(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
01 November 2018 (01.11.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/200072 A1

(51) International Patent Classification:
G06K 9/62 (2006.01) *G06N 3/02* (2006.01)

(21) International Application Number:
PCT/US2018/020101

(22) International Filing Date:
28 February 2018 (28.02.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/489,529 25 April 2017 (25.04.2017) US
15/906,710 27 February 2018 (27.02.2018) US

(71) Applicant: **NEC LABORATORIES AMERICA, INC.** [US/US]; 4 Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

(72) Inventors: **CHOI, Wongun**; 16 Willow Street, Lexington, Massachusetts 02421 (US). **SCHULTER, Samuel**; 900 Pepper Tree Lane, Santa Clara, California 95051 (US). **SOHN, Kihyuk**; 37168 Meadowbrook Common, Fremont, California 94536 (US). **CHANDRAKER, Manmohan**; 444 Saratoga Avenue, Santa Clara, California 95050 (US).

(74) Agent: **KOLODKA, Joseph**; NEC Laboratories America, Inc., 4 Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,

(54) Title: CYCLIC GENERATIVE ADVERSARIAL NETWORK FOR UNSUPERVISED CROSS-DOMAIN IMAGE GENERATION

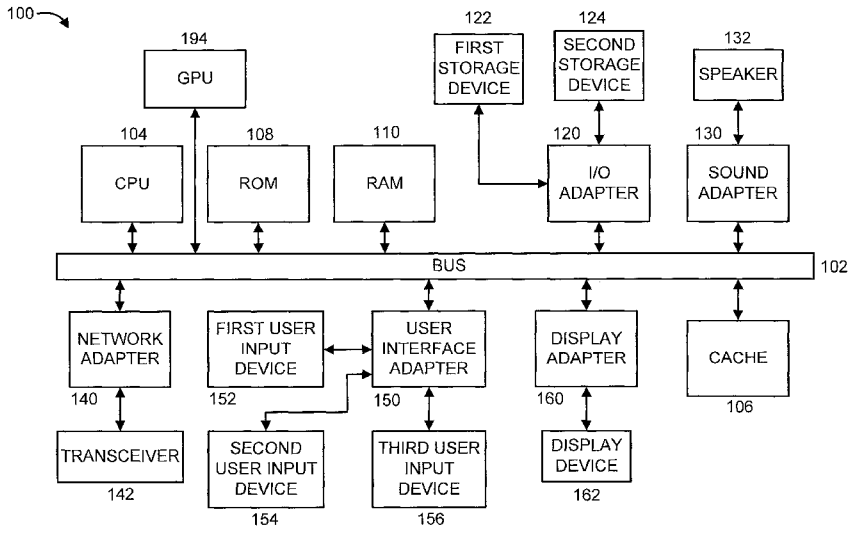


FIG. 1

(57) Abstract: A system is provided for unsupervised cross-domain image generation relative to a first and second image domain that each include real images. A first generator generates synthetic images similar to real images in the second domain while including a semantic content of real images in the first domain. A second generator generates synthetic images similar to real images in the first domain while including a semantic content of real images in the second domain. A first discriminator discriminates real images in the first domain against synthetic images generated by the second generator. A second discriminator discriminates real images in the second domain against synthetic images generated by the first generator. The discriminators and generators are deep neural networks and respectively form a generative network and a discriminative network in a cyclic GAN framework configured to increase an error rate of the discriminative network to improve synthetic image quality.

WO 2018/200072 A1

[Continued on next page]

OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

CYCLIC GENERATIVE ADVERSARIAL NETWORK FOR UNSUPERVISED CROSS-DOMAIN IMAGE GENERATION

RELATED APPLICATION INFORMATION

[0001] This application claims priority to provisional application serial number 62/489,529 filed on April 25, 2017 and U.S. utility application serial number 15/906,710 filed February 27, 2018, incorporated herein by reference.

BACKGROUND

Technical Field

[0002] The present invention relates to image recognition, and more particularly to a cyclic generative adversarial network for unsupervised cross-domain image generation.

Description of the Related Art

[0003] Generating images in target domains while having labels only in the source domain allows learning image recognition classifiers in the target domain without need for target domain labels. In fields that can involve image generation, the test (target) and source (training) domains for the image generator can often vary in a multitude of ways. As such, the quality of the images generated by the image generator may be lacking and paired training data of corresponding images from the two domains may not be available. Accordingly, there is a need for domain adaptation in order to reduce such variations and provide enhanced classification accuracy.

SUMMARY

[0004] According to an aspect of the present invention, a system is provided for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images. The system includes a first image generator for generating synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain. The system further includes a second image generator for generating synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain. The system also includes a first discriminator for discriminating the real images in the first image domain against the synthetic images generated by the second image generator. The system additionally includes a second discriminator for discriminating the real images in the second image domain against the synthetic images generated by the first image generator. The discriminators and the generators are deep neural networks and respectively form a generative network and a discriminative network in a cyclic Generative Adversarial Network (GAN) framework that is configured to increase an error rate of the discriminative network to improve a quality of the synthetic images.

[0005] According to another aspect of the present invention, a computer-implemented method is provided for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images. The method includes generating, by a first image generator, synthetic images having a similar appearance to one or more of the real images in the second image domain while including

a semantic content of one or more of the real images in the first image domain. The method further includes generating, by a second image generator, synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain. The method also includes discriminating, by a first discriminator, the real images in the first image domain against the synthetic images generated by the second image generator. The method additionally includes discriminating, by a second discriminator, the real images in the second image domain against the synthetic images generated by the first image generator. The generators are each neural network based and respectively form a generative network and a discriminative network in a cyclic Generative Adversarial Network (GAN) framework. The method further includes increasing an error rate of the discriminative network to improve a quality of the synthetic images.

[0006] According to yet another aspect of the present invention, a computer program product is provided for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images. The computer program product includes a non-transitory computer readable storage medium having program instructions embodied therewith. The program instructions are executable by a computer to cause the computer to perform a method. The method includes generating, by a first image generator of the computer, synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain. The method further includes generating, by a second image generator of the computer, synthetic images

having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain. The method also includes discriminating, by a first discriminator of the computer, the real images in the first image domain against the synthetic images generated by the second image generator. The method additionally includes discriminating, by a second discriminator of the computer, the real images in the second image domain against the synthetic images generated by the first image generator. The discriminators and the generators are each neural network based and respectively form a generative network and a discriminative network in a cyclic Generative Adversarial Network (GAN) framework. The method further includes increasing an error rate of the discriminative network to improve a quality of the synthetic images.

[0007] These and other features and advantages will become apparent from the following detailed description of illustrative embodiments thereof, which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

[0008] The disclosure will provide details in the following description of preferred embodiments with reference to the following figures wherein:

[0009] FIG. 1 shows an exemplary processing system to which the present principles may be applied, according to an embodiment of the present principles;

[0010] FIG. 2 shows an exemplary cyclic Generative Adversarial Network (GAN) framework of the present invention, in accordance with an embodiment of the present invention;

[0011] FIG. 3 shows a portion of the cyclic GAN framework of FIG. 2 during a testing phase, in accordance with an embodiment of the present invention; and

[0012] FIGs. 4-5 show an exemplary method for unsupervised cross-domain image generation relative to a first image domain and a second image domain, in accordance with an embodiment of the present principles.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[0013] The present invention is directed to a cyclic generative adversarial network for unsupervised cross-domain image generation.

[0014] In an embodiment, a cyclic generative adversarial network is proposed that takes images from a source domain to generate images in a different target domain and then back to the source domain, without having any corresponding pairs of images in the source and target domains. This is used for image recognition applications such as object detection and semantic segmentation, with labels available in the source domain, but no labels in the target domain, whereby the generated images become available as training data with labels preserved across the source and target domains while the image properties change.

[0015] In an embodiment, the present invention provides an image generation algorithm that can transfer an image in one domain to another domain. For example, the domain transfer can involve, but is not limited to, for example, generating realistic images from synthetic images, generating night images from day images, and so forth. In an embodiment, the generation process keeps the high level semantic concepts in the

input image while transforming the image characteristics to make it indistinguishable from the images in the target domain.

[0016] In an embodiment, an unsupervised domain-to-domain translation model is provided that could be learned without any supervision in the training dataset. This enables us to learn a high quality image generation model for many valuable applications, such as real image generation from synthetic images, rainy scene generation from bright day images, and so forth, where it is not possible to have corresponding images in both domains (that is, supervised as in image-to-image translation).

[0017] In an embodiment, the present invention employs a cyclic Generative Adversarial Network (GAN) framework that enforces recovering the entire image content when applied to two domain transfers cyclically such as, for example, but not limited to, rainy image to bright image and back to a rainy image. Enforcing such cyclic consistency aids in domain transfer model learning which keeps the semantic contents consistent across the generation process while adapting image properties to the target domain.

[0018] FIG. 1 shows an exemplary processing system 100 to which the invention principles may be applied, in accordance with an embodiment of the present invention. The processing system 100 includes at least one processor (CPU) 104 operatively coupled to other components via a system bus 102. A cache 106, a Read Only Memory (ROM) 108, a Random Access Memory (RAM) 110, an input/output (I/O) adapter 120, a sound adapter 130, a network adapter 140, a user interface adapter 150, and a display adapter 160, are operatively coupled to the system bus 102. At least one Graphics Processing Unit (GPU) 194 is operatively coupled to the system bus 102.

[0019] A first storage device 122 and a second storage device 124 are operatively coupled to system bus 102 by the I/O adapter 120. The storage devices 122 and 124 can be any of a disk storage device (e.g., a magnetic or optical disk storage device), a solid state magnetic device, and so forth. The storage devices 122 and 124 can be the same type of storage device or different types of storage devices.

[0020] A speaker 132 is operatively coupled to system bus 102 by the sound adapter 130. A transceiver 142 is operatively coupled to system bus 102 by network adapter 140. A display device 162 is operatively coupled to system bus 102 by display adapter 160.

[0021] A first user input device 152, a second user input device 154, and a third user input device 156 are operatively coupled to system bus 102 by user interface adapter 150. The user input devices 152, 154, and 156 can be any of a keyboard, a mouse, a keypad, an image capture device, a motion sensing device, a microphone, a device incorporating the functionality of at least two of the preceding devices, and so forth. Of course, other types of input devices can also be used, while maintaining the spirit of the present invention. The user input devices 152, 154, and 156 can be the same type of user input device or different types of user input devices. The user input devices 152, 154, and 156 are used to input and output information to and from system 100.

[0022] Of course, the processing system 100 may also include other elements (not shown), as readily contemplated by one of skill in the art, as well as omit certain elements. For example, various other input devices and/or output devices can be included in processing system 100, depending upon the particular implementation of the same, as readily understood by one of ordinary skill in the art. For example, various types of wireless and/or wired input and/or output devices can be used. Moreover,

additional processors, controllers, memories, and so forth, in various configurations can also be utilized as readily appreciated by one of ordinary skill in the art. These and other variations of the processing system 100 are readily contemplated by one of ordinary skill in the art given the teachings of the present invention provided herein.

[0023] Moreover, it is to be appreciated that framework 200 described below with respect to FIG. 2 is a framework for implementing respective embodiments of the present invention. Part or all of processing system 100 may be implemented in one or more of the elements of framework 200.

[0024] Further, it is to be appreciated that processing system 100 may perform at least part of the method described herein including, for example, at least part of method 400 of FIGs. 4-5. Similarly, part or all of framework 200 may be used to perform at least part of method 400 of FIGs. 4-5.

[0025] FIG. 2 shows an exemplary cyclic Generative Adversarial Network (GAN) framework 200 of the present invention, in accordance with an embodiment of the present invention.

[0026] The cyclic GAN framework (hereinafter "framework" in short) 200 includes a first domain input (hereinafter "input A" in short) 201 and a second domain input (hereinafter "input B" in short) 251 corresponding to a first image domain (hereinafter "domain A" in short) 291 and a second image domain (hereinafter "domain B" in short) 292, respectively. Domain A 291 and domain B 292 are respective image domains that include real images. Hence, input A 201 and input B 251 are implemented as respective real images. Thus, the two domains are not required to be supervised (include the same images).

[0027] The framework 200 further includes a neural network based discriminator (hereinafter "discriminator A" in short) 210, a neural network based discriminator (hereinafter "discriminator B" in short) 220, a neural network based image generator (hereinafter "generator A2B" in short) 230, a neural network based image generator (hereinafter "generator B2A" in short) 240, and a cyclic consistency loss (also referred to herein as "L2 loss") 250. The generator A2B 230 can be implemented as a generative model that is trained with domain A 291. In an embodiment, the generators can be implemented by convolutional neural networks, while the discriminators can be implemented by de-convolutional neural networks. Of course, other types of neural networks can also be used in accordance with the teachings of the present invention, while maintaining the spirit of the present invention.

[0028] The generator A2B 230 generates images ABA 277 that are looks similar to the images in domain B 292 but including the semantic contents of the input image from domain A 291. The generator B2A 240 generates images AB 278 based on the output of the generator A2B 230. The discriminator A 210 and discriminator B 220 are trained to discriminate real image from domain A 291 (or domain B 292) against the generated images for domain A 291 (or domain B 292). That is, the discriminator A 210 discriminates a real image from the domain A 291 against a generated image for the domain A 291, while the discriminator B 220 discriminates a real image from domain B 292 against a generated image for domain B 292.

[0029] The Generative Adversarial Network (GAN) framework 200, together with cyclic consistency loss L2 250, learn the neural network based elements (210, 220, 230, and 240). The GAN loss encourages the generated outputs to look similar to the images

in the corresponding target domain, which is achieved by the gradient coming from the discriminators. On the other hand, the cyclic consistency loss (L2 in our case) helps keep the semantic contents of the image. Note that we also learn the image cyclic GAN for the BAB directions (that is, from B to A and back to B) simultaneously. The GAN loss encourages images synthesized by the generator to have similar statistics as images from the target domain, which is achieved by gradients from the discriminator. The L2 loss results from the cyclic approach used for the GAN framework and represents the cyclic consistency loss in the cross-domain image generation. It compares the original image with the image synthesized using the output of the first generator as input. While an L2 loss is used in this embodiment, equivalent formulation may be derived using alternative losses such as L1, SSIM, perceptual loss, or other objectives that impose consistency for certain image statistics such as edge distributions.

[0030] In an embodiment the GAN framework 200 is configured to learn both the generative model and the discriminator at the same using the combination of GAN and L2 loss functions. The resulting training dynamics are usually described as a game between a generator(s) (i.e., the generative model(s)) and a discriminator(s) (i.e., the loss function(s)).

[0031] The discriminators (210 and 220) and the generators (230 and 240) respectively form a generative network and a discriminative network in the Generative Adversarial Network (GAN) framework 200, where the GAN framework is configured to increase an error rate of the discriminative network (i.e., "fool" the discriminator network by producing novel synthetic images that appear to have come from the true data distribution). That is, the goal of the generators (230 and 240) is to produce realistic

samples that fool the discriminators (210 and 220), while the discriminators (210 and 220) are trained to distinguish between the true training data and samples from the generators (230 and 240).

[0032] In an embodiment, the framework 200 can be based on deep learning and is trainable, rather than relying on a handcrafted generation algorithm. Thus, it can be applied to many different domain transfer tasks as long as there exists suitable datasets. Also, since the present invention does not require any supervision, it is widely applicable to many different image generation tasks.

[0033] FIG. 3 shows a portion of the cyclic GAN framework 200 of FIG. 2 during a testing phase 300, in accordance with an embodiment of the present invention.

[0034] Once the generator A2B 230 is trained with domain A 291, the generator A2B 230 can be deployed to produce images in domain B 292 with any images from domain A 291.

[0035] FIGs. 4-5 show an exemplary method 400 for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images, in accordance with an embodiment of the present principles.

[0036] The method 400 is performed by a cyclic Generative Adversarial Network (GAN) having a first image generator (e.g., generator A2B 230), a second image generator (e.g., generator B2A 240), a first discriminator (e.g., discriminator A 210), and a second discriminator (e.g., discriminator B 220). The discriminators and the generators are each neural network based and respectively form a generative network and a discriminative network in the cyclic Generative Adversarial Network (GAN) framework. The cyclic GAN framework is configured to increase an error rate of the discriminative

network to improve a quality of the synthetic images. In an embodiment, blocks 410-440 can correspond to a training phase of the cyclic GAN framework, and blocks 450 and 460 can correspond to a test phase of the cyclic GAN framework.

[0037] At block 410, generate, by the first image generator, synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain.

[0038] At block 420, generate, by the second image generator, synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain.

[0039] At block 430, discriminate, by the first discriminator, the real images in the first image domain against the synthetic images generated by the second image generator.

[0040] In an embodiment, block 430 can include block 430.

[0041] At block 430A, obtain gradients from a discrimination process applied to the real images in the first image domain versus the synthetic images generated by the second image generator.

[0042] At block 440, discriminate, by the second discriminator, the real images in the second image domain against the synthetic images generated by the first image generator.

[0043] In an embodiment, block 440 can include block 440.

[0044] At block 440A, obtain gradients from a discrimination process applied to the real images in the second image domain versus the synthetic images generated by the first image generator.

[0045] At block 450, generate, by the generative network (now trained per blocks 410-440), one or more additional synthetic images using an input image from the first image domain. The one or more additional synthetic images are generated to appear similar to at least a subset of the images in the second image domain while including a semantic content of the input image from the first image domain. The additional synthetic images will be of a higher quality than the previously generated synthetic images due the learning process implemented by the training of the cyclic GAN framework. For example, the additional synthetic images can use the obtained gradients from blocks 430A and 440A in order to exploit the GAN loss to obtain similar appear, while a cyclic consistency loss (L2) is exploited to preserve semantic content from a source domain.

[0046] The additional synthetic images can be used for a myriad of applications, as readily appreciated by one ordinary skill in the art. For example, other applications to which the present invention can be applied include, but are not limited to, training other supervised learning elements in an object category detection or other type of detection/classification network (see, e.g., block 450A), the generation of datasets for different weather conditions (see, e.g., block 450B), a cyclic domain transfer (see, e.g., block 450C), annotation extraction and corresponding response action performance (e.g., block 450D), and so forth.

[0047] In an embodiment, block 450 can include blocks 450A-C.

[0048] At block 450A, train another supervised learning element in an object category detection or other type of detection/classification network using one or more of the additional synthetic images.

[0049] At block 450B, generate the additional synthesized images for different weather and/or other environmental conditions.

[0050] At block 450C, perform a cyclic domain transfer with respect to the first image domain and the second image domain using the additional synthesized images.

[0051] In an embodiment, block 450C can include block 450C1.

[0052] At block 450C1, enforce, by the cyclic GAN framework, cyclic consistency across the cyclic domain transfer while adapting the image properties from one of the domains to another one of the domains.

[0053] At block 450D, perform an annotation operation using the additional synthetic images, perform a matching operation between the resultant annotations and a predetermined set of action words, and initiate a response action when one or more matches occur.

[0054] A further description will now be given regarding various aspects of the present invention, in accordance with an embodiment of the present invention.

[0055] The present invention incorporates a principled deep generative model to learn high quality image generation models.

[0056] The present invention introduces a novel cyclic GAN framework that does not require supervised datasets, e.g., same image in the two domains.

[0057] At test time, new images can be efficiently generated using the already trained generative network.

[0058] In an embodiment, the generated images could be used for training other supervised learning modules such as semantic segmentation or object category detection networks. Particularly, using the image generation network trained for domain transfer

from synthetic to real, a dataset with detailed annotations can be obtained almost for free. The annotations can be obtained and/or otherwise derived from the semantic content, as preserved by the L2 loss. In an embodiment, a processor (e.g., CPU 104) can be used to receive the annotations and perform matching against the annotations such that matches between the annotations and a predetermined set of actions can be used to initiate a response action. For example, in the case of classifying an action as dangerous (e.g., due to the presence of a weapon (e.g., a firearm or knife), an action can be initiated by the processor such as locking a door to keep the person holding the weapon out of an area or contained within an area.

[0059] In an embodiment, the present invention can be applied to train an image generation network for different weather conditions. It is to be appreciated that having a large dataset for all possible weather condition can be prohibitively costly. However, using the present invention, datasets can be generated for different weather conditions without additional labors.

[0060] These and other applications to which the present invention can be applied are readily determined by one of ordinary skill in the art, given the teachings of the present invention provided herein, while maintaining the spirit of the present invention.

[0061] Some of the many advantages and/or contributions made by the present invention include, but are not limited to, the following.

[0062] The present invention does not require a supervised dataset to train the model, noting that a supervised dataset is often not available in many important application domains.

[0063] Moreover, the present invention can generate higher quality images that conventional approaches.

[0064] Additionally, the present invention can be used to generate image data for other supervised learning methods, such as object detection, semantic segmentation, and so forth. This can significantly reduce the cost for data acquisition.

[0065] Embodiments described herein may be entirely hardware, entirely software or including both hardware and software elements. In a preferred embodiment, the present invention is implemented in software, which includes but is not limited to firmware, resident software, microcode, etc.

[0066] Embodiments may include a computer program product accessible from a computer-usable or computer-readable medium providing program code for use by or in connection with a computer or any instruction execution system. A computer-usable or computer readable medium may include any apparatus that stores, communicates, propagates, or transports the program for use by or in connection with the instruction execution system, apparatus, or device. The medium can be magnetic, optical, electronic, electromagnetic, infrared, or semiconductor system (or apparatus or device) or a propagation medium. The medium may include a computer-readable medium such as a semiconductor or solid state memory, magnetic tape, a removable computer diskette, a random access memory (RAM), a read-only memory (ROM), a rigid magnetic disk and an optical disk, etc.

[0067] It is to be appreciated that the use of any of the following "/", "and/or", and "at least one of", for example, in the cases of "A/B", "A and/or B" and "at least one of A and B", is intended to encompass the selection of the first listed option (A) only, or the

selection of the second listed option (B) only, or the selection of both options (A and B). As a further example, in the cases of "A, B, and/or C" and "at least one of A, B, and C", such phrasing is intended to encompass the selection of the first listed option (A) only, or the selection of the second listed option (B) only, or the selection of the third listed option (C) only, or the selection of the first and the second listed options (A and B) only, or the selection of the first and third listed options (A and C) only, or the selection of the second and third listed options (B and C) only, or the selection of all three options (A and B and C). This may be extended, as readily apparent by one of ordinary skill in this and related arts, for as many items listed.

[0068] Having described preferred embodiments of a system and method (which are intended to be illustrative and not limiting), it is noted that modifications and variations can be made by persons skilled in the art in light of the above teachings. It is therefore to be understood that changes may be made in the particular embodiments disclosed which are within the scope and spirit of the invention as outlined by the appended claims. Having thus described aspects of the invention, with the details and particularity required by the patent laws, what is claimed and desired protected by Letters Patent is set forth in the appended claims.

WHAT IS CLAIMED IS:

1. A system for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images, comprising:

a first image generator for generating synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain;

a second image generator for generating synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain;

a first discriminator for discriminating the real images in the first image domain against the synthetic images generated by the second image generator; and

a second discriminator for discriminating the real images in the second image domain against the synthetic images generated by the first image generator;

wherein the discriminators and the generators are deep neural networks and respectively form a generative network and a discriminative network in a cyclic Generative Adversarial Network (GAN) framework that is configured to increase an error rate of the discriminative network to improve a quality of the synthetic images.

2. The system of claim 1, wherein the cyclic GAN framework employs a cyclic consistency loss in order to preserve the semantic contents for inclusion in the generated synthetic images.

3. The system of claim 1, wherein the first image domain and the second image domain include at least some different real images relative to each other.

4. The system of claim 1, wherein the generators are implemented by respective convolutional neural networks, and the discriminators are implemented by respective de-convolutional neural networks.

5. The system of claim 1, wherein the generators generate the synthetic images using gradients provided by the discriminators.

6. The system of claim 1, wherein the generative network is configured to train another supervised learning element in an object category detection network.

7. The system of claim 1, wherein the cyclic GAN framework is configured to perform a cyclic domain transfer with respect to the first image domain and the second image domain.

8. The system of claim 7, wherein the cyclic GAN framework is configured to enforce cyclic consistency across the cyclic domain transfer while adapting the image properties from one of the domains to another one of the domains.

9. The system of claim 1, wherein the generative network of the cyclic GAN framework is configured to generate the synthetic images for different weather conditions from the real images in any of the first domain and the second domain.

10. The system of claim 1, wherein the cyclic GAN framework forms an unsupervised domain-to-domain translation model configured for unsupervised learning from a training dataset from among the domains.

11. A computer-implemented method for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images, comprising:

generating, by a first image generator, synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain;

generating, by a second image generator, synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain;

discriminating, by a first discriminator, the real images in the first image domain against the synthetic images generated by the second image generator; and

discriminating, by a second discriminator, the real images in the second image domain against the synthetic images generated by the first image generator,

wherein the discriminators and the generators are each neural network based and respectively form a generative network and a discriminative network in a cyclic

Generative Adversarial Network (GAN) framework, and the method further comprises increasing an error rate of the discriminative network to improve a quality of the synthetic images.

12. The computer-implemented method of claim 11, further comprising utilizing a cyclic consistency loss in the cyclic GAN framework in order to preserve the semantic contents for inclusion in the generated synthetic images.

13. The computer-implemented method of claim 11, wherein the first image domain and the second image domain include at least some different real images relative to each other.

14. The computer-implemented method of claim 11, further comprising:

configuring each of a pair of convolutional neural networks as a respective one of the generators; and

configuring each of a pair of de-convolutional neural networks as a respective one of the discriminators.

15. The computer-implemented method of claim 11, wherein said generating steps generate the synthetic images using gradients provided by the discriminators.

16. The computer-implemented method of claim 11, further comprising training, by the generative network, another supervised learning element in an object category detection network.

17. The computer-implemented method of claim 11, further comprising performing, by the cyclic GAN framework, a cyclic domain transfer with respect to the first image domain and the second image domain.

18. The computer-implemented method of claim 17, further comprising configuring the cyclic GAN framework to enforce cyclic consistency across the cyclic domain transfer while adapting the image properties from one of the domains to another one of the domains.

19. The computer-implemented method of claim 11, further comprising configuring the generative network of the cyclic GAN framework to generate the synthetic images for different weather conditions from the real images in any of the first domain and the second domain.

20. A computer program product for unsupervised cross-domain image generation relative to a first image domain and a second image domain that each include real images, the computer program product comprising a non-transitory computer readable storage medium having program instructions embodied therewith, the program

instructions executable by a computer to cause the computer to perform a method comprising:

generating, by a first image generator of the computer, synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain;

generating, by a second image generator of the computer, synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain;

discriminating, by a first discriminator of the computer, the real images in the first image domain against the synthetic images generated by the second image generator; and

discriminating, by a second discriminator of the computer, the real images in the second image domain against the synthetic images generated by the first image generator,

wherein the discriminators and the generators are each neural network based and respectively form a generative network and a discriminative network in a cyclic Generative Adversarial Network (GAN) framework, and the method further comprises increasing an error rate of the discriminative network to improve a quality of the synthetic images.

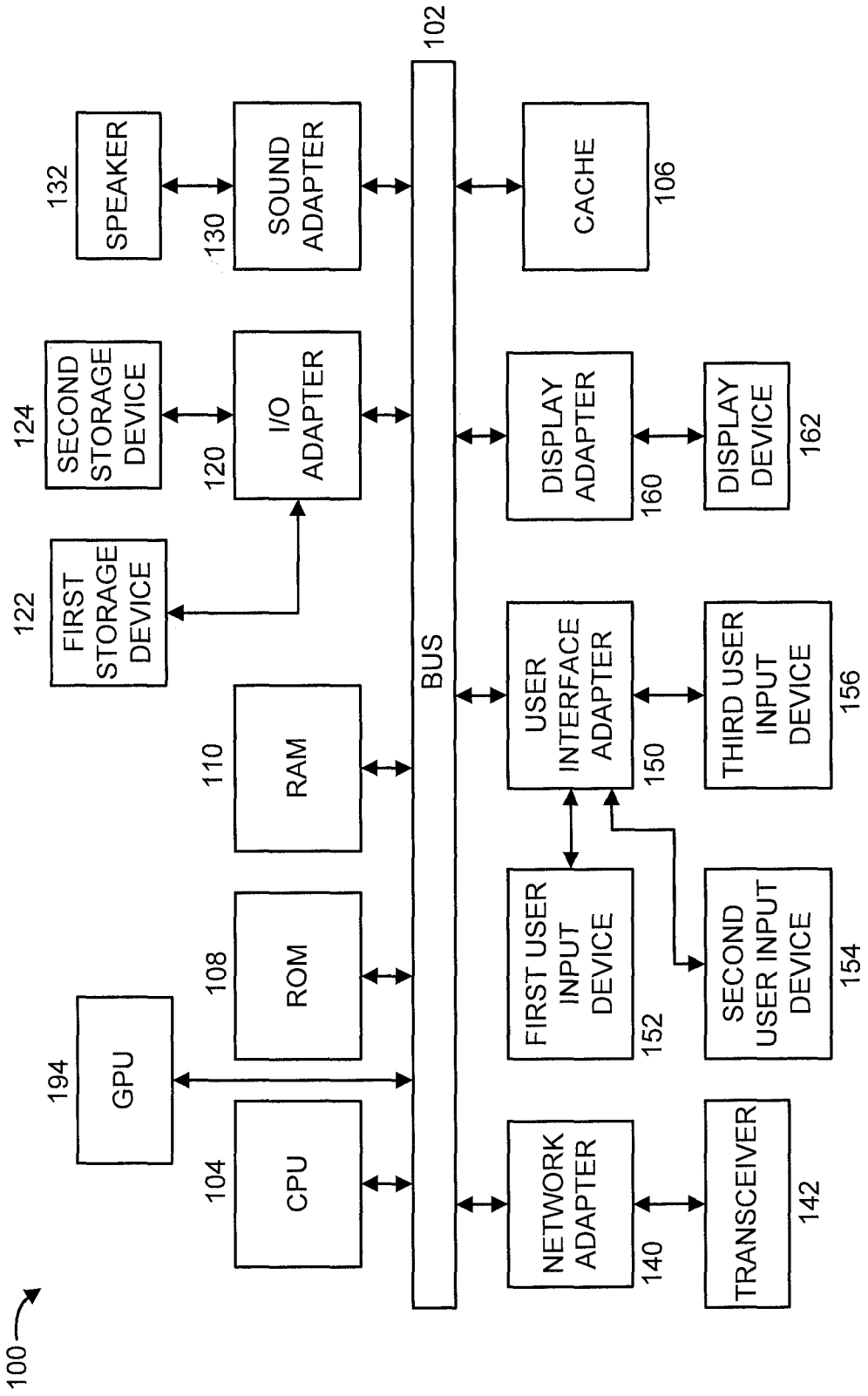

100
102
BUS
104
CPU
106
CACHE
108
ROM
110
RAM
120
I/O ADAPTER
122
FIRST STORAGE DEVICE
124
SECOND STORAGE DEVICE
130
SOUND ADAPTER
132
SPEAKER
140
NETWORK ADAPTER
142
TRANSCEIVER
150
USER INTERFACE ADAPTER
152
FIRST USER INPUT DEVICE
154
SECOND USER INPUT DEVICE
156
THIRD USER INPUT DEVICE
160
DISPLAY ADAPTER
162
DISPLAY DEVICE
194
GPU

FIG. 1

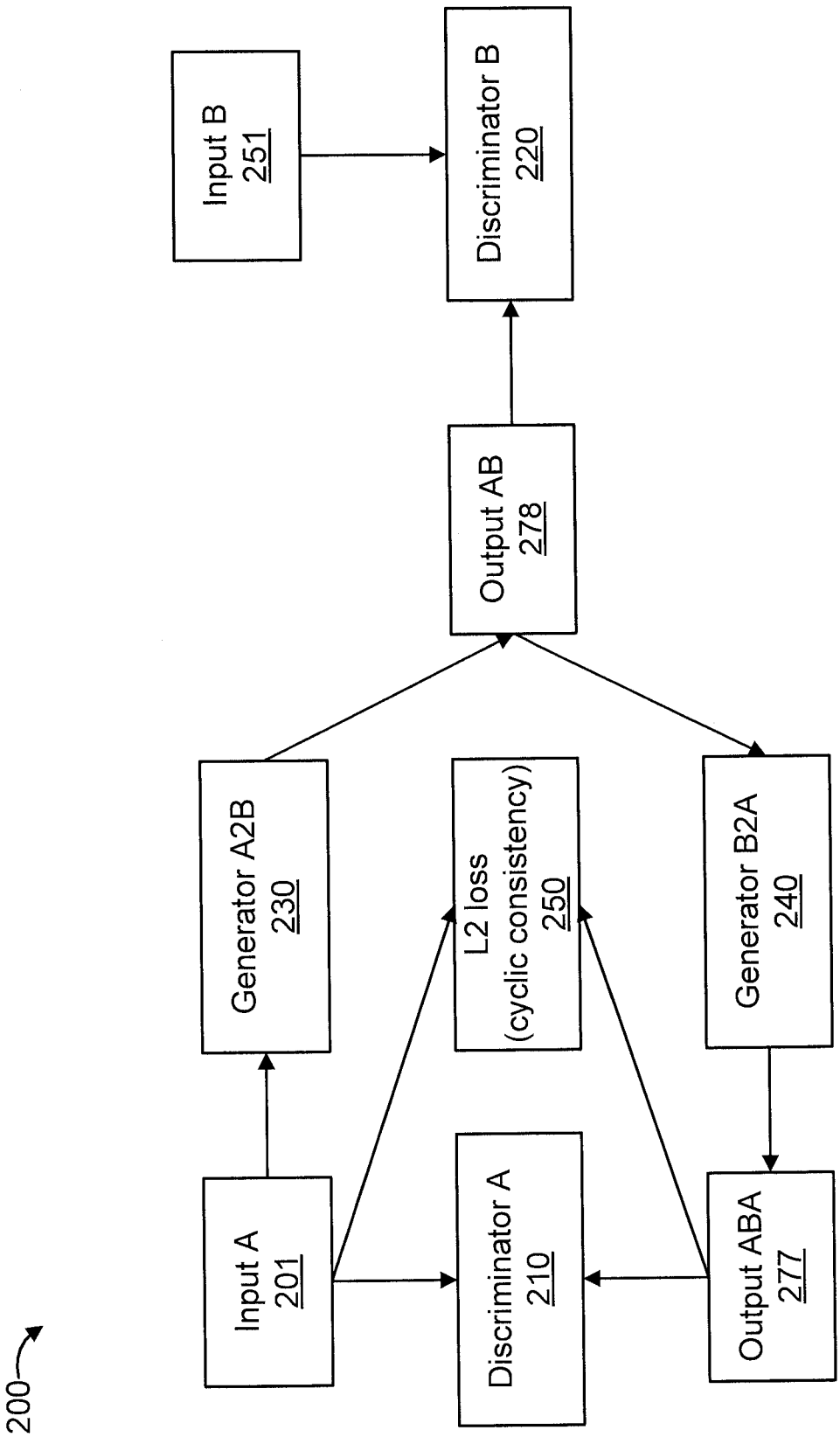

200
Input A
201
Input B
251
Generator A2B
230
Generator B2A
240
L2 loss
(cyclic consistency)
250
Discriminator A
210
Discriminator B
220
Output AB
278
Output ABA
277

FIG. 2

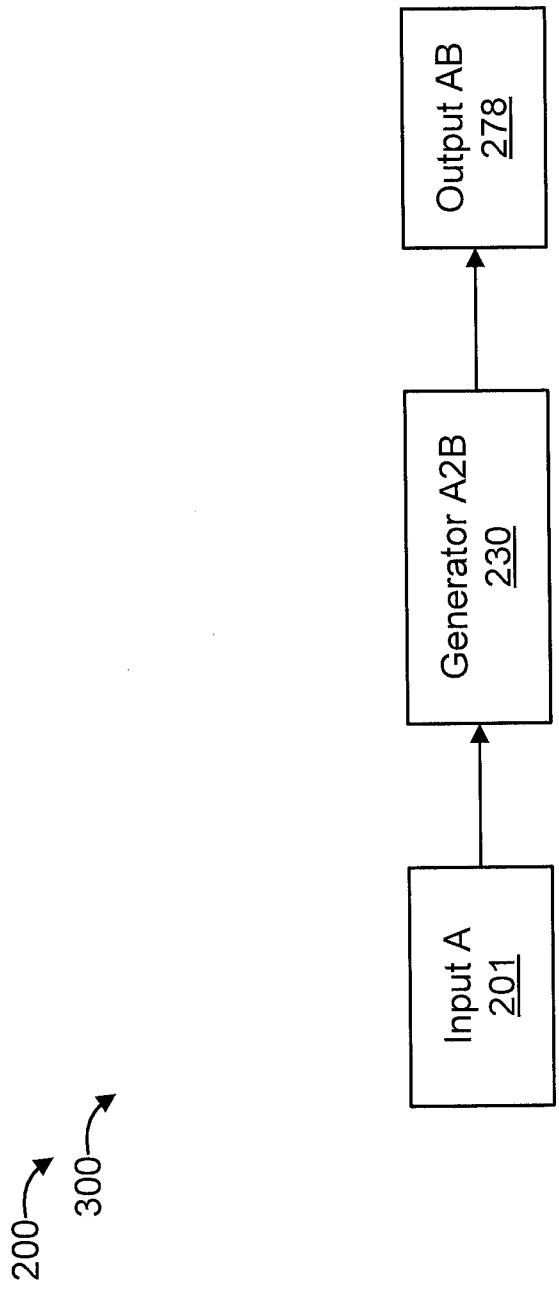


FIG. 3

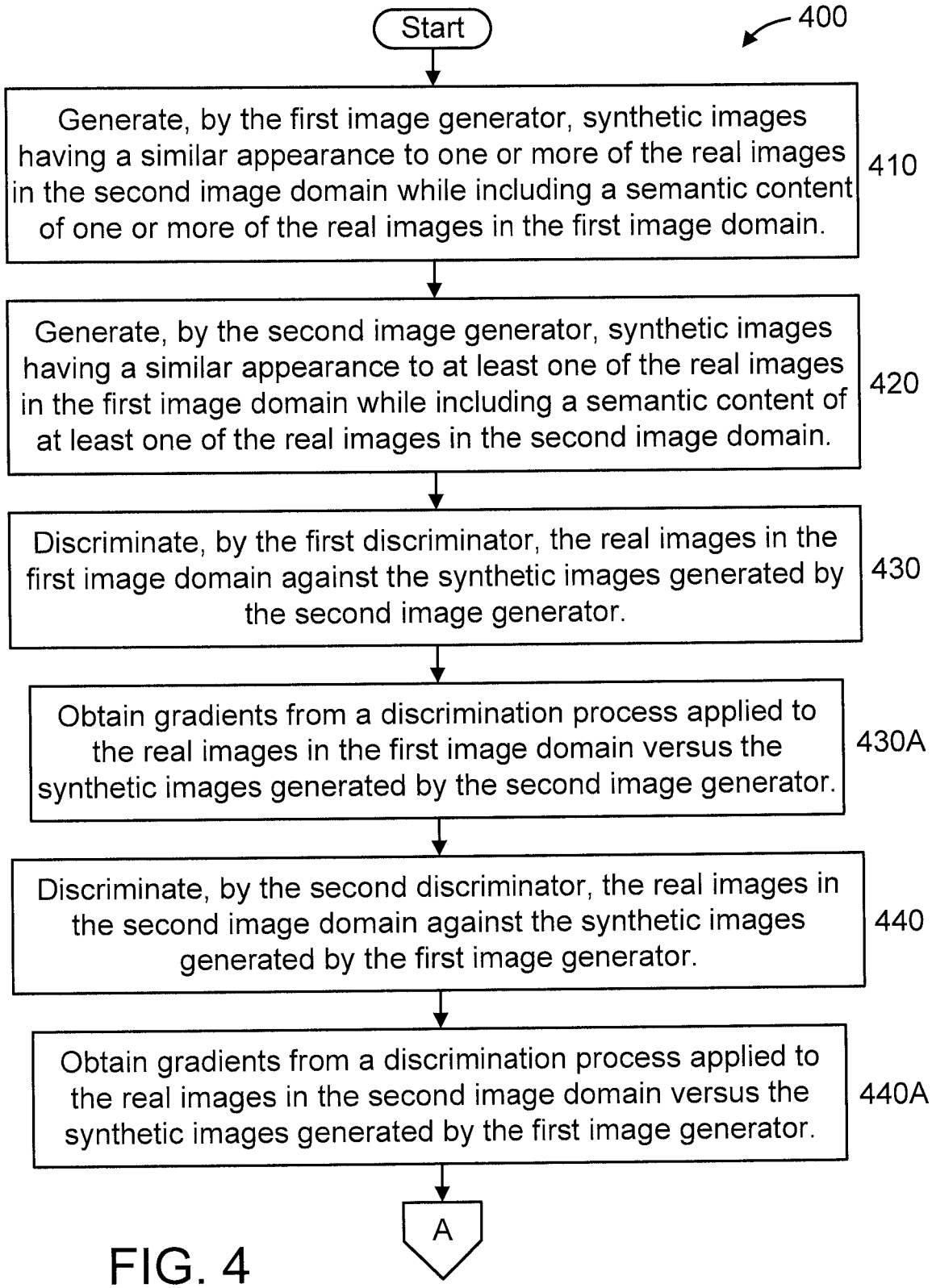

Start
400
Generate, by the first image generator, synthetic images having a similar appearance to one or more of the real images in the second image domain while including a semantic content of one or more of the real images in the first image domain.
410
Generate, by the second image generator, synthetic images having a similar appearance to at least one of the real images in the first image domain while including a semantic content of at least one of the real images in the second image domain.
420
Discriminate, by the first discriminator, the real images in the first image domain against the synthetic images generated by the second image generator.
430
Obtain gradients from a discrimination process applied to the real images in the first image domain versus the synthetic images generated by the second image generator.
430A
Discriminate, by the second discriminator, the real images in the second image domain against the synthetic images generated by the first image generator.
440
Obtain gradients from a discrimination process applied to the real images in the second image domain versus the synthetic images generated by the first image generator.
440A
A

FIG. 4

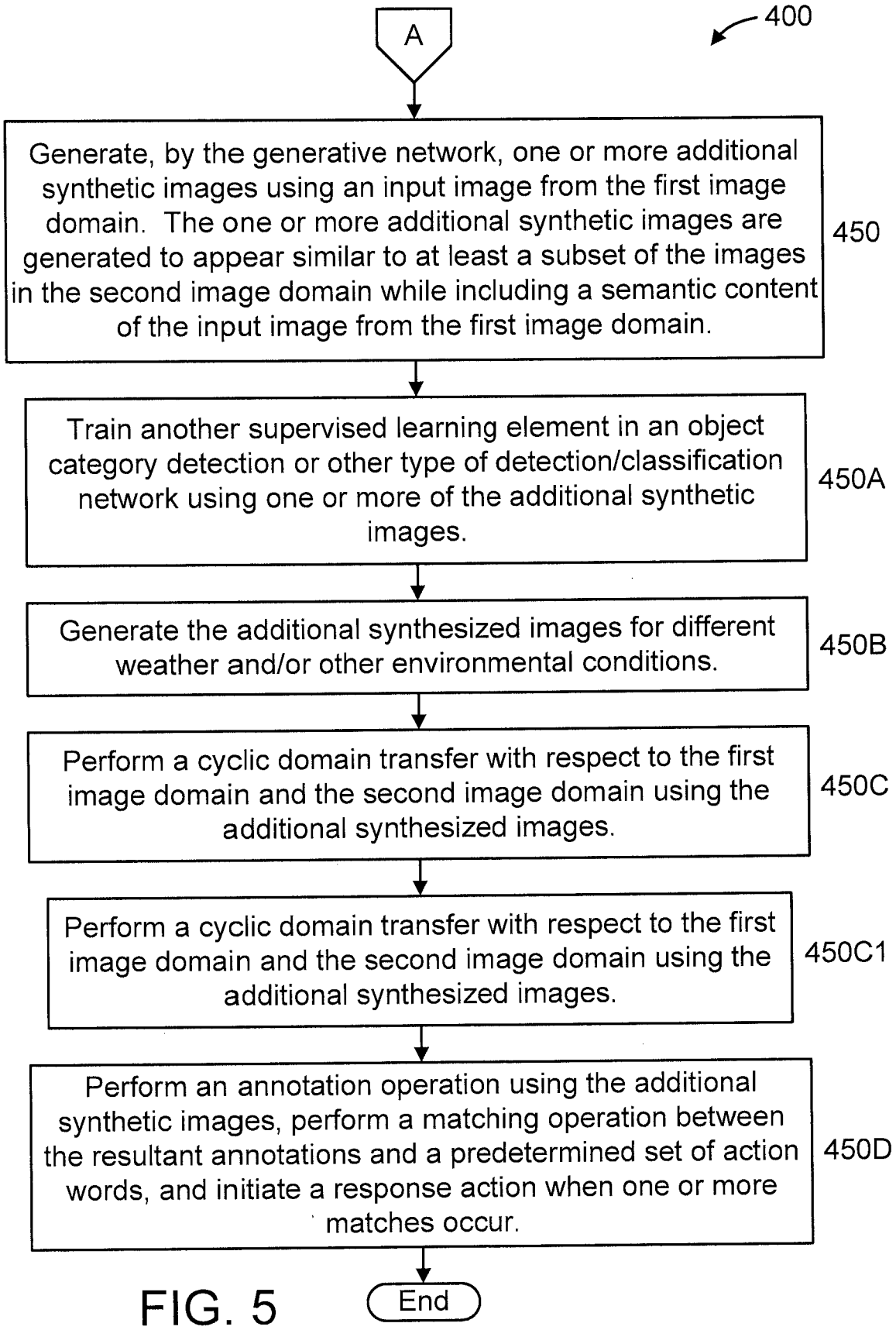

400
A
Generate, by the generative network, one or more additional synthetic images using an input image from the first image domain. The one or more additional synthetic images are generated to appear similar to at least a subset of the images in the second image domain while including a semantic content of the input image from the first image domain.
450
Train another supervised learning element in an object category detection or other type of detection/classification network using one or more of the additional synthetic images.
450A
Generate the additional synthesized images for different weather and/or other environmental conditions.
450B
Perform a cyclic domain transfer with respect to the first image domain and the second image domain using the additional synthesized images.
450C
Perform a cyclic domain transfer with respect to the first image domain and the second image domain using the additional synthesized images.
450C1
Perform an annotation operation using the additional synthetic images, perform a matching operation between the resultant annotations and a predetermined set of action words, and initiate a response action when one or more matches occur.
450D
End

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/020101

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)i, G06N 3/02(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06K 9/62; H04N 7/18; G06F 17/00; G06K 9/32; G06T 7/20; G06T 7/00; G06N 3/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & Keywords: first domain, second domain, synthetic image, real image, semantic content, discriminate, neural network, error rate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2014-0143310 A (SAMSUNG ELECTRONICS CO., LTD. et al.) 16 December 2014 See paragraphs [0027]-[0045] and figure 5.	1-20
A	US 7814040 B1 (ZHANG et al.) 12 October 2010 See column 9, line 51 - column 11, line 4 and figure 2.	1-20
A	JP 2016-058079 A (XEROX CORP.) 21 April 2016 See paragraphs [0016]-[0048] and figures 1-2.	1-20
A	US 2016-0155020 A1 (XEROX CORPORATION) 02 June 2016 See paragraphs [0006]-[0011], claim 1 and figure 1.	1-20
A	US 2016-0070969 A1 (KAVANAU) 10 March 2016 See paragraphs [0196]-[0197], claim 1 and figures 1A, 3-4.	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 June 2018 (12.06.2018)	**14 June 2018 (14.06.2018)**

Name and mailing address of the ISA/KR
International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea
Facsimile No. +82-42-481-8578

Authorized officer
LEE, Myung Jin
Telephone No. +82-42-481-8474

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2018/020101

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
KR 10-2014-0143310 A	16/12/2014	EP 2811424 A2	10/12/2014
		EP 2811424 A3	03/06/2015
		KR 10-1758064 B1	14/07/2017
		US 2014-0363047 A1	11/12/2014
		US 2014-0363076 A1	11/12/2014
		US 9311713 B2	12/04/2016
		US 9449392 B2	20/09/2016
US 7814040 B1	12/10/2010	US 8204842 B1	19/06/2012
		US 9280562 B1	08/03/2016
JP 2016-058079 A	21/04/2016	EP 2993618 A1	09/03/2016
		US 2016-0070986 A1	10/03/2016
		US 9710729 B2	18/07/2017
US 2016-0155020 A1	02/06/2016	EP 3035239 A1	22/06/2016
		JP 2016-110635 A	20/06/2016
		US 9607245 B2	28/03/2017
US 2016-0070969 A1	10/03/2016	CA 2342462 A1	30/09/2001
		CA 2342515 A1	30/09/2001
		CA 2343475 A1	30/09/2001
		CA 2565870 A1	15/12/2005
		EP 1138398 A2	04/10/2001
		EP 1138398 A3	14/05/2003
		EP 1138399 A2	04/10/2001
		EP 1138399 A3	14/05/2003
		EP 1147724 A2	24/10/2001
		EP 1147724 A3	04/06/2003
		EP 1751696 A2	14/02/2007
		JP 2007-536644 A	13/12/2007
		US 2002-0010975 A1	31/01/2002
		US 2002-0020349 A1	21/02/2002
		US 2003-0005533 A1	09/01/2003
		US 2009-0274375 A1	05/11/2009
		US 2014-0072207 A1	13/03/2014
		US 2017-0249523 A9	31/08/2017
		US 2018-0053065 A1	22/02/2018
		US 8503825 B2	06/08/2013
		US 9064188 B2	23/06/2015
		US 9792518 B2	17/10/2017
		WO 2005-117541 A2	15/12/2005
		WO 2005-117541 A3	23/03/2006