26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Surface Quality Augmentation for Metalworking Industry with Pix2Pix

Dirk Hölscher[a], Christoph Reich[a], Martin Knahl[a], Frank Gut[a], Nathan Clarke[b]

[a]*Institute for Data Science, Cloud Computing and IT Security, Furtwangen University, 78120 Furtwangen, Germany*
[b]*Centre for Cyber Security, Communications and Network Research, Plymouth University, Plymouth, U.K*

Abstract

Image augmentation has become an important part of the data preprocessing pipeline, helping to acquire more samples by altering existing samples by cutting, shifting, etc.. For some domains, augmenting existing images is not sufficient, due to missing samples in the domains (e.g., faulty work pieces or events that occur infrequently). In such a case, new samples must be generated, since images with surface quality defects are often rare occurrence in metalworking and the amount of samples even with standard augmentation techniques does not meet requirements to train a Convolutional Neural Network (CNN) for fault detection.

This paper utilizes Pix2Pix for image augmentation to generate new images with surface quality defects. The approach allows specifying the kind of defect, location, and size and transforms images by adding new defects. Furthermore, metrics to evaluate the augmented images are discussed and a recommendation of the best performing metric within the domain of metalworking is given.

Keywords: Augmentation, Pix2Pix, Metrics, Metalworking, Generative Models

1. Introduction

The production of metal workpieces is subject to a precise, previously defined production tolerance. Deviations in the manufacturing tools can cause damage to the workpieces and the machine. Some contaminations such as micro scratches cannot be avoided and always occur. Other defects only occur during production or by mishandling the workpiece and reduce the overall quality. Scratches have a certain characteristic, depth and length. With all the possible variances and other defects which can occur, it is important to have a well-balanced dataset with enough representations per class to train a classification model.

---

* Corresponding author.

E-mail address: dirk.hoelscher@hs-furtwangen.de

Augmentation is an important task for machine learning and helps to enrich datasets with additional samples, mitigating the absence of samples while enabling training with neural networks with usable output. The absence of samples creates various problems like overfitting, low accuracy or problems with generalization, This lack of data leads to imbalanced data and makes it harder to train deep learning classification models [1]. In some domains there are possible events which occur infrequently and are the exception but have an high impact within the domain. Such events produce only a fraction of the required samples and thus are in dire need of augmentation. With traditional augmentation methods, sample size can be increased but is limited to the existing samples. Furthermore, within the metalworking industry, the detection of surface defects is important to prevent damage and to analyze the cause of damage. Timely detection is key to limit the occurrence of surface defects [2]. On metal workpieces scratches can occur with different size, depth, length and angle all over the workpiece. Some of these might be extremely rare. Some samples can be created through basic augmentation steps (e.g., rotation, mirroring or cropping) but other instances cannot be replicated with these tools.

Augmentation with the help of generative models are able to mitigate such problems. Although a lot of GAN architectures rely on a lot of training data themselves to create new images for example, Deep Convolutional Generative Adversarial Network (DCGAN) [3], StyleGAN [4] or Copy-Pasting GAN [5], Image-to-Iamge translation GANS like CycleGAN [6] and Pix2Pix [7] are able to work with less data and learn the distribution between domains in an unpaired (CycleGAN) or paired (Pix2Pix) manner. The difference between paired and unpaired lies in the structure of the required data and how the model learns the transformation. In a paired approach the models requires an image of the same structure in both domains (e.g., winter to summer translation of a tree, requires an image of the same tree in winter and summer). In contrast to that unpaired approaches don't have the requirement of pairs. In this case the summer and winter image can have different locations with different trees and learns the translation itself instead of showing the model what has to be learned through pairs. Through these translations, the network is able to generate new samples based on the input of domain A (existing images of surface defects) and translate them into samples of domain B (samples of rare occurrences through added scratches and changed angles). Pix2Pix is used to learn to place scratches at certain locations depending on the position of the input's scratch.

The paper is organized as follows: Section 2 presents related work about the domain of metalworking, augmentation techniques to create balanced datasets. Section 3 describes basic augmentation process. The next Section 4 describes the domain as well as the requirements placed on the domain. Section 5 presents the results of the augmentation process using Pix2Pix, with Section 6 evaluating the achieved results using image quality metrics. The last Section 7 concludes this paper.

2. Related Work

The following section will give a brief overview of related work regarding metalworking domain and augmentation mechanisms.

2.1. Scratch Detection

In [2] Konovalenko et al, present a classifier for metal surface defects using a Convolutional Neural Network. The model is able to recognize scratches and abrasions. Surface defects are a regularity when working with metal with various sources these defects can originate from such as the original workpiece, chemical composition, or the utilized equipment. To detect such defects, the authors build a classifier based on ResNet50 and ResNet152. The dataset contained a total of 3938 images collected from various sources and was split into 10% test, 15% validation and 75% training. RestNet was chosen because it is easy to optimize while increasing model depth. In addition, transfer learning was used to prevent random detection. The weights of the network were obtained by training on the ImageNet dataset. The build classifier is binary and has a range between 0 and 1 with a score closer to one indicating the presence of a defect. For augmentation, random crop was used to expand the sample size. 12 models were trained with an uneven dataset as most samples did not contain surface defects and achieved an overall accuracy of 97%.

Huang et al., [8] present a slim Convolutional Neural Network for Automated Optical Inspection of metal workpieces to detect surface defects. The designed network was compared against LeNet, VGG-19, ResNet-34, DarkNet-19, and DarkNet-53 and showed that a neural network without pooling and fully connected layers can reliably detect

surface defects. The dataset contained 1895 samples with 593 showing defects. The compared models achieved an accuracy between 93.68% and 98.64%. The customized CNN contains 8 convolutional layers with Batch Normalization/ReLU activation between the first 6 convolutional layers. After the convolutional layer 7 and 8 Softmax function is used. The sixth layer has an output size of 24 x 24 to reach an output size of 1 x 1 in the 7 layer with a filter size of 512 replacing the fully connected layer and reducing parameters. The network achieved an accuracy of 99,36%.

2.2. Augmentation

Hammami, Friboulet and Kechichian [9] propose a Cycle-GAN based augmentation method to detect multiple organs in CT images using YOLO. In their domain paired images are hard to obtain as it is impossible to obtain images for one patient under the same condition, making a mapping between source and target domain impossible. Their aim is to have segment the liver without any ground-truth for the target domain. First Cycle-GAN is used to synthesize images and afterwards YOLO is used to detect multiple organs. MRI images were used as the source domain an the Cycle-GAN was trained to transform the images to an CT image. The translated images show characteristics of CT images with brighter organs and darker muscle tissue. To evaluate the quality of the approach, SSIM was used to calculate the score between the reference MRI image and the transformed CT image. The approach achieved a mean SSIM score of 0.97.

In [10] Zhang, Kinoshita and Kiya present an augmentation method for exposure based on luminance. The proposed method uses Automatic exposure compensation for single-image-based multi-exposure fusion to create dark, bright and normally illuminated images. First, the luminance of the input image is calculated. Second, the local contrast of a pixel is enhanced by taking the average luminance around the pixel. Third, an area of the image is separated using luminance distribution. Fourth, is the calculation of the scaled luminance. The fifth step is a tone map to increase luminance. The last step is to generate a pseudo ME images with the desired luminance. Using the GTSRB and CIFAR-10 datasets and a ResNet-18. To simulate the absence of data, only a quarter or an eighth of the data was used. First, the model was trained without any augmentation and scored an accuracy of 97.48% for GTSRB and 85.92% for CIFAR-10. Using the proposed augmentation, the accuracy increased to 97.88% for GTSRB and 85.94% for CIFAR-10.

Shorten and Khoshgoftaar [11] conducted a survey on augmentation techniques, summarizing all the different augmentation techniques. The summarized techniques include geometric transformations, colour space augmentations, kernel filters, mixing images, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning. Furthermore, the authors discuss test-time augmentation, impact on resolution or final dataset size. A taxonomy of the different techniques is presented, divided into basic manipulations such as geometric transformations or kernel filters; deep learning techniques applying adversarial training neural style transfer or GANs and Meta learning with neural, auto and smart augmentation. They set the definition why augmentation is required by describing the problem of classification and problems (viewpoint, lighting, occlusion or background) the model is facing when training to differentiate between classes. The survey is very detailed and list the individual augmentation steps and explains in detail how well they work and the implementation effort.

3. Augmentation Techniques

Image augmentation is the process of manipulating images by altering their characteristics. Characteristics such as contrast, blur, sharpness, noise or brightness can be added to the image individually or combined. Further geometric augmentation steps are translation, rotation, scaling, perspective distortion, pixelation or random erasing, where random pixels in a defined space are deleted [12]. In addition, image synthesis has become increasingly popular for augmentation. For example, pairing samples by overlapping two images to increase the overall sample size lead to improvements [13] or neural style transfer techniques[14]. For each augmentation step, the severity of the added effect can be defined. The change of contrast effects the overall brightness of the image and can lead to brighter or darker areas in an image. Sharpness is responsible for how detailed features in an image appear. With decreasing sharpness, the edges around features degrade and become harder to spot. Blur is similar to sharpness and reduces information around the edges of objects while smoothing out the transition between colours. Fog and clouds adds bright spots in an image, limiting visibility in the area. Furthermore, augmentation can not only alter characteristics of an image, but

also its geometry. An image can be cropped, zoomed in, mirrored, distorted, rotated or remove pixel all to change the perception of the original image.

Augmentation and image quality are closely linked with CNN based object detection. Standard augmentation techniques work well, if you want to get a more general data set and generate variations of existing images. The augmentation through generation as a first step, that helps to increase sample size by creating new plausible images, which in return can be further altered via traditional augmentation. To achieve a more balanced dataset both techniques are required. All this helps to train more robust and detailed neural networks.

4. Domain Specific Requirements for Data Augmentation

Smart factories allow choosing the best machine with the best available production precision to produce the desired workpiece. This highly individualized manufacturing processes requires, depending on the customer's wishes, varying quality levels of materials and product precision, based on the potential use of the workpiece. Wear and tear of tools is common and depending on the workpiece, tool condition and other factors with production impact, results can vary. Failure and errors in manufacturing can occur, but usually specific cases very rare.

Metal surfaces have functional requirements, such as hardness and abrasion resistance, but also optical requirements. Manufacturing processes often causes kinds of impurities and micro scratches on metal surfaces, which might be negligible in the overall process, whereas deeper and longer scratches are a problem in higher quality tiers.

With the help of image generation as augmentation technique, surface scratches of different sizes, depth, width and angles can be created to train a more robust model which can identify and classify manufacturing errors in a more reliable way. For detection and classification tasks, a larger pool of different scratches having a variety of the aforementioned characteristics helps to assess plates and determine if they meet the required quality attributes. There are certain requirements for data augmentation to be met:

Domain Specification: Within a domain are certain thresholds specifying if the captured image of a workpiece or a situation is within the allowed domain boundaries. This separates defect from non-defect objects. For example, this can be the size, condition of the workpiece itself but also concerns the image's characteristics such as contrast, saturation and luminance.

Defect Specification: The next step is to define what counts as a defect in the domain and if so, if there are different level of defects. As previously stated, everything outside the expected range is a defect.

Analysis of missing data: Missing samples must be identified. Within a domain a certain level of errors can occur. Some of these are very rare. For a successful augmentation process such events or errors must be identified.

Specification of Field of Interest: After defect definition and analysis of what is missing in the data, it must be determined if every error is worth detecting or if some of them are neglectable for the overall process, because they are above the defined threshold.

5. Domain Specific Data Augmentation with Pix2Pix

The following section will present the Pix2Pix's ability to perform transformations of higher complexity and shows the achieved augmentation results with different test sets. The ability to add new surface scratches, as well as the possibility to alter existing scratches by changing the angle or the length of scratches, helps to further increase sample size and the variety of occurring scratches. Some tested image transformations are more out of scope for Pix2Pix's intended purpose of style transfer's by dramatically changing the styles between domain A and B for the example of adding scratches.

Pix2Pix by Isola et al. [7] is a Generative Adversarial Network for image-to-image translations. The generator is responsible to create new synthetic images, with the discriminator classifying if the sample is plausible/real or fake. Both models are trained simultaneously, with the discriminator being updated directly and the generator being updated through the discriminator. Pix2Pix is a conditional GAN with the output image requiring to satisfy the condition of the source image (domain A). The discriminator receives the source image as well as the target image and must determine

if the received target image is a transformation from domain A to domain B. Adversarial loss is used to train the generator and L1 loss (difference between true (source image) and predicted (output image) value). As Pix2Pix is a paired approach, each image of domain A requires a corresponding image from domain B. Unlike Cycle-GAN which does not require pairs to learn the distribution. Pix2Pix allows to specify the precise nature of a defect and eliminates the uncertainty of letting the model itself trying to find the distribution. This is achieved by handing the model pairs of both domains, showing how the transformation is supposed to be.

Adding new scratches follows a random pattern of already existing scratches and newly added ones at random locations of the used aluminium plate. Therefore, there is no fixed pattern Pix2Pix can learn for transformations. For example, a 2 centimetre long scratch in the top right corner of the plate in domain A has a pair with an added scratch in the lower middle of 3 centimetres. Normally Pix2Pix would be able to learn the pair's distribution but in our experiments there is a high chance of another sample with a top right scratch and a different scratch added in domain B, to represent a more realistic data distribution. This behaviour of mixed samples is used across all experiments and forces Pix2Pix to learn a different distribution than it normally does. The outcome of these experiments could lead to the Pix2Pix not being able to learn a meaningful distribution and simply ignoring domain B, a trade-off where it does get some of the distribution right but cannot translate the input into the expected output. Pix2Pix is able to fully learn the distribution and translate images from domain A to B. The dataset is based on a squared aluminium plate with micro scratches and other minor impurities. Scratches are added in white at random. This is the case for all experiments. Each dataset, contains 1500 pairs of images for the translation. The first conducted experiment as shown in Figure 1 was to add new surface scratches to the aluminium plate. Pix2Pix was trained on 100 epochs

Fig. 1. Experiment 1 - Add scratches - Left Source Domain, Right Target Domain

with a batch size of 1 and each 10 epochs the best model is saved. Training showed varying results during training rounds, sometimes failing to learn any distribution and simply ignoring the changes present in domain B or converging faster than expected with marginal improvements. Figure 2 and Figure 3 show the most promising training results of adding another scratch to a plate. As previously discussed, Pix2Pix wasn't able to learn the exact distribution to transfer between domain A and B but learned a combined distribution of the samples to add new scratches. Adding new scratches to previously unseen images work reliable and well, although the scratches are rather short. The next
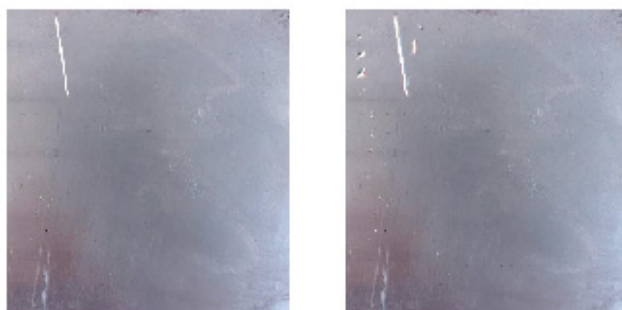

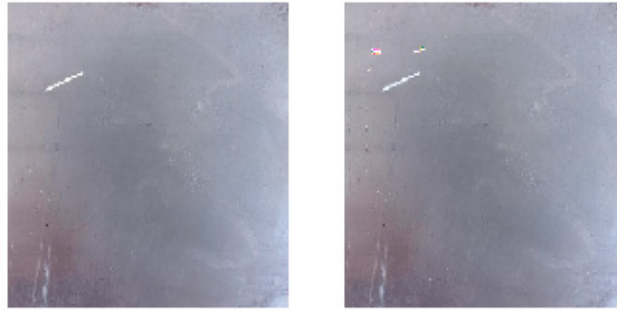Fig. 2. Experiment 1 - Add scratches - Result 1 - Left Source, Right Generated

Fig. 3. Experiment 1 - Add scratches - Result 2 - Left Source, Right Generated

experiment's dataset involved shifting the angle of a scratch as shown in Figure 4 with one fix point to rotate the scratch in any direction. The succeeding Figures 5 and 6 illustrate the training success of the approach. As with the first experiment a direct mapping from one scratch to another is not completely possible due to the randomness while creating the dataset, there is a high chance of a duplicate scratch in domain A with a different angled counterpart in domain B. Still Pix2Pix was able to learn the distribution for a reliable transformation allowing to change the angle of scratches. Furthermore, the achieved quality of the transformation is higher as with the previous experiment.

Fig. 4. Experiment 2 - Change of Angle - Left Source Domain, Right Target Domain
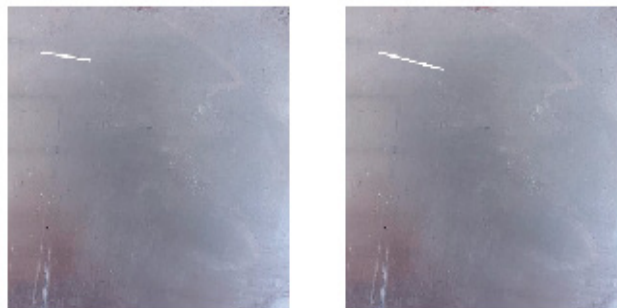

Fig. 5. Experiment 2 - Change of Angle - Result 1 - Left Source, Right Generated

The third experiment is similar to the first experiment and differs only by using a different darker aluminium plate, hiding some of the previous recorded impurities. The overall process of generating images was the same as described in experiment 1. A total of 1500 pairs were randomly created to simulate real world deviations. Figure 7 shows an example pair of Pix2Pix's training input using the aforementioned darker aluminium plate, while Figure 8 illustrates the training success of an unseen input. Similar to the first experiment of adding scratches a darker surface with less structural details on the base plate also showed promising results, although Pix2Pix still was not able to fully learn

Fig. 6. Experiment 2 - Change of Angle - Result 2 - Left Source, Right Generated

the desired transformation, but is able to generate new images with added scratches. Although the problem of an exact translation between both domains still persists new images can be created meeting the required parameters, thus solving the problem of sparse images.
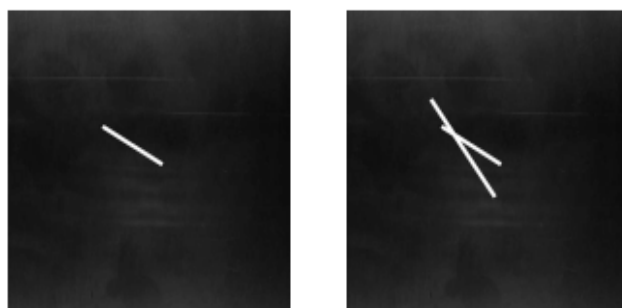

Fig. 7. Experiment 3 - Add scratches on darker surface - Left Source Domain, Right Target Domain
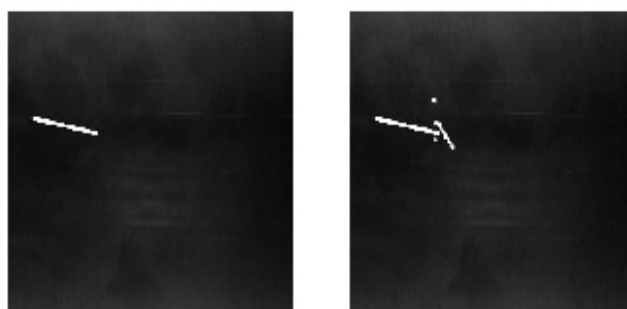

Fig. 8. Experiment 3 - Add scratches - Result - Left Source, Right Generated

6. Evaluation of the Approach

The following section will evaluate the achieved results presented in Section 5. For the evaluation process different metrics for quality assessment were selected. The reference images for the experiments are shown in Figure 9. Among the selected metrics are Root Mean Squared Error (RMSE), Peak signal-to-noise ratio (PSNR), Structural Similarity Index (SSIM), Feature-based similarity index (FSIM) and Universal image quality index (UIQ). The first chosen metric is RMSE a similarity measurement between the reference and generated image and is the rooted summed squared
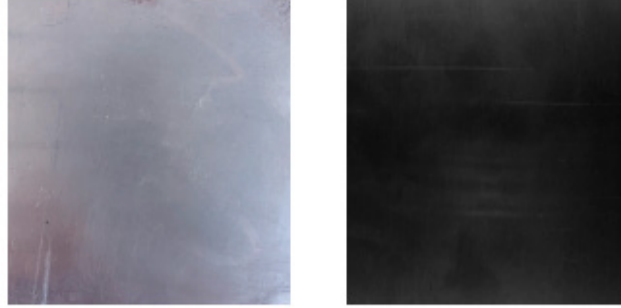
Fig. 9. Reference Images

difference between both images. RMSE has some problems with scale variation which can lead to differentiating results when using images of 8 and 10-Bit depth. High quality samples in RMSE score close to 0.0.

$$\sqrt{RMSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [g(i,j) - I(i,j)]^2} \ [15] \tag{1}$$

Therefore, the second metric chosen is PSNR (see [15]) mitigating the above-mentioned problems by scaling the error calculated with RMSE / MSE. $g(i,j)$ be the resulting image of a transformation process and let $I(i,j)$ be the original image. PSNR utilizes MSE as shown in equation without calculating the square root.

$$PSNR = -10 \log_{10} \frac{e_{MSE}}{S^2} \ [15] \tag{2}$$

In PSNR, S indicates light intensity of the image and the given result is in decibels. Usually PSNR is used for reconstruction purposes and high quality sample scores are undefined or can score to infinity due to the metrics structure of MSE reaching close to 0 for high quality samples.

Instead of estimating the changes in an image SSIM (see [16]) the third used metric models the changes in the image's information while also using sub-samples instead of the whole image. X and y are the location of sub-sample used with size NxN, with μ being the average pixel intensity within the frame of x and y. σ is the varying intensity of x and y with c being the covariance. The results of SSIM are between -1 and 1, where 1 is the perfect score.

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \ [16] \tag{3}$$

Next UIQ is calculated by using the signals of the original images x and test images y with $\mathbf{x} = \{x_i \mid i = 1, 2, \ldots, N\}$ and $\mathbf{y} = \{y_i \mid i = 1, 2, \ldots, N\}$. The first part of the equation calculates the degree of linear correlation of x and y and lies between -1 and 1 (best). The second part calculates the mean luminance of x and y and lies between 0 and 1, scoring 1 if x and y are equal. The third part calculates the contrast of x and y and how similar the contrast between both on a scale from 0 to 1 (same contrast). The score of UIQ lies between -1 and 1, with 1 being a perfect score.

$$Q = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \ [17] \tag{4}$$

The most complex metric FSIM is based on phase congruency (PC), which measures the significance of a local structure while ignoring contrast, and gradient magnitude (GM) a measurement how rapidly the colour and intensity changes in a certain direction.

$$FSIM = \frac{\sum_{\mathbf{x}\in\Omega} S_L(\mathbf{x}) \cdot PC_m(\mathbf{x})}{\sum \mathbf{x} \in \Omega PC_m(\mathbf{x})} \ [18] \tag{5}$$

Calculating the similarity between images, first the PC and GM feature maps must be calculated. For each conducted experiment ten images were created and evaluated with each metric. The average of each metric is shown in the following Tables 1, 2, 3

Table 1. Experiment 1- Add Scratches

Metric	Score (t))
SSIM - [-1,1(best)]	0.9983066917864712
RMSE - [0.0(best)	0.0016183736734092236
FSIM - [0,1(best)]	0.48860036274459956
PSNR - [0.0,∞(best)]	55.81842425992923
UIQ - [-1,1(best)]	0.39679456965473264

Table 2. Experiment 2 - Change of Angle

Metric	Score (t))
SSIM - [-1,1(best)]	0.998854581602763
RMSE - [0.0(best)	0.0012314228806644678
FSIM - [0,1(best)]	0.5151540127044197
PSNR - [0.0,∞(best)]	58.191854988841584
UIQ - [-1,1(best)]	0.405105805156216

Table 3. Experiment 3 - Add Scratches Dark Plate

Metric	Score (t))
SSIM - [-1,1(best)]	0.9866701288305512
RMSE - [0.0(best)	0.00512992637231946
FSIM - [0,1(best)]	0.6230707060790953
PSNR - [0.0,∞(best)]	45.797777210979035
UIQ - [-1,1(best)]	0.508555331389446

The chosen metrics calculate the difference in quality between the reference image and the newly generated images. The achieved scores show a concordance between the samples, although the changes of the output are rather obvious, the overall change is minor when assessed with the previous mentioned metrics. This is due to the scale of the image, the total surface area of the aluminium plate in relation to the changed angles or added scratches is quite large. For a human observer, the changes are more striking. Furthermore, evaluation of the images show that high-level or less complex metrics like RMSE and SSIM score significantly higher and achieve better results as come complex metrics like the considered FSIM and UIQ metrics. With such subtle changes pixel comparison metrics where pixel values are compared, the added changes are only within a few pixels leading to high results, the same can be said for SSIM where a sub-sample of the image is used for evaluating the score. Depending on the chosen parameters like sample size the achieved score is more likely to score closer the perfect score of 1, whereas the more complex metrics consider a lot more characteristics of the image to calculate the score and thus achieve a more desirable and realistic score. Benchmarking the metrics performance regarding run time can be neglected for our particular use case. The reference image as well as the generated images are fairly simple in structure and contrast, luminance or saturation is constant across the images. Thus even more complex metrics calculate the score in an average time of 45 seconds, with RMSE and SSIM scoring nearly instant results. Within the chosen domain the recommended metrics are FSIM or UIQ, showing a better representation of the changes and augmentation.

7. Conclusion

This paper proposes an augmentation process for adding new scratches and changing the angle of scratches using Pix2Pix's style transfer. Pix2Pix is able to learn a distribution managing to add scratches onto aluminium plates or change their angle. The scope of this work was defined using the domain of Industry 4.0 and metalworking industry to show the novelty of the approach by generating new images based on real input utilizable for augmentation and

further processing. Furthermore, an explanation for the process of identifying and define images for augmentation was given. Pix2Pix's ability to learn the distribution for transformation even if the mapping is not completely identical can be further utilized for further augmentation steps not present in this work. For example, further experiments showed satisfying results to make scratches thicker, longer, shorter or a combination of the trained scenarios. In addition, we evaluated the generated images using quality assessment metrics to show which metrics perform best in our chosen scenario.

In our future work, we will train Pix2Pix to learn additional augmentation steps for scratches, evaluate the results by comparing accuracy of CNNs. Futhermore, we will explore Cycle-GAN as a another way of augmentation an compare the results between Pix2Pix and Cycle-GAN.

References

[1] S. Rajcomar, A. W. Pillay, and E. Jembere, "Paired Augmentation for Improved Image Classification using Neural Network Models," *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020*, dec 2020.

[2] I. Konovalenko, P. Maruschak, V. Brevus, and O. Prentkovskis, "metals Recognition of Scratches and Abrasions on Metal Surfaces Using a Classifier Based on a Convolutional Neural Network," 2021. [Online]. Available: https://doi.org/10.3390/met11040549

[3] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2016, [Online] http://arxiv.org/abs/1511.06434 [retrieved: 05-2020].

[4] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June. IEEE Computer Society, jun 2019, pp. 4396–4405.

[5] R. Arandjelović and A. Zisserman, "Object Discovery with a Copy-Pasting GAN," may 2019, [Online] http://arxiv.org/abs/1905.11369 [retrieved: 05-2020].

[6] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2017-Octob. Institute of Electrical and Electronics Engineers Inc., dec 2017, pp. 2242–2251.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976.

[8] Y.-C. Huang, K.-C. Hung, C.-C. Liu, T.-H. Chuang, and S.-J. Chiou, "Customized Convolutional Neural Networks Technology for Machined Product Inspection," *Applied Sciences 2022, Vol. 12, Page 3014*, vol. 12, no. 6, p. 3014, mar 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/6/3014/htmhttps://www.mdpi.com/2076-3417/12/6/3014

[9] M. Hammami, D. Friboulet, and R. Kechichian, "Cycle GAN-Based Data Augmentation for Multi-Organ Detection in CT Images Via Yolo," *Proceedings - International Conference on Image Processing, ICIP*, vol. 2020-October, pp. 390–393, oct 2020.

[10] W. Zhang, Y. Kinoshita, and H. Kiya, "Image-Enhancement-Based Data Augmentation for Improving Deep Learning in Image Classification Problem," *2020 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-Taiwan 2020*, sep 2020.

[11] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, dec 2019.

[12] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random Erasing Data Augmentation," *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, pp. 13 001–13 008, aug 2017. [Online]. Available: https://arxiv.org/abs/1708.04896v2

[13] H. Inoue, "Data Augmentation by Pairing Samples for Images Classification," jan 2018. [Online]. Available: https://arxiv.org/abs/1801.02929v2

[14] X. Zheng, T. Chalasani, K. Ghosal, S. Lutz, and A. Smolic, "STaDA: Style Transfer as Data Augmentation," *VISIGRAPP 2019 - Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 4, pp. 107–114, sep 2019. [Online]. Available: http://arxiv.org/abs/1909.01056http://dx.doi.org/10.5220/0007353401070114

[15] A. Distante and C. Distante, *Handbook of image processing and computer vision: Volume 1: From energy to image*, 2020. [Online]. Available: https://www.springer.com/de/book/9783030381479

[16] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[17] Zhou Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.

[18] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.