



(51) International Patent Classification:

G06K 9/00 (2006.01) G06K 9/46 (2006.01)
G06N 3/04 (2006.01) G06K 9/62 (2006.01)

(21) International Application Number:

PCT/EP2019/061618

(22) International Filing Date:

07 May 2019 (07.05.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

20180100477 23 October 2018 (23.10.2018) GR

(71) Applicant: **NOKIA TECHNOLOGIES OY** [FI/FI];
Karakaari 7, 02610 Espoo (FI).

(72) Inventor: **MILIORIS, Dimitrios**; 4 rue Madeleine Bres,
75013 Paris (FR).

(74) Agent: **BERTHIER, Karine**; Alcatel-Lucent International,
Site de Nokia Paris-Saclay, Route de Villejust, 91620
NOZAY (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,
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))

(54) Title: METHODS, APPARATUS AND COMPUTER-READABLE MEDIUM TO REALIZE INDOOR LOCALIZATION

(57) Abstract: Method, computer-readable medium and apparatus to realize indoor localization is described. This method comprises a step of determining a four-dimensional tensor based on an input image and a flatten step configured to reshapes the four-dimensional tensor and a first fully connected step and a second fully connected step and a softmax function step.

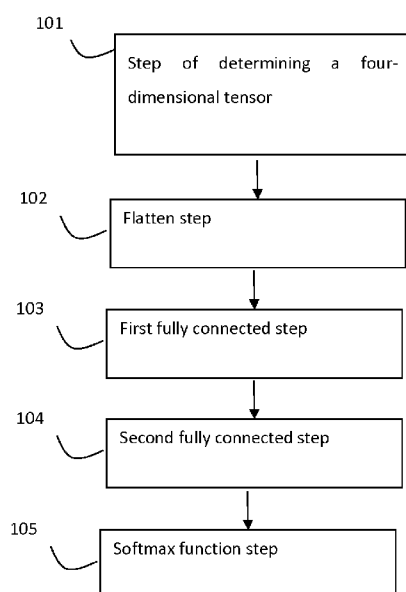


Figure 1



METHODS, APPARATUS AND COMPUTER-READABLE MEDIUM TO REALIZE INDOOR LOCALIZATION

FIELD OF INVENTION

[0001] The present subject matter relates to method and apparatus to realize indoor localisation. More precisely this method relies on deep learning and in some embodiments could contain a refinement step.

BACKGROUND

[0002] It is known that humans heavily rely on visual cues to position themselves and find their way in space. Indeed, visual diversity has for long been considered essential in building the character of a place, helping people to locate themselves in a physical space, and even improving the correlation of restorative experiences of spaces. Thus, knowing to identify and correlate spatial visual cues is a critical asset of way finding.

[0003] To realize indoor localization, it is known to use Global Positioning Systems (GPS). This system allows the user to find our way in outdoor environments. However, GPS technologies do not solve the problem of way finding in indoor environments for several reasons, including non-line-of-sight conditions and relatively smaller-scale spaces requiring spatial decisions to be taken with higher accuracy.

SUMMARY

[0004] This summary is provided to introduce concepts related to method and apparatus to realize indoor localisation.

[0005] In one implementation, a method to realize indoor localisation is described. This method comprises a step of determining a four-dimensional tensor based on an input image and a flatten step configured to reshapes the four-dimensional tensor and a first fully connected step and a second fully connected step and a softmax function step.

[0006] In one implementation another method to realize indoor localization is described. This method comprises a refinement step configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

[0007] In another implementation, apparatus to realize indoor localisation is described. This apparatus comprises means for determining a four-dimensional tensor based on an input image and means for flatten, these means being configured to reshapes the four-dimensional tensor and means for realizing a first fully connected layer and means for realizing a second fully connected layer and means for realizing a softmax function.

[0008] In another implementation, computer-readable medium is described. This computer-readable medium comprises a method to realize indoor localization. This method comprises a step of determining a four-dimensional tensor based on an input image and a flatten step configured to reshapes the four-dimensional tensor and a first fully connected step and a second fully connected step and a softmax function step.

BRIEF DESCRIPTION OF THE FIGURES

[0009] The detailed description is given with reference to the accompanying figures. In the figures, the left-most digit(s) of a reference number identifies the figure in which the reference number first appears. The same numbers are used throughout the figures to reference like features and components. Some embodiments of systems and/or methods in accordance with embodiments of the present subject matter are now described, by way of example only, and with reference to the accompanying figures, in which:

[0010] Figure 1 presents one embodiment of a method to realize indoor localization.

[0011] Figure 2 presents one embodiment of a method to realize indoor localization.

[0012] Figures 3-a & 3-b present two embodiments of a method to realize indoor localization.

[0013] Figures 4-a & 4-b present two embodiments of a method to realize indoor localization.

[0014] In the present document, the word "exemplary" is used herein to mean "serving as an example, instance, or illustration." Any embodiment or implementation of the present subject matter described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other embodiments.

[0015] It should be appreciated by those skilled in the art that any block diagrams herein represent conceptual views of illustrative systems embodying the principles of the present subject matter. Similarly, it will be appreciated that any flow charts, flow diagrams, state transition diagrams, pseudo code, and the like represent various processes which may be substantially represented in computer readable medium and so executed by a computer or processor, whether or not such computer or processor is explicitly shown.

DESCRIPTION OF EMBODIMENTS

[0016] The method and apparatus described in the present subject matter can for example be used in the robotics area

[0017] In one embodiment illustrated figure 1, a method is proposed, this method comprises

- A step 101 of determining a four-dimensional tensor based on an input image
- A flatten step 102 configured to reshape the four-dimensional tensor
- A first fully connected step 103
- A second fully connected step 104
- A softmax function step 105

[0018] In a convolutional neural network, a fully connected layer is a layer having connections to all activations in the previous layer, as seen in regular neural networks. Their activations can hence be computed with a matrix multiplication followed by a bias offset.

[0019] The methods of the present divulgation used a neural network that has been previously trained with a set of images each image being associated with a localization.

[0020] In an embodiment, the first fully connected step 103 is configured to learn the non-linear relations and determine the right class of the input image.

[0021] In an embodiment, the second fully connected step 104 configured to learn the non-linear relations and determine the right class of the input image.

[0022] In an embodiment of the present subject matter the step 101 of determining a four-dimensional tensor is configured to realize

- A first convolutional step 201

- A second convolutional step 202
- A third convolutional step 203

[0023] In an embodiment of the present subject matter at least one output of the first convolutional step 201 and/or the second convolutional step 202 and/or the third convolutional step 203 is a tensor with four dimensions. In an embodiment those four dimensions tensors are reduced to two dimensions tensors, in order to be fed to the fully connected layer.

[0024] In an embodiment, the first convolutional step 201 and/or the second convolutional step 202 and/or the third convolutional step 203 are configured to use 2x2 max-pooling. This 2x2 max-pooling is used to down sample the image at each step.

[0025] In an embodiment of the method of the present subject matter, biases are added to the output of the first convolutional step 201 and/or the second convolutional step 202 and/or the third convolutional step 203.

[0026] In an embodiment of the present subject matter the first fully connected step 103 and/or the second fully connected step 104 are using a model built on TensorFlow, and used 2x2 pooling with strides set to one in all dimensions.

[0027] In an embodiment, the second fully connected step 103 has as output vectors of the length equal to the number of classes. The aim of the second fully connected step 103 is to determine in which class the input image belongs.

[0028] In embodiments presented figure 3-a and figure 3-b, the method of the present subject matter also comprises a rectified linear (ReLU) step 301. This rectified linear step is used to learn the non-linear relations in the fully connected layer.

[0029] In an embodiment, the rectified linear step 301 is using the following formula:

$$f(x) = \log(1 + \exp(x))$$

One of the advantage of this formula is to add non-linearity.

[0030] In an embodiment, the rectified linear step 301 is realized between the first fully connected step 102 and the second fully connected step 103. One of the advantages of this localization of the rectified linear step is to reduce cost of the method. Indeed, the ReLU is

executed before the pooling process, but we save time and the cost of operations of the ReLU when we perform max-pooling first. In this situation, the cost is eliminated by 25%.

[0031] In embodiments presented figure 4-a and figure 4-b global average pooling step 401 follows the second fully connected step 104. In of the aims of this global average pooling step is to regularize structures, therefore avoids overfitting when compared to the fully connected layers introduced to the model.

[0032] In an embodiment and in order to predict the right class, we get the likelihood of the images which belong to each one of the classes, after a normalization routine has been applied to limit the probability between 0 and 1, and use the softmax function. The largest number is pointed in the vector and shows the predicted class. We obtain the cost by calculating the cross entropy for the total classification of images.

[0033] The output of the softmax function step 105 being the localization of the input image. The localization being the one associated to one of the images used to train the neural network model with the maximum likelihood, as determine in output of the soft max function step 105, of the input image.

[0034] In an embodiment of the present subject matter, the method is using the parameters $w(t)$ and a loss function $L(t)$, and t indexes the training iteration, starting at 1. Then, the main parameters of Adam Optimiser are given by:

$$\begin{aligned}
 m_w^{(t+1)} &= \beta_1 m_w^{(t)} + (1 - \beta_1) \nabla_w L^{(t)} \\
 v_w^{(t+1)} &= \beta_2 v_w^{(t)} + (1 - \beta_2) (\nabla_w L^{(t)})^2 \\
 \hat{m}_w &= \frac{m_w^{(t+1)}}{1 - \beta_1^t} \\
 \hat{v}_w &= \frac{v_w^{(t+1)}}{1 - \beta_2^t} \\
 w^{(t+1)} &= w^{(t)} - \eta \frac{\hat{m}_w}{\sqrt{\hat{v}_w} + \epsilon}
 \end{aligned}$$

where ϵ prevents division by 0, and β_1, β_2 are the forgetting factors for gradients and second moments of gradients, respectively.

[0035] In other words, in the previously described embodiment a hierarchical system is proposed. In this system, an image is received and classified in a building level segment. Then, we increase the area of interest by dividing the large segments into smaller ones, which leads to a higher physical space resolution. During the training phase, we use the images extracted from videos captured in large segments in order to train our model, and then in a hierarchical way of thinking, we perform the same for the room level segments, to narrow down the area and receive a more precise location.

[0036] Another object of the present subject matter is a refinement step. This step can follow the different embodiments of the method presented above but could also follows other methods of indoor localization.

[0037] If this refinement step follows the methods previously described this refinement step will be used instead of the softmax function step 105. The aim of this refinement step is to increase the localization accuracy at the vicinity of corners. This refinement step avoids the misclassification of images and therefore give better location estimation This refinement step is configured to normalize the n th components of a vector X_i of size N ($1 \leq n \leq N$) as the probability that the image is at location n .

[0038] In another embodiment and if this refinement step follows the methods previously described the refinement step can be realize in addition to the softmax function step so that it can check the probabilities and re-classify more properly.

[0039] In an embodiment this refinement step is configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

[0040] In other words for the refinement step of distance pictures or pictures taken around corners, assume we have N locations that we train the neural network. Then every image i after passing through the trained neural network generates an N vector X_i just prior to the softmax function. With normalization, we can interpret the n -th component of X_i ($1 \leq n \leq N$) as the probability that the image is at location n .

[0041] The softmax just takes the maximum value and locates at that place. Having a vector of probabilities can be used to make corrections based on a prior distribution of the ground truth (pictures which are correctly classified according to our training).

[0042] Example: If we take a picture 10 meters away from the path, then this vector is not going to be like (0.001, 0.01, ..., 0.93, 0.02) but more like (0.25, ..., 0.4, .. 0.15) and taking the max this time is not the best thing to do. That can be used for detecting the usage of the patent along with the system.

[0043] Another caveat is that unless the landmark is in the image, the space is not identified. So, in our case the goal is not simply find the most efficient system to locate someone is space, but propose a system that regain the human capability of locating oneself in space by reading visual cues and spatial features and matching such inputs with previous visual experiences. In the case of computer-vision techniques, this visual experience of space takes the form of large image datasets.

[0044] In another embodiment, an apparatus to realize indoor localization is proposed. This apparatus comprises:

Means for determining a four-dimensional tensor based on an input image

Means for flatten, these means being configured to reshapes the tensors

Means for realizing a first fully connected layer

Means for realizing a second fully connected layer

Means for realizing a softmax function.

[0045] In another embodiment, an apparatus comprising means for realizing a refinement, these means being configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

[0046] In an embodiment the apparatus described in the previous section comprises prior to means for realizing a refinement, means for determining a four-dimensional tensor based on an input image, means for flatten, these means being configured to reshapes the tensors, means for realizing a first fully connected layer, means for realizing a second fully connected layer.

[0047] In another embodiment, a second apparatus is proposed. This second apparatus comprises:

- at least one processor; and
- at least one memory including computer program code, the at least one memory and computer program code configured to, with the at least one processor, cause the apparatus:
- 5 • determining a four-dimensional tensor based on at least one input image
- Flattening the four-dimensional tensor configured to reshape the tensors of the output of the second and third layer
- A first fully connected step 103
- 10 • A second fully connected step 104
- A softmax function step 105.

[0048] In other embodiments, these previously presented apparatus can be organized as a specific circuit, or computer or can be a functional circuit, part of a common server or any partition in between.

[0049] As used in this application, the term “circuitry” may refer to one or more or all of the following:

- hardware-only circuit implementations (such as implementations in only analog and/or digital circuitry) and
- combinations of hardware circuits and software, such as (as applicable):
- 20 • a combination of analog and/or digital hardware circuit(s) with software/firmware and
- any portions of hardware processor(s) with software (including digital signal processor(s)), software, and memory(ies) that work together to cause an apparatus, such as a mobile phone or server, to perform various functions) and
- hardware circuit(s) and or processor(s), such as a microprocessor(s) or a portion of a
- 25 microprocessor(s), that requires software (e.g., firmware) for operation, but the software may not be present when it is not needed for operation.

[0050] This definition of circuitry applies to all uses of this term in this application, including in any claims. As a further example, as used in this application, the term circuitry also covers an implementation of merely a hardware circuit or processor (or multiple processors) or

portion of a hardware circuit or processor and its (or their) accompanying software and/or firmware. The term circuitry also covers, for example and if applicable to the particular claim element, a baseband integrated circuit or processor integrated circuit for a mobile device or a similar integrated circuit in server, a cellular network device, or other computing or network device.

[0051] A person skilled in the art will readily recognize that steps of the methods, presented above, can be performed by programmed computers. Herein, some embodiments are also intended to cover program storage devices, for example, digital data storage media, which are machine or computer readable and encode machine-executable or computer-executable programs of instructions, where said instructions perform some or all the steps of the described method. The program storage devices may be, for example, digital memories, magnetic storage media, such as magnetic disks and magnetic tapes, hard drives, or optically readable digital data storage media.

[0052] Another embodiment of the present subject matter is a computer-readable medium having embodied thereon a computer program configured to realize the method, presented figure 1, to realize indoor localization. The method comprises the following steps:

- A step 101 of determining a four-dimensional tensor based on an input image
- A flattening step 102 configured to reshape the four-dimensional tensor
- A first fully connected step 103
- A second fully connected step 104
- A softmax function step 105.

[0053] Another embodiment computer-readable medium having embodied thereon a computer program configured to realize a method to realize indoor localization, the method comprising a refinement step configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

[0054] In another embodiment of the computer-readable medium of the previous section also comprises, prior to the refinement step, the step (101) of determining a four-dimensional tensor based on an input image a flatten step (102) configured to reshapes the four-dimensional tensor a first fully connected step (103) a second fully connected step (104) as described within the previous sections.

CLAIMS

1. Method to realize indoor localisation comprising:

- 5
- a step (101) of determining a four-dimensional tensor based on an input image
 - a flatten step (102) configured to reshapes the four-dimensional tensor
 - a first fully connected step (103)
 - a second fully connected step (104)
 - a softmax function step (105).

10

2. Method according to claim 1 wherein the step (101) of determining the four-dimensional tensor is configured to realize

- A first convolutional step (201)
- A second convolutional step (202)
- 15 • A third convolutional step (203).

3. Method according to the claim 1 or 2 also comprising a rectified linear step (301).

4. Method according to claim 1 or 2 or 3 wherein the rectified linear step (301) is realized between
20 the first fully connected step (102) and the second fully connected step (103).

5. Method according to claim 1 or 2 or 3 or 4 also comprising global average pooling step (401) following the second fully connected step (104).

25 6. Method to realize indoor localization comprising

- a refinement step configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

7 Method according to claim 6 also comprising, prior to the refinement step, the step (101) of
 5 determining a four-dimensional tensor based on an input image a flatten step (102) configured to
 reshapes the four-dimensional tensor a first fully connected step (103) a second fully connected
 step (104) as described within the claim 1 to 5.

8. Apparatus comprising:

10 Means for determining a four-dimensional tensor based on an input image
 Means for flatten, these means being configured to reshapes the four-dimensional tensor
 Means for realizing a first fully connected layer
 Means for realizing a second fully connected layer
 Means for realizing a softmax function.

15

9. Apparatus comprising means for realizing a refinement, these means being configured to
 determine for an input image the probability of being at a set of different locations from which
 images as been used to train a neural network model.

20 10 Apparatus according to claim 9 comprising prior to means for realizing a refinement, means for
 determining a four-dimensional tensor based on an input image, means for flatten, these means
 being configured to reshapes the tensors, means for realizing a first fully connected layer, means
 for realizing a second fully connected layer as described claim 8.

25 11. Computer-readable medium having embodied thereon a computer program configured to
 realize a method to realize and indoor localization, the method comprising:

- A step (101) of determining a four-dimensional tensor based on an input image
- A flatten step (102) configured to reshapes the four-dimensional tensor
- A first fully connected step (103)
- A second fully connected step (104)
- 5 • A softmax function step (105)

12 Computer-readable medium having embodied thereon a computer program configured to realize a method to realize and indoor localization, the method comprising:

- 10 • a refinement step configured to determine for an input image the probability of being at a set of different locations from which images as been used to train a neural network model.

13 Computer-readable medium to claim 12 also comprising, prior to the refinement step, the step (101) of determining a four-dimensional tensor based on an input image a flatten step (102) configured to reshapes the four-dimensional tensor a first fully connected step (103) a second fully
15 connected step (104) as described within the claim 11.

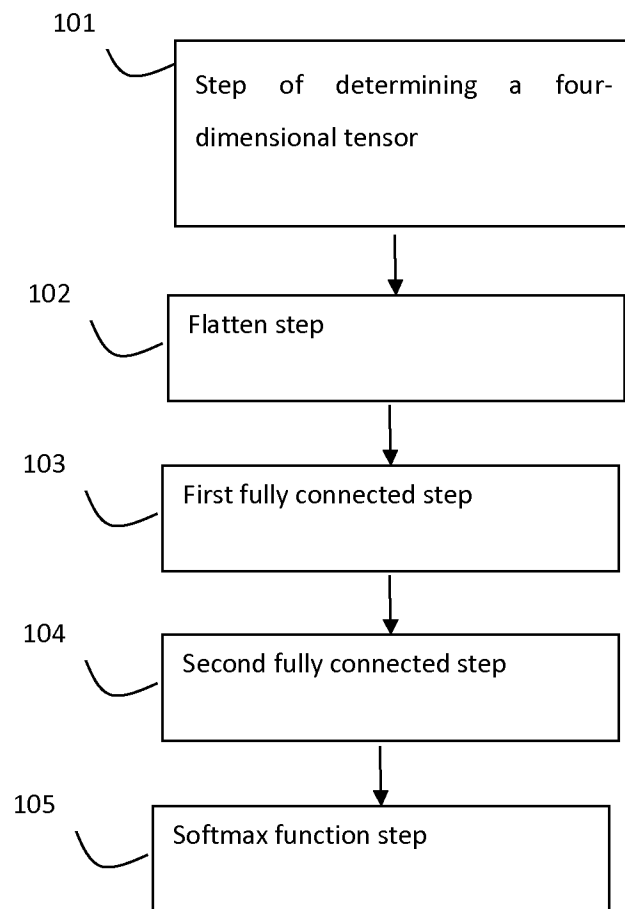


Figure 1

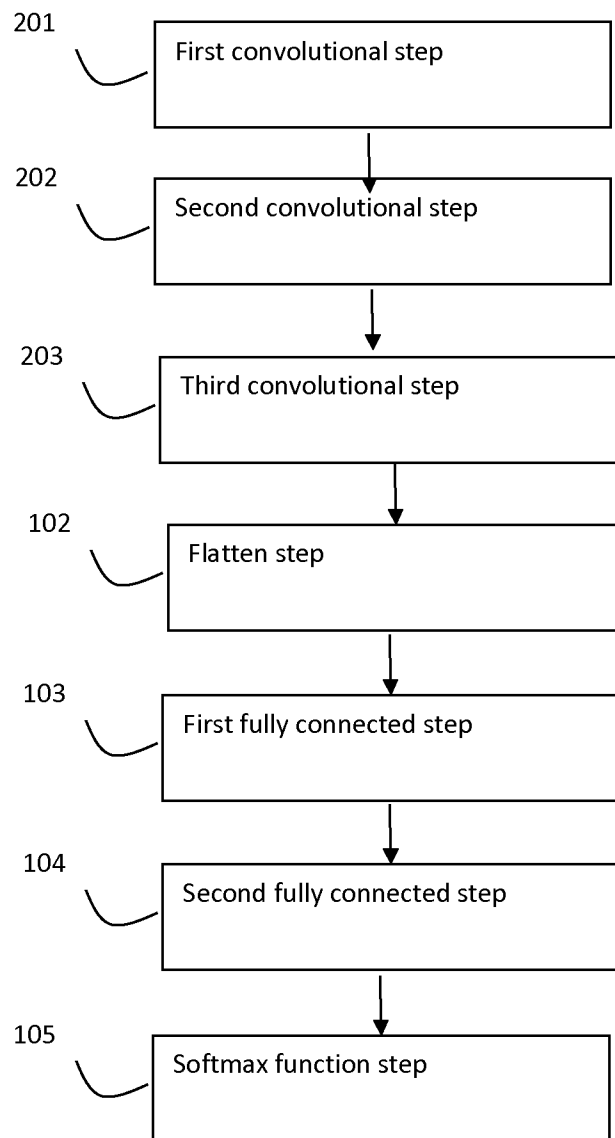


Figure 2

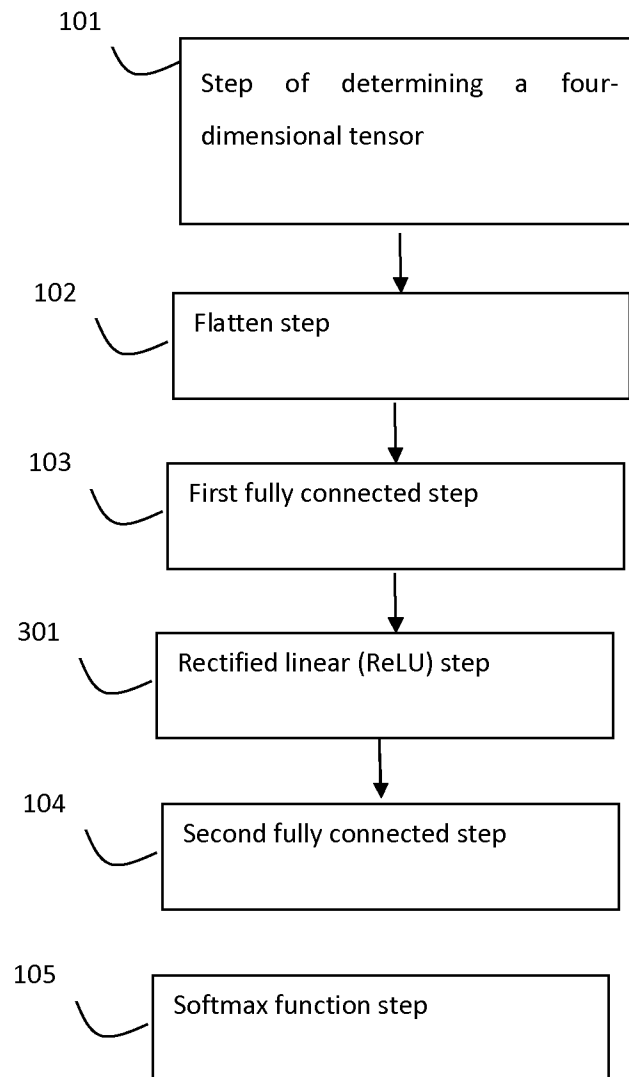


Figure 3-a

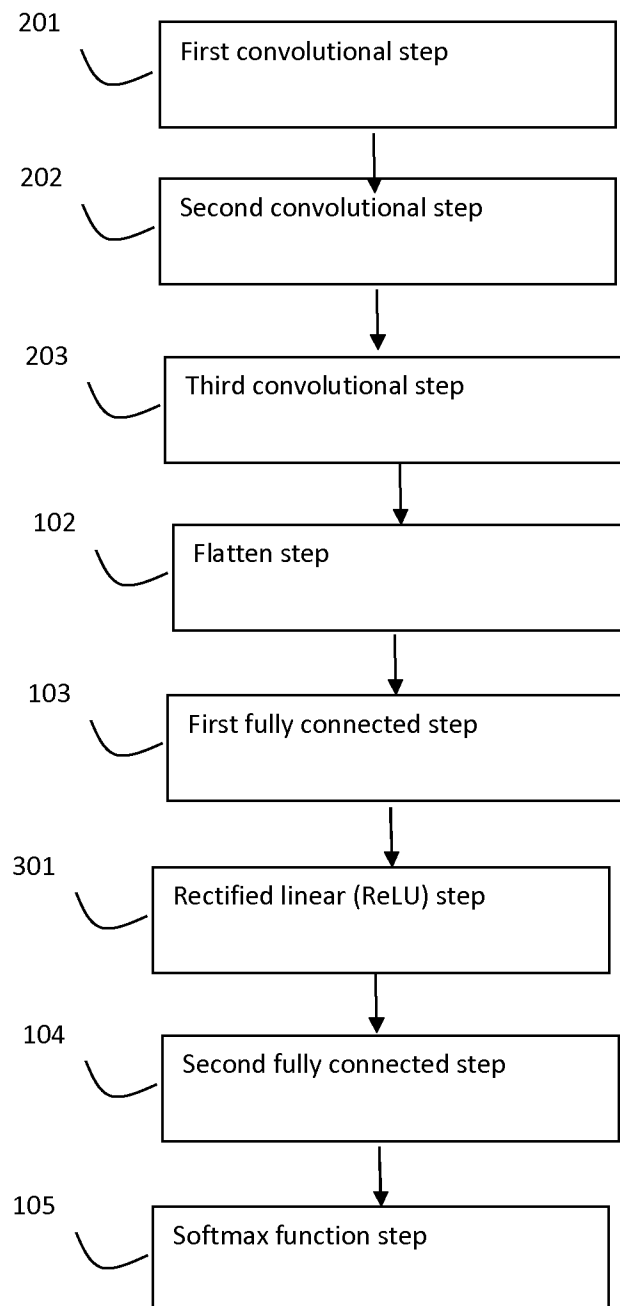


Figure 3-b

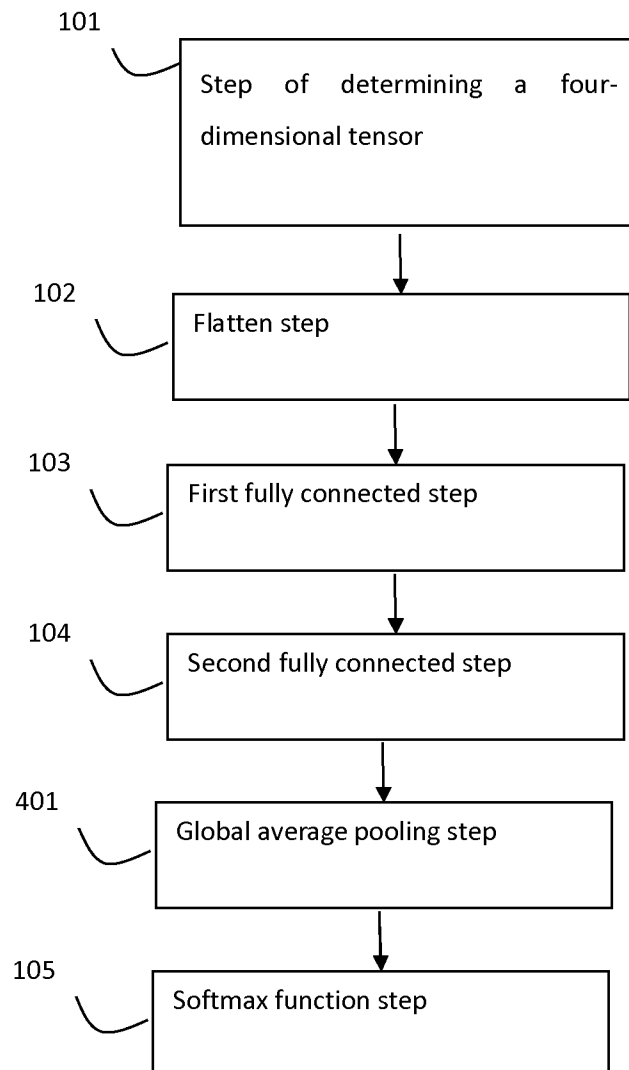


Figure 4-a

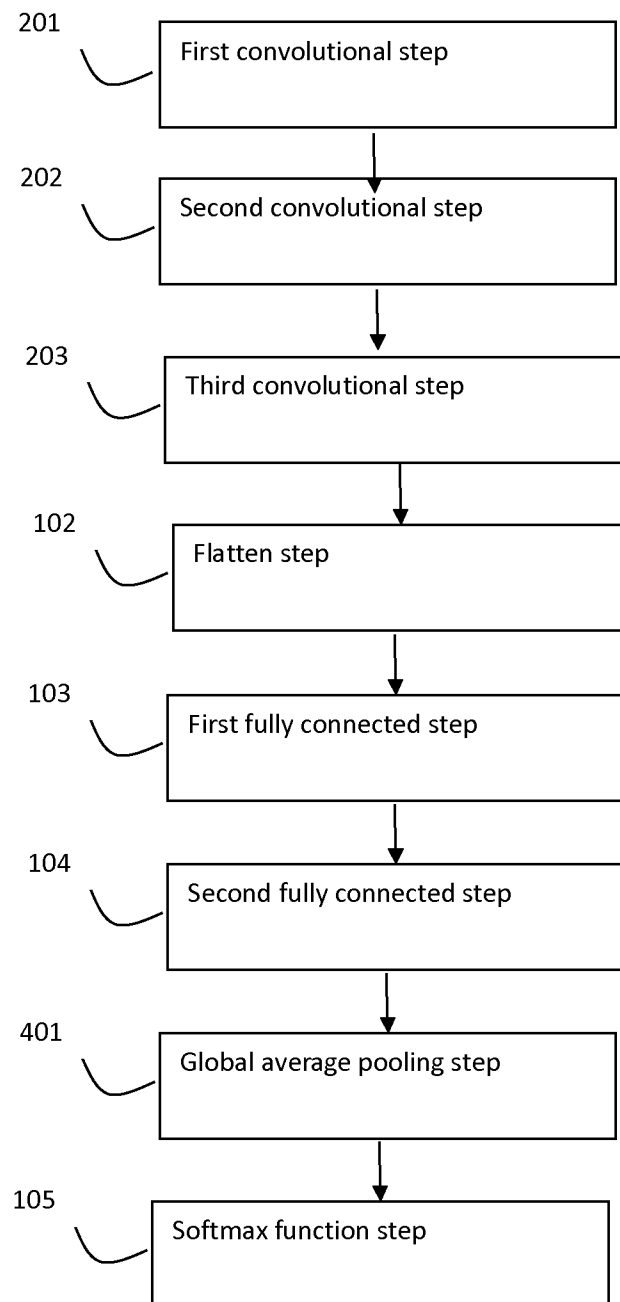


Figure 4-b

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/061618

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06K9/00

ADD. G06N3/04 G06K9/46 G06K9/62

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 G06T 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	Fan Zhang ET AL: "Indoor Space Recognition using Deep Convolutional Neural Network: A Case Study at MIT Campus", arXiv (Cornell University Library), 7 October 2016 (2016-10-07), XP055374046, Retrieved from the Internet: URL:https://arxiv.org/ftp/arxiv/papers/1610/1610.02414.pdf [retrieved on 2017-05-18] abstract figures 2,3 section 2 -----	1-13



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

17 July 2019

Date of mailing of the international search report

24/07/2019

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

Martinez, Francis