

WHAT IS CLAIMED IS:

1. A method for training a set of parameters of a predictive computer model, the method comprising:

identifying a set of images captured by a set of cameras while affixed to one or more image collection systems;

for each image in the set of images, identifying a training output for the image;

for one or more images in the set of images, generating an augmented image for a set of augmented images by:

generating an augmented image for a set of augmented images by modifying the image with an image manipulation function that maintains camera properties of the image, and

associating the augmented training image with the training output of the image;

training the set of parameters of the predictive computer model to predict the training output based on an image training set including the images and the set of augmented images.

2. The method of claim 1, wherein the training output is an object in the image.

3. The method of claim 1, wherein the image training set does not include images generated by image manipulation functions that modify camera properties of an image.

4. The method of claim 3, wherein the image manipulation functions that modify camera properties include crop, pad, horizontal or vertical flip, or affine transformations.

5. The method of claim 1, wherein the image manipulation function is cutout, hue, saturations, value jitter, salt and pepper, domain transfer or any combination thereof.

6. The method of claim 1, wherein the image manipulation function is a cutout applied to the image based on a location of the training output in the image.

7. The method of claim 1, wherein the image manipulation function is a cutout applied to a portion of the image smaller than a bounding box of the training output.

8. The method of claim 1, wherein the image manipulation function is a cutout applied to a portion of the image that partially overlaps with the location of the training output in the image.

9. A system comprising one or more processors and non-transitory computer storage media storing instructions that when executed by the one or more processors, cause the processors to perform operations comprising:

- identifying a set of images captured by a set of cameras while affixed to one or more image collection systems;

- for each image in the set of images, identifying a training output for the image;

- for one or more images in the set of images, generating an augmented image for a set of augmented images by:

- generating an augmented image for a set of augmented images by modifying the image with an image manipulation function that maintains camera properties of the image, and

- associating the augmented training image with the training output of the image;

- training the set of parameters of the predictive computer model to predict the training output based on an image training set including the images and the set of augmented images.

10. The system of claim 9, wherein the image training set does not include images generated by image manipulation functions that modify camera properties of an image.

11. The system of claim 10, wherein the image manipulation functions that modify camera properties include crop, pad, horizontal or vertical flip, or affine transformations.

12. The system of claim 9, wherein the image manipulation function is cutout, hue, saturations, value jitter, salt and pepper, domain transfer or any combination thereof.

13. The system of claim 9, wherein the image manipulation function is a cutout applied to a portion of the image that partially overlaps with the location of the training output in the image.

14. A non-transitory computer-readable medium having instructions for execution by a processor, the instructions when executed by the processor causing the processor to:

identify a set of images captured by a set of cameras while affixed to one or more image collection systems;

for each image in the set of images, identify a training output for the image;

for one or more images in the set of images, generate an augmented image for a set of augmented images by:

generate an augmented image for a set of augmented images by modifying the image with an image manipulation function that maintains camera properties of the image, and

associate the augmented training image with the training output of the image;

train the computer model to learn to predict the training output based on an image training set including the images and the set of augmented images.

15. The non-transitory computer-readable medium of claim 14, wherein the image training set does not include images generated by image manipulation functions that modify camera properties of an image.

16. The non-transitory computer-readable medium of claim 15, wherein the image manipulation functions that modify camera properties include crop, pad, horizontal or vertical clip, or affine transformations.

17. The non-transitory computer-readable medium of claim 14, wherein the image manipulation function is cutout, hue, saturations, value jitter, salt and pepper, domain transfer or any combination thereof.

18. The non-transitory computer-readable medium of claim 14, wherein the image manipulation function is a cutout applied to a portion of the image smaller than a bounding box of the training output.

19. The non-transitory computer-readable medium of claim 14, wherein the image manipulation function is a cutout applied to a portion of the image smaller than a bounding box of the training output.

20. The non-transitory computer-readable medium of claim 14, wherein the image manipulation function is a cutout applied to a region that partially overlaps with the location of the training output in the image.