

# Experiments on Anomaly Detection in Autonomous Driving by Forward-Backward Style Transfers

Daniel Bogdoll<sup>†‡\*</sup>, Meng Zhang<sup>†\*</sup>, Maximilian Nitsche<sup>‡</sup>, J. Marius Zöllner<sup>†‡</sup>  
<sup>†</sup>FZI Research Center for Information Technology, Germany. Email: bogdoll@fzi.de  
<sup>‡</sup>Karlsruhe Institute of Technology, Germany.

**Abstract**—Great progress has been achieved in the community of autonomous driving in the past few years. As a safety-critical problem, however, anomaly detection is a huge hurdle towards a large-scale deployment of autonomous vehicles in the real world. While many approaches, such as uncertainty estimation or segmentation-based image resynthesis, are extremely promising, there is more to be explored. Especially inspired by works on anomaly detection based on image resynthesis, we propose a novel approach for anomaly detection through style transfer. We leverage generative models to map an image from its original style domain of road traffic to an arbitrary one and back to generate pixelwise anomaly scores. However, our experiments have proven our hypothesis wrong, and we were unable to produce significant results. Nevertheless, we want to share our findings, so that others can learn from our experiments.

**Index Terms**—autonomous driving, anomaly detection, style transfer, generative models

## I. INTRODUCTION

Driven by the technology of deep neural networks for computer vision, tremendous advances have been seen in the field of autonomous driving. For instance, various deep learning models [1] are nowadays able to provide outstanding performances for the task of object detection. However, in order to achieve effective deployment for real-world scenarios, reliability issues must be overcome. That is to say, under rare or unknown conditions, an autonomous vehicle is required not only to be able to identify the object classes from the training dataset, but also to detect atypical objects that have not been included in the training set. Anomaly detection, therefore, is an active topic in the research field of autonomous driving. As shown in [2], there are five main strategies to detect anomalies: Confidence scores, reconstructions, generative approaches, feature extraction methods and prediction errors. The approach presented here falls into the category of reconstructive methods, in which an attempt is made to produce a reconstruction of an original image via a certain process, with which, differences from the original image can be classified as anomalies.

Inspired by approaches in this category, we came up with the idea to use generative models for style transfer based reconstructions, as shown in Figure 1. We have formulated our

This work results from the KIGLIS project supported by the German Federal Ministry of Education and Research (BMBF), grant number 16KIS1231.

\*These authors contributed equally

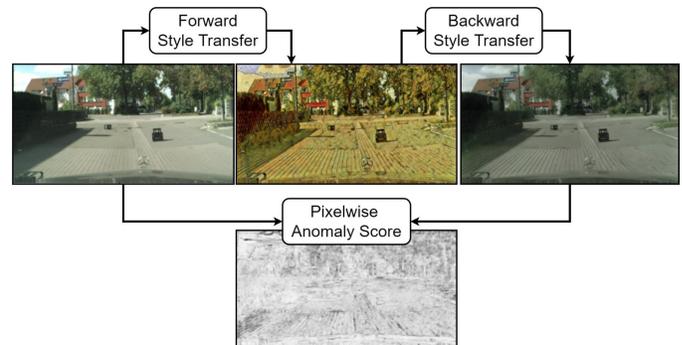


Fig. 1. Overview of the novel anomaly detection approach. An image with anomalies is transferred into a stylized version of it with a forward style transfer, based on a target domain of choice. The resulting image is then transformed back to its original domain. Finally, a difference map between the original and the re-created image is computed, which can be interpreted as a map of pixel-wise anomaly scores.

hypothesis as follows: Transferring an image from the road traffic domain, which includes anomalies, such as unknown objects, to another domain and back to the original domain, will lead to a resulting image that is similar to the original in the domain-typical aspects and deviates from it in the anomalous regions, which allows for the detection of anomalies. This hypothesis was made under the assumption that anomalies might disappear or get distorted through two successive style changes, which would allow for the detection of such. This hypothesis to detect anomalies is based on an assumption with respect to reconstruction errors with generative models, where “...normal samples are located on a manifold and all anomalous samples are located outside.” [3].

For the experiments, we used the KITTI [4] and Cityscapes [5] datasets, and qualitatively evaluated the detection of anomalies with the Fishyscapes [6] and Lost and Found [7] datasets. However, our experiments have proven our hypothesis wrong and have shown that the concept as we have implemented it does not seem suitable for anomaly detection. To share this finding with the scientific community, we present our hypothesis, experiments, and negative results, as is often suggested [8]–[13], but rarely practiced.

arXiv:2207.06055v2 [cs.CV] 1 Sep 2022

## II. RELATED WORK

Most of the current research in anomaly detection for autonomous driving focuses on camera-based approaches [2], which we will present in more detail in the following section. Subsequently, we will give an overview of style transfer methods.

### A. Camera-based Anomaly Detection

Modern anomaly detection methods that work on camera data can be split into generative, reconstructive, and confidence-based approaches [2]. The latter detect anomalies by thresholding a derived uncertainty measure of the model [14], [15]. Moreover, Du et al. [16] focus on the calibration of the neural network’s confidence score to differentiate in- and outliers more thoroughly. For this purpose, they shape the decision boundary via virtual outliers sampled from a learned feature distribution [16]. Reconstructive approaches on the other hand try to resynthesize the normality of a driving scene. Here, normality can be understood as what a model learned, given certain training data. Anomalous objects or events are then detected by comparing the real and reconstructed scene, resulting in a reconstruction error. However, modern generative adversarial networks (GANs) have shown outstanding reconstruction performance even on unseen data [17]. In order to predict normality in autonomous driving, various machine learning algorithms are being used [2], like restricted Boltzmann machines [18], autoencoder architectures [19] [20], or normalizing flows [21]. The NFlowJS model proposed by Grcić et al. [22], however, uses normalizing flow to synthesize outliers during training and to sensitize the model for potential anomalies. Other generative approaches use GANs to resynthesize the input image [23] [24] and detect anomalous objects on the road, e.g., by the perceptual loss [25]. While there also exist methods that are based on feature extraction [2], these are typically older, which is why we excluded them from this overview.

### B. Image-to-Image Style Transfer

Benefiting from the powerful feature extraction abilities of deep neural networks, image style transfer has been an active research area in image translation. Specifically, the aim of style transfer is to render an input image into a new image, which has a similar style and texture as a style image from a target domain. As pioneers of style transforms, Gatys et al. [26] firstly proposed their method Neural Style Transfer (NST) to synthesize an image to match the style of a painting by an artist. Notably, separate control of the content and style of the picture is achieved by introducing content and style loss functions. However, due to its application to single image pairs, this approach cannot be used for large-scale style conversions. Motivated by Neural Style Transfer, another style transfer model, introduced by Johnson et al. [27], performs much more efficiently by optimizing an image transform net.

In addition to ordinary neural network models, the emergence of GANs [28], [29] opens up more possibilities in the area of style transfer. As the name suggests, the core idea of

GANs is to create desired images by simultaneously training two opposing neural networks. While a generator is responsible for synthesizing more realistic images under the guidance of a discriminator, the discriminator is trained to distinguish fake images from real images more accurately. Specifically, there are two well-known generative models for the task of image translation. With the help of adversarial and perceptual losses, the framework Pix2Pix [30] learns the mapping from a source domain to the target domain by training a model with paired training data. Differently, CycleGAN [31] employs a cycle consistency loss [32], as also applied by DualGAN [33], to conduct image-to-image translation under the unsupervised setting, making style transfers more flexible. Similar works include DiscoGAN [34], GANimorph [35], ACL-GAN [36], and AttentionGAN [37]. Besides, Park et al. [38] firstly applied contrastive learning [39] for unpaired image-to-image translation. Interestingly, the model is trained to learn the invariance between the patch of the generated image and its corresponding patch in the input image by using the InfoNCE loss [40].

## III. APPROACH

Our hypothesis to detect anomalies is based on an assumption in respect to reconstruction errors with generative models, where “...normal samples are located on a manifold and all anomalous samples are located outside. Since the manifold can be learned only where the training data lie...” [3], we were curious if generative models for style transfers might perform worse in areas of data points outside of their learned distribution. Thus, the core idea of our approach is to perform anomaly detection by leveraging the impact of successive style transformations, as visualized in Figure 1. The forward style transfer is responsible for the conversion of an input image into a stylized version of it. The backward style transfer is then tasked to transform the image back to its original style domain. Finally, a pixel level comparison between the input and re-constructed image is conducted to recognize anomalies in the input image.

**Forward Style Transfer** By forward style transfer, we refer to the process of transferring an input image to an arbitrary style domain. Thereby, the input image will be rendered in the style of the target image, while preserving its content.

**Backward Style Transfer** The aim of backward style transfer is to transform the stylized image, which is the output from the forward style transfer, back to the original style of the input image.

**Pixelwise Anomaly Detection** Finally, a pixel-by-pixel comparison is made between the input image and the newly synthesized image. Based on our hypothesis, we expected to detect anomalies in the input image through comparison. While we are aware that such a direct comparison between the re-synthesized image and the original is not sufficient for anomaly detection due to noise [41], spikes in such difference maps are a sufficient indicator for the general suitability of the approach.

## IV. EVALUATION

In this section, we present qualitative results of our approach. We performed three different sets of experiments and show the resulting pixelwise anomaly scores for multiple input images. As we performed the experiments as a first step to determine whether the methodology is suitable for anomaly detection, we omitted a quantitative analysis afterwards.

### A. Data Selection

Since the training data should reflect normality in road traffic, we first chose to utilize the well established KITTI [4] dataset and also performed experiments with the much larger Cityscapes dataset [5]. For the following evaluation, datasets with anomalies were necessary. While there exist many datasets in the field of autonomous driving [42], only a few of them are suited for the task of anomaly detection. We utilize data samples from the Fishyscapes [6] and Lost and Found [7] datasets to include both real and augmented anomalies. As the Lost and Found dataset is part of the Fishyscapes suite, we might refer to both as Fishyscapes. Both provide ground-truth labels for out-of-domain pixels, which we consider anomalies. In autonomous driving, these are also often called corner cases [43] in the category unknown object [44]. These unknown objects can typically also be seen clearly in the images directly. Figure 2 shows a selection of animals and small objects. For the art styles used for the forward passes, we selected a wide variety of styles, including art from Van Gogh, Giles and pieces in the Ukiyo-e style.

### B. Model Selection

We have performed experiments with two different image style transfer methods. First, we used CycleGAN [31], which is an established framework for unpaired image translations. For model-based style transfers, it is important to choose a model that can be trained with unpaired images as training data, contrary to approaches such as Pix2Pix [30], where paired images are necessary. Among those, CycleGAN [31] is the most prominent representative. While there are newer models, CycleGAN is optimal for experiments, since high quality open-source code [45] is available and many works have used it already, making it a mature method.

Second, we utilized the Neural Style Transfer approach [26], which is regarded as one of the early milestones in the field of style transformations. While CycleGAN is able to generate stylized images with a high quality, we have no dynamic control over how much an image should be transformed. This becomes possible with NST, so that we were able to examine the influence of conversion degrees. Based on convolutional neural networks (CNN), the Neural Style Transfer Algorithm is able to render a random input image in the style of a target image by leveraging the VGG19 model [46], which has been widely used for feature extraction in various tasks. Moreover, the highlight of the algorithm is controlling the transfer degrees through two weight parameters of the loss functions: content weight and style weight. On the one hand,

a content image would be always compared with the generated image, in order to make sure the semantic content is retained in the process of stylization. Specifically, the content structure can be preserved through minimizing the mean squared error of feature maps between a content image and the input image. On the other hand, through computation of the Gram matrix, a style image in the target domain would be always compared with the generated image, such that the synthesized image will have a similar style as the given style image. Thus, it is a tradeoff problem for the task of combining the content image with another style image, where it is impossible to meet the two constraints perfectly at the same time. For instance, the output image could have the texture of the style image, but lose some content when a larger style weight is set. On the contrary, a much larger content weight would result in that the synthesized image is further away from the target style.

### C. Style Transfer with CycleGAN

In this first experiment, we have utilized CycleGAN for both the forward and backward style transfers. First, we will give an overview of the details of the forward style transfer.

1) *Forward Style Transfer*: For this experiment, instead of training a CycleGAN model from scratch, a pre-trained model was directly used for the forward style transfer. Although the pre-trained CycleGAN model [45] was trained on paintings by Van Gogh as the target domain and ordinary photos as the source domain, it has shown to have good generalization capabilities on various datasets, such as KITTI, Cityscapes, as well as Fishyscapes and Lost and Found.

2) *Backward Style Transfer*: For the backward style transfer direction, however, the observed performance of the pre-trained GAN was not sufficient. Therefore, we chose to train the model from scratch. Notably, in order to perform the task of style transfer with a CycleGAN model successfully, images from both the source and target domain are required to be fed into the model for training. Specifically, paintings by famous artists are commonly used as source images and ordinary photos are taken for the target domain. For achieving better performance of the backward transfer, we trained our model on several data sources, which are presented in Table I.

Group	Source Domain	Target Domain
A	Paintings by Van Gogh	KITTI
B	Paintings by Van Gogh	KITTI + photos
C	Ukiyo-e style paintings	KITTI
D	Ukiyo-e style paintings	KITTI + photos
E	Van Gogh stylized KITTI	KITTI
F	Van Gogh stylized Cityscapes	Cityscapes

TABLE I  
DATA SOURCES USED FOR THE TRAINING OF CYCLEGAN MODELS

We firstly conducted four sets of trainings, shown as groups A to group D. Paintings by Van Gogh were adopted as the source domain images in both groups A and B. In group A, we used images from the KITTI dataset as the target domain, whereas a combination of KITTI data and ordinary photos

served as target style images in group B. Additionally, we employed the Japanese painting style Ukiyo-e [47] as the source style in group C and D. For the target domain, group C used KITTI data as done for group A, and group D combined KITTI with ordinary photos as in group D. However, this set of experiments did not perform well, even though we observed slightly better results when using Van Gogh style images and KITTI data. Here, two problems occurred. On the one hand, some noise was constantly present, which we attribute to the special texture of Van Gogh paintings. On the other hand, the performance of the model was erratic, resulting in the occasional lack of much difference between transformations and the corresponding input images. After this analysis, we found two directions to resolve the problem. Firstly, rather than directly using paintings from artists, we took stylized images, which were the output from the forward style transfer network, to train our model with more domain-specific data. Secondly, there were too few images from the source domain for training. With the newly selected data source, we were able to increase the number of stylized images significantly and performed another two trainings, as shown in groups E and F. In group E, Van Gogh stylized KITTI samples were used for the source domain, and KITTI samples were used for the target domain. In group F, we followed the same approach, but used Cityscapes data.

3) *Pixelwise Anomaly Detection*: For the evaluation, an input image chosen from the Fishyscapes dataset was used as the input. With the final approach as shown in group F of Table I, we were able to generate images with the backward style transfer pass that were consistently close to the original ones, as shown in Figure 2.

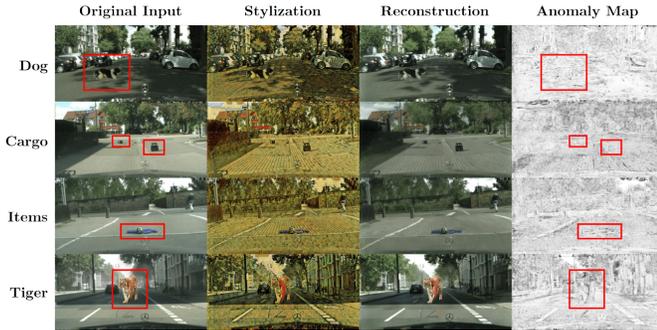


Fig. 2. Evaluation results of the model trained with Van Gogh stylized Cityscapes images as the source domain and Cityscapes samples as the target domain. Four different types of anomalies are included in the Fishyscapes scenes, which are highlighted in red.

This observed performance improvement on the Fishyscapes dataset is largely due to the fact, that Fishyscapes is an extension of the original Cityscapes datasets, so the general domain remains unchanged. After the conversion of the input into a stylized version of it, based on the forward style transfer, we reconstructed the input image based on the backward style transfer. The re-synthesized images are generally very close to the original ones, which was a wanted effect. However, as

visible in the last column, we did not observe any deformations in the area of the anomaly.

#### D. Style Transfer with NST

For the second experiment, we used the Neural Style Transfer method for both the forward and backward style transformations. In the CycleGAN framework, weight parameters are updated iteratively, until the model converges. Contrary, with NST, it is a single image that is optimized in the training process. Thus, only two images are necessary for the process: a content image and a style image. Therefore, we directly utilize Fishyscapes data samples for the style transfers.

1) *Forward Style Transfer*: For the forward style transfer, a sample from Fishyscapes is used as the content image, and a painting from Van Gogh is taken as the style image in the optimization process. Content weight and style weight are two important parameters, that can be set to control how much the image will be stylized. As recommended by reference implementations, [48] [49], the content weight is kept unchanged at  $1e5$ , such that the process of style transfer can be easily controlled by only changing the style weight.

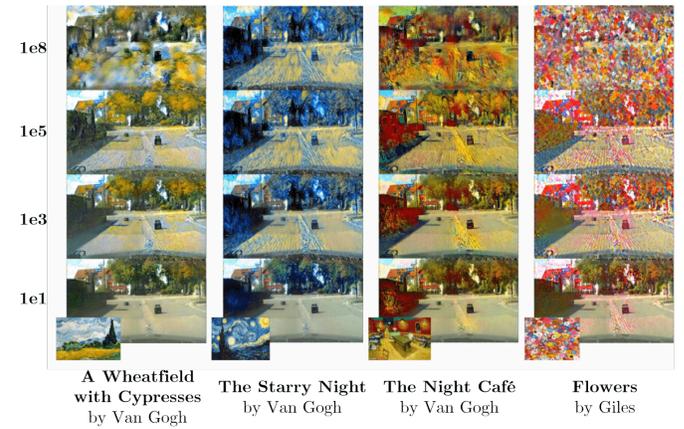


Fig. 3. Style transfers with NST with a scene from the Fishyscapes dataset with lost cargo as anomalies. Each column represents a different style and each row a different style weight.

To determine, which degree of style transfer is most likely to lead to distortions of the content, especially anomalies, during the forward style transfer, we performed a set of small experiments with the style weight. As visualized in Figure 3, we performed experiments with four different styles [50]–[53], and four different style weights. The larger the style weight, the better the generated image matches the style image. Since a style weight of  $1e8$  often led to a nearly complete loss of the content information, we utilized a value of  $1e5$  for the further experiment to maximize the effect of the style transfer, while retaining the content.

2) *Backward Style Transfer*: To convert the stylized Fishyscapes images back to their original style of a regular street scenario, we chose an image from the Cityscapes dataset as the style image. Importantly, this image from Cityscapes does not contain any anomalies, such that it only works as

a style guidance, avoiding an unexpected influence on the content part.

3) *Pixelwise Anomaly Detection*: Illustrative results of the second experiment are shown in Figure 4. Here, we utilized the same styles as introduced in Figure 3. To highlight the effects of the different styles on the reconstruction and anomaly detections, we show only one scene from the Fishyscapes dataset.

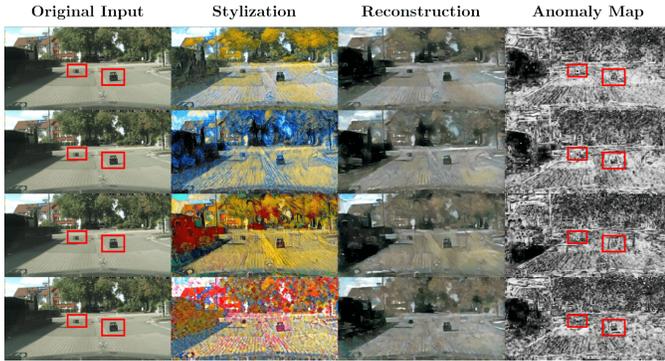


Fig. 4. Evaluation of Neural Style Transfer with a content weight and a style weight of  $1e5$  each. A sample scene from the Fishyscapes dataset is used as the input. Each row shows the process with a different style. Anomalies are highlighted in red.

Here, it becomes visible that the results of the backwards style transfers differ less than the differences of the stylized images might suggest. Again, the anomalies remain well visible in the reconstruction. The anomaly map in the last column shows no spikes in these regions, but shows a high level of difference values in general.

#### E. Hybrid Style Transfer

For the third and final experiment, we combined both approaches. Here, we performed the forward style transfer with the Neural Style Transfer method, while we used our CycleGAN model, trained with Cityscapes data, for the backward style transfer. For this experiment, we followed the intuition that the NST allows for a better stylization control for the forward pass, while the CycleGAN enables higher quality results for the backward pass. As we have explained both passes in the previous Sections IV-C and IV-D, here we describe only the results, which are visualized in Figure 5. It can be seen, that with such a variety of styles, the GAN, which was only trained with one style, was not capable to perform the backward style transfer properly. Here, a leakage of the applied style to the reconstruction occurs with styles, which were not used for the training of the GAN. Regarding the anomaly detections, we observe a similar behavior as in the previous experiments. The amount of noise is, compared to the other two experiments, in a medium range.

With this presented set of three comprehensive experiments, we examined the influence of training data, style data, anomaly data, style weights, and style transfer models for the purpose of anomaly detection based on successive style

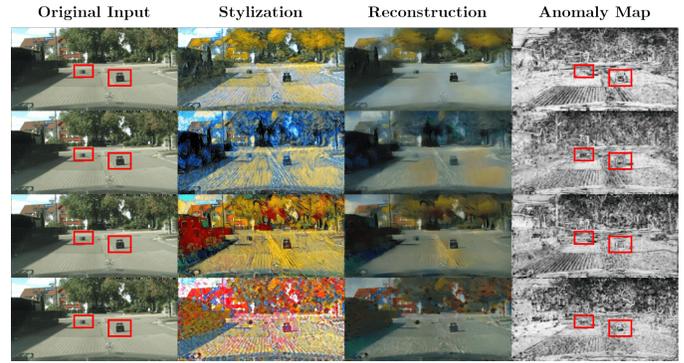


Fig. 5. Evaluation of the hybrid approach with a sample scene from the Fishyscapes dataset. Here, NST is responsible for the forward style transfer. The backward style transfer was carried out by our CycleGAN model. Each row represents a different style. Anomalies are highlighted in red.

transfers. While the results showed high-quality reconstructions, the style transfers unfortunately reconstructed the areas with anomalies similarly to the rest of the images. Thus, the difference maps showed primarily noise, which makes the detection of anomalies with this method unlikely. Specifically, our experiments have placed a focus on the comparison of CycleGAN, Neural Style Transfer, and a hybrid combination of both for the Forward and Backward Style Transfers parts. The results differed in the value distribution of the difference maps and the amount of noise, but showed no sufficient differences for the pixels of the anomalies.

## V. CONCLUSION

The aim of our work was to investigate whether our proposed approach of successive style transfers is suitable as a novel method for anomaly detection in the domain of autonomous driving. Therefore, we performed a set of three extensive experiments. Within these, we transferred images from the road traffic domain to multiple art domains with a set of forward style transfer methods and then transferred the results back to the original domain. We compared these reconstructions with the original images to calculate a pixelwise anomaly score. Our experiments revealed, that anomalies were still consistently visible in the reconstructed image, which did not lead to high anomaly scores in the difference maps within the areas of the anomalies. Based on our analysis visualizations, these negative results should serve as inspiration and knowledge to support following research in the field of anomaly detection in autonomous driving.

## REFERENCES

- [1] Z. Zou, Z. Shi, Y. Guo, and J. Ye, “Object detection in 20 years: A survey,” *arXiv:1905.05055*, 2019.
- [2] D. Bogdoll, M. Nitsche, and J. M. Zöllner, “Anomaly Detection in Autonomous Driving: A Survey,” *arXiv:2204.07974*, 2022.
- [3] V. Šmídl, J. Bím, and T. Pevný, “Anomaly scores for generative models,” 2019.
- [4] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

- [5] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *IEEE conference on Computer Vision and Pattern Recognition*, 2016.
- [6] H. Blum, P.-E. Sarlin, J. Nieto, R. Siegwart, and C. Cadena, "Fishyscapes: A benchmark for safe semantic segmentation in autonomous driving," in *IEEE/CVF International Conference on Computer Vision Workshops*, 2019.
- [7] P. Pinggera, S. Ramos, S. Gehrig, U. Franke, C. Rother, and R. Mester, "Lost and found: detecting small road hazards for self-driving vehicles," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016.
- [8] D. Fanelli, "Negative results are disappearing from most disciplines and countries," *Scientometrics*, vol. 90, 2012.
- [9] V. Rodrigues, "Publication and reporting biases and how they impact publication of research," <https://www.editage.com/insights/publication-and-reporting-biases-and-how-they-impact-publication-of-research>, 2013, accessed: 2022-05-03.
- [10] L. Goodchild van Hilten, "Why it's time to publish research 'failures'," <https://www.elsevier.com/connect/scientists-we-want-your-negative-results-too>, 2015, accessed: 2022-05-03.
- [11] D. Mehta, "Highlight negative results to improve science," <https://www.nature.com/articles/d41586-019-02960-3>, 2019, accessed: 2022-05-03.
- [12] G. Conroy, "Three reasons to share your research failures," <https://www.natureindex.com/news-blog/three-reasons-to-share-your-research-science-failures>, 2020, accessed: 2022-05-03.
- [13] IEEE Access, "Submission guidelines," <https://ieeaccess.ieee.org/guide-for-authors/submission-guidelines/>, 2022, accessed: 2022-05-03.
- [14] V. B. Alex Kendall and R. Cipolla, "Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding," in *British Machine Vision Conference (BMVC)*. BMVA Press, 2017.
- [15] F. Heidecker, A. Hannan, M. Bieshaar, and B. Sick, "Towards Corner Case Detection by Modeling the Uncertainty of Instance Segmentation Networks," in *Pattern Recognition. ICPR International Workshops and Challenges*, 2021.
- [16] X. Du, Z. Wang, M. Cai, and S. Li, "Vos: Learning what you don't know by virtual outlier synthesis," in *International Conference on Learning Representations*, 2022.
- [17] D. Bogdoll, J. Jesträm, J. Rauch, C. Scheib, M. Wittig, and J. M. Zöllner, "Compressing Sensor Data for Remote Assistance of Autonomous Vehicles using Deep Generative Models," *NeurIPS Conference on Neural Information Processing Systems Workshop on Machine Learning for Autonomous Driving (ML4AD)*, 2021.
- [18] C. Creusot and A. Munawar, "Real-time small obstacle detection on highways using compressive RBM road reconstruction," in *IEEE Intelligent Vehicles Symposium*, 2015.
- [19] T. Ohgushi, K. Horiguchi, and M. Yamanaka, "Road Obstacle Detection Method Based on an Autoencoder with Semantic Segmentation," *Computer Vision*, vol. 12627, 2021.
- [20] T. Vojir, T. Šipka, R. Aljundi, N. Chumerin, D. O. Reino, and J. Matas, "Road anomaly detection by partial image reconstruction with segmentation coupling," in *IEEE/CVF International Conference on Computer Vision*, 2021.
- [21] H. Blum, P.-E. Sarlin, J. Nieto, R. Siegwart, and C. Cadena, "The Fishyscapes Benchmark: Measuring Blind Spots in Semantic Segmentation," *International Journal of Computer Vision*, vol. 129, 2021.
- [22] M. Grcić, P. Bevandić, Z. Kalafatić, and S. Šegvić, "Dense anomaly detection by robust learning on synthetic negative data," *arXiv:2112.12833*, 2021.
- [23] D. Haldimann, H. Blum, R. Siegwart, and C. Cadena, "This is not what I imagined: Error Detection for Semantic Segmentation through Visual Dissimilarity," *arXiv:1909.00676*, 2019.
- [24] J. Nitsch, M. Itkina, R. Senanayake, J. Nieto, M. Schmidt, R. Siegwart, M. J. Kochenderfer, and C. Cadena, "Out-of-Distribution Detection for Automotive Perception," in *IEEE International Intelligent Transportation Systems Conference*, 2021.
- [25] K. Lis, K. Nakka, P. Fua, and M. Salzmann, "Detecting the unexpected via image resynthesis," in *IEEE/CVF International Conference on Computer Vision*, 2019.
- [26] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *IEEE conference on Computer Vision and Pattern Recognition*, 2016.
- [27] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European Conference on Computer Vision*, 2016.
- [28] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, 2014.
- [29] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv:1701.00160*, 2016.
- [30] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *IEEE conference on computer vision and pattern recognition*, 2017.
- [31] J.-Y. Zhu, T. Park, P. Isola, and A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [32] Z. Kalal, K. Mikolajczyk, and J. Matas, "Forward-backward error: Automatic detection of tracking failures," in *International Conference on Pattern Recognition*, 2010.
- [33] Z. Yi, H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," in *IEEE international conference on computer vision*, 2017.
- [34] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *International Conference on Machine Learning (ICML)*, 2017.
- [35] A. Gokaslan, V. Ramanujan, D. Ritchie, K. I. Kim, and J. Tompkin, "Improving shape deformation in unsupervised image to image translation," in *European Conference on Computer Vision*, 2018.
- [36] Y. Zhao, R. Wu, and H. Dong, "Unpaired image-to-image translation using adversarial consistency loss," in *ECCV*, 2020.
- [37] H. Tang, H. Liu, D. Xu, P. H. Torr, and N. Sebe, "Attentiongan: Unpaired image-to-image translation using attention-guided generative adversarial networks," *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2021.
- [38] T. Park, A. A. Efros, R. Zhang, and J.-Y. Zhu, "Contrastive learning for unpaired image-to-image translation," in *European Conference on Computer Vision*, 2020.
- [39] J. Li, P. Zhou, C. Xiong, and S. C. Hoi, "Prototypical contrastive learning of unsupervised representations," in *ICLR*, 2021.
- [40] A. Van den Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv:1807.03748*, 2018.
- [41] G. Di Biase, H. Blum, R. Siegwart, and C. Cadena, "Pixel-wise anomaly detection in complex driving scenes," in *IEEE/CVF conference on computer vision and pattern recognition*, 2021.
- [42] D. Bogdoll, F. Schreyer, and J. M. Zöllner, "ad-datasets: a meta-collection of data sets for autonomous driving," in *International Conference on Vehicle Technology and Intelligent Transport Systems*, 2022.
- [43] D. Bogdoll, J. Breitenstein, F. Heidecker, M. Bieshaar, B. Sick, T. Fingscheidt, and M. Zöllner, "Description of Corner Cases in Automated Driving: Goals and Challenges," in *IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2021.
- [44] J. Breitenstein, J.-A. Termöhlen, D. Lipinski, and T. Fingscheidt, "Systematization of corner cases for visual perception in automated driving," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020.
- [45] J.-Y. Zhu, "CycleGAN and pix2pix in PyTorch," <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>, 2022, accessed: 2022-05-19.
- [46] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations ICLR*, 2015.
- [47] Wikipedia, "Ukiyo-e," <https://en.wikipedia.org/wiki/Ukiyo-e>, 2022, accessed: 2022-05-05.
- [48] Malikanhar, "Neural-Style-Transfer," <https://github.com/Malikanhar/Neural-Style-Transfer>, 2021, accessed: 2022-05-20.
- [49] A. Gordić, "Neural Style Transfer (optimization method)," <https://github.com/gordicaleksa/pytorch-neural-style-transfer>, 2020, accessed: 2022-05-20.
- [50] V. Van Gogh, "The Starry Night," [https://en.wikipedia.org/wiki/The\\_Starry\\_Night](https://en.wikipedia.org/wiki/The_Starry_Night), 1889, accessed: 2022-05-19.
- [51] —, "The Night Café," [https://en.wikipedia.org/wiki/The\\_Night\\_Caf%C3%A9](https://en.wikipedia.org/wiki/The_Night_Caf%C3%A9), 1888, accessed: 2022-05-19.
- [52] —, "Wheat Field with Cypresses," [https://en.wikipedia.org/wiki/Wheat\\_Field\\_with\\_Cypresses](https://en.wikipedia.org/wiki/Wheat_Field_with_Cypresses), 1889, accessed: 2022-05-19.
- [53] B. L. Giles, "2011–2014 Cover," <https://benlewisgiles.format.com/collageg>, 2014, accessed: 2022-05-19.