Autonomous Robotic Inspection based on Active Vision and Deep Reinforcement Learning

WEN TANG and MOHAMMAD R. JAHANSHAHI

ABSTRACT

Over the years, various methods based on computer vision have been proposed for the problem of damage detection and segmentation. Recently, remarkable progress has been made in this field owing to the emergence of deep neural networks. However, there is a main assumption among these works that the data collection (e.g., taking photos) is usually carried out by human inspectors, so there is little occlusion or bad lighting conditions. With that being said, the uncertainties that could occur in data collection are handled manually to ensure the dataset is clean and free of confusing or occluded damages. This assumption limits the applicability of autonomous robotic inspection to real-world settings due to the uncertainties in data collection and data interpretation. To bridge this gap, this study integrates the concept of active perception into damage detection and proposes a framework based on the Partially Observable Markov Decision Process (POMDP) and Deep Reinforcement Learning (DRL). The proposed framework facilitates the learning process for robotic agents to explore the 3D environment and intelligently select informative viewpoints to reduce uncertainty and minimize confusion, which leads to more reliable decision-making. Besides uncertainty reduction, the DRL agent can also inspect the workspace more efficiently compared with traditional raster scanning. The trained DRL agent is evaluated for the autonomous assessment of cracks on metallic surfaces. Results show that the agent equipped with the active perception module outperforms the raster scanning inspection by 57% in terms of crack IoU. In addition, the DRL agent can reduce the total inspection time by two times while the prediction accuracy is on par with the raster scanning.

INTRODUCTION

Structural defect detection is an important aspect of structural health monitoring (SHM). In-time detection of cracks provides important information regarding the condition of the structure that can prevent destructive events from happening. With the success of deep learning and computer vision, many models can detect and segment cracks in images reasonably well [1–4]. However, existing models are mostly passive visual in-

Wen Tang, PhD Student, Email: tang493@purdue.edu. Smart Informatix Laboratory (SIL), Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA.

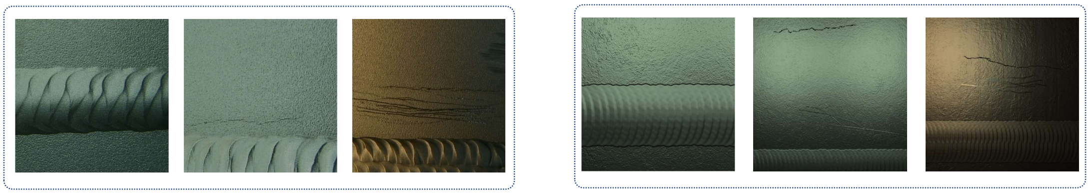
spection systems that typically fail to detect the damage or produce excessive amounts of false positives due to bad lighting conditions or viewing angles. Whereas a human inspector has the ability to move in the 3D environments and actively control the viewpoint to gain a better interpretation of the damage. The problem with the passive visual system is that it treats the image as independent and identically distributed (i.i.d.) data points, which ignores the potential correlation between the images and throws away important information like the spatio-temporal consistency in the underlying 3D environment. To overcome the above-mentioned issue, a deep reinforcement learning (DRL) agent that leverages the power of active perception is proposed in this study. The proposed DRL agent exhibits the capability to systematically explore the environment while selectively directing its attention toward viewpoints that provide more informative cues. To the authors' knowledge, this is the first work in the field of structural health monitoring that combines damage detection and active vision, which is one step closer to the realization of fully autonomous robotic inspection systems.

The main contribution of this work can be summarized as follows:

- Introduce the concept of active damage detection by defining a new task, called Active Damage Segmentation, where an agent can move in the 3D environment to perform damage segmentation on a metallic surface.
- To tackle the active damage segmentation task, a DRL model is proposed to select informative viewpoints to improve the prediction.
- An interactive photo-realistic 3D simulator based on computer graphics is built to train the DRL model.
- Results show that the proposed RL agent leveraging active perception consistently outperforms the passive visual system. Moreover, the learned behavior leads to much more efficient data collection schemes as opposed to raster scanning.

SIMULATION ENVIRONMENT AND DATASET

To train and evaluate the DRL model, photo-realistic steel plates with cracks, welding seams, and scratches are rendered in a computer graphic tool (Houdini) with texture and lighting conditions similar to the metallic surfaces that have been inspected in real-life. Similar to the field scenario, the camera and the light source are mounted on the robotic arm so the light condition changes frequently as the robotic agent moves around in the 3D scenes. To this end, images are rendered from simulated robotic inspections of 20 damaged metallic surfaces to create a dataset for the training of RL agents. In each simulated inspection, a dense raster scanning is performed to make sure every part of the surface is inspected with various lighting conditions. Each synthetic image is associated with annotations of scratch and crack, which can be extracted automatically from the scene. Out of the 20 scenes, 10 of the scenes are used for training and the rest of the scenes are used for testing. Figure 1 shows some examples of the real images captured during a field inspection of the metallic surface of a nuclear power plant reactor and the images rendered from the simulation environments.



(a) Real images of metallic surfaces (b) Synthetic images of metallic surfaces

Figure 1. Comparison between images captured from field inspections and images rendered from the simulation environment.

ACTIVE DAMAGE SEGMENTATION TASK

To better implement the concept of active damage detection, a new task called Active Damage Segmentation is introduced in this study. In this task, 10 unseen metallic surfaces in the simulation environment are used as the test set. The goal is to segment out all the cracks and reduce the false positives as much as possible. The steel surfaces in the simulation environment are carefully constructed to ensure a close resemblance to the steel surfaces being inspected in the field. The area to be inspected is $219 \times 153\, mm^2$, the camera moving speed is $25\, mm/s$, and the width of the crack varies from 0.1 to $0.5\, mm$. To cover the entire metallic surface, the agent starts with regular raster scanning to ensure that every corner of the surface is inspected. When the current observation meets the criteria of an ambiguous frame, the agent switches to active perception mode. In this study, the frame is deemed to be "ambiguous" if the count of pixels with Softmax scores above 0.6 exceeds 200 in the predicted mask of the current frame. It is important to note that more advanced and complex criteria, such as uncertainty quantification (UQ) based on Bayesian networks or other techniques, can also be employed to determine if a frame is ambiguous or not. After the activation of the "active perception" mode, the frame that triggers the mode becomes the initial frame $I_{t=1}$ for the interactive process of active perception. The active perception loop is terminated once the agent chooses the $Terminate$ action or the time horizon allotted for the episode is exceeded. The agent will resume the raster scanning pattern after the termination of the "active perception" mode until it encounters the next frame that activates active perception.

PROBLEM FORMULATION

The goal of active damage segmentation is to propose a sequence of actions (viewpoints) and acquire useful new information to enhance the initial prediction mask $M_{t=1}$ of the first frame $I_{t=1}$. By aggregating information from new viewpoints, a fused mask $M_{t=T}$ is generated at the final time step, which serves as the final prediction mask for the first frame $I_{t=1}$. To solve t his task, an approach based on the Partial Observable Markov Decision Process (POMDP) can be adopted.

A discrete-time POMDP is defined as a tuple $\{S, A, \mathcal{T}, R, \Omega, O\}$, where $S = \{s_1, s_2, ..., s_n\}$ is a set of partially observable states of the world, $A = \{a_1, a_2, ..., a_m\}$ is a set of actions available to the agent, $\mathcal{T}$ is a set of conditional transition probabilities from state s to state s': $P(s'|s, a)$, $R : S \times A \to \mathbb{R}$ is the reward function,

$\Omega = \{o_1, o_2, ..., o_k\}$ is a set of observations, and O is a set of observation probabilities $O(o|s', a)$ conditioned on the reached state and the action taken.

At each time step, the environment is in some unknown state $s \in S$. The agent chooses an action $a \in A$, which causes the environment to transit to state $s' \in S$ with probability $\mathcal{T}(s'|s, a)$. At the same time, the agent receives an observation $o \in \Omega$ that depends on the new state s' with probability $O(o|s', a)$. Finally, the agent receive a reward $r \in R(s, a)$. This process repeats until it terminates in an episodic setup. Let τ be the trajectory that contains a sequence of (o_t, a_t, r_t), where $a_t \sim \pi(\cdot|o_t)$, and $S_{t+1} \sim \mathcal{T}(S_t, a_t)$. Given a discount factor γ, the optimal policy π^* can be expressed as follow:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{\tau\sim\pi}[R_T],\ where\ R_T = \sum_{t=1}^{T} \gamma^{t-1} r_t \tag{1}$$

The objective is to find a policy π that maximizes the discounted accumulative return R_T over an episode. One technique to find such policy π is Proximal Policy Optimization (PPO) [5], which is an on-policy algorithm belonging to the gradient-policy family.

ACTIVE PERCEPTION MODEL

To tackle the active damage segmentation task, a DRL model is proposed. The observations space of the ACS-DRL model consists of RGB images of size 448×448, the current viewpoint $C_t \in \{0, 1\}^{N\times N}$, visited viewpoints $V_t \in \{0, 1\}^{N\times N}$, and fused crack mask $M_t \in [0, 1]^{448\times448}$ with $N = 3 \times 448$. Each pixel in C_t is 0 or 1, indicating whether the corresponding area is covered by the current viewpoint, and each pixel in V_t indicates if the corresponding area has been visited. Note that the centers of C_t, V_t, and M_t share the same global coordinate as the center of $I_{t=1}$. The action space A is comprised of a set of discrete 7×7 viewpoints around the current viewpoint of the agent and an additional $Terminate$ action. The reward r_t is designed as follows:

$$r_t = \begin{cases} IoU(M_t) - IoU(M_{t-1}) - cost_{a_t}, & \text{if } case\,1 \\ +\alpha, & \text{if } case\,2 \\ -\alpha, & otherwise \end{cases} \tag{2}$$

In Eq.(2), $case\,1$ refers to scenarios where $a_t \neq Terminate$; $case\,2$ refers to scenarios where $a_t = Terminate, FP(M_t) \leq \beta * FP(M_{t=1}), Recall(M_t) \geq \eta$; and all other cases fall into the $otherwise$ case. The values of $\alpha, \beta, \eta, andcost_{a_t}$ are assigned as follows: $\alpha = 0.5, \beta = 0.9, \eta = 0.9$, and $cost_{a_t} = 0.01$. The fused mask M_t averages the softmax score of the overlapping area between current prediction P_t and the previous fused prediction mask M_{t-1}.

The detailed architecture of the DRL agent is shown in Fig. 2. The proposed model consists of two modules, a crack segmentation network and a policy network. The model generates a fused crack segmentation mask and outputs the next best view based on the current RGB image, the current viewpoint, and the visited viewpoints. At each time step from $t = 1$ to $t = T$, given an observed RGB image I_t, the agent predicts a damage mask P_t, which may or may not be right. Then the agent takes action a_t that is specified by the policy network π_θ. The fused mask M_{t-1} is defined as $M_{t=t} = f_{agg}(M_{t-1}, P_t)$ where

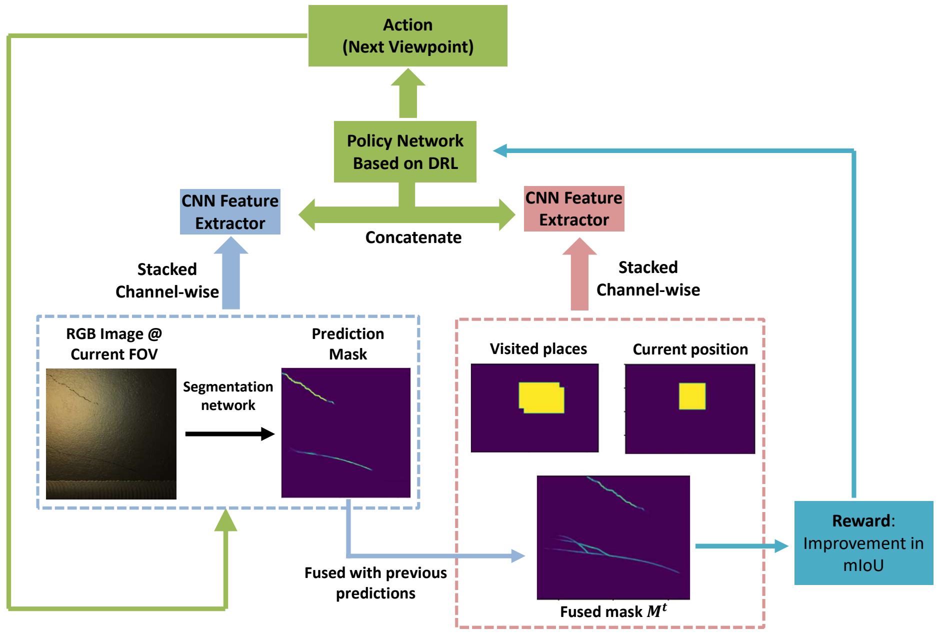


Figure 2. The network architecture of the DRL model for active damage segmentation.

$f_{agg}(\cdot)$ is a function that takes in previous fused mask M_{t-1} and current prediction P_t and fuses them together to generate fused mask M_t. A simple average function is employed in this study. However, more sophisticated functions, such as Bayesian update, could be utilized as alternative aggregation functions.

The segmentation network predicts the damage mask P_t given a single RGB image I_t at each time step. The architecture of the segmentation network is based on U-Net++ [6], which utilizes the ResNet-101 [7] model pre-trained on ImageNet as the backbone. The segmentation network is fine-tuned on an online crack dataset [8] to identify cracks on metallic surfaces. It should be noted that the segmentation network is not trained on images from the simulation environment, as it can easily overfit to the simulation environment, which hinders it from learning any useful viewpoint selection policies. This study emphasizes the importance of selecting informative viewpoints and observing the same area from different perspectives to correct the final predictions. Therefore, it is reasonable to start with an imperfect segmentation network. The weights of the segmentation network are frozen after the fine-tuning and during the training of the DRL agent.

The policy network is an actor-critic style network that takes in the current RGB frame T_t, current prediction P_t, current viewpoint C_t, visited viewpoints V_{t-1}, and fused mask M_{t-1}, and outputs a probability distribution over the action space A. During test time, the action with the largest probability is taken by the agent. The CNN feature extractors in Figure 2 uses EfficientNet-B0 to extract the features and form the state embeddings of the DRL agent. The embeddings are then concatenated and fed into two separate branches called actor-network and action-network, which follow the convention of the actor-critic network [9]. The actor-network and the critic-network employs the same architecture, which contains a Gated Recurrent Unit (GRU) layer followed by two fully connected layers. The Generalized Advantage Estimation (GAE) is used to stabilize the variance of the expected rewards. The entire policy network is trained using the PPO algorithm.

EXPERIMENTS AND RESULT ANALYSIS

To evaluate the performance of the DRL agent, the proposed method is compared with raster scanning on 10 unseen metallic surfaces. To inspect the metallic surfaces, the agent starts with regular raster scanning and switches to active perception whenever the count of pixels with Softmax scores above 0.6 exceeds 200 in the current frame. The frame that activates the active perception serves as the starting point $I_{t=1}$ for the active perception process. Once the $Termination$ action is selected, the agent resumes raster scanning until it encounters the next frame that activates active perception.

TABLE I. Quantitative comparison between raster scanning and active damage segmentation at different overlap ratios

	Pure Raster Scanning				**Active Damage Segmentation**			
Frames Overlap	**Crack IoU**	**mIoU**	**F1 Score**	**FP Cracks**	**Crack IoU**	**mIoU**	**F1 Score**	**FP Cracks**
None	0.2629	0.6315	0.4164	4	0.4129	0.7110	0.5935	0
25%	0.3648	0.6824	0.5346	7	0.5230	0.7615	0.6868	1
45%	0.4026	0.7013	0.5741	7	0.5257	0.7629	0.6891	1
63%	0.4618	0.7309	0.6318	5	0.5763	0.7882	0.7312	0
81%	0.4611	0.7306	0.6312	6	**0.5770**	**0.7885**	**0.7317**	0

The performance of the DRL model is evaluated using the above-mentioned active damage segmentation task. The agent inspects the entire surface, and the mIoU score of the predicted mask, the IoU score of the crack, the count of false positive crack instances (FP Cracks column in Table I), and the duration of inspection are reported. If the number of connected pixels with softmax scores higher than a certain threshold is more than 1000, then the blob on the predicted segmentation mask is determined as a crack instance.

As shown in Table I, the active damage segmentation framework can achieve a notable improvement in the crack IoU, with an increase of up to 57% compared to raster scanning. Furthermore, the proposed method consistently outperforms the raster scanning approach across all other cases. The proposed method also shows promising results in terms of data collection efficiency. A comparison between the 25% overlapping case in active damage segmentation and the 81% overlapping case in raster scanning reveals that the proposed method is able to perform a rapid inspection that reduces the total inspection time by more than two times while yielding a 12% higher crack IoU. The count of false positive crack instances is also reported in the ”FP Cracks” column in Table I. It can be observed that the false positive crack instances are reduced to very low numbers when DRL models are used, which is another strong support for formulating the damage detection tasks as active perception problems. Table II shows the detailed breakdown of

TABLE II. Detailed breakdown of total inspection time

Frames Overlap	Pure Raster Scanning None	25%	45%	63%	81%	Active Damage Segmentation None	25%	45%	63%	81%
Total Time (Sec)	23.2	31.8	47.0	63.7	101.9	33.6	48.0	69.1	106.5	138.9
Data Coll. (Sec)	23.2	31.8	47.0	63.7	101.9	28.8	42.0	62.3	94.9	110.9
Computation (Sec)	-	-	-	-	-	4.8	6.0	6.8	11.6	28.0
Time Step	1.0	8.0	15.0	24.0	120.0	2.3	2.3	2.3	2.3	2.3

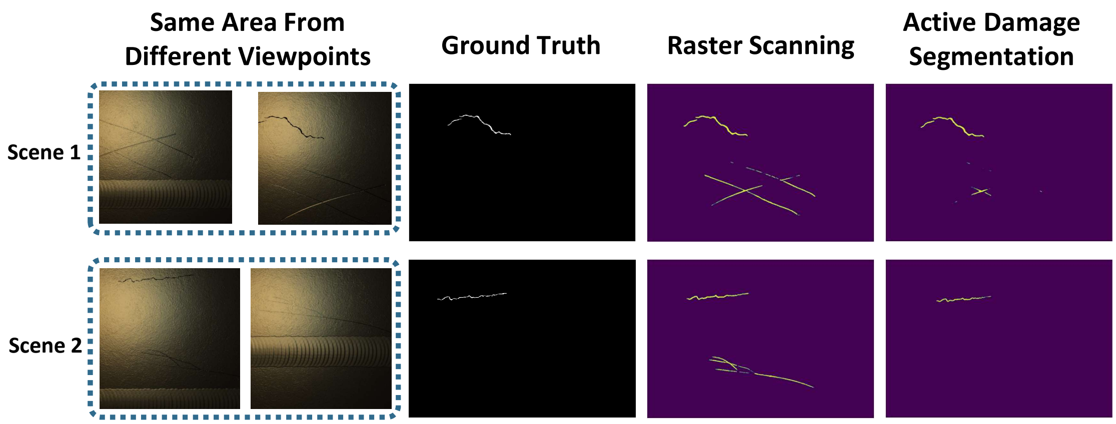


Figure 3. Sample predictions obtained from raster scanning and active damage segmentation.

the total inspection time of active damage segmentation in different cases. For conventional raster scanning, the "Time Step" row is equivalent to the number of frames that are fused with the initial frame and overlap with the initial frame, provided the viewpoints are discretized with an 81% overlap between the neighboring frames. It shows that the DRL agent learns to select fewer but more informative viewpoints to improve the mIoU of the crack segmentation mask. Figure 3 illustrates a set of samples comparing the prediction masks obtained through raster scanning and active damage detection. It can be observed that the crack prediction is accurately persevered, while a majority of the false positive scratches are eliminated when active perception is incorporated.

CONCLUSIONS

This study introduces a deep reinforcement-learning-based active vision model for damage detection. To train and evaluate the RL agent, photo-realistic synthetic 3D scenes were constructed, and a dataset was generated. The agent can move freely in the 3D scenes and improve the accuracy of the predicted masks by selecting informative viewpoints and fusing the information from those viewpoints. The DRL agent also learns to terminate the episode early, which leads to efficient data collection. Evaluation on metallic surfaces shows that the agent can increase the crack IoU by up to 57% when compared to pure raster scanning. Additionally, the agent can conduct a rapid inspection that reduces the overall inspection time by more than two times while achieving a 12% higher crack IoU than that of the dense raster scanning approach.

REFERENCES

1. Cha, Y., W. Choi, and O. Büyüköztürk. 2017. "Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks," *Computer-Aided Civil and Infrastructure Engineering*, 32.
2. Choi, W. and Y. Cha. 2020. "SDDNet: Real-Time Crack Segmentation," *IEEE Transactions on Industrial Electronics*, 67:8016–8025.
3. Liu, Y., J. Yao, X. Lu, R. Xie, and L. Li. 2019. "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation," *Neurocomputing*, 338:139–153.
4. Mondal, T. G., M. R. Jahanshahi, R.-T. Wu, and Z. Y. Wu. 2020. "Deep learning-based multi-class damage detection for autonomous post-disaster reconnaissance," *Structural Control and Health Monitoring*, 27.
5. Schulman, J., F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. 2017. "Proximal Policy Optimization Algorithms," *ArXiv*, abs/1707.06347.
6. Zhou, Z., M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang. 2018. "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support : 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, S...*, 11045:3–11.
7. He, K., X. Zhang, S. Ren, and J. Sun. 2015. "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*:770–778.
8. Bianchi, E. and M. H. Hebdon. 2022. "Development of Extendable Open-Source Structural Inspection Datasets," *J. Comput. Civ. Eng.*, 36.
9. Mnih, V., A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu. 2016. "Asynchronous methods for deep reinforcement learning," in *International conference on machine learning*, PMLR, pp. 1928–1937.