DEEP LEARNING PLATFORMS FOR AUTOMATED VISUAL INSPECTION

FIELD OF DISCLOSURE

[0001] The present application relates generally to automated visual inspection, and more specifically to techniques for training, testing and utilizing deep learning models to detect defects (e.g., container defects and/or foreign particles) in pharmaceutical or other applications.

BACKGROUND

[0002] In certain contexts, such as quality control procedures for manufactured drug products, it is necessary to examine samples (e.g., containers such as syringes or vials, and/or their contents such as fluid or lyophilized drug products) for defects, with any sample exhibiting defects being rejected, discarded, and/or further analyzed. To handle the quantities typically associated with commercial production of pharmaceuticals, the defect inspection task has increasingly become automated (automated visual inspection, or AVI). Some manufacturers have developed specialized equipment that can detect a broad range of defects, including container integrity defects such as cracks, cosmetic container defects such as scratches or stains on the container surface, and defects associated with the drug product itself such as atypical liquid colors or the presence of foreign particles. However, specialized equipment of this sort occupies a large footprint within a production facility, and is very complex and expensive. As just one example, the Bosch® 5023 commercial line equipment, which is used for the fill-finish inspection stage of drug-filled syringes, includes 15 separate visual inspection stations with a total of 23 cameras (i.e., one or two cameras per station). The high number of camera stations is dictated not only by the range of perspectives required for good coverage of the full range of defects, but also by processing limitations. In particular, the temporal window for computation can be relatively short at high production speeds. This can limit the complexity of individual image processing algorithms for a given station, which in turn necessitates multiple stations that each run image processing algorithms designed to look for only a specific class of defect. In addition to being large and expensive, such equipment generally requires substantial investments in manpower and other resources to qualify and commission each new product line. Maintenance of these AVI systems, and transitioning to a new product line, generally requires highly trained and experienced engineers, and often incurs substantial additional costs when assistance is required from field engineers associated with the AVI system vendor.

SUMMARY

[0002a] According to a first aspect, there is provided a method for reducing usage of processing resources when training a plurality of neural networks to perform automated visual inspection for a plurality of respective defect categories, with each of the plurality of respective defect categories being associated with a respective feature of containers or container contents, the method comprising:

obtaining, by one or more processors, a plurality of container images;

generating, by one or more processors processing the plurality of container images, a plurality of training image sets each corresponding to a different one of the plurality of container images, wherein for each training image set,

generating the training image set includes generating a different training image for each of the plurality of respective defect categories,

generating a different training image for each of the plurality of respective defect categories includes generating a first training image for a first defect category associated with a first feature, and

generating the first training image includes

identifying the first feature in the container image that corresponds to the training image set, and

generating the first training image such that the first training image (i) encompasses only a subset of the container image that corresponds to the training image set, and (ii) depicts the identified first feature; and

training, by one or more processors and using the plurality of training image sets, the plurality of neural networks to perform automated visual inspection for the plurality of defect categories.

[0002b] According to a second aspect, there is provided a method of using an efficiently trained neural network to perform automated visual inspection for detection of defects in a plurality of defect categories, with each of the plurality of defect categories being associated with a respective feature of containers or container contents, the method comprising:

obtaining, by one or more processors, a plurality of neural networks each corresponding to a different one of the plurality of defect categories, the plurality of neural networks having been trained using a plurality of training image sets each corresponding to a different one of a plurality of container images, wherein for each training image set and corresponding container image,

the training image set includes a different training image for each of the plurality of respective defect categories, and

at least some of the different training images within the training image set consist of different portions of the corresponding container image, with the different portions depicting different features of the corresponding container image;

obtaining, by one or more processors, a plurality of additional container images; and

performing automated visual inspection on the plurality of additional container images using the plurality of neural networks.

[0002c] There is provided a method of training an automated visual inspection (AVI) neural network to more accurately detect defects, the method comprising:

obtaining, by one or more processors, a plurality of container images;

for each container image of the plurality of container images, generating, by one or more processors, a corresponding set of new images, wherein generating the corresponding set of new images includes, for each new image of the corresponding set of new images, moving a portion of the container image that depicts a particular feature to a different new position; and

training, by one or more processors, the AVI neural network using the sets of new images corresponding to the plurality of container images.

[0002d] Tthere is provided a method of performing automated visual inspection for detection of defects, the method comprising:

obtaining, by one or more processors, a neural network trained using a plurality of container images and, for each container image, a plurality of augmented container images, wherein for each of the augmented container images, a portion of the container image that depicts a particular feature was moved to a different new position;

obtaining, by one or more processors, a plurality of additional container images; and

performing automated visual inspection on the plurality of additional container images using the neural network.

[0002e] There is provided a method of training an automated visual inspection (AVI) neural network to more accurately detect defects, the method comprising:

obtaining, by one or more processors, a plurality of images depicting real containers;

training, by one or more processors and using the plurality of images depicting real containers, a deep generative model to generate synthetic container images;

generating, by one or more processors and using the deep generative model, a plurality of synthetic container images; and

training, by one or more processors and using the plurality of synthetic container images, the AVI neural network.

[0002f] There is provided a method of performing automated visual inspection for detection of defects, the method comprising:

obtaining, by one or more processors, a neural network trained using synthetic container images generated by a deep generative model;

obtaining, by one or more processors, a plurality of additional container images; and

performing automated visual inspection on the plurality of additional container images using the neural network.

[0002g] According to a third aspect, there is provided a system comprising one or more processors and one or more memories, the one or more memories storing instructions that, when executed by the one or more processors, cause the one or more processors to perform the method of any one of preceding aspects.

[0003] Embodiments described herein relate to systems and methods that implement deep learning to reduce the size/footprint, complexity, cost, and/or required maintenance for AVI equipment, to improve defect detection accuracy of AVI equipment, and/or to simplify the task of adapting AVI equipment for use with a new product line. One potential advantage of deep learning is that it can be trained to simultaneously differentiate "good" products from products that exhibit any of a number of different defects. This parallelization, combined with the potential for deep learning algorithms to be less sensitive to nuances of perspective and illumination, can also allow a substantial reduction in the number of camera stations. This in turn allows a substantial reduction in the required amount of mechanical conveyance/handling (e.g., via starwheels, carousels, etc.), thereby further reducing the size of the AVI system and removing or reducing a potential source of variability and/or malfunctions. As one example, commercial AVI equipment with a footprint on the order of 3 x 5 meters may be reduced to a footprint on the order of 1 x 1.5 meters or less. Deep learning may also reduce the burden of transitioning to a new product line. For example, previously trained neural networks and the associated image libraries may be leveraged to reduce the training burden for the new product line.

[0004] While there have been recent, generalized proposals for using deep learning in the visual inspection task, the implementation of deep learning in this context gives rise to a number of significant technical issues, any of which can prevent the advantages listed above from being realized in practice. For example, while it may be relatively straightforward to determine whether a particular deep learning model provides sufficient detection accuracy in a specific use case (e.g., for a particular drug

What is claimed is:

1. A method for reducing usage of processing resources when training a plurality of neural networks to perform automated visual inspection for a plurality of respective defect categories, with each of the plurality of respective defect categories being associated with a respective feature of containers or container contents, the method comprising:

obtaining, by one or more processors, a plurality of container images;

generating, by one or more processors processing the plurality of container images, a plurality of training image sets each corresponding to a different one of the plurality of container images, wherein for each training image set,

generating the training image set includes generating a different training image for each of the plurality of respective defect categories associated with a respective feature of container or container contents,

generating a different training image for each of the plurality of respective defect categories includes generating a first training image for a first defect category associated with a first feature, and

generating the first training image includes

identifying the first feature in the container image that corresponds to the training image set using an object detection technique, and

cropping the container image that corresponds to the training image set to generate the first training image such that the first training image (i) encompasses only a portion of the container image that corresponds to the training image set, (ii) depicts the identified first feature, and (iii) excludes areas of the container image not relevant to the first feature; and

training, by one or more processors and using the plurality of training image sets, the plurality of neural networks to perform automated visual inspection for the plurality of defect categories.

2. The method of claim 1, wherein training the plurality of neural networks includes training each of the plurality of neural networks to infer a presence or absence of defects in a different one of the plurality of respective defect categories.

3. The method of claim 2, wherein training each of the plurality of neural networks includes, for each of the plurality of training image sets, using a different training image to train a different one of the plurality of neural networks.

4. The method of any one of claims 1 through 3, wherein identifying the first feature includes (i) identifying the first feature using template matching, or (ii) identifying the first feature using blob analysis.

5. The method of any one of claims 1-4, wherein

1) the first feature is a meniscus of a fluid within a container, and wherein the first defect category is a presence of one or more particles in or near the meniscus , or

2) the plurality of container images is a plurality of syringe images, and wherein either:

the first feature is a syringe plunger and the first defect category is one or both of (i) a plunger defect or (ii) a presence of one or more particles on the syringe plunger;

the first feature is a syringe barrel and the first defect category is one or both of (i) a barrel defect or (ii) a presence of one or more particles within the syringe barrel;

the first feature is a syringe needle shield and the first defect category is one or both of (i) an absence of the syringe needle shield or (ii) misalignment of the syringe needle shield; or

the first feature is a syringe flange and the first defect category is one or both of (i) a malformed flange or (ii) a defect on the syringe flange ; or

3) the plurality of container images is a plurality of cartridge images, and wherein either:

the first feature is a cartridge piston and the first defect category is one or both of (i) a piston defect or (ii) a presence of one or more particles on the cartridge piston;

the first feature is a cartridge barrel and the first defect category is one or both of (i) a barrel defect or (ii) a presence of one or more particles within the cartridge barrel; or

the first feature is a cartridge flange and the first defect category is one or both of (i) a malformed flange or (ii) a defect on the cartridge flange ; or

4) the plurality of container images is a plurality of vial images, and wherein either:

the first feature is a vial body and the first defect category is one or both of (i) a body defect or (ii) a presence of one or more particles within the vial body;

the first feature is a vial crimp and the first defect category is a defective crimp; or

the first feature is a lyophilized cake and the first defect category is a crack or other defect of the lyophilized cake.

6. The method of any one of claims 1 through 5, wherein:

generating a different training image for each of the plurality of respective defect categories further includes generating a second training image for a second defect category associated with a second feature; and

generating the second training image includes generating the second training image such that the second training image depicts the second feature.

7. The method of claim 6, wherein generating the second training image includes generating the second training image by down-sampling at least a portion of the container image that corresponds to the training image set to a lower resolution.

8. The method of claim 6, wherein generating the second training image includes identifying the second feature in the container image that corresponds to the training image set.

9. The method of any one of claims 1 through 8, wherein:

generating the plurality of training image sets each corresponding to a different one of the plurality of container images includes digitally aligning at least some of the plurality of container images, at least in part by resampling the at least some of the plurality of container images; and

digitally aligning at least some of the plurality of container images includes (i) detecting an edge within one or more of the plurality of container images, and (ii) comparing a position of the detected edge to a position of a reference line.

10. A method of using an efficiently trained neural network to perform automated visual inspection for detection of defects in a plurality of defect categories, with each of the plurality of defect categories being associated with a respective feature of containers or container contents, the method comprising:

obtaining, by one or more processors, a plurality of neural networks each corresponding to a different one of the plurality of defect categories, the plurality of neural networks having been trained using a plurality of training image sets each corresponding to a different one of a plurality of container images, wherein for each training image set and corresponding container image,

the training image set includes a different training image for each of the plurality of respective defect categories associated with a respective feature of containers or container contents, and

at least some of the different training images within the training image set consist of different portions of the corresponding container image, with the different portions depicting different features of the corresponding container image and excluding areas of the container image not relevant to the respective feature;

obtaining, by one or more processors, a plurality of additional container images; and

performing automated visual inspection on the plurality of additional container images using the plurality of neural networks.

11. The method of claim 10, wherein (i) the plurality of container images is a plurality of syringe images, and the different features include a syringe plunger, a syringe barrel, a syringe needle shield, and/or a syringe flange, (ii) the plurality of container images is a plurality of cartridge images, and the different features include a cartridge piston, a cartridge barrel, a barrel defect, and/or a cartridge flange, or (iii) the plurality of container images is a plurality of vial images, and the different features include a vial body, a vial crimp, and/or a lyophilized cake.

12. The method of claim 10 or 11, wherein the different training images include images down-sampled to different resolutions.

13. A system comprising one or more processors and one or more memories, the one or more memories storing instructions that, when executed by the one or more processors, cause the one or more processors to perform the method of any one of claims 1-12.