), as well as some of those planned for ITER make use of robotically deployed inspection systems, which typically collect data for offline analysis. This can be a slow, laborious process with subjective or error-prone results. There are significant benefits to be gained through automation or user assistance, for example through prioritisation of samples for analysis.

Automated visual anomaly detection is a highly challenging problem due to high dimensionality of the input data, meaning that the normal statistical distribution cannot be directly modeled. We provide a robotic and algorithmic framework that utilizes Generative Adversarial Networks (GANs) to indirectly model this distribution, and hence provide a mechanism to quantify the anomalousness of given image data samples from a tokamak environment.

This paper presents an approach to visual anomaly detection that combines multiple deep neural network architectures in order to extract individual components and then classify anomalies. An overview of the architecture and algorithms employed as well as quantitative and qualitative assessments of the performance against data from both a benchmark dataset, and real data gathered from JET components is provided.

I. INTRODUCTION

Visual anomaly detection is a task which is of significant relevance to quality control processes and inspection tasks in various domains. Anomaly detection can be described as the identification of the presence of out-of-the-ordinary content in a given data sample. In visual anomaly detection, data samples are typically images, and anomalies typically represent features present in the scene which are not normally expected or occur with relatively low frequency.

In industrial applications, anomaly detection is important for detecting defects in products, machines, and infrastructure such as cracks in concrete, delamination in steels, or corrosion on metal surfaces.

A common application of computer vision techniques in industrial inspection is in quality control through the automated detection of surface defects. Surface defect detection is, however, not only invaluable for manufacturing quality control, but also in inspection of equipment in operation. Surface defect detection has been applied to wood surfaces [1], metallic surfaces [2], [3]. Common techniques include frequency analysis methods, Gabor filters [4], and more recently, methods based on deep learning techniques [5].

¹Robert Skilton is with Remote Applications in Challenging Environments, UK Atomic Energy Authority, UK, as well as the STAR Lab, Surrey Space Centre, University of Surrey, UK robert.skilton@ukaea.uk

²Yang Gao is head of the STAR Lab, Surrey Space Centre, University of Surrey, UK yang.gao@surrey.ac.uk

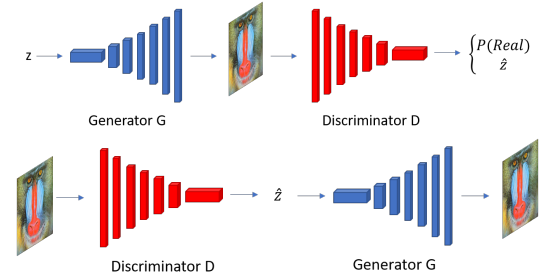


Fig. 1. The GAN is first trained in a standard configuration (top), to simultaneously discriminate real from generated samples whilst predicting z through a latent regressor. At runtime, the network is inverted (bottom), the discriminator becoming an encoder, predicting the latent representation which is then used to regenerate the sample image based on the learned generative model.

A large volume of research has been directed towards techniques for detection of cracks in concrete [6], [7], steel, and other structures [8]. This is of obvious importance for a wide range of civil infrastructure inspection tasks, including bridges, tunnels [9], [10], and pipelines [11].

One machine inspection task that is of particular importance to regulatory aspects of machine build and operation is that of weld inspection. Pressure vessels and pipelines are usually required to conform to codes and standards such as ASME B31 [12], which involve inspection prior to first use as well as routine inspection during the lifetime of the equipment. Elements of the problem include detection of the weld bead [13], and detection of weld seam profiles [14], as well as the subsequent characterisation of weld quality. Sometimes these tasks are required to be carried out in confined spaces with limited scope for access. As such there is an active interest in development of robotic systems that can deploy sensors into challenging environments for weld inspection [15], [16] and development of crawlers for in-pipe inspection [17]. Other work has looked at inspection of various types of machines and components such as storage tanks [18], solar power plant equipment [19], railway equipment [20], and ship structures [21]. In the world of nuclear fusion inspection, although there have been several experimental studies in automated inspection, e.g. [22], the large majority of in-vessel inspections are conducted manually which can be extremely slow, laborious, and error-prone.

Numerous examples can be found in the literature of industrial inspection and visual anomaly detection for identification of specific defects, however the identification of unspecified, general defects is somewhat more challenging. Anomaly detection in other domains often relies on modeling

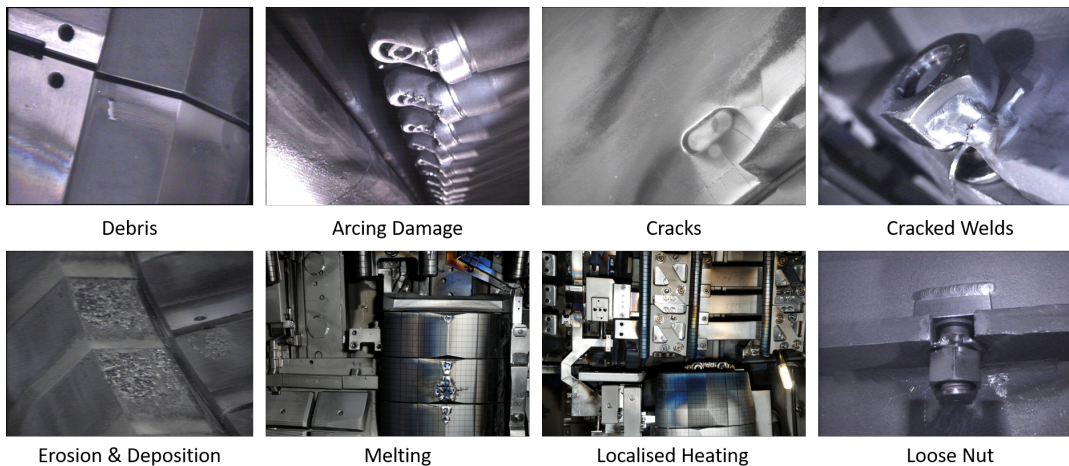


Fig. 2. Common types of damage that occur in tokamak (Images taken from the Joint European Torus, JET).

the probability distribution of normal data samples, and detecting where query samples fall outside this normal distribution. With visual data, directly modeling this distribution is computationally intractable, however deep learning techniques such as GANs [23] may allow us to indirectly model this distribution by providing a mapping from a simple, uniform random distribution, into the complex structured domain of real-world images [24].

This work looks at applying GAN-based techniques to general visual anomaly detection, and enhancing the techniques through a novel workflow. This is illustrated with both a reproducible dataset and an example industrial datasets.

II. GAN-BASED ANOMALY DETECTION

GANs [23] are a recent approach to training generative models based on an unsupervised, adversarial approach. They have been demonstrated to have capability in generating complex natural images [25].

GANs consist of two main elements, the Generator G , and a Discriminator D . The role of the discriminator is to estimate the probability that a given data sample (e.g. an image) is a natural image as opposed to an artificial generated image. The role of the generator is to attempt to randomly generate realistic data samples that are able to fool the discriminator. As such, the Generator and the Discriminator are playing adversarial roles in a 2-player game, which can be described in the minimax function:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}} [\log(D(x))] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \quad (1)$$

The Generator G operates on a simple random prior, $z \sim p_z$, and implicitly defines a probability distribution p_g of generated samples $G(z)$. In order for the generated samples to match the real-world data, it is desirable for p_g to converge to p_{data} , the distribution of natural training samples.

With $D(x)$, the output of the Discriminator being defined as the probability that sample x came from the data (as

opposed to being generated), the optimisation problem can be defined as backpropagation by ascending the gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log(D(x^{(i)})) + \log(1 - D(G(z^{(i)})))] \quad (2)$$

and descending the gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^{(i)}))) \quad (3)$$

In other words, maximising the log probability that $x^{(i)}$ is marked as coming from the training data plus the log probability that $z^{(i)}$ is flagged as being generated by adjusting the Discriminator parameters, whilst simultaneously minimising the log probability that $z^{(i)}$ is marked as being artificially generated by adjusting the Generator parameters.

GAN-based anomaly detection can be considered as directly equivalent to standard probability distribution-based anomaly detection methods, where low-probability samples represent likely anomalies. With natural images, the dimensionality of the data is so high that directly modeling the distribution is computationally intractable, however GANs provide a means for indirectly modeling the complex structured distribution as a highly nonlinear function, which maps from a smaller, simple, latent distribution, into the target domain.

GANs have been used in [24] for anomaly detection in medical imaging data. A GAN was first trained on healthy samples, and then used to predict anomalies based on methods in [26] for finding the closest Generated data to the real, query data.

Anomalies are then detected by adapting the coefficients of the latent distribution from which images are generated (z) by backpropagation [26], and an anomaly score $A(x)$ is produced, which can be used for detection of anomalous regions within an image. The final residual image can be used to identify anomalous regions. The backpropagation process, however requires many iterations and is extremely slow.

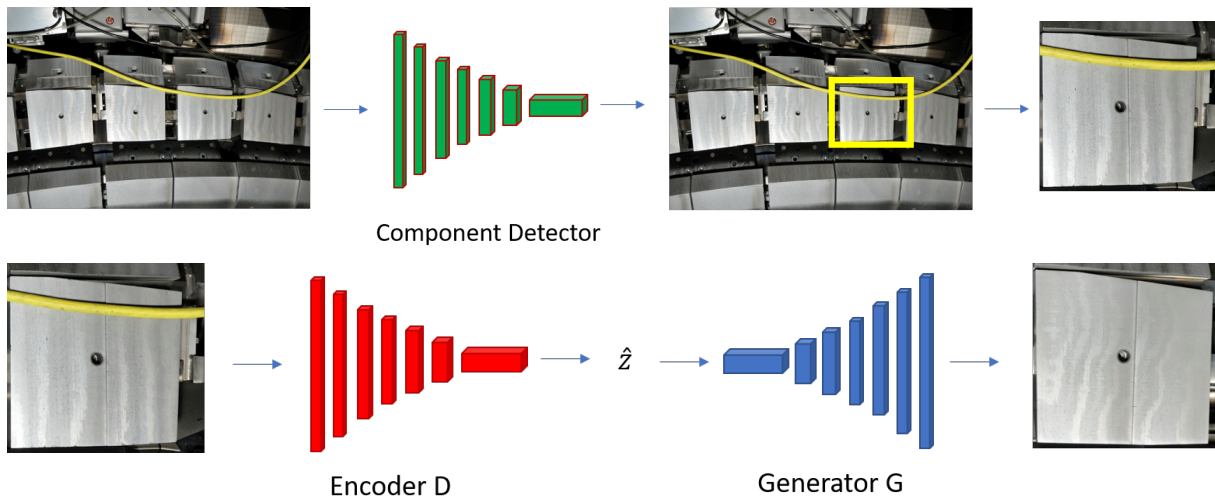


Fig. 3. Architecture of the combined object detector and anomaly detector. Firstly, individual components are identified by the object detector (top). These object detections are then used to make a new set of sub-images, each of which corresponds to an individual component. The component sub-images are then fed into the discriminator network, which operated in Encoder mode in order to estimate the latent space representation. The generator is subsequently used to regenerate the component image based on the learned semantic model of typical component images, in order that anomalies can be identified by comparison between the query (component) sub-image, and its regeneration. Regenerated images are further improved by backpropagation.

Our work applies GAN models, with the capability of predicting latent representations for the task of image regeneration, giving a much closer starting point for backpropagation than a random initial starting point. Regenerations are compared with original images in a number of different ways, assessing the resulting anomaly detection performance. Methods of comparison include a residual score, summing absolute differences in pixel values between the two images, Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) score, and the AnoGAN metric.

Previous attempts [27] at using GANs for industrial anomaly detection have tried using randomly cropped and centered images to avoid the computational or operational complexity associated with finding components of interest and then automatically performing these operations. However, historic work involving GANs, for example in generation of new realistic images have focused on datasets containing relatively well-structured images, centered about the object of interest, and cropped to the region containing that object in a consistent manner. Typical datasets such as MNIST [28] and CelebA Celebrity Faces [29] comprise images which are centered and cropped about the object or type of object in question.

In order to simplify the visual problem, this work similarly operates on data which is centered and cropped about individual objects, in our case, individual components from within the JET tokamak. To facilitate the use of this in real-world scenarios, a method is presented for automatically centering and cropping query images such that they are suitable for this process.

Additionally, previous work in GAN-based anomaly detection has made use of either backpropagation to find a close approximation to z , or alternatively used some kind of encoder or latent regressor. Our work combines the two

methods in order to seed the optimisation process with a very good starting point.

In summary, the novel contribution of this work are twofold:

- Combining object detection with GAN-based anomaly detection for effective in-component anomaly detection.
- Combining latent space prediction using an encoder or latent regressor with backpropagation for refinement.

III. OBJECT DETECTION OF IN-VESSEL COMPONENT IN FUSION FACILITIES

In order to identify sub-images relating to individual components, an object detector network is employed as the first stage of the in-component anomaly detection pipeline. In this work we used the YOLO [30], [31] network as the object detector, which is used to create multiple proposed object bounding boxes within each given query image relating to individual components. Our YOLO network was trained on an annotated dataset taken from a collection of high resolution full vessel survey images, taken in 2014.

The training data consists of an augmented set of 969 images, containing 7372 annotated components in 25 classes. The classes were selected from some of the more common component types that are visible in the JET vessel including 7 types of divertor tile, 9 other types of tile, various antennas, and other components.

Training was conducted over 22 epochs of the dataset. Average time for regression of bounding boxes at run time on CPU was 0.23 seconds per image, or 4.44 images per second.

IV. COMBINING OBJECT DETECTION WITH ANOMALY DETECTION

The automated centering and cropping process makes use of the object detection pipeline described above in

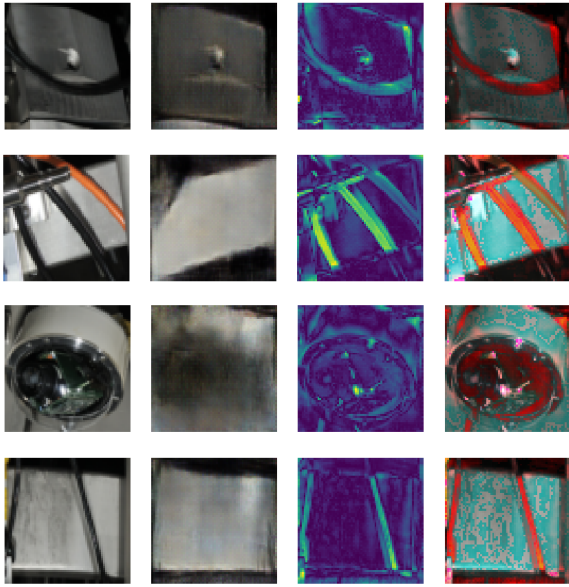


Fig. 4. Anomaly detection example results for images containing anomalies. From left-to-right, the columns show: 1. the original query image, 2. The GAN regeneration, 3. The resulting residual image, 4. A Thresholded residual image overlaid over the original query image in red. The images clearly show successful detection of the anomalies at pixel-level. Note, in row 3, the component that has been detected (a camera) is itself an anomaly as it is not part of the in-vessel component set. It is successfully detected as such by the GAN system.

order to identify components of interest in new images, and provide bounding boxes for each detected component in the image. The object detector is used to propose multiple component bounding boxes within each given query image. The bounding boxes are then used to create multiple sub-images which each correspond to a detected object, in our case, tokamak components. Each sub-image is created by taking the x , y coordinates of the bounding box, sampling a new image centered about these coordinates, with width and height as detected, and then reshaping this sub-image to a 128x128 square image compatible with the GAN regenerator. This entire pipeline is illustrated in Figure 3.

The GAN is used firstly in latent encoder-regenerator mode, and then regenerated images are subsequently refined using 1000 iterations of backpropagation. The entire process for optimising the estimate of the latent space representation is provided in Algorithm 1.

Once the query image has been regenerated, three methods are used to compare the original component sub-image with the regenerated version. These include taking the direct residual between the two images, taking the Peak Signal-to-Noise Ratio (PSNR), and using the Structural Similarity index (SSIM). An optimal threshold is then applied to this anomalously score, based on the highest validation set f_1 score, which is used to classify components into anomalous or not-anomalous categories. The scores can similarly be used to rank or prioritise components.

Algorithm 1 Latent space optimisation using contextual and perceptual losses and backpropagation, give a trained generator, G , and query image x .

Estimate initial latent space representation $\hat{z}^{(0)}$ using Encoder $\hat{z}^{(0)} = E(x) p_g(\hat{z})$.

for $i = 1 \dots k$ steps **do**

 Calculate loss for current estimate:

$$\mathcal{L}_c^i(z|x, M) + \mathcal{L}_p(z)$$

 Update z estimate by backpropagating loss:

$$\hat{z}^{i+1} = \hat{z}^i + \frac{\partial \mathcal{L}_c^i(z|x, M) + \mathcal{L}_p(z)}{\partial z}$$

end for

Transform the optimised latent representation into the data space:

$$\hat{x}^{(*)} = G(\hat{z}^{(k)})$$

The gradient-based updates can use any standard gradient-based learning rule.

V. PERFORMANCE EVALUATION AND RESULTS

The component-centric GAN-based anomaly detector was trained on a dataset which simulates outputs of the component detector subsystem. Bounding boxes from the object detection training set were extracted from the training set annotations, and used to automatically generate sub-images corresponding to components within the system. This provided a simple and effective method for generating training data in a fully automated way. Training data consisted entirely of non-anomalous samples (in this case, samples that did not include any maintenance equipment or other non-native objects).

Once the GAN was trained, 64 non-anomalous test samples and 28 anomalous test samples were regenerated using the above described process, and compared with the original query test samples using each of the three metrics in order to produce anomalously scores for each component image.

Anomaly detection performance is evaluated using standard classification metrics, including the classification accuracy, precision and recall, and associated F1 score, specificity, and areas under the ROC curve and Precision-Recall curves. The scores for each of the three comparison methods is provided in Table I.

If accuracy scores for the residual, PSNR, and SSIM metrics are compared with those of the previous work, where randomly cropped and centered images have been used, the accuracy improvement through use of object detection is significant, indicating that the centering and cropping of query images about detected components plays a significant role in the GANs ability to distinguish anomalous components from non-anomalous components.

One interesting observation from the results is that the highest scoring metric has actually changed from residual score to SSIM, between previous results and the present work. This could perhaps be explained by the inability of

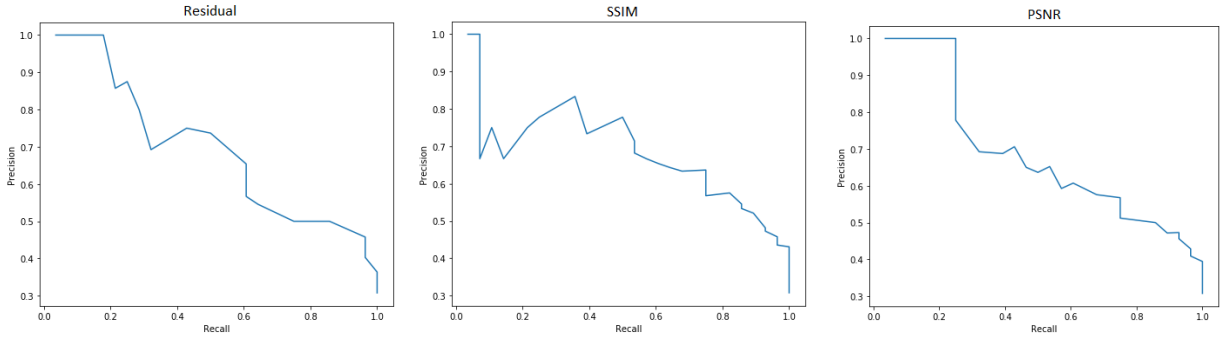


Fig. 5. Precision-Recall curves for each of the three methods.

TABLE I
RESULTS FOR ANOMALY DETECTION ON COMPONENTS AT THRESHOLD VALUE WITH HIGHEST F_1 SCORE.

Comparison Method	Accuracy	Accuracy (previous work)	Precision	Recall	F_1 Score	Specificity	ROC AUC	PR AUC
Residual	79%	63%	0.50	0.86	0.63	0.63	0.81	0.66
PSNR	77%	53%	0.57	0.75	0.65	0.75	0.81	0.66
SSIM	80%	52%	0.64	0.75	0.69	0.81	0.84	0.65

the Generator to find a suitable representation resulting in a bias towards the numerical optimisation process in the regeneration stage. This would lead to regenerations being closer in terms of pixel-wise distance to the query image than in terms of semantic distance. As this new method is able to better leverage the GANs generative capability, resulting regenerations are more structurally representative, and therefore the SSIM metric is more appropriate.

A. Example failure cases

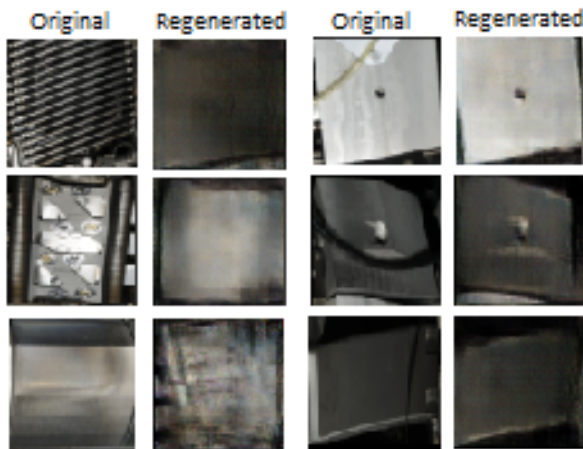


Fig. 6. Some example failure cases. The left-hand two columns show original images that contain no anomalies, and their corresponding regenerations, where the images have been misclassified (i.e. are false positives). The right-hand two columns show original images that contain anomalies, and their corresponding regenerations, where the images have been misclassified (i.e. are false negatives).

A small selection of noteworthy failure cases including both false-positives, and false-negatives can be seen in Figure 6. In the case of false positives, or incorrectly classified non-

anomalous images, the examples can be seen to indicate that the high anomalousness score results from a failure to regenerate the original image. This may be as a result of the optimisation algorithm reaching a local minimum, or a failure of the Generator to learn generalisable mappings (e.g. through mode collapse, or biased training data).

In the case of the three examples of false negatives, or incorrectly classified images containing anomalies, it can clearly be seen that the regeneration process has worked well. As expected, the generator has successfully regenerated the component image elements, without the anomalous items in the foreground. One clear limitation of this work is that the anomalousness scores for images are all based on the sums of pixel-wise scores. This means that the size of the anomaly within the image will directly impact the likelihood of detection. Further work is required in order to determine better methods of decoupling the anomalousness scoring of the image from the pixel-size of the anomaly.

B. Pixel-wise anomaly segmentation

Although not quantitatively assessed here, the various anomalousness scores can be used to classify individual pixels within the image using the same method as for entire sub-images. This results in a pixel-wise delineation of the anomalous element. Figure 4 shows some examples of this being applied with the residual score.

VI. CONCLUSIONS

It has been shown, not only that GANs are a powerful tool for visual inspection and anomaly detection in machine components, but also that this can be greatly enhanced by using component-centric, cropped data to reinforce the statistical strengths of the GAN. This can be achieved in runtime practice through the combination of an object detector

network with the regenerative GAN in order to produce a real-time automated process for doing so.

It is worth noting that whilst there is still room for improvement and optimisation of the presented techniques, this work can already add value in terms of automatically scoring and prioritising test cases for human examination. In this case, we are not so concerned about false-positives, and so the F_1 score is suitable. When considering a more fully automated anomaly detection and alert system, there is clearly a trade-off to be considered between false positives and false negatives. In this case, different threshold values (e.g. highest $F_{0.5}$ score) may be more appropriate. This is clearly a topic for further studies involving human experts.

The GAN-based method employed in this work made use of a single network trained across all sub-images. A possible future direction of this work might be to investigate the performance differences between per-class regenerators and across-all-class regenerators.

ACKNOWLEDGMENTS

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

- [1] Gonzalo A Ruz, Pablo A Estevez, and Claudio A Perez. A neurofuzzy color image segmentation method for wood surface defect detection. *Forest products journal*, 55(4):52–58, 2005.
- [2] Hong Zheng, Ling Xue Kong, and Saied Nahavandi. Automatic inspection of metallic surface defects using genetic algorithms. *Journal of materials processing technology*, 125:427–433, 2002.
- [3] Timo Piironen, Olli Silven, Matti Pietikäinen, Toni Laitinen, and Esko Strömmer. Automated visual inspection of rolled metal surfaces. *Machine Vision and Applications*, 3(4):247–254, 1990.
- [4] D-M Tsa and S-K Wu. Automated surface inspection using gabor filters. *The International Journal of Advanced Manufacturing Technology*, 16(7):474–482, 2000.
- [5] Ruoxu Ren, Terence Hung, and Kay Chen Tan. A generic deep-learning-based approach for automated surface inspection. *IEEE transactions on cybernetics*, 48(3):929–940, 2018.
- [6] Tomoyuki Yamaguchi and Shuji Hashimoto. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Machine Vision and Applications*, 21(5):797–809, 2010.
- [7] Zhenhua Zhu, Stephanie German, and Ioannis Brilakis. Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation. *Automation in Construction*, 20(7):874–883, 2011.
- [8] Eduardo Zalama, Jaime Gómez-García-Bermejo, Roberto Medina, and José Llamas. Road crack detection using visual features extracted by gabor filters. *Computer-Aided Civil and Infrastructure Engineering*, 29(5):342–358, 2014.
- [9] Elisabeth Menendez, Juan G Victores, Roberto Montero, Santiago Martínez, and Carlos Balaguer. Tunnel structural inspection and assessment using an autonomous robotic system. *Automation in Construction*, 87:117–126, 2018.
- [10] Seung-Nam Yu, Jae-Ho Jang, and Chang-Soo Han. Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel. *Automation in Construction*, 16(3):255–261, 2007.
- [11] Sunil K Sinha and Paul W Fieguth. Automated detection of cracks in buried concrete pipe images. *Automation in Construction*, 15(1):58–72, 2006.
- [12] ASME. Process piping b31.3. 2016.
- [13] Yuan Li, You Fu Li, Qing Lin Wang, De Xu, and Min Tan. Measurement and defect detection of the weld bead based on online vision inspection. *IEEE Transactions on Instrumentation and Measurement*, 59(7):1841–1849, 2010.
- [14] Yinshui He, Yuxi Chen, Yanling Xu, Yiming Huang, and Shanben Chen. Autonomous detection of weld seam profiles via a model of saliency-based visual attention for robotic arc welding. *Journal of Intelligent & Robotic Systems*, 81(3-4):395–406, 2016.
- [15] Jianzhong Shang, Bryan Bridge, Tariq Sattar, Shyamal Mondal, and Alina Brenner. Development of a climbing robot for inspection of long weld lines. *Industrial Robot: An International Journal*, 35(3):217–223, 2008.
- [16] Carlos Balaguer, Antonio Giménez, José Manuel Pastor, VM Padron, and Mohamed Abderrahim. A climbing autonomous robot for inspection applications in 3d complex environments. *Robotica*, 18(3):287–297, 2000.
- [17] Fathi Hassan Ghorbel, James Bruster Dabney, J Ryan Steger, Cliff Avery Thomas, Demetri Polychronis Spanos, Nathan Christopher Lowry, and Betty Wingyi Seto. Autonomous robotic crawler for in-pipe inspection, May 1 2007. US Patent 7,210,364.
- [18] Nicole Barry, Erin Fisher, and Joshua Vaughan. Modeling and control of a cable-suspended robot for inspection of vertical structures. In *Journal of Physics: Conference Series*, volume 744, page 012071. IOP Publishing, 2016.
- [19] Mayorkinos Papaalias, Liang Cheng, Maria Kogia, Abbas Mohimi, Vassilios Kappatos, Cem Selcuk, Louis Constantinou, Carlos Quiterio Gómez Muñoz, Fausto Pedro García Marquez, and Tat-Hean Gan. Inspection and structural health monitoring techniques for concentrated solar power plants. *Renewable Energy*, 85:1178–1191, 2016.
- [20] Xavier Gibert, Vishal M Patel, and Rama Chellappa. Robust fastener detection for autonomous visual railway track inspection. In *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*, pages 694–701. IEEE, 2015.
- [21] Emilio Garcia-Fidalgo, Alberto Ortiz, Francisco Bonnin-Pascual, and Joan P Company. A mosaicing approach for vessel visual inspection using a micro-aerial vehicle. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 104–110. IEEE, 2015.
- [22] Emil T Jonasson, Jonathan Boeuf, Stephen Kyberd, Robert Skilton, Guy Burroughes, Paul Amayo, and Steve Collins. Reconstructing jet using lidar-vision fusion. *Fusion Engineering and Design*, 146:110952, 2019.
- [23] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- [24] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging*, pages 146–157. Springer, 2017.
- [25] Emily L Denton, Soumith Chintala, Rob Fergus, et al. Deep generative image models using a laplacian pyramid adversarial networks. In *Advances in neural information processing systems*, pages 1486–1494, 2015.
- [26] Raymond A Yeh, Chen Chen, Teck-Yian Lim, Alexander G Schwing, Mark Hasegawa-Johnson, and Minh N Do. Semantic image inpainting with deep generative models. In *CVPR*, volume 2, page 4, 2017.
- [27] Robert Skilton and Yang Gao. Visual detection of generic defects in industrial components using generative adversarial networks. In *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 489–494. IEEE, 2019.
- [28] Yann LeCun, Corinna Cortes, and CJ Burges. Mnist handwritten digit database. *AT&T Labs [Online]*. Available: <http://yann.lecun.com/exdb/mnist>, 2:18, 2010.
- [29] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Large-scale celebfaces attributes (celeba) dataset. *Retrieved August*, 15:2018, 2018.
- [30] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640, 2015.
- [31] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *CoRR*, abs/1804.02767, 2018.