

Segmentation and Removal of Markings in Metal Inspection Images

Luciane Soares*, Eduardo Evangelista*, Vinicius Maurente*, Matheus M. Dos Santos*, Igor Maurell*,
Marcelo Pias*, Paulo Drews-Jr* and Silvia S. C. Botelho*

*Center for Computational Science - C3

Universidade Federal do Rio Grande - FURG, Rio Grande, Brazil
Emails: lucianebaldassari, marcelopias, paulodrews, silvia@furg.br

Abstract—The inspection process of metallic surfaces, especially FPSO tanks, is still heavily reliant on manual methods, requiring long production downtime and posing health risks to inspectors. Automating this analysis step will provide significant benefits to the management of these vessels' integrity, reducing expenses, downtime, and, most importantly, the exposure time of employees to hazards associated with inspection activities. During manual inspections, inspectors make annotations using paint, typically in white and yellow colors, directly on the tank walls, hindering the automation of the inspection process as it complicates the segmentation and identification of potential flaws on the tank wall using techniques such as neural network models. Recognizing this problem, this work presents a proposal for the identification and segmentation of these markings by segmenting them in the images, followed by the removal of the segmented markings using image texture-filling techniques.

I. INTRODUCTION

Steel has been the most widely used material in most industrial segments of the production of society's basic goods. In the last decades, considerable progress has occurred both in the manufacture of new ferrous alloys and non-ferrous alloys and in the development of new composite materials. On the other hand, given the breadth of use of carbon steel, the field of exposure to deterioration also occurs more widely [1]. The wider corrosive process brings attention to several topics in specific sectors of production. The corrosive process becomes part, directly or indirectly, of diverse professional life, starting with the design itself, the assemblies, the production operation, the inspection, and ending with the industrial effluent [2].

In the case of the petroleum industry, such as refineries, production platforms, and petrochemical plants, the study of corrosion processes has a greater place, considering that about 50% of material failures are related to corrosion. Thus, the process of gaining knowledge of both the principles of corrosion and of corrosion protection, as well as the rules of practical suitability, has been a challenge in the field of equipment engineering and systematic monitoring.

FPSO units - Floating Production Storage and Offloading, a type of vessel used for oil exploration, production, storage, and offloading through oil tankers, are particularly susceptible to corrosion, whether galvanic, chemical, or anaerobic. This susceptibility is due to factors such as long-term anchoring over the production basin, infrequent inspections while anchored, and the fact that many of these structures are often constructed

from old and naturally worn-out vessels. The corrosion of deep storage tanks that occurs beneath the oil sludge deposits is also a problem due to the pumping of seawater with the oil [3].

Inspection operations inside the tanks of FPSO units involve complex activities both from the logistical point of view and the exposure of human beings to the risk of the activity, bringing the high risk of accidents. Thus, tank inspections are an important aspect in the life cycle of FPSOs to the operational reliability of this equipment throughout the ship's useful life. The tank inspection process (ballast or cargo) consists of the complete emptying of the tank; primary cleaning (Crude Oil Washer) and cleaning of the tank bottom; atmospheric inertization; preparation of the area and assembly of structures for the inspectors (lighting, ventilation, and climbing ropes) and visual inspection and thickness measurement by an inspector-escalator.

This is considered an unhealthy activity, and the time required for a complete inspection is between ten to fifteen days. The inspection must be carried out in accordance with international standards that define precise aspects and physical points of the tank that must be checked and inspected in their physical integrity and are listed and described in the tank's Inspection Plan.

Image capture inside the tanks can be done manually, where the inspector-climber captures the image to be inspected, or autonomously using drones, for example. Depending on the assessment of the inspectors, the tank may undergo a maintenance cycle to address the identified anomalies or be released for operation. After the maintenance and repairs cycle [4], the tank undergoes a new local inspection of the treated areas [1], [5]. Therefore, automating this inspection process brings significant gains to the management of the integrity of these vessels, reducing costs, downtime, and, most importantly, the exposure hours of the employees involved to the risks associated with the inspection activity.

When trying to apply neural network models to the images of FPSO tanks to segment the corrosion, in order to automate part of the inspection process, another problem was found. Despite the cleaning performed on the tanks, the walls have paint markings (yellow and/or white) from previous inspections, which makes it difficult for the neural network models to learn the corrosion aspects, as can be seen in Figure 1.

For the inspection to occur in an efficient way, the work



Fig. 1. Sample image of an FPSO tank with markings of previous inspections.

proposed here aims to apply segmentation and inpainting techniques to remove the markings from the images. Thus, restoring the original shape of the images and then applying corrosion inspection methods.

II. LITERATURE REVIEW

This section describes the techniques that have inspired or been used in this work for ink mark removal from metal plate images. This removal can be achieved in two main steps: the identification and segmentation of the ink markings, and the removal of the segmented markings through image texture filling.

Just like the problem, which needs to be solved in two main steps, namely the identification and removal of the markings, the main techniques that underlie this work can be divided into two groups associated with each of these steps: segmentation and inpainting.

A. Segmentation

There are various image segmentation techniques and classical techniques such as threshold-based segmentation aim to segment elements in the image based on their color. Usually, the image is transformed into the hue, saturation, and value (HSV) color space to reduce the influence of lighting on the color tone of the image elements [6]–[8]. Another classical method involves detecting the edges of the markings in the image by exploiting contrast variations. The Canny method [9] is well-known for utilizing this approach, where contrast variation or gradient is identified using the Sobel operator. The main advantage of edge-based segmentation is its reduced dependence on the lighting and color of the image objects. Preprocessing techniques are often used in conjunction with classical segmentation methods, including the use of Gaussian filters to reduce image noise and morphological operations such as dilation and erosion to eliminate small segments resulting from noise. However, classical techniques lack generalization and adaptation capabilities for the various situations encountered in the real environment [8], [10]–[12].

Deep neural network-based segmentation methods, due to their models with hundreds of thousands of parameters, are

capable of learning and adapting to diverse situations and have greater generalization capabilities, often achieving more robust results than classical segmentation methods [13]. Networks such as U-Net [14], Mask R-CNN [15], [16], Yolo [17], Detectron2 [18] and SAM [19] are examples of object detection, segmentation, and semantic instance segmentation networks that have become popular due to their excellent results in various applications [20]–[22].

B. Inpainting

The inpainting process is the art of reconstructing images, either by recovering lost, deteriorated, or unwanted parts using information from surrounding areas [23]. In this work, we aim to recover the texture hidden beneath the ink markings in the images of metal plates.

In 2004, Alexandru Telea proposed a classical inpainting approach known as Telea [24]. The technique uses information from neighboring pixels to estimate the values of missing pixels and performs an iterative inpainting optimization. The TELEA approach has been widely used in various applications, such as unwanted object removal, restoration of damaged photographs, and reconstruction of lost or corrupted areas of interest in an image.

According to Yu et al. [25], there are several inpainting proposals based on deep neural network, such as Lama [26], Repaint [27], MAT [28], and ZITS [29]. However, for the SAM merging method [19], the LaMa network was chosen because it is a single-stage mask-oriented approach that can generalize repetitive structures and is trained using intensive random mask generation [26]. In the present work, segmentation, and inpainting refer to the removal of unwanted ink markings on metal plates in order to expand a corrosion image dataset.

III. METHODOLOGY

The aim of the article is to propose a method for removing ink markings from images of metal plates. The method consists of two stages: i) segmentation of the markings using classical techniques or neural network models, and ii) removal of the markings using classical techniques or neural network models. The proposed pipeline can be seen in Figure 2. We compare the performance of both neural network approaches and classical computer vision approaches.

A. Dataset

The dataset used consists of 258 randomly selected images from online repositories and search engines, such as Google Images. The dataset is composed of a variety of image sizes, which is why a resize is performed to make all images have a configuration of 1024×1024 pixels. Since the images lack ground truth segmentation for the ink markings, they were manually annotated using the VGG Image Annotator (VIA) software. To train the method using neural network models, 70% of the images were used for training, 15% for validation, and 15% for testing.

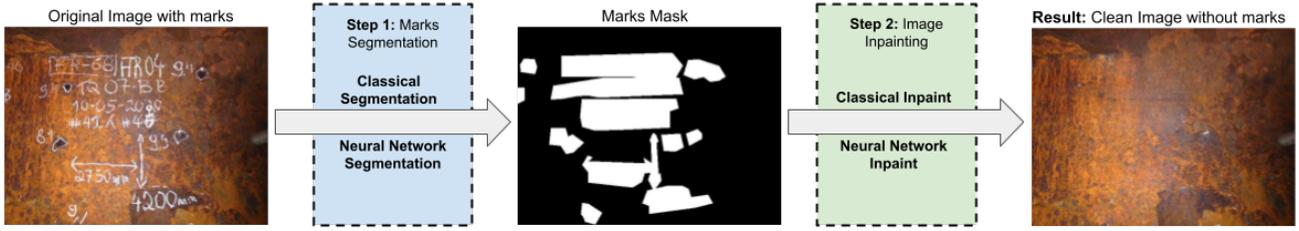


Fig. 2. Pipeline for markings removal.

B. Segmentation of the markings

Segmentation is employed to identify the paint markings in the tank images for subsequent removal. Two segmentation methods are evaluated, one based on classical techniques and the other based on neural networks.

1) *Using classical technique:* In the classical segmentation approach, the images are first converted to grayscale. Then, a Gaussian blur filter [30] is applied to smooth out details and reduce noise by applying a blur based on the Gaussian distribution, using a matrix size of 7×7 . Next, the Canny Edge Detection technique [9] is applied to detect edges. Morphological operations, such as dilation and erosion, are then performed to eliminate small segments [31]. Finally, the contours of the segments are identified and filled, resulting in a binary image mask of the identified segments [32].

2) *Using neural network models:* To perform the segmentation of paint markings in the images, two neural network models were used: Mask R-CNN [15] and Detectron2 [18].

The architecture of Mask R-CNN combines the object detection stages of R-CNN with instance segmentation capability, allowing for the identification of pixels corresponding to each detected object. The network consists of three main components: a convolutional network for feature extraction, a region of interest (RoI) pooling layer, and two subsequent branches—one for object detection and another for instance segmentation. An implementation of Mask R-CNN was used, where the model generates a bounding box and a mask segmentation for each instance. It is based on the Feature Pyramid Network (FPN) and utilizes the ResNet101 backbone. The architecture of the network can be seen in Figure 3.

Detectron2 is an open-source library developed by Facebook AI Research (FAIR) for object detection, semantic segmentation, and other computer vision tasks [18]. It utilizes a convolutional neural network architecture to extract relevant features from images.

Regions of interest (ROI) are classified using algorithms such as Fast R-CNN, Faster R-CNN, or Mask R-CNN. These algorithms employ a convolutional neural network to process the image and generate region proposals. Fast R-CNN and Faster R-CNN algorithms use these region proposals to classify and localize objects within the image, while Mask R-CNN extends their functionality by also predicting pixel-level segmentation masks for each object.

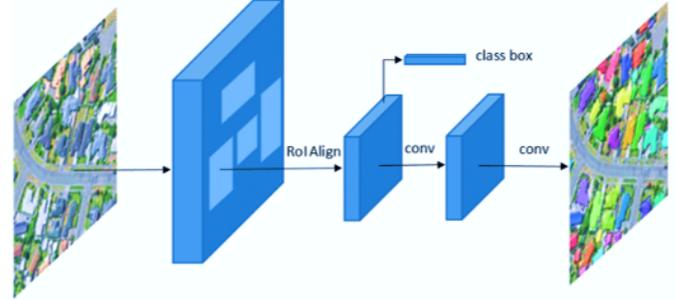


Fig. 3. Mask R-CNN neural network architecture, adapted of [33].

C. Removal of image markings

For the removal of markings, an inpainting technique was employed [34]. This technique aims to restore or reconstruct parts of an image based on the background information of the image [23].

1) *Using classical techniques:* In the process of removing paint markings through classical methods, the method proposed by [24] was applied using the OpenCV library. The method utilizes the mask obtained in the segmentation step and the original tank image as inputs. With the mask outlining the undesired areas, the method fills in the indicated parts of the image, using information from neighboring pixels to estimate the values of the missing pixels in the area to be filled.

2) *Using neural network methods:* Most texture inpainting methods can only fill small areas and struggle to perform inpainting on larger areas. Therefore, the Resolution-robust Large Mask Inpainting with Fourier Convolutions (LaMa) [35] was employed.

LaMa stands out for its ability to effectively fill large mask areas, surpassing the limitations of previous methods. The utilization of fast Fourier convolutions (FFCs) allows LaMa to consider the global context of the image from the early layers of the network. This feature is crucial for inpainting high-resolution images.

Additionally, a dilation process with fifteen iterations is applied to the masks obtained in the segmentation step. This approach is adopted to expand the inpainting area and prevent residual markings at the edges of the segments from interfering with the inpainting process of the image. Inference on the images is performed using the LaMa network model with pre-trained weights from the original model.

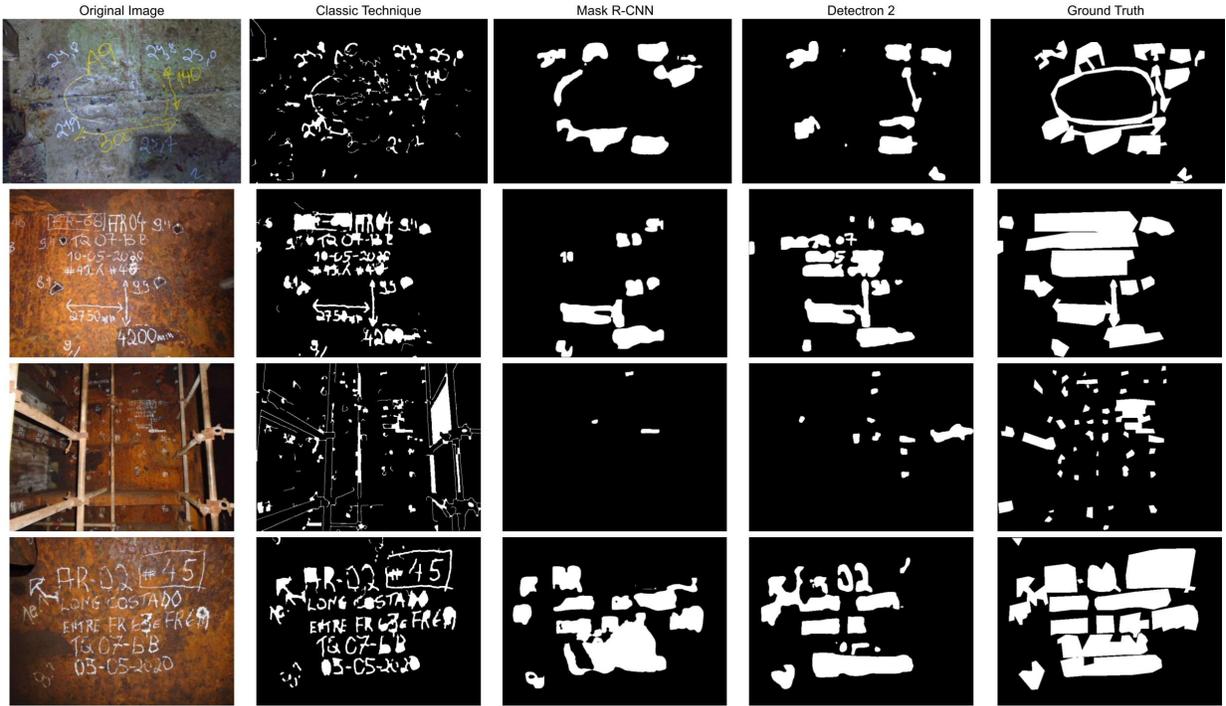


Fig. 4. Comparison of classical segmentation and neural network segmentation. The first column displays the original images, the second column shows the result of classical segmentation, the third column shows the result of segmentation using the Mask R-CNN neural network model, the fourth column shows the segmentation with the Detectron 2 neural network model, and the last column shows the manual segmentation, which serves as the ground truth.

IV. EXPERIMENTS AND RESULTS

The presented results are divided into the segmentation of the markings to be removed from the images, using classical techniques and neural networks, and the restoration of images using classical and neural network-based inpainting methods. Finally, a comparison of the methods is provided.

A. Segmentation Results

In the segmentation step, both classical segmentation and neural network-based segmentation approaches were evaluated.

The training of the neural network methods was carried out using specific parameter configurations. These configurations included a total of 3000 steps, distributed across 30 epochs, implying that the neural network would iterate through the entire dataset 30 times during training, with 100 steps for each mini-batch of data. A batch size of 1 was selected due to memory restrictions of the Graphical Processing Unit (GPU), indicating that individual data samples were processed independently. The learning rate used for training was set at 0.001.

The classical segmentation method utilized the operations described in Section III-B1. The Canny edge detector was configured with a primary threshold of 80 and a secondary threshold of 140. To remove segmentation noise, morphological operations of dilation and erosion with a 3x3 kernel and two iterations were applied.

Figure 4 presents a comparison of the segmentation results. It can be observed that the results using neural networks were superior in quality for both images.

The metric for comparing the effectiveness of the presented segmentation methods is the intersection over union (IoU) distribution, according to Equation 1.

$$IoU = \frac{AreaOfOverlap}{AreaOfUnion} \quad (1)$$

When comparing the ground truth mask with the results obtained from classical methods, the metric showed an average of 27%. However, the results obtained using neural network models were 40% with Mask RCNN and 76% with Detectron2.

B. Results of image restoration using inpainting

Figure 5 presents the original image, a comparison of the segmentation performed by the Detectron2 model, and the ground truth of the original image. Lastly, the result of the LaMa inpainting method in reconstructing the image is shown. It can be observed that the better the segmentation result, the better the inpainting result.

Figure 6, on the other hand, demonstrates the result of the segmentation of paint markings performed by the classical technique and after inpainting performed by the classical TELEA method and the LaMa neural network model.

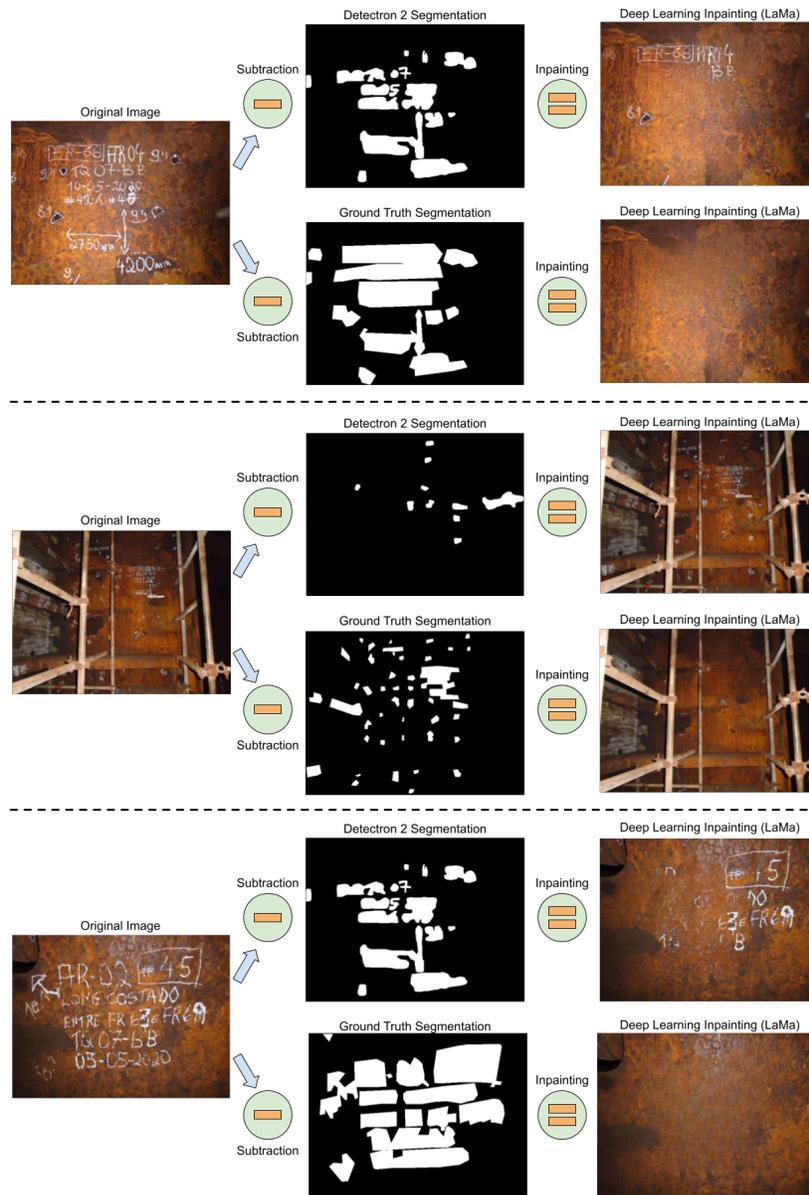


Fig. 5. Comparison of deep learning inpainting results with LaMa, using ground truth segmentation versus deep learning segmentation Detectron 2.

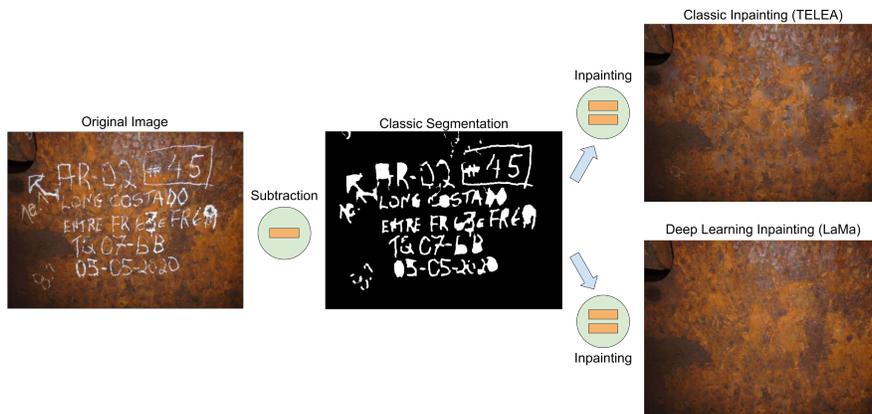


Fig. 6. Comparison of Telea's classical inpainting method and LaMa's deep learning inpainting method utilizing classical segmentation.

V. CONCLUSION

The article proposed the segmentation and recovery of images using classical segmentation and inpainting techniques, as well as the application of a neural network model, to remove paint markings that hinder the inspection of the images.

The results obtained so far demonstrate the effectiveness of the proposed method in removing image markings and preparing the data for training neural networks focused on automatic corrosion identification. Particularly noteworthy are the results of the paint marking segmentation methods using the Detectron2 network, achieving an IoU of 76% compared to ground truth. As for the restoration techniques using inpainting, the results are qualitative, as shown in Figure 6.

The results reinforce the relevance of this research in the field of developing techniques for analyzing corrosion on image inspection task and the importance of future work to improve the segmentation methods of paint markings for achieving increasingly accurate restoration.

ACKNOWLEDGMENT

The authors would like to thank the Brazilian National Agency of Petroleum, Natural Gas and Biofuels (ANP) for their financial support through the Human Resources Program of ANP to the Petroleum and Gas Sector - PRH-ANP. The authors also would like to thank National Council for Scientific and Technological Development – CNPq.

REFERENCES

- [1] L. Soares, S. Botelho, R. Rodrigues, and P. Drews-Jr, "A visual inspection proposal to identify corrosion levels in marine vessels using a deep neural network," in *Latin American Robotics Symposium (LARS)*. IEEE, 2021, pp. 222–227.
- [2] R. Terzi and F. B. Mainier, "Monitoramento da corrosão interna em plataformas offshore," *Tecno-Lógica*, vol. 12, no. 1, pp. 14–21, 2008.
- [3] L. L. Menegaldo, G. A. N. Ferreira, M. F. Santos, and R. S. Guerato, "Development and navigation of a mobile robot for floating production storage and offloading ship hull inspection," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 9, pp. 3717–3722, 2009.
- [4] L. Soares, A. Weis, R. Rodrigues, P. Drews-Jr, B. Guterres, S. Botelho, and N. Filho, "Seam tracking and welding bead geometry analysis for autonomous welding robot," in *Latin American Robotics Symposium (LARS)*, 2017, pp. 1–6.
- [5] L. B. Soares, H. Costa, S. Botelho, D. Souza, R. Rodrigues, and P. Drews-Jr, "A robotic passive vision system for texture analysis in weld beads," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 45, no. 1, p. 44, 2023.
- [6] C. Steffens, P. L. J. Drews, and S. S. Botelho, "Deep learning based exposure correction for image exposure correction with application in computer vision for robotics," in *Latin American Robotics Symposium (LARS)*. IEEE, 2018, pp. 194–200.
- [7] C. Steffens, V. Huttner, L. Messias, P. Drews-Jr, S. Botelho, and R. da Guerra, "Cnn-based luminance and color correction for ill-exposed images," in *IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 3252–3256.
- [8] C. R. Steffens, L. R. Messias, P. J. Drews-Jr, and S. S. d. C. Botelho, "Cnn based image restoration: Adjusting ill-exposed srgb images in post-processing," *Journal of Intelligent & Robotic Systems*, vol. 99, pp. 609–627, 2020.
- [9] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.
- [10] C. Steffens, L. Messias, P. Drews-Jr, and S. da Botelho, "Can exposure, noise and compression affect image recognition? an assessment of the impacts on state-of-the-art convnets," in *Latin American Robotics Symposium (LARS)*, 2019, pp. 61–66.
- [11] C. Steffens, L. Messias, P. Drews-Jr, and S. Botelho, "On robustness of robotic and autonomous systems perception: An assessment of image distortion on state-of-the-art robotic vision model," *Journal of Intelligent & Robotic Systems*, vol. 101, pp. 1–17, 2021.
- [12] L. Messias, P. Drews-Jr, and S. Botelho, "Attention-based neural network for ill-exposed image correction," in *IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 3913–3917.
- [13] P. Bharati and A. Pramanik, "Deep learning techniques—r-cnn to mask r-cnn: a survey," *Computational Intelligence in Pattern Recognition*, pp. 657–668, 2020.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.
- [15] W. Abdulla, "Mask r-cnn for object detection and instance segmentation on keras and tensorflow," https://github.com/matterport/Mask_RCNN, 2017.
- [16] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *IEEE ICCV*, 2017, pp. 2961–2969.
- [17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *IEEE CVPR*, 2016, pp. 779–788.
- [18] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2," <https://github.com/facebookresearch/detectron2>, 2019.
- [19] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment anything," 2023.
- [20] P. Ganesh, K. Volle, T. Burks, and S. Mehta, "Deep orange: Mask r-cnn based orange detection and segmentation," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 70–75, 2019.
- [21] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE TPAMI*, vol. 44, no. 7, pp. 3523–3542, 2021.
- [22] Ö. Kaya, M. Y. Çodur, and E. Mustafaraj, "Automatic detection of pedestrian crosswalk with faster r-cnn and yolov7," *Buildings*, vol. 13, no. 4, p. 1070, 2023.
- [23] B. H. Patil and P. Patil, "A comprehensive review on state-of-the-art image inpainting techniques," *Scalable Computing: Practice and Experience*, vol. 21, no. 2, pp. 265–276, 2020.
- [24] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of graphics tools*, vol. 9, no. 1, pp. 23–34, 2004.
- [25] T. Yu, R. Feng, R. Feng, J. Liu, X. Jin, W. Zeng, and Z. Chen, "Inpaint anything: Segment anything meets image inpainting," *arXiv preprint arXiv:2304.06790*, 2023.
- [26] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," in *IEEE/CVF WACV*, 2022, pp. 2149–2159.
- [27] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, "Repaint: Inpainting using denoising diffusion probabilistic models," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 11 461–11 471.
- [28] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang, and J. Jia, "Mat: Mask-aware transformer for large hole image inpainting," in *IEEE/CVF CVPR*, 2022, pp. 10 758–10 768.
- [29] Q. Dong, C. Cao, and Y. Fu, "Incremental transformer structure enhanced image inpainting with masking positional encoding," in *IEEE/CVF CVPR*, 2022, pp. 11 358–11 368.
- [30] R. A. Hummel, B. Kimia, and S. W. Zucker, "Deblurring gaussian blur," *Computer Vision, Graphics, and Image Processing*, vol. 38, no. 1, pp. 66–80, 1987.
- [31] P. Soille and P. Soille, "Erosion and dilation," *Morphological Image Analysis: Principles and Applications*, pp. 63–103, 2004.
- [32] G. Bradski, A. Kaehler et al., "Opencv," *Dr. Dobb's journal of software tools*, vol. 3, no. 2, 2000.
- [33] D. A. J. Quispe and J. Sulla-Torres, "Automatic building change detection on aerial images using convolutional neural networks and handcrafted features," *IJACSA*, vol. 11, no. 6, 2020.
- [34] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *IEEE CVPR*, 2018, pp. 8759–8768.
- [35] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, N. Kong, H. Goka, K. Park, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," in *IEEE/CVF WACV*, 2022, pp. 2149–2159.