



(51) International Patent Classification:  
G06K 9/00 (2006.01) G06K 9/62 (2006.01)

(21) International Application Number:  
PCT/EP2019/065629

(22) International Filing Date:  
14 June 2019 (14.06.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10 2018 116 036.4  
03 July 2018 (03.07.2018) DE

(71) Applicant: CONNAUGHT ELECTRONICS LTD.  
[IE/IE]; Dunmore Road, Tuam, County Galway (IE).

(72) Inventor: SHIVAMURTHY, Swaroop; c/o Connaught Electronics Ltd, Dunmore Road, County Galway, Tuam (IE).

(74) Agent: JAUREGUI URBahn, Kristian; Valeo Schalter und Sensoren GmbH, Laiernstr. 12, 74321 Bietigheim-Bissingen (DE).

(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.

(54) Title: TRAINING A DEEP CONVOLUTIONAL NEURAL NETWORK FOR INDIVIDUAL ROUTES

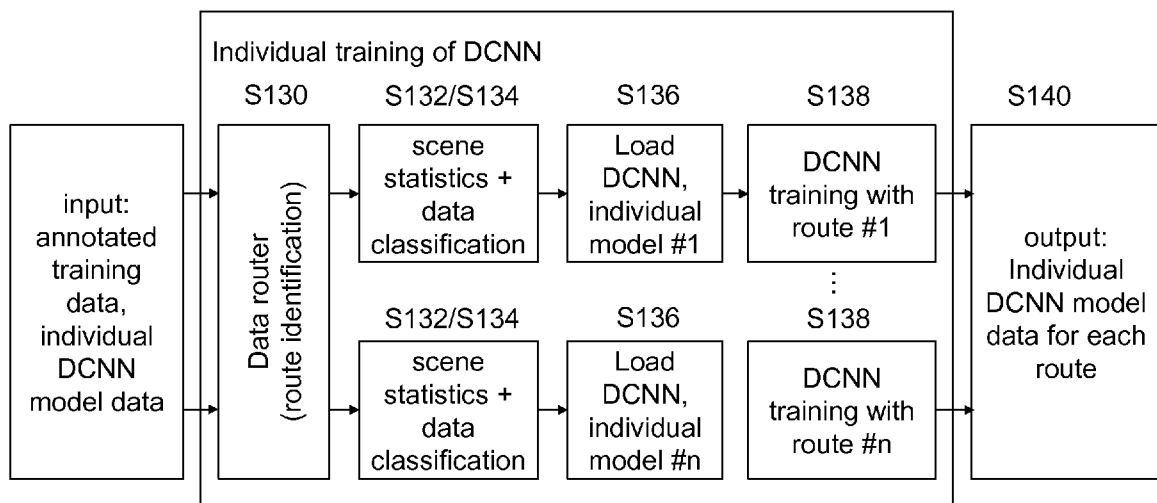


FIG. 3

(57) Abstract: The present invention refers to a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information, comprising the steps of providing an initially trained deep convolutional neural network with general model data, providing a set of annotated training data including route information, setting up the deep convolutional neural network with individual model data associated to the route based on the general model data, performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data, and storing the individual model data of the deep convolutional neural network data associated to the route. The present invention also refers to a method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route, comprising the steps of identifying the route, setting up the deep convolutional neural network with individual model data associated to the route, and processing the image



(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))*

## Training a deep convolutional neural network for individual routes

The present invention refers to a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information, comprising the steps of providing an initially trained deep convolutional neural network with general model data, providing a set of annotated training data including route information,

The present invention also refers to a method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route.

Furthermore, the present invention refers to a driving support system for performing any of the above methods.

Autonomous and semi-autonomous driving is becoming a more and more important issue in automotive industry. Prototypes for autonomous driving are already developed and tested, in some places even under real driving situations. Autonomous driving is considered as a disruptive technology in the automotive sector.

Autonomous and semi-autonomous driving depends on knowledge about the environment around the vehicle. Based on this knowledge, the driving support system can e.g. identify dangers for the vehicle or third parties and act in an appropriate way to resolve dangerous driving situations. Hence, different types of environment sensors are employed in vehicles to monitor the environment around the vehicle. Such environment sensors can include any combination of sensors out of ultrasonic sensors, LiDAR based sensors, radar sensors and optical cameras.

Processing image data from optical cameras, in particular video data comprising a sequence of multiple frames per second, is very challenging. Huge amounts of data have to be processed in real time in order to reliably detect the environment of the vehicle. However, resources of the vehicle for processing the data are limited in respect to space for housing processing devices and also in respect to available electrical

power. Even when the technical issues were resolved, in order to provide vehicles at an affordable price, the resources keep limited to their price.

One powerful means in order to process image data are neural networks. Applications of neural networks for image processing are typically based on convolutional neural networks and, in particular, on deep convolutional neural networks. Usage of such network types has shown promising results at an affordable price.

When using deep convolutional neural networks, a set of questions arises while developing suitable structures of neural networks, in particular when developing them from the scratch. This includes, just to name a few, a definition of a suitable type of neural network, determining a number and a location of inputs, determining a number of hidden layers to be used in case of a deep convolutional neural network, an determining a number of output neurons required. These questions are important because in case the neural network is too large / too small, the neural network could potentially overfit / underfit the image data. Consequently, the neural network would not learn properly to generalize well out of provided training data. In such cases, even when the error on the training data is driven to a very small value, the error be too big, when new image data, i.e. image data different to the training data previously used to train the neural network, is presented to the neural network. In this case, the neural network has memorized the training data, but it has not learned to generalize out of the provided training data, which would be required for such new situations with new image data. Therefore, most important for using such neural networks is a proper training with proper training data, which can be applied to different neural networks in order to evaluate and compare the performance of the different neural networks.

Many companies are constantly involved in development of their methods and systems to enable autonomous and semi-autonomous driving applications, e.g. an autopilot feature to be provided in upcoming hardware generations in their vehicles. However, usage of neural networks in autonomous and semi-autonomous driving applications depends on providing a properly trained neural network, which can be used in the vehicles. Hence, availability of suitable training data is an essential basis for autonomous and semi-autonomous driving applications. Therefore, companies involved in developing such autonomous and semi-autonomous driving applications are constantly gathering image data through their fleet of vehicles circulating on roads under

different conditions, e.g. environment conditions, traffic conditions, and others. Unfortunately, in order to use the image data as training data, annotations are required, which is typically done by a human. Therefore, annotating the image data is very time consuming and expensive.

Furthermore, data collection plays vital role to build neural networks. If the training data is overly large, the neural network might have problems in learning up to a desired accuracy level. Hence, a designer may want to identify e.g. most significant pieces of training data from the overall set of training data. Such significant training data may refer to images containing a high number of features to be detected. Accordingly, only those most significant pieces of training data are used for training the neural network, and insignificant training data is removed from the set of training data.

Driving support is most desired for most boring paths to drive. This refers in particular to driving routes, which are frequently driven, sometimes even on a daily basis. Such frequent commute paths comprise by way of example a route from a home to an office or from the office back home, a route from home to a supermarket, a route from home to a park etc. Since these are the routes most frequently driven, most accidents occur on these routes, and there is an increased need also to safely drive the frequent commute paths. Furthermore, these frequent commute paths potentially bore a driver or owner of the vehicle, so that autonomous driving functions, in particular an autopilot feature, for driving the vehicle are highly desired.

Therefore, it is essentially important to enable driving these commute paths autonomously and safely.

In this context, document EP 2 213 980 B1 refers to a method for operating a navigation device of a motor vehicle. The navigation device determines the position and/or travel route of the motor vehicle with the aid of geographic base data available on the motor vehicle side. At least one geographic information about a region of the motor vehicle environment is determined from the sensor data of at least one sensor device on the vehicle side. The least one geographic information is stored in a storage device as an additional geographical data and is used to supplement and/or update the geographical base data.

Document DE 100 30 932 A1 refers to a method, in which a vehicle with a positioning system records its position over a route and records it in memory. The measured route is transferred to a database and a computer processes the data, ideally with other vehicle data, to check and statistically analyze the data to yield a reworked street map. The latter can then be transmitted over telecommunications to users.

According to document DE 102 58 470 B4, a navigation system for a motor vehicle comprises a database with a digital map and a position detection unit. The vehicle also has instrumentation for sensing vehicle operating parameters and or its surroundings. These parameters can be used to derive information about the type of road and or its direction. This information is then compared with available data derived from the digital map data and can be used to update the database if necessary. The motor vehicle navigation database is updated using a learning method in which sensed vehicle data is used to update current database data.

Document DE 103 54 910 B4 refers to a self-control traveling system for an expressway and a method for controlling the same. A plurality of road signs are consecutively installed at left and right sides of the expressway. An imaging device images the road signs arranged along the expressway. An image processor analyzes at least one image, discriminates the roadside signs and reads sign contents contained in the roadside signs. A navigation database stores map information of the expressway. An artificial intelligence-electronic control unit determines a current position and a road state on the basis of the map information stored in the navigation database and the sign contents read by the image processor. An integrated ECU controls a steering operation, a speed decrement/increment operation and others associated with the self-control traveling operation according to a result of the determination.

Document DE 199 16 967 C1 refers to a method for updating a data-processing-supported traffic route network map that includes route data about a traffic-route network and associated attribute data. By means of at least one test vehicle, current route and attribute data are taken from the respective trafficable section of the route, and the traffic route network map is updated on the basis of this recorded data.

Document DE 10 2013 211 696 A1 relates to a method for completing and / or updating a digital road map, comprising the steps of: a) acquiring a digital image of a route

section by means of a camera of a motor vehicle located on the route section; b) determining a current position of the motor vehicle in the width direction of the track section and / or at least one property of the track section based on the image by means of a computing device of the motor vehicle; c) transmission of a data set, which includes information about a geographical position of the capture of the image and information about the at least one determined property of the link and / or the position in the width direction, to a server device. The steps a) to c) are performed by a plurality of motor vehicles in the region and the digital road map is completed and / or updated on the basis of the received data records by means of the server device.

Document DE 10 2015 010 542 A1 relates to a method for generating a digital environment map for a vehicle, wherein by means of environmental sensors of the vehicle and/or at least one further vehicle, an environmental detection is performed. Based on acquired sensor data, map information is generated which is stored in the environment map. The environment map is updated automatically, whereby objects and/or obstacles located in the environment are detected by means of at least one environmental sensor of the vehicle and/or of the at least one further vehicle. During the detection of the objects and/or obstacles, a route actually traveled by the vehicle and/or by the at least one further vehicle is detected. Depending on a comparison of positions of the actually traveled route with positions of a route in the area map, updated map information is generated and stored in the area map.

It is an object of the present invention to provide a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information, a method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route, and a driving support system for performing any of the above methods of the above-mentioned type, which enable an efficient training of a deep convolutional neural network and improving performance when applying the deep convolutional neural network. It is a further object to enable efficient training of a DCNN installed in a vehicle or a driving support system.

This object is achieved by the independent claims. Advantageous embodiments are given in the dependent claims.

In particular, the present invention provides a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information, comprising the steps of providing an initially trained deep convolutional neural network with general model data, providing a set of annotated training data including route information, setting up the deep convolutional neural network with individual model data associated to the route based on the general model data, performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data, and storing the individual model data of the deep convolutional neural network data associated to the route.

The present invention also provides a method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route, comprising the steps of identifying the route, setting up the deep convolutional neural network with individual model data associated to the route, and processing the image data acquired when driving the route using the deep convolutional neural network with individual model data associated to the route.

The present invention further provides a driving support system for performing any of the above methods.

The basic idea of the invention is to provide a simple but efficient improvement for deep convolutional neural networks (DCNN), in particular for every-day driving situations, which refer to repeatedly driving the same routes, also referred to as commute paths. Such commute paths comprise by way of example a route from a home to an office or from the office back home, a route from home to a supermarket, a route from home to a park etc. Hence, individual training of the deep convolutional neural network is performed for these routes, so that the training step provides respective model data for the deep convolutional neural network, which is highly adapted to driving these routes. With the particular training of the DCNN for these routes, which cover a major part of driving the vehicle, accidents occurring on these routes can be improved. This is in particular valid, since driving these routes is considered as boring by the people, e.g. the owner or the driver of the vehicle. Attention of the driver or owner of the vehicle is frequently reduced when being bored, so that also autonomous functions, in particular an autopilot feature for autonomous driving, are highly desired to increase traffic safety.

The neural network trained with the proposed method is a deep convolutional neural network (DCNN) for processing image data. DCNN are a particular implementation of more general convolutional neural networks (CNN). CNNs comprise an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers. The convolutional layers apply a convolution operation to the input, passing the result to the next layer. The convolution emulates the response of an individual neuron to visual stimuli.

Such DCNNs are used more and more in a context of driving applications in vehicles. Hence, the DCNNs are employed e.g. in driving support systems of vehicles. The driving support systems comprise driver assistance systems, which assist a human driver of the vehicle in different driving situations, or systems, which support autonomous driving of the vehicle by providing input to reliably deal with different traffic situations.

The step of providing an initially trained deep convolutional neural network with general model data refer to providing a previously trained DCNN. The DCNN can be trained previously to being transferred to the vehicle with standard training data. This standard training data is independent from routes driven by a driver or owner of the vehicle. The previously trained DCNN is ready for use in the driving support system.

The step of providing a set of annotated training data including route information refers to providing the training data with labels in respect to shown features. The labels can be added manually, e.g. by a human person, or automatically using e.g. an already trained DCNN or any other technology. The route information refers to any kind of identification of the route. The identification can be added manually, e.g. by the owner or driver of the vehicle, or automatically, e.g. based on position data.

The step of setting up the deep convolutional neural network with individual model data associated to the route based on the general model data refers to providing the DCNN for training the DCNN for a particular route.

Performing the training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data

refers to training the deep convolutional neural network as set-up with the individual model data using the respective set of annotated training data, which refers to this route. Accordingly, a particular training covering possible features along this route can be performed, thereby providing a highly efficient DCNN when driving this route.

The step of storing the individual model data of the deep convolutional neural network data associated to the route refers to storing the individual model data for later use when using the vehicle to drive the particular route. Preferably, a verification is performed to verify if an error on the training data using the individual model data is lower than the error of previous individual model data.

The method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route refers to using the vehicle with the driving support system employing the DCNN.

The step of identifying the route can be performed in advance, e.g. when performing a route planning using a navigation system of the vehicle. Alternatively, the route can be identified when driving the vehicle, e.g. based on position data, in particular satellite position data, based on a recognition of landmarks, or based on any other recognition of the route.

The step of setting up the deep convolutional neural network with individual model data associated to the route refers to activating the DCNN with the individual model data of the respective route.

The step of processing the image data acquired when driving the route using the deep convolutional neural network with individual model data associated to the route refers to an application of the DCNN in the driving support system.

According to a modified embodiment of the invention, the step of providing a set of annotated training data including route information comprises identifying a driving route, providing image data as training data from at least one camera when driving the route, and automatically annotating the image data to provide the set of annotated training data. The driving route can be identified based on a user input or automatically based e.g. on location information of at least a starting point and an end point of the route. The

route can be identified prior to driving the route, or afterwards. The image data can be provided when driving the route with a particular vehicle, e.g. the vehicle having implemented the DCNN. However, the image data can also be provided when driving the route with different vehicles, even without the particular vehicle having implemented the DCNN. Furthermore, the step of automatically annotating the image data to provide the set of annotated training data enables generation of the set of annotated training data without human interaction. Hence, creation of valid training data can be improved, and time and money typically spent for manual annotation of training data can be reduced. The step of automatically annotating the image data to provide the set of annotated training data can be performed at any place, e.g. using the driving support system of the vehicle. Hence, the entire step of providing the set of annotated training data including route information can be performed with the vehicle and within the vehicle. No data transmission or data transfer between the driving support system and an external server device is required.

According to a modified embodiment of the invention, the step of providing image data from at least one camera when driving the route comprises tagging the provided image data with position information, in particular with satellite based position information. Apart from identifying the route driven with the vehicle when generating the image data, the position information can be further used e.g. for the purpose of composing particular routes based on combinations of chunks of routes, which have been driven previously in order to gather training data and to provide the set of annotated training data. Accordingly, it is not even required to drive a particular route in order to provide a set of annotated training data including route. With a sufficiently large data base of image data, a set of annotated training data can be provided for essentially any route. The set of annotated training data can be provided immediately, already at the time when using the vehicle for a first time. Still further, the position information can be used to process the image data to determine different conditions for providing the image data from the same location. Hence, image data provided at similar locations can be identified and filtered to select image data with different characteristics, e.g. different lighting condition, different weather conditions, different traffic conditions, or others. This will reduce data redundancy in the set of annotated training data, and the training step of the DCNN can be performed in an efficient way. The image data can be tagged with position information based on location information provided from e.g. GPS, Galileo, GLONASS, Beidou satellite location systems.

According to a modified embodiment of the invention, the step of automatically annotating the image data comprises the steps of processing a set of not annotated training data with the deep convolutional neural network to generate a set of automatically labeled data, calculating confidence metrics for the set of automatically labeled data, and automatically relabeling the set of automatically labeled data based on the calculated confidence metrics to generate the set of annotated training data.

The proposed usage of the DCNN to automatically annotate the image data is simple and runtime efficient, since the usage of the initially trained DCNN for processing the image data to generate the set of automatically labeled data already takes advantage of the DCNN. The confidence metrics can also be calculated automatically and in an efficient way. The confidence metrics for the set of automatically labeled data enables a simple evaluation of the relevance of the labels created by the DCNN. In case the confidence metrics indicates a high confidence level in the automatically labeled data generated by the DCNN, the automatic relabeling step does not change the labels. Otherwise, it is required to change the labeling and to automatically relabel the set of automatically labeled data based. This can refer to modifying all labels, or only to modifying some labels assigned to the set of automatically labeled data. The relabeling is not as efficient as processing the image data using the DCNN. However, the effort of relabeling, i.e. using the less efficient labeling, is only required in cases, where the DCNN cannot process the set of not annotated training data in a sufficiently reliable way, as indicated by the confidence metrics. The deep convolutional neural network can be the initially trained deep convolutional neural network with the general model data, or the deep convolutional neural network with the individual model data associated to the route, to which the image data corresponds.

According to a modified embodiment of the invention, the step of calculating confidence metrics for the set of automatically labeled data comprises applying a computer vision based confidence metric algorithm. Computer vision is in general powerful and has already achieved a high quality level for recognizing features thereof. In this area, different methods are known in order to calculate confidence metrics, which are applied to calculate the confidence metrics for the set of automatically labeled data.

According to a modified embodiment of the invention, the step of automatically relabeling the set of automatically labeled data based on the calculated confidence metrics comprises applying a computer vision based labeling algorithm. Computer vision is in general powerful and has already achieved a high quality level for recognizing features thereof. Hence, in case the confidence of the automatic labeling by the DCNN is not sufficient, computer vision can be used to add necessary labels to the respective image data.

According to a modified embodiment of the invention, the step of setting up the convolutional neural network with individual model data associated to the route comprises generating the individual model data associated to the route as a copy of the general model data or loading the individual model data of the deep convolutional neural network data associated to the route as previously stored. When the method is performed for the first time, no individual model data is available. Since the DCNN has received a suitable training in order to provide the general model data, this general model can be taken as starting point for training the DCNN for the particular route. Once, such a training for a particular route has been performed, individual model data of the deep convolutional neural network data associated to the route has been generated and stored, so that it is available. Hence, this individual model data can be used as basis for a further training of the DCNN for the particular route.

According to a modified embodiment of the invention, performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data comprises training last layers of the deep convolutional neural network, in particular only the last layer. Training only the last layers, in particular the last layer of the DCNN reduces a number of parameters to be stored for the individual model data. Given the general model data of the DCNN, which has been generated based on a training, the DCNN can be assumed to be working properly, at least in general. However, given the specific routes, further improvement can be achieved for these specific routes without modification of the upper layers of the DCNN, i.e. with only minor adaptations of the DCNN based on the general model data. Accordingly, the training step can even be performed in systems having processing constraints, either processing power or memory. Further preferred, the step of storing the individual model data of the deep convolutional neural network data associated to the route comprises storing only the last layers of the deep convolutional neural network,

in particular only the last layer. The remaining model data can be taken from the general model data provided for the DCNN. Accordingly, there is only few storage memory required to store the individual model data for each route, which can be important in systems having memory constraints. Hence, also when setting up the convolutional neural network with individual model data associated to the route, it is only required e.g. to generate the individual model data or to load the individual model data of the DCNN for the only the last layers thereof, in particular only for the last layer.

For systems without such constraints, it can be preferred to perform the training step for all layers of the DCNN in order to further improve the overall performance of the DCNN.

According to a modified embodiment of the invention, the method is implemented as an in-vehicle method, whereby the training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data is performed for training the deep convolution neural network as provided for a particular vehicle. Hence, the vehicle, in particular its driving support system, can be adapted e.g. to an individual owner or driver of the vehicle and to particular routes repeatedly driven with the vehicle. Accordingly, the DCNN will be best trained for the needs of the owner or driver of the vehicle to provide a highest possible level of safety and comfort for each owner or driver of the vehicle. The set of annotated training data can be provided from externally, e.g. when the owner or driver of the vehicle defines preferred routes and receives the set of annotated training data for these routed from an external provider. This enables an immediate training of the DCNN. In this case, the set of annotated training data can be provided from a storage device connected to the vehicle or via a data connection between a server device and the vehicle. The data connection is a wired or wireless data connection. The server device can be provided e.g. as cloud server via the internet.

Performing the training step is time consuming and can require full processing power of the driving support system of the vehicle. Hence, when performing the training step, it can be required that the vehicle is driven manually to reuse the same hardware of the driving support system. No dedicated hardware for training the DCNN is required. To avoid such situations, it is preferred that the step of training the DCNN is performed, when the vehicle is not moving to ensure full support of the driving support system.

According to a modified embodiment of the invention, the step of providing a set of annotated training data including route information comprises providing the set of annotated training data including route information using at least one camera of the vehicle. Hence, the annotated training data is provided based on image data generated using the camera of the vehicle, for which the DCNN is trained. Hence, the method can be performed entirely within the vehicle, i.e. by the driving support system of the vehicle. Accordingly, the method can be performed as an in-vehicle method. This has the additional advantage that image data for training the DCNN is provided using the same hardware like when using the DCNN. Hence, influences of the hardware on the performance of the DCNN can be excluded. A satellite receiver of the vehicle for receiving satellite position data of the vehicle can be used to identify the route and/or to tag the image data.

According to a modified embodiment of the invention, the step of providing the set of annotated training data including route information using at least one camera of the vehicle comprises accumulating training data and evaluating when the set of annotated training data comprises sufficient training data. Training data from driving the route only once typically does not provide a suitable set of annotated training data. Hence, it is required to provide training data from driving the route multiple times in order to obtain a suitable set of annotated training data. The training data can be accumulated e.g. when the owner of the vehicle drives a route multiple times using his vehicle, i.e. when the vehicle is used in typical driving conditions. The method can comprise an additional step of requesting a driver or owner of the vehicle to start performing the training step of the deep convolutional neural network with the accumulating training data upon successful evaluation of the set of annotated training data comprising sufficient training data. The step of requesting the driver or owner of the vehicle to start performing the training step of the deep convolutional neural network with the accumulating training data upon successful evaluation of the set of annotated training data comprising sufficient training data can be performed independently for each route. Hence, e.g. on a frequency of driving a route, the training step can be performed at different moments of time.

According to a modified embodiment of the invention, the step of providing the set of annotated training data including route information using at least one camera of the vehicle comprises filtering the training data based on a statistical analysis of the training data. It is helpful to provide the training data to the DCNN at different conditions on order

to obtain an efficiently trained DCNN. Hence, the training data can be provided e.g. at traffic congestions, with differently textured road surfaces, at different lighting conditions, or any different traffic scenarios like overtaking, lane changing, roundabout, parking areas, or others. Hence, based on the statistical analyses, filtering the training data enables an efficient training of the DCNN, since the different conditions assure that the DCNN is properly trained. The statistical analyses can be performed e.g. using a histogram of individual features of the respective training data. Further preferred, the step of filtering the training data based on a statistical analysis of the training data.

According to a modified embodiment of the invention, the step of performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data comprises the steps of determining scene statistics for the set of annotated training data, classifying the set of annotated training data based on the scene statistics, and automatically selecting training data out of the set of annotated training data based on the classification. The scene statistics refers to a statistical determination of features shown in the training data. The features can be any kind of relevant features, which can be detected. The features can comprise the annotations of the training data, i.e. labels added thereto. The features can refer to static features without movement information, or dynamic features including movement information, e.g. pedestrian walking or, even more detailed, pedestrian walking from left to right. The classification of the training data refers to a classification based on occurrences of features within the training data. The classification can e.g. identify training data showing a particular type of features and/or a particular number of features. By way of example, such classifications can include training data with a high density of features, training data comprising a particular type of features, e.g. pedestrians or vehicles, with or without dynamic information, e.g. not moving, moving, moving from right to left. Furthermore, the training data can be classified and selected to be separated into training data for training, for testing and for validation of the DCNN.

According to a modified embodiment of the invention, the step of determining scene statistics for the set of annotated training data comprises determining label statistics of at least one out of a type of labels, a position of labels, and/or properties of labelled objects, preferably determining label combination statistics. Hence, the labels are evaluated to determine the scene statistics. The label statistics of at least one out of a

type of labels, a position of labels, and/or properties of labelled objects refers to different aspects of the label, which can all be used to statistically identify common features of different training data. Preferably, the label statistics are determined for multiple or all out of available types of labels, positions of the labels, and/or properties of labelled objects. Determining label combination statistics can refer to statistics of a combination of multiple labels e.g. within the same image, e.g. a certain number of pedestrians or a certain number of vehicles within the image. In addition or alternatively, the label combination statistics can refer to a combination of multiple out of the type of labels, the position of labels, and/or properties of labelled objects.

According to a modified embodiment of the invention, the step of determining label statistics of a type of labels comprises performing a histogram analysis of the type of labels from the set of annotated training data. Hence, label statistics are provided as histograms in respect to different kinds of labels. This enables simple identification of similar training data by histogram analysis. Furthermore, for each label type, a histogram of luminance and color properties can be provided in order to enable a further distinction of label types. Hence, it can be possible to identify not only simple label type like e.g. pedestrian or vehicle, but to distinguish between complex label types like e.g. a pedestrian/vehicle in shadow or sunlight or at night.

According to a modified embodiment of the invention, the step of determining label statistics of a position of labels comprises dividing images into equally sized blocks. The blocks are used as basis to identify the position of the respective label. Preferably, the equally sized blocks have a size of  $N \times N$  pixel, where  $N$  is typically a big number, e.g.  $N=64$ . Based on the equally sized blocks, the position of a label can be defined e.g. by a simple block number. Preferably, a histogram analysis of labels overlapping with each block is performed to extract position statistics of each annotated label type.

According to a modified embodiment of the invention, the step of classifying the set of annotated training data based on the scene statistics comprises classifying the set of annotated training data under additional consideration of odometry information of the vehicle. The odometry information refers to the vehicle generating the underlying video frames for providing the training data at the time the underlying video frames are recorded. Hence, additional information in respect to the video frames is provided, which can help to efficiently classify the annotated training data. The odometry information

refers to information in respect to a movement of the vehicle, e.g. velocity, steering angle, or others. This information enables further classification of the training data, which extends beyond a classification performed merely based on the image data.

According to a modified embodiment of the invention, the step of classifying the set of annotated training data based on the scene statistics comprises identifying a hardware used when obtaining set of annotated training data. The hardware refers in particular to the camera used to obtain the set of annotated training data, which can directly influence an appearance of the set of annotated training data. Hence, a camera type and/or a camera resolution, or other camera properties can be evaluated. Furthermore, also the used vehicle can influence the appearance of the set of annotated training data, e.g. in case of appearance of the vehicle within the images or a position of the camera at the vehicle.

According to a modified embodiment of the invention, the step of automatically selecting training data out of the set of annotated training data, respectively, comprises balancing training data out of the set of annotated training data based on the classification between different classes. The balancing enables training the DCNN reliably with all classes of training data. Hence, a similar performance of the DCNN for all traffic situations can be achieved, even under consideration of changing conditions, e.g. environmental conditions. Such balancing of the training data helps to generalize the performance of the DCNN for a particular route for different traffic situations including different lighting conditions. The different traffic situations can include traffic lights, sharp turns, roundabouts, frames with high pedestrians, supermarket parking slot, just to name a few. Even though these scenarios can be not very frequent scenarios or even seldom, it is important to train the DCNN also to handle these scenarios, which may occur when driving a route, in a reliable way. Preferably, the balancing of the training data comprises removing training data, which refers to common driving situations, e.g. driving on a highway without many objects around and similar lighting conditions. Accordingly, high performance of the DCNN can be achieved for different possible scenarios for the particular route to travel.

According to a modified embodiment of the invention, the method comprises a step of filtering unsuitable training data out of the set of annotated training data based on the scene statistics for the set of annotated training data and/or based on the classification

of the set of annotated training data. Unsuitable training data refers to training data e.g. without labels or with an unreasonably low number of labels. Training the DCNN with such training data does not essentially help to improve its performance. Hence, omitting training with such training data improves training efficiency.

These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter. Individual features disclosed in the embodiments can constitute alone or in combination an aspect of the present invention. Features of the different embodiments can be carried over from one embodiment to another embodiment.

In the drawings:

Fig. 1 shows a flow chart of a method step for providing an initially trained deep convolutional neural network with general model data in accordance with a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information according to a first, preferred embodiment,

Fig. 2 shows a flow chart of a method step for providing a set of annotated training data including route information in accordance with the method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information according to the first embodiment,

Fig. 3 shows a flow chart of an individual training of the DCNN in accordance with the method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information according to the first embodiment, and

Fig. 4 shows a flow chart of a method for applying a deep convolutional neural network for processing image data in a driving support system of a

vehicle based on route information of a desired route according to a second embodiment.

Figures 1 to 3 together show a method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information according to a first, preferred embodiment. The method is executed using the driving support system of the vehicle.

The neural network trained with the proposed method is a deep convolutional neural network (DCNN) for processing image data. The DCNN is a particular implementation of a convolutional neural network (CNN), which comprises an input and an output layer together with multiple hidden layers arranged between the input layer and the output layer. The hidden layers of a CNN typically comprise convolutional layers, pooling layers, fully connected layers and normalization layers. The convolutional layers apply a convolution operation to the input, passing the result to the next layer. The convolution emulates the response of an individual neuron to visual stimuli.

The DCNN is used in the driving application in a vehicle. Hence, the DCNNs is employed in a driving support system of the vehicle. The driving support systems is a system, which assists a human driver of the vehicle in different driving situations, or a system, which supports autonomous driving of the vehicle by providing input to reliably deal with different traffic situations.

The method starts with step S100. Step S100 refers to providing an initially trained deep convolutional neural network using an initial set of annotated training data. The annotated training data is manually annotated training data, i.e. a human has added labels to the provided image data.

Method step S100 comprises in this embodiment three individual sub-steps S102, S104, and S106, as discussed below in detail.

Step S102 refers to determining scene statistics for the initial set of annotated training data. The scene statistics refers to a statistical determination of features shown in the training data. The features include static features without movement information and dynamic features including movement information, e.g. pedestrian walking or even

pedestrian walking from left to right. The features can comprise the annotations of the training data, i.e. labels added thereto.

The scene statistics are performed including types of labels, positions of labels, and properties of labelled objects. The label statistics are evaluated for a type of labels, a position of labels, and properties of labelled objects. Additionally, label combination statistics, i.e. statistics of a combination of multiple labels e.g. within the same image, e.g. a certain number of pedestrians or a certain number of vehicles within the image, are determined.

Label statistics here comprises performing a histogram analysis of the type of labels from the set of annotated training data in respect to different kinds of labels. This includes for each label type providing a histogram of luminance and color properties.

Furthermore, in order to determine label statistics of a position of labels, the images are divided into equally sized blocks. The blocks are used as basis to identify the position of the respective label. The equally sized blocks have a size of  $N \times N$  pixel, where  $N=64$ . Based on the equally sized blocks, the position of a label is defined e.g. by a simple block number. A histogram analysis of labels overlapping with each block is performed to extract position statistics of each annotated label type.

Step S104 refers to classifying the set of annotated training data based on the scene statistics.

The classification of the annotated training data refers to a classification based on occurrences of features within the training data based on the histogram. The classification enables identifying training data showing a particular type of features and/or a particular number of features. Such classifications include training data with a high density of features, training data comprising a particular type of features, e.g. pedestrians or vehicles, with or without dynamic information, e.g. not moving, moving, moving from right to left.

Step S106 refers to automatically selecting training data out of the set of annotated training data based on the classification. This comprises balancing training data out of the set of annotated training data based on the classification between different classes.

Step S106 includes filtering unsuitable training data out of the set of annotated training data based on the scene statistics for the set of annotated training data and based on the classification of the set of annotated training data. Unsuitable training data refers to training data e.g. without labels or with an unreasonably low number of labels.

The training of the deep convolutional neural network is known from common DCNNs. The chosen training data from the initial set of annotated training data of the DCNN is fed to the DCNN to perform the training. Step S106 stops, when the DCNN shows a sufficient performance.

In subsequent step S108, general DCNN model data is provided, which is transferred to the driving support system of the vehicle and stored therein.

Step S110, which is indicated in figure 2, refers to providing a set of automatically annotated training data including route information, as discussed below in detail. Hence, the training data is provided with labels in respect to shown features. Step S110 comprises individual sub-steps S112, S114, S116, S118, and S120, as discussed below in detail.

According to step S110, a camera of the vehicle is used to provide image data when driving the route. The image data is tagged with satellite based position information using a receiver for GPS signals. Further, a driving route is identified based on the location information when driving the route. Step S110 is performed using the driving support system of the vehicle.

Step S112 refers to processing a set of not annotated training data with the initially trained deep convolutional neural network to generate a set of automatically labeled data. The stored model data of the initially trained DCNN is used to process the set of not annotated training data to generate a set of automatically labeled data.

Step S114 refers to calculating confidence metrics for the set of automatically labeled data. The confidence metrics is calculated applying a computer vision based confidence metric algorithm. The confidence metric algorithm provides a confidence value as indication of the confidence level. The confidence value is a probability indicating a

probability for a valid detection of features and labeling of the data based on the detected features. An evaluation of the confidence metrics for the automatically labeled data is performed. Based on the evaluation result of the confidence metrics, the annotation is indicated or not.

Step S116 refers to filtering the training data based on a statistical analysis of the training data. Hence, traffic congestions, differently textured road surfaces, different lighting conditions, or any different traffic scenarios like overtaking, lane changing, roundabout, parking areas, or others can be identified and used to filter the training data. Furthermore, based on the position information, the image data is processed to determine different conditions for the image data from the same location. Image data provided at similar locations is identified and filtered to select image data with different characteristics, e.g. different lighting condition, different weather conditions, different traffic conditions, or others at the same location or at similar locations.

Step S118 refers to automatically relabeling the set of automatically labeled data based on the calculated confidence metrics to generate the set of annotated training data. In case the confidence metrics indicates a high confidence level in the automatically labeled data generated by the DCNN, the labels are not changed. Otherwise, the labels are modified using a computer vision based labeling algorithm.

Step S120 refers to storing the set of annotated training data for further usage, e.g. for subsequent training the DCNN. The labelled training data is accumulated each time the owner of the vehicle drives a particular route using the vehicle, i.e. when the vehicle is used in typical driving conditions. The training data is accumulated in respect to the respective route.

Step S120 further comprises evaluating when the set of annotated training data comprises sufficient training data. A notification to the driver or owner of the vehicle can be generated as soon as the set of annotated training data with sufficient annotated training data for training the DCNN for an individual route to proceed further to step S130.

Figure 3 illustrates an individual training of the DCNN based on the set of annotated training data for a particular route.

In detail, step S130 refers an identification of the route of the set of annotated training data. Depending on the route, different subsequent parallel path of the method are performed based on the identified route. The paths are performed identically but using different model data for routes #1 to #n of the DCNN and training data based on the identified route. Each parallel path comprises steps S132 to S138.

Step S132 refers to a step of determining scene statistics for the set of annotated training data. The scene statistics refers to a statistical determination of features shown in the training data. Hence, label statistics of at least one out of a type of labels, a position of labels, and/or properties of labelled objects, preferably determining label combination statistics are determined for multiple or all out of available types of labels, positions of the labels, and/or properties of labelled objects. The details specified with respect to step S102 apply. Accordingly, a histogram analysis of the type of labels from the set of annotated training data is performed. Furthermore, for each label type, a histogram of luminance and color properties is provided in order to enable a further distinction of label types.

Step S134 refers to classifying the set of annotated training data based on the scene statistics, and automatically selecting training data out of the set of annotated training data based on the classification. The details specified with respect to step S104 apply.

Step S136 refers to setting up the deep convolutional neural network with individual model data associated to the route identified in step S130 based on the general model data. Accordingly, the DCNN is provided for training for a particular route. When the method is performed for the first time, the individual model data associated to the route is provided as a copy of the general model data. Otherwise, the individual model data of the deep convolutional neural network data associated to the route as previously stored is loaded.

Step S138 refers to a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data. This comprises in this embodiment training only the last layer of the DCNN. Step S138 is performed in accordance with the above description of method step S106.

Step S140 refers to storing the individual model data of the deep convolutional neural network data associated to the route. Hence, only the model data for the last layer of the DCNN is stored, the remaining layers of the DCNN remain unchanged.

In this embodiment, all of method steps S130 to S140 are performed using the driving support system of the vehicle having employed the DCNN. Hence, the method is implemented as an in-vehicle method.

The present invention also provides a method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route, which is shown in the flow chart of figure 4. This method refers to using the vehicle with the driving support system employing the DCNN.

The method starts with method step S200. Method step S200 refers to identifying the route. Step S200 is performed as described in detail with reference to step S130.

Step S210 refers setting up the deep convolutional neural network with the individual model data associated to the route. Step S210 is performed as described in detail with reference to S136. Accordingly, the DCNN is activated with the individual model data of the respective route.

Step S220 refers to processing the image data acquired when driving the route with the vehicle and using the driving support system employing the deep convolutional neural network with the individual model data associated to the route. Hence, an application of the DCNN in the driving support system is provided.

## Patent claims

1. Method for training a deep convolutional neural network for processing image data for application in a driving support system of a vehicle using training data including route information, comprising the steps of
  - providing an initially trained deep convolutional neural network with general model data,
  - providing a set of annotated training data including route information,
  - setting up the deep convolutional neural network with individual model data associated to the route based on the general model data,
  - performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data, and
  - storing the individual model data of the deep convolutional neural network data associated to the route.
2. Method according to claim 1, characterized in that
  - the step of providing a set of annotated training data including route information comprises identifying a driving route, providing image data as training data from at least one camera when driving the route, and automatically annotating the image data to provide the set of annotated training data.
3. Method according to preceding claim 2, characterized in that
  - the step of providing image data from at least one camera when driving the route comprises tagging the provided image data with position information, in particular with satellite based position information.
4. Method according to any of preceding claims 2 or 3, characterized in that
  - the step of automatically annotating the image data comprises the steps of processing a set of not annotated training data with the deep convolutional neural network to generate a set of automatically labeled data,
  - calculating confidence metrics for the set of automatically labeled data, and

- automatically relabeling the set of automatically labeled data based on the calculated confidence metrics to generate the set of annotated training data.
5. Method according to preceding claim 4, characterized in that  
the step of calculating confidence metrics for the set of automatically labeled data comprises applying a computer vision based confidence metric algorithm.
  6. Method according to any of preceding claims 4 or 5, characterized in that  
the step of automatically relabeling the set of automatically labeled data based on the calculated confidence metrics comprises applying a computer vision based labeling algorithm.
  7. Method according to any preceding claim, characterized in that  
the step of setting up the convolutional neural network with individual model data associated to the route comprises generating the individual model data associated to the route as a copy of the general model data or loading the individual model data of the deep convolutional neural network data associated to the route as previously stored.
  8. Method according to any preceding claim, characterized in that  
performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data comprises training last layers of the deep convolutional neural network, in particular only the last layer.
  9. Method according to any preceding claim, characterized in that  
the method is implemented as an in-vehicle method, whereby the training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data is performed for training the deep convolution neural network as provided for a particular vehicle.
  10. Method according to preceding claim 9, characterized in that

- the step of providing a set of annotated training data including route information comprises providing the set of annotated training data including route information using at least one camera of the vehicle.
11. Method according to preceding claim 10, characterized in that  
the step of providing the set of annotated training data including route information using at least one camera of the vehicle comprises accumulating training data and evaluating when the set of annotated training data comprises sufficient training data.
  12. Method according to any of preceding claims 10 or 11, characterized in that  
the step of providing the set of annotated training data including route information using at least one camera of the vehicle comprises filtering the training data based on a statistical analysis of the training data.
  13. Method according to any preceding claim, characterized in that  
the step of performing a training step of the deep convolutional neural network with the individual model data associated to the route using the respective set of annotated training data comprises the steps of  
determining scene statistics for the set of annotated training data,  
classifying the set of annotated training data based on the scene statistics,  
and  
automatically selecting training data out of the set of annotated training data based on the classification.
  14. Method according to preceding claim 13, characterized in that  
the step of determining scene statistics for the set of annotated training data comprises determining label statistics of at least one out of a type of labels, a position of labels, and/or properties of labelled objects, preferably determining label combination statistics.
  15. Method for applying a deep convolutional neural network for processing image data in a driving support system of a vehicle based on route information of a desired route, comprising the steps of

- identifying the route,
  - setting up the deep convolutional neural network with individual model data associated to the route, and
  - processing the image data acquired when driving the route using the deep convolutional neural network with individual model data associated to the route.
16. Driving support system for performing any of the methods according to any of preceding method claims 1 to 15.

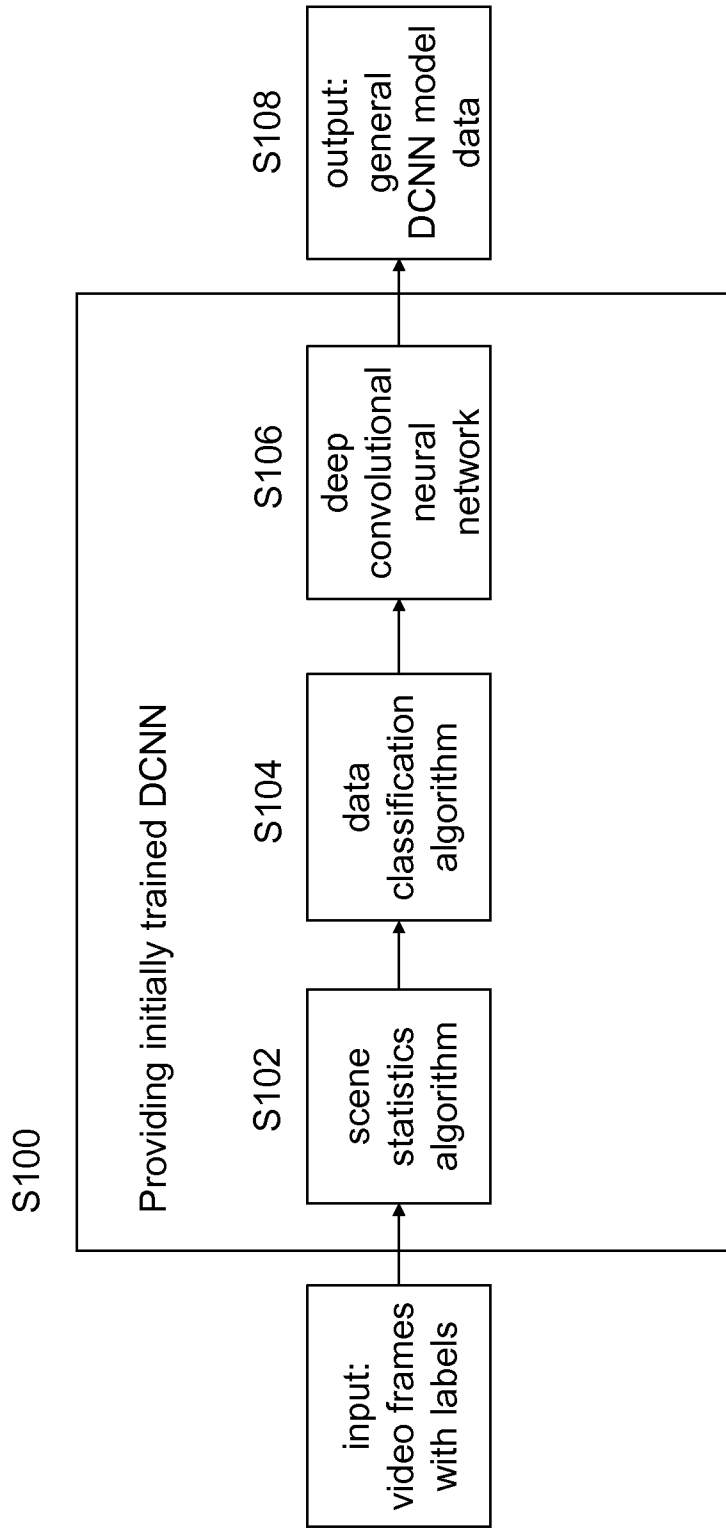


FIG. 1

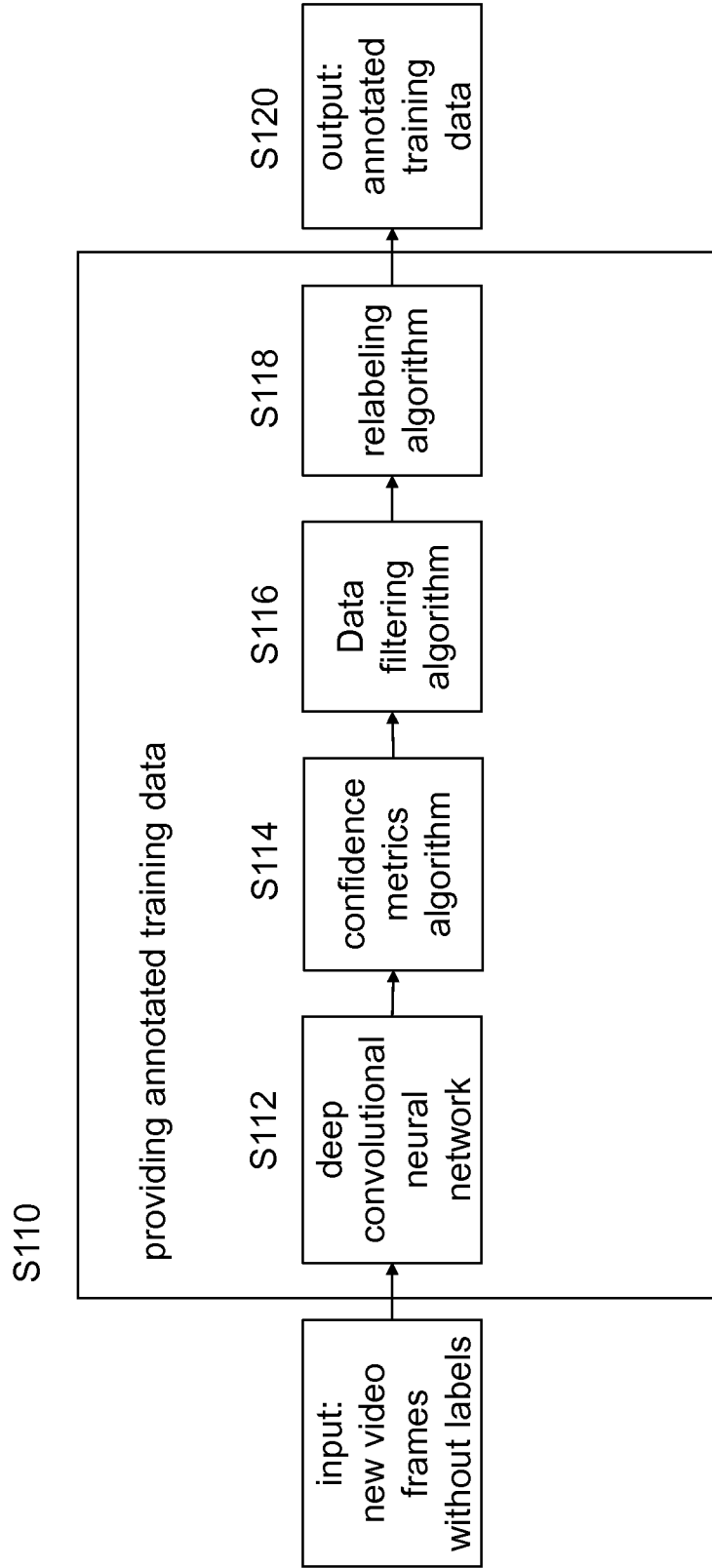


FIG. 2

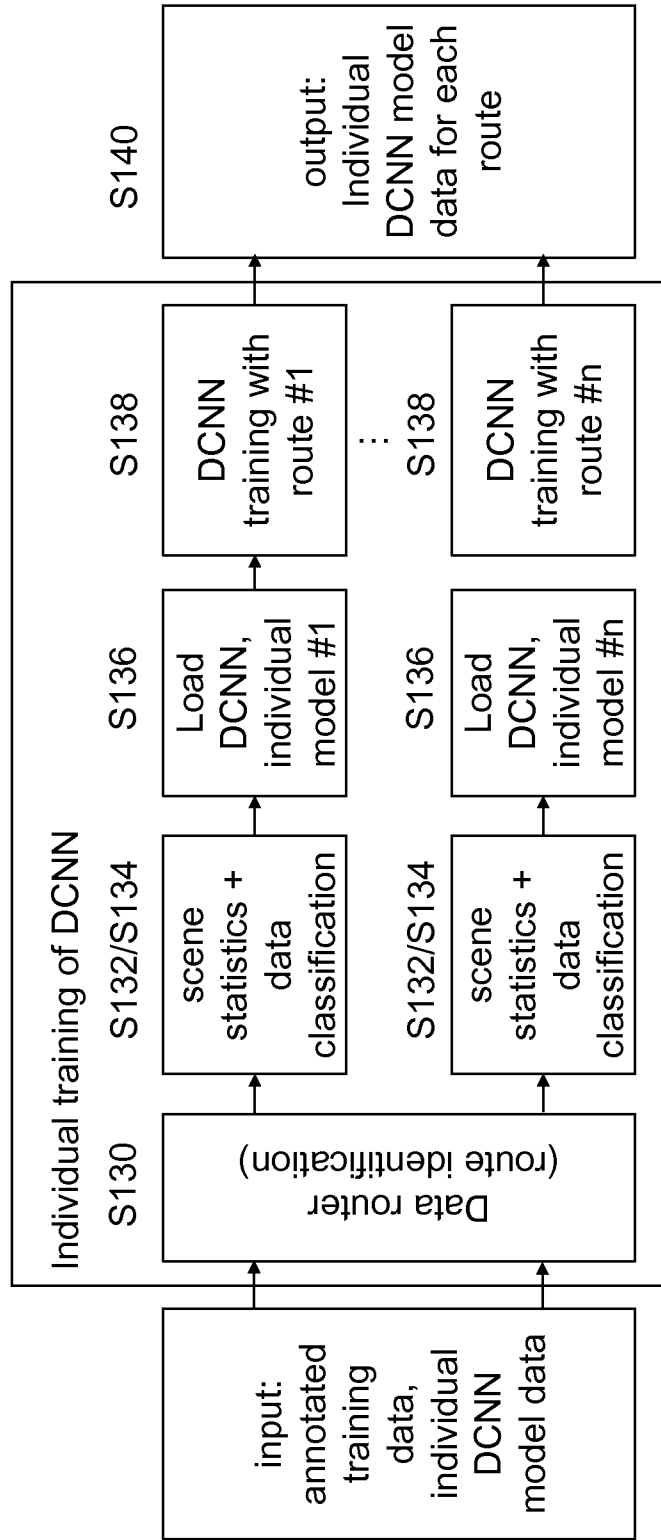


FIG. 3

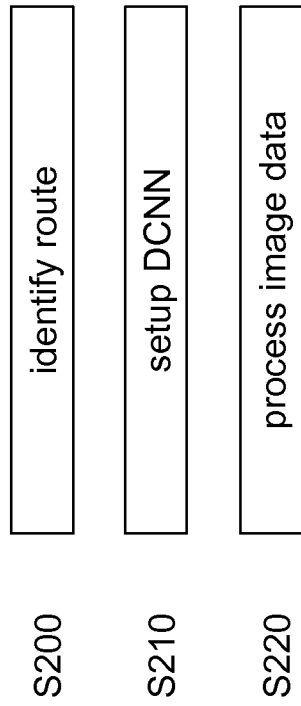


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2019/065629

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G06K9/00 G06K9/62  
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)  
G06K  
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	HOLDER CHRISTOPHER J ET AL: "From On-Road to Off: Transfer Learning Within a Deep Convolutional Neural Network for Segmentation and Classification of Off-Road Scenes", 18 September 2016 (2016-09-18), INTERNATIONAL CONFERENCE ON COMPUTER ANALYSIS OF IMAGES AND PATTERNS. CAIP 2017: COMPUTER ANALYSIS OF IMAGES AND PATTERNS, SPRINGER, BERLIN, HEIDELBERG, XP047354225, ISBN: 978-3-642-17318-9 pages 149-162, [retrieved on 2016-09-18]	1-3,7-16
Y	Section 1. Sections 2.1-2.2 Sections 3.1-3.2 figures 1, 3,4 ----- -/--	4-6

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

27 August 2019

Date of mailing of the international search report

02/09/2019

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

Moreno, Marta

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2019/065629

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 2018/055378 A1 (UNIV OXFORD INNOVATION LTD [GB]) 29 March 2018 (2018-03-29)  page 1, line 3 - line 5  page 5, line 2 - line 27  page 6, line 18 - line 33  page 9, line 10 - page 11, line 31  page 12, line 15 - page 13, line 2  page 14, line 1 - page 15, line 4  page 20, line 24 - line 34  page 26, line 18 - line 26  figure 1A</p> <p style="text-align: center;">-----</p>	15,16
Y	<p>LU XIQUN: "Self-supervised road detection from a single image",  2015 IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP),  27 September 2015 (2015-09-27), pages 2989-2993, XP032827020,  IEEE,Piscataway, NJ, USA  DOI: 10.1109/ICIP.2015.7351351  [retrieved on 2015-12-09]  Section 2.3</p> <p style="text-align: center;">-----</p>	4-6
A	<p>LADDHA ANKIT ET AL: "Map-supervised road detection",  2016 IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV),  19 June 2016 (2016-06-19), - 22 June 2016 (2016-06-22), pages 118-123, XP032938952,  IEEE, Piscataway, NJ, USA  DOI: 10.1109/IVS.2016.7535374  [retrieved on 2016-08-05]  Section IV.</p> <p style="text-align: center;">-----</p>	1-16
A	<p>LI SUN ET AL: "Weakly-supervised DCNN for RGB-D Object Recognition in Real-World Applications Which Lack Large-scale Annotated Training Data",  ARXIV.ORG, CORNELL UNIVERSITY LIBRARY,,  19 March 2017 (2017-03-19), XP080757958,  NY201 OLIN LIBRARY CORNELL UNIVERSITY  ITHACA, NY 14853  Section IV.B</p> <p style="text-align: center;">-----</p>	1-16

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2019/065629

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018055378	A1	29-03-2018	
		EP 3516582 A1	31-07-2019
		GB 2554481 A	04-04-2018
		WO 2018055378 A1	29-03-2018
-----			