Claims

1. A method for training a neural network, comprising:

transforming each training image of a plurality of training images by, for each training image:

detecting a hole in the training image; and

transforming the training image into a transformed image, to suppress non-crack information, by detecting crack features by reference to the hole detected in the respective training image; and

training a neural network using the transformed images, to detect cracks in images,

wherein transforming each training image further comprises, for each training image:

applying a set of filters, the set of filters being designed to enhance image features based on a shape of each said feature, segmenting the training image to produce a segmentation mask, and combining the segmentation mask with responses from the set of filters;

identifying, based on the segmentation mask and the responses from the set of filters, a surrounding region; and

applying a filter to the surrounding region to enhance crack features to produce a filter response, applying Gaussian blur to the surrounding region to produce a Gaussian blurred segmentation mask of the hole, and attributing to the transformed image a maximum of the filter response and Gaussian blurred segmentation mask.

2. The method of claim 1, further comprising:

receiving at least one further image; and

applying the trained neural network to the at least one further image, to detect cracks in the at least one further image.

3. The method of claim 1, wherein segmenting comprises performing active contour-based segmentation of the hole in the image.

4. The method of claim 1, wherein applying a set of filters comprises applying Jerman's 2D vesselness filter and a multi-scale line filter to the image to enhance crack information in the training image.

5. The method of claim 1, wherein a plurality of the training images form a series of temporally spaced images of a common object, and detecting crack features by reference to the hole detected in the respective training image comprises comparing crack features detected in a first image in the series with crack features detected in a second image in the series to determine crack progression.

6. The method of claim 5, wherein each image comprises an object, the method further comprising:

receiving a further image of an object;

applying the trained neural network to the further image, to detect crack features in the further image; and

comparing the crack features in the further image to crack features in a temporally previous image of the object of the further image to determine crack progression.

7. The method of 1, wherein transforming each training image into a transformed image, to suppress non-crack information in the training image, comprises disregarding crack features that are not associated with a hole detected in the respective training image.

8. The method of claim 1, further comprising producing one or more further training images by combining a hole segmentation mask with a crack segmentation mask.

9. The method of claim 1, wherein training the neural network comprises training a conditional generative adversarial network (cGAN) using the transformed images to detect cracks, the cGAN comprising an adversarial discriminator for distinguishing between real ground-truth (GT) images and generated images, and a generator for generating the generated images.

10. The method of claim 9, comprising incorporating a perceptual discriminator into the cGAN such that an ability of the perceptual discriminator to discriminate between GT images and generated images increases as an ability of the adversarial discriminator to discriminate between GT images and generated images decreases.

11. The method of claim 10, wherein the perceptual discriminator outputs an estimated probability p of the training image containing a crack, wherein a perceptual discrimination loss is defined a loss of binary classification according to:

$$\boldsymbol{L_{dis_p}(x) = -\log p_g(x) = -\, g \log p(x) - (1-g) \log (1-p(x))}$$

where g indicates a ground-truth class.

12. The method of claim 8, wherein parameters of the generator are obtained by optimizing a loss function being a sum of the adversarial loss and a perceptual loss produced by the adversarial discriminator and perceptual discriminator, according to:

$$L_G(G) = L_a(G) + L_{dis_p}(G)$$

where

$$L_a(G) = -\log D_a\big(y, G(y,z)\big).$$

13. A system for training a neural network, comprising:

memory; and

at least one processor, the memory storing instructions that, when executed by the at least one processor, cause the at least one processor to:

transform each training image of a plurality of training images by, for each training image, by:

applying a set of filters, the set of filters being designed to enhance image features based on a shape of each said feature, segmenting the training image to produce a segmentation mask, and combining the segmentation mask with responses from the set of filters,

detecting a hole in the training image;

transforming the training image into a transformed image, to suppress non-crack information, by detecting crack features by reference to the hole detected in the respective training image;

identifying, based on the segmentation mask and the responses from the set of filters, a surrounding region; and

applying a filter to the surrounding region to enhance crack features to produce a filter response, applying Gaussian blur to the surrounding region to produce a Gaussian blurred segmentation mask of the hole, and attributing to the transformed image a maximum of the filter response and Gaussian blurred segmentation mask; and

train a neural network using the transformed images, to detect cracks in images.

14. The system of claim 13, wherein the at least one processor is configured to perform the method of claim 1.

15. The system of claim 13, wherein the at least one processor transforms each training image into a transformed image, to suppress non-crack information in the training image, by detecting crack features by reference to the hole detected in the respective training image.

16. The system of claim 15, wherein either:

the at least one processor is further configured to:

receive at least one further image; and

apply the trained neural network to the at least one further image, to detect cracks in the at least one further image,

wherein transforming each training image into a transformed image, to suppress non-crack information in the training image, comprises detecting crack features by reference to the hole detected in the respective training image; and/or

wherein a plurality of the training images form a series of temporally spaced images of a common object, and the at least one processor is configured to detect crack features by reference to the hole detected in the respective training image by comparing crack features detected in a first image in the series with crack features detected in a second image in the series to determine progression of a crack, and wherein each image comprises an object, the at least one processor being further configured to:

receive a further image of an object;

apply the trained neural network to the further image, to detect crack features in the further image; and

compare the crack features in the further image to crack features in a temporally previous image of the object of the further image to determine crack progression.