

CLAIMS

1. A method for evaluating damages to an article, comprising:
 - obtaining at least two images that are time sequentially related and show the article at different angles;
 - providing the at least two images as input to a detection model in time order, wherein the detection model comprises a deep convolutional neural network and a long short-term memory (LSTM) network that have been jointly trained on a set of training samples, each training sample comprising multiple training images associated with labels indicating respective article damage degrees of an article shown in the multiple training images;
 - processing the at least two images using the deep convolutional neural network to output a feature processing result for each image based on respective features identified from the image;
 - processing the feature processing result using the long short-term memory (LSTM) network receive the feature processing results and to output a damage detection result based on performing time series analysis on the feature processing results, wherein the damage detection result comprises a classification result of each of one or more types of damage; and
 - obtaining, as output from the detection model, the damage detection result.
2. The method of claim 1, wherein the long short-term memory (LSTM) network is an attention-based LSTM network.
3. The method of claim 1, wherein the at least two images that are time sequentially related and show the article at different angles comprise at least one of:
 - photos of the article in motion captured by a stationary camera, videos of the article in motion recorded by a stationary camera, photos of the article in a stationary state captured by a moving camera, or videos of the article in a stationary state recorded by a moving camera.
4. The method of claim 1, wherein the classification result of each of one or more types of damage comprises a binary variable indicating whether the article bears the type of damage.
5. The method of claim 1, wherein processing the at least two images using the deep

convolutional neural network to output the feature processing result comprises, for each image:

performing feature extraction, damage discovery, and feature fusion to determine the feature processing result.

6. The method of claim 1, wherein obtaining the at least two images comprises:
moving the article in accordance with a first set of movement instructions; or
moving a camera in accordance with a second set of movement instructions.

7. The method of claim 1, wherein obtaining the at least two images further comprises:
illuminating the article in accordance with a set of illumination instructions.

8. A non-transitory, computer-readable medium storing one or more instructions executable by a computer system to perform one or more operations for evaluating damages to an article, wherein the operations comprise:

obtaining at least two images that are time sequentially related and show the article at different angles;

providing the at least two images as input to a detection model in time order, wherein the detection model comprises a deep convolutional neural network and a long short-term memory (LSTM) network that have been jointly trained on a set of training samples, each training sample comprising multiple training images associated with labels indicating respective article damage degrees of an article shown in the multiple training images;

processing the at least two images using the deep convolutional neural network output a feature processing result for each image based on respective features identified from the image;

processing the feature processing result using the long short-term memory (LSTM) network to receive the feature processing results and to output a damage detection result based on performing time series analysis on the feature processing results, wherein the damage detection result comprises a classification result of each of one or more types of damage; and

obtaining, as output from the detection model, the damage detection result.

9. The non-transitory, computer-readable medium of claim 8, wherein the long short-term memory (LSTM) network is an attention-based LSTM network.

10. The non-transitory, computer-readable medium of claim 8, wherein the at least two images that are time sequentially related and show the article at different angles comprise at least one of:

photos of the article in motion captured by a stationary camera, videos of the article in motion recorded by a stationary camera, photos of the article in a stationary state captured by a moving camera, or videos of the article in a stationary state recorded by a moving camera.

11. The non-transitory, computer-readable medium of claim 8, wherein the classification result of each of one or more types of damage comprises a binary variable indicating whether the article bears the type of damage.

12. The non-transitory, computer-readable medium of claim 8, wherein processing the at least two images using deep convolutional neural network to output the feature processing result comprises, for each image:

performing feature extraction, damage discovery, and feature fusion to determine the feature processing result.

13. A computer-implemented system, comprising:

one or more computers; and

one or more computer memory devices interoperably coupled with the one or more computers and having tangible, non-transitory, machine-readable media storing one or more instructions that, when executed by the one or more computers, perform one or more operations for evaluating damages to an article, wherein the operations comprise:

obtaining at least two images that are time sequentially related and show the article at different angles;

providing the at least two images as input to a detection model in time order, wherein the detection model comprises a deep convolutional neural network and a long short-term memory (LSTM) network that have been jointly trained on a set of training samples, each training sample comprising multiple training images associated with labels indicating respective article damage degrees of an article shown in the multiple training images;

processing the at least two images using the deep convolutional neural network to output a feature processing result for each image based on respective features identified from the image;

processing the feature processing result using the long short-term memory (LSTM) network receive the feature processing results and to output a damage detection result based on performing time series analysis on the feature processing results, wherein the damage detection result comprises a classification result of each of one or more types of damage; and

obtaining, as output from the detection model, the damage detection result.

14. The computer-implemented system of claim 13, wherein the long short-term memory (LSTM) network is an attention-based LSTM network.

15. The computer-implemented system of claim 13, wherein the at least two images that are time sequentially related and show the article at different angles comprise at least one of:

photos of the article in motion captured by a stationary camera, videos of the article in motion recorded by a stationary camera, photos of the article in a stationary state captured by a moving camera, or videos of the article in a stationary state recorded by a moving camera.

16. The computer-implemented system of claim 13, wherein the classification result of each of one or more types of damage comprises a binary variable indicating whether the article bears the type of damage.

17. The computer-implemented system of claim 13, wherein processing the at least two images using the deep convolutional neural network output the feature processing result comprises, for each image:

performing feature extraction, damage discovery, and feature fusion to determine the feature processing result.