



(51) International Patent Classification:

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

(21) International Application Number:

PCT/EP2019/063650

(22) International Filing Date:

27 May 2019 (27.05.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicants: **TOYOTA MOTOR EUROPE** [BE/BE]; Avenue du Bourget 60, 1140 BRUSSELS (BE). **TECHNISCHE UNIVERSITÄT MÜNCHEN** [DE/DE]; Arcisstrasse 21, 80333 MÜNCHEN (DE).

(72) Inventors: **MEIER, Sven**; c/o Toyota Motor Europe, European Technical Administration, Division, Avenue du Bourget 60, 1140 BRUSSELS (BE). **KOBORI, Norimasa**; c/o Toyota Motor Europe, European Technical Administration, Division, Avenue du Bourget 60, 1140 BRUSSELS (BE). **MANHARDT, Fabian**; c/o TECHNISCHE UNIVERSITÄT MÜNCHEN, Chair of Computer Aided Medical Procedures and Augmented Reality, Boltzmannstr. 3, 85478 GARCHING (DE). **TOMBARI, Federico**; c/o TECHNISCHE UNIVERSITÄT MÜNCHEN, Chair

of Computer Aided Medical Procedures and Augmented Reality, Boltzmannstr. 3, 85478 GARCHING (DE). **RUPPRECHT, Christian**; c/o TECHNISCHE UNIVERSITÄT MÜNCHEN, Chair of Computer Aided Medical Procedures and Augmented Reality, Boltzmannstr. 3, 85478 GARCHING (DE).

(74) Agent: **HEWEL, Christoph** et al.; CABINET BEAU DE LOMENIE, 158 Rue de l'Université, 75340 PARIS CEDEX 07 (FR).

(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, 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,

(54) Title: SYSTEM AND METHOD FOR TRAINING A GENERATIVE ADVERSARIAL MODEL GENERATING IMAGE SAMPLES OF DIFFERENT BRIGHTNESS LEVELS

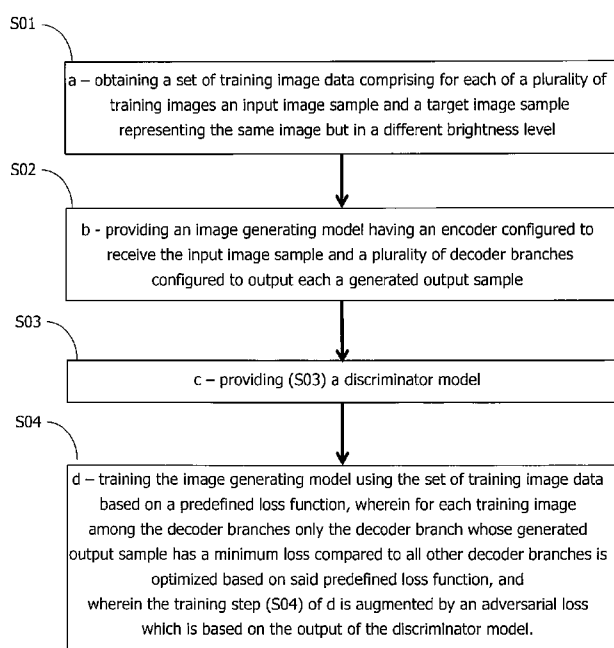


Fig. 1

(57) Abstract: The invention relates to a method and system for training a generative adversarial model generating image samples of different brightness levels. The method comprises the steps of: a - obtaining (S01) a set of training image data comprising for each of a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level, b - providing (S02) an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches configured to output each a generated output sample, c - providing (S03) a discriminator model, d - training (S04) the image generating model using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is optimized based on said predefined loss function, and wherein the training step (S04) of d is augmented by an adversarial loss which is based on the output of the discriminator model.

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))*

SYSTEM AND METHOD FOR TRAINING A GENERATIVE ADVERSARIAL MODEL GENERATING IMAGE SAMPLES OF DIFFERENT BRIGHTNESS LEVELS

5 FIELD OF THE DISCLOSURE

[0001] The present disclosure is related to the field of image processing, in particular to a method for training a generative adversarial model generating image samples of different brightness levels.

BACKGROUND OF THE DISCLOSURE

10 [0002] Machine learning techniques require sample data for training. For
example in the field of computer vision, such training data may be image
samples. Computer vision tasks comprise e.g. the extraction of high-
dimensional data from the real world in order to produce numerical or symbolic
information, e.g., in the forms of decisions. In particular, deep learning
15 architectures using a trained model, e.g. in the form of an artificial neural
network, can be applied in the field of computer vision. The model can be
trained in a supervised, semi-supervised or unsupervised manner.

[0003] Examples of deep learning architectures in the computer vision field comprise e.g.:

20 Fabian Manhardt et.al. (ECCV2018) "Deep Model-Based 6D Pose Refinement
in RGB", and

Rupprecht, Christian & Laina, Iro & DiPietro, Robert & Baust, Maximilian & Tombari, Federico & Hager, Gregory & Navab, Nassir (2017). Learning in an Uncertain World: Representing Ambiguity Through Multiple Hypotheses.

25 10.1109/ICCV.2017.388.

[0004] A further specific class of machine learning systems comprise generative adversarial network (GAN). In this class two neural networks contest with each other in a zero-sum game framework, cf. e.g.:

30 Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-
Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio (2014) "Generative
Adversarial Nets", Advances in neural information processing systems, Pages
2672-2680, and

Alec Radford, Luke Metz, Soumith Chintala, 2015 "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks" (arXiv:1511.06434).

[0005] When training a model, it may be desirable that the model learns to perform a task in different lighting conditions, i.e. to increase their robustness to lighting. For example, in case the computer vision tasks are in the context of automated driving, the trained model may be required to extract high-dimensional data from images and/or to semantically segment these images in different lighting conditions, e.g. during different day time and/or in tunnels.

[0006] The mentioned exemplary task of semantic image segmentation is a method to automatically determine the semantic labels of the objects which appear in an image. For example, the image may be acquired by a camera mounted in a vehicle. Semantic segmentation of such an image allows recognizing cars, pedestrians, traffic lanes, etc. Therefore, semantic segmentation is the backbone technique for autonomous driving systems or other automated systems. Semantic image segmentation typically uses models such as neural networks to perform the segmentation. These models need to be trained.

[0007] Training such a model typically comprises inputting known images to the model. For these images, a predetermined semantic segmentation is already known (an operator may have prepared the predetermined semantic segmentations of each image by annotating the images). The output of the model is then evaluated in view of the predetermined semantic segmentation, and the parameters of the model are adjusted if the output of the model differs from the predetermined semantic segmentation of an image. It follows that in order to train a semantic segmentation model, a large number of image samples and predetermined semantic segmentations are necessary.

[0008] It has been observed that the illumination condition at e.g. nighttime (in particular when there is no direct solar illumination but e.g. only electrical street lights, i.e. nocturnal artificial lighting) creates visibility problems for drivers and for automated systems. While sensors and computer vision algorithms are constantly getting better, the improvements are usually benchmarked with images taken during daylight time. Those methods often fail to work in nighttime condition. This prevents the automated systems from actually being used: it is not conceivable for a vehicle to avoid nighttime, and the vehicle has to be able to distinguish different objects during in daytime and nighttime conditions.

[0009] Compared to daylight, nocturnal artificial lighting degrades the visibility of a scene significantly, according to the darkness of the driving scene.

[0010] It may thus be desirable to train semantic segmentation models with nighttime images (images taken at night time with no direct solar illumination but e.g. only electrical street lights, i.e. nocturnal artificial lighting). Obtaining e.g. nighttime image samples (taken by a camera) or image samples of other brightness (e.g. of a relatively high brightness) can be difficult and time-consuming, especially if for each scene images at different brightness levels, e.g. during different day times have to be taken. Furthermore it can be difficult to annotate night-time images due to their light conditions (i.e. for the task of semantic segmentation).

[0011] It is noted though that the described segmentation task is merely an example for image processing. There may be other image processing tasks where it is desirable to train a model with image samples of different brightness levels, in order to increase robustness of the trained model with regard to varying lighting.

SUMMARY OF THE DISCLOSURE

[0012] Currently, it remains desirable to obtain image samples suitable for training an algorithm (e.g. another model) such that the robustness of the trained model with regard to varying lighting is increased. In particular, it remains desirable to provide a method and a system for training a model to learn generating image samples of different brightness levels, in particular such that the trained model can generate from one input image sample a set of different brightness output image samples, while preserving the main content of the input image sample.

[0013] Therefore, according to the embodiments of the present disclosure, a (desirably computer-implemented) method for training a generative adversarial model generating image samples of different brightness levels is provided. The method comprises the steps of:

a – obtaining a set of training image data comprising for each of a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level,

- b – providing an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches configured to output each a generated output sample,
- c – providing a discriminator model,
- 5 d – training the image generating model using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is optimized based on said predefined loss function, and
- 10 wherein the training step of d is augmented by an adversarial loss which is based on the output of the discriminator model.

[0014] Accordingly, it is proposed to use a model in the form of a generative adversarial model which is able to learn generating realistic output image samples which preserve the main content of the input image sample. In
15 particular, the generative adversarial model desirably reduces blur effects in the generated images.

[0015] Furthermore, due to the multiple hypothesis approach (i.e. the fact that for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all
20 other decoder branches is optimized based on said predefined loss function) the trained model becomes able to generate multiple images of consistently different lighting levels from a single input image.

[0016] By training such a model it becomes possible to use the generated image samples of the trained model to train other algorithms, e.g. other
25 models, for robustness to lighting.

[0017] For example, the generated image samples of the trained model may be used to extend the dynamic range (like e.g. a pre-processor) for an existing algorithm (e.g. another model) by feeding every output image into the existing algorithm.

30 [0018] For example, the standard L1 and/or L2 loss may be augmented by an adversarial loss which is based on a discriminator model trained to distinguish machine generated image samples and (realistic) target image samples.

[0019] For each training image the adversarial loss may be used to optimize
35 all decoder branches. Accordingly, at each training iteration all decoder

branches may be updated, in order to learn generating more realistic output image samples.

[0020] The step d of training the image generating model using the set of training image data may comprise a training according to a supervised learning method. In other words, the training uses a target image sample.

[0021] The predefined loss function may determine a loss between an output image sample generated by the image generating model and the target output image.

[0022] The method may further comprise the step of: e – training the discriminator model using the set of training image data and generated output image samples generated by the image generating model, in order to learn discriminating between real image samples and machine generated image samples.

Accordingly, the discriminator model is trained together with the image generating model.

[0023] The discriminator model and the image generating model may be trained alternately.

[0024] The step of training the discriminator model may comprise, for each generated image sample of a training image, inputting the generated image sample and the corresponding target image sample (which relates to the same input image sample as the generated image sample) and computing the adversarial loss for said generated image sample.

Accordingly, the discriminator model may be trained based on all generated image samples, i.e. on image samples with different brightness levels.

[0025] The adversarial loss used for augmenting the training step of the image generating model may be the mean of the adversarial losses for said generated image samples of the respective training image (i.e. those image samples which have been generated based on the same input image sample representing the training image).

[0026] The image generating model may be a deep neural network.

[0027] The discriminator model may be a deep neural network.

Each decoder branch may be a separate decoder of the model.

[0028] Alternatively the model may comprise a decoder with at least one decoder layer, in which the decoder is branched such the last decoder layer comprises the plurality of decoder branches. This alternative has the advantage

that during each training iteration of the decoder the common parts of decoder branches (including the decoder branch having the minimum loss) are optimized based on the predefined loss function.

- [0029] The present disclosure further relates to a system for training a generative adversarial model generating image samples of different brightness level, comprising:
- a module A for obtaining a set of training image data comprising for each of a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level,
 - 10 a module B for providing an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches configured to output each a generated output sample,
 - a module C for providing a discriminator model, and
 - a model D for training the image generating model using the set of training
 - 15 image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is optimized based on said predefined loss function, and
 - wherein the training in module D is augmented by an adversarial loss which is
 - 20 based on the output of the discriminator model.

[0030] The system may comprise further (sub-) modules and features corresponding to the features of the method described above.

- [0031] The present disclosure further relates to a (computer) system for generating image samples of different brightness levels, comprising the trained
- 25 model of step d of the method described above or of the trained module D of the system described above.

- [0032] The present disclosure further relates to a computer vision system for processing an input image sample (e.g. to semantically segment the input image sample, and/or to acquire information from said the input image sample), comprising a pre-processor module configured to generate image
- 30 samples of different brightness level based on the input image sample, the pre-processor module comprising the trained model of step d of the method described above or of the trained module D of the system described above.
- Accordingly, said computer vision system may comprise a module including an

algorithm (e.g. another model) configured to receive and process the output of the pre-processor module (i.e. including the generated image samples).

[0033] Furthermore the present disclosure relates to a computer program including instructions for executing the steps of a method, as described above, when said program is executed by a computer.

[0034] This program can use any programming language and take the form of source code, object code or a code intermediate between source code and object code, such as a partially compiled form, or any other desirable form.

[0035] Finally, the present disclosure relates to a recording medium readable by a computer and having recorded thereon a computer program including instructions for executing the steps of a method, as described above.

[0036] The information medium can be any entity or device capable of storing the program. For example, the medium can include storage means such as a ROM, for example a CD ROM or a microelectronic circuit ROM, or magnetic storage means, for example a diskette (floppy disk) or a hard disk.

[0037] Alternatively, the information medium can be an integrated circuit in which the program is incorporated, the circuit being adapted to execute the method in question or to be used in its execution.

[0038] It is intended that combinations of the above-described elements and those within the specification may be made, except where otherwise contradictory.

[0039] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the disclosure, as claimed.

[0040] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate embodiments of the disclosure and together with the description, and serve to explain the principles thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0041] Fig. 1 shows a schematic flow chart of the steps of a method for training a generative adversarial model generating image samples of different brightness levels according to embodiments of the present disclosure; and

[0042] Fig. 2 shows a schematic block diagram of a system according to embodiments of the present disclosure.

DESCRIPTION OF THE EMBODIMENTS

[0043] Reference will now be made in detail to exemplary embodiments of the disclosure, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the
5 drawings to refer to the same or like parts.

[0044] Fig. 1 shows a schematic flow chart of the steps of a method for training a generative adversarial model generating image samples of different brightness levels according to embodiments of the present disclosure.

[0045] Accordingly, the model being desirably a neural network is trained to
10 take as input an image sample and produces as output a multitude of image samples of same size but varying lighting conditions. Other than the brightness and clipping introduced by the limited dynamic range of the image format the scene is desirably not modified by the network.

[0046] In a first step S01 a set of training image data is obtained comprising
15 for each of a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level.

[0047] In a further step S02 (which may be carried out before, after or at the same time as step S01) an image generating model is provided having an encoder configured to receive the input image sample and a plurality of
20 decoder branches configured to output each a generated output sample.

[0048] In a further step S03 (which may be carried out before, after or at the same time as step S01 and/or S02) a discriminator model is provided.

[0049] Accordingly, the network consists of a single encoder and multiple decoders, each producing one of the multiple output images. In particular, a
25 global discriminator is used to provide the loss for each of the output images.

[0050] In a subsequent step S04 the image generating model is trained using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all
30 other decoder branches is optimized based on said predefined loss function.

[0051] In other words, the final loss is the minimum value of all losses associated with the respective decoder branches. Only the encoder and decoder branch associated with the lowest loss is updated on each training cycle.

[0052] The training step S04 is augmented by an adversarial loss which is
35 based on the output of the discriminator model.

[0053] In other words, the training objective is composed of two individual loss terms, i.e. the discriminator loss and the predefined (e.g. L1) loss. Thereby each of the n output image samples will be evaluated by the discriminator, which ensures that all images fit the target domain in terms of structure and realism.

[0054] Accordingly, each of the (predicted, i.e. generated) image samples is paired with the corresponding target image sample, and is sent through the discriminator. Afterwards, an (adversarial) loss value is obtained for each prediction and the mean over the individual losses is taken. Hence, the loss for the real input-target pair and all prediction-target pairs is computed and the mean is taken.

[0055] The image generating model and discriminator more are trained alternating. The discriminator model computes the loss for all predicted samples (each paired with the target) and then computes the mean over all loss terms, as expressed by the following simplified equations

for the generator step:

$$L = \min(L1(I, T)) + \frac{1}{M} \sum_{j=1}^M \text{Discriminator}(I_j, T)$$

and for the discriminator step:

$$L = \frac{1}{M} \sum_{j=1}^M \text{Discriminator}(I_j, T) + \text{Discriminator}(O, T)$$

wherein:

O = Generator Input Image

T = Target Image

I = All predicted Images

I_j = Predicted image j

[0056] The predefined loss will be also computed on all output images, but inspired by e.g. C. Rupprecht et al (2017). Learning in an Uncertain World: Representing Ambiguity Through Multiple Hypotheses. 10.1109/ICCV.2017.388,

only the minimum contributes to the final loss. This enforces that the different hypotheses specialize on the different target lightning conditions.

[0057] If for example the hypothesis (i.e. generated image sample) of the nth decoder branch is the closest of the target image sample, only its loss will contribute to the final loss to let it further optimize on the given sampled lightning condition.

[0058] In other words, the discriminator loss is the average over all generated image samples with respect to the discriminator. Thereby each of the n generated image samples will be evaluated by the discriminator, which ensures that all images fit the target domain in terms of structure and realism and in our case also lighting. The L1 loss will be also computed on all generated image samples, but only the minimum contributes to the final loss. This enforces that the different hypotheses specialize on the different target lightning conditions.

[0059] At runtime the trained model will output a multitude of generated image samples based on a given input image sample, each with a different brightness level but similar content to the original input image sample. The network can be used to augment an existing dataset to extend the lighting range while preserving labels associated to the original image. Additionally or alternatively it can be used to pre-process existing images to achieve a high dynamic range sequence of images for detection of objects or other existing perception applications.

[0060] Fig. 2 shows a schematic block diagram of a system according to embodiments of the present disclosure.

[0061] In this figure, a system 200 for training a model has been represented. This system 200, which may be a computer, comprises a processor 201 and a non-volatile memory 202. The system 200 may not only be configured for training a generative adversarial model generating image samples of different brightness levels. It may also apply the trained model to another algorithm 400. For example the trained model may be applied to a computer vision system 400. In other words, a computer vision system for processing an input image sample 400 may comprise a pre-processor module configured to generate image samples of different brightness level based on the input image sample, the pre-processor module comprising said trained model.

[0062] The system 200 may further be connected to a (passive) optical sensor 300, in particular a digital camera (e.g. integrated into the vehicle and being oriented to at least one of the front, the sides and the back). The digital camera 300 is configured such that it can take pictures which may be used as
5 input image samples provided to the model. The camera 300 desirably generates image data comprising a 2D or 3D image of the environment. There may also be provided a set of monocular cameras which generate a panoramic 2D or 3D image.

[0063] In the non-volatile memory 202, a set of instructions is stored and
10 this set of instructions comprises instructions to perform a method for training a model.

[0064] In particular, these instructions and the processor 201 may respectively form a plurality of modules:

a module A for obtaining a set of training image data comprising for each of
15 a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level,

a module B for providing an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches configured to output each a generated output sample,

20 a module C for providing a discriminator model, and

a model D for training the image generating model using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is
25 optimized based on said predefined loss function, and

wherein the training in module D is augmented by an adversarial loss which is based on the output of the discriminator model.

[0065] Throughout the description, including the claims, the term
30 "comprising a" should be understood as being synonymous with "comprising at least one" unless otherwise stated. In addition, any range set forth in the description, including the claims should be understood as including its end value(s) unless otherwise stated. Specific values for described elements should be understood to be within accepted manufacturing or industry tolerances known to one of skill in the art, and any use of the terms "substantially" and/or

"approximately" and/or "generally" should be understood to mean falling within such accepted tolerances.

[0066] Although the present disclosure herein has been described with reference to particular embodiments, it is to be understood that these
5 embodiments are merely illustrative of the principles and applications of the present disclosure.

[0067] It is intended that the specification and examples be considered as exemplary only, with a true scope of the disclosure being indicated by the following claims.

CLAIMS

1. A method for training a generative adversarial model generating image samples of different brightness levels, comprising the steps of:
 - 5 a – obtaining (S01) a set of training image data comprising for each of a plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level,
 - b – providing (S02) an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches
 - 10 c – providing (S03) a discriminator model,
 - d – training (S04) the image generating model using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is
 - 15 optimized based on said predefined loss function, and wherein the training step (S04) of d is augmented by an adversarial loss which is based on the output of the discriminator model.
- 20 2. The method according to the claim 1, wherein for each training image the adversarial loss is used to optimize all decoder branches.
3. The method according to the claim 1, wherein
 - 25 d – training (S04) the image generating model using the set of training image data comprises a training according to a supervised learning method.
4. The method according to claim 1 or 2, wherein the predefined loss function determines a loss between an output image sample
- 30 generated by the image generating model and the target output image.
5. The method according to any one of the preceding claims, further comprising the step of:
 - e – training (S06) the discriminator model using the set of training image data
 - 35 and generated output image samples generated by the image generating

model, in order to learn discriminating between real image samples and machine generated image samples.

6. The method according to any the preceding claim, wherein
5 the discriminator model and the image generating model are trained alternately.

7. The method according to any the preceding claim, wherein
training (S06) the discriminator model comprises, for each generated image
10 sample of a training image, inputting the generated image sample and the corresponding target image sample and computing the adversarial loss for said generated image sample, wherein
the adversarial loss used for augmenting the training step (S04) of the image
generating model is the mean of the adversarial losses for said generated
15 image samples of the respective training image.

8. The method according to the preceding claim, wherein
the image generating model is a deep neural network, and/or
the discriminator model is a deep neural network.
20

9. The method according to the preceding claim, wherein
each decoder branch is a separate decoder of the model, or
the model comprises a decoder with at least one decoder layer, in which the
decoder is branched such the last decoder layer comprises the plurality of
25 decoder branches.

10. A system for training a generative adversarial model generating
image samples of different brightness level, comprising:
a module A for obtaining a set of training image data comprising for each of a
30 plurality of training images an input image sample and a target image sample representing the same image but in a different brightness level,
a module B for providing an image generating model having an encoder configured to receive the input image sample and a plurality of decoder branches configured to output each a generated output sample,
35 a module C for providing a discriminator model, and

a model D for training the image generating model using the set of training image data based on a predefined loss function, wherein for each training image among the decoder branches only the decoder branch whose generated output sample has a minimum loss compared to all other decoder branches is
5 optimized based on said predefined loss function, and
wherein the training in module D is augmented by an adversarial loss which is based on the output of the discriminator model.

11. A system for generating image samples of different brightness levels,
10 comprising the trained model of step d of any one of claims 1 to 9 or of the trained module D of claim 10.

12. A computer vision system for processing an input image sample,
comprising a pre-processor module configured to generate image samples of
15 different brightness level based on the input image sample, the pre-processor module comprising the trained model of step d of any one of claims 1 to 9 or of the trained module D of claim 10.

13. A computer program including instructions for executing the steps of
20 a method according to any one of claims 1 to 9 when said program is executed by a computer.

14. A recording medium readable by a computer and having recorded thereon a computer program including instructions for executing the steps of a
25 method according to any one of claims 1 to 9.

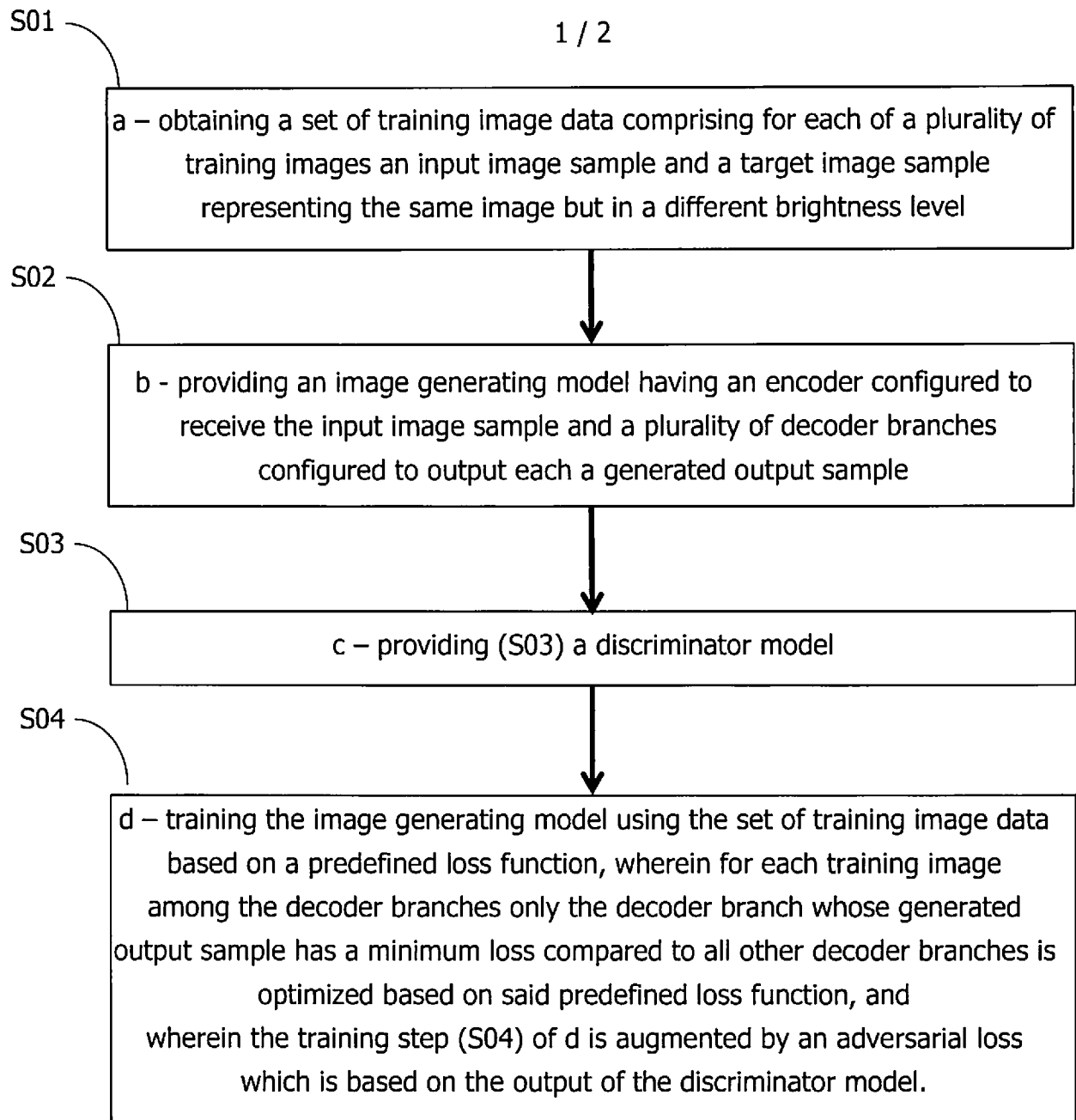


Fig. 1

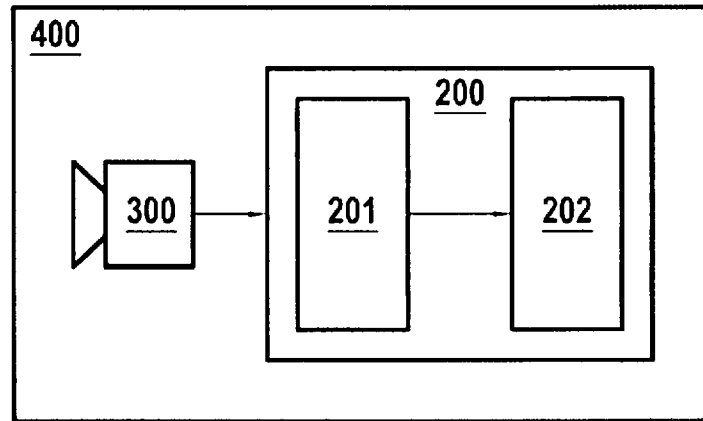


Fig. 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/063650

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06N3/04 G06N3/08
ADD.

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)
G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Wei-Sheng Lai ET AL: "Semi-Supervised Learning for Optical Flow with Generative Adversarial Networks Jia-Bin Huang 2", 1 January 2017 (2017-01-01), XP055676590, Retrieved from the Internet: URL:https://faculty.ucmerced.edu/mhyang/papers/nips17_flowGAN.pdf [retrieved on 2020-03-13] the whole document ----- -/--	1-14



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

18 March 2020

Date of mailing of the international search report

03/04/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Baldan, Marco

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2019/063650

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NGUYEN VU ET AL: "Shadow Detection with Conditional Generative Adversarial Networks", 2017 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV), IEEE, 22 October 2017 (2017-10-22), pages 4520-4528, XP033283326, DOI: 10.1109/ICCV.2017.483 [retrieved on 2017-12-22] the whole document -----	1-14
A	JOOST VAN AMERSFOORT ET AL: "Frame Interpolation with Multi-Scale Deep Loss Functions and Generative Adversarial Networks", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 16 November 2017 (2017-11-16), XP081289078, the whole document -----	1-14