

【DESCRIPTION】

【Invention Title】

METHOD FOR TRAINING NEURAL NETWORK BY USING DE-IDENTIFIED IMAGE AND SERVER PROVIDING SAME

【Technical Field】

[0001] The disclosure relates to a neural network training method, and more particularly, to a neural network training method using a de-identified image.

【Background Art】

[0002] As big data-based analysis technology is applied to various industrial fields, it is necessary to separately protect and manage personal information including information on an individual among collected data.

[0003] Simultaneously, artificial intelligence technology is also replacing conventional data analysis technology to expand the scope of its application. For example, in more cases, the artificial intelligence uses biometric information such as a fingerprint, a face or the like to verify the identity of a user as a method for user authentication and provides a personalized service based thereon.

[0004] Here, for the user authentication using personal biometric information, a neural network used for the artificial intelligence may need to learn various biometric information, and training data itself used for this training may include the personal information as the

biometric information. It is thus necessary to be cautious about the usage and management of the personal information.

[0005] Accordingly, it is necessary to secure a basis of the usage in order to perform the training by using big data including the personal information in the field of artificial intelligence and provide the service based thereon. Recently prepared is a law to de-identify such personal information, exclude de-identified information from personal credit information protection, and the like.

[0006] In a financial industry, there are increasing needs for protecting a user's facial image, an identity card image or the like, which is inevitably required, in order to open an account and provide a financial service in a non-face-to-face manner by launching an internet bank or the like and for de-identifying the personal information in order for the artificial intelligence to perform the training.

[0007] However, the current training method used for the artificial intelligence may be about data in a general format, and does not consider the protection and de-identification of the personal information.

[0008] It is necessary to secure a lot of training data for the training of the neural network used for the artificial intelligence, and a more efficient method of de-identifying the training data and a training method based thereon are required in order to provide an efficient service based on

the artificial intelligence.

【Disclosure】

【Technical Problem】

[0009] An object of the disclosure is to provide a neural network training method using a de-identified image.

[0010] Another object of the disclosure is to provide a method to de-identify an image by considering prediction accuracy of a neural network.

【Technical Solution】

[0011] In one general aspect, a neural network training method using a de-identified image includes: encoding a first image represented by an n-th dimensional vector into a predetermined p-th dimensional second image; decoding the second image into a q-th dimensional third image; inputting the third image to the neural network and extracting object information included in the third image; and training at least one parameter information used for a computation in the neural network by using the extracted object information.

[0012] The training may include training the parameter information used for the encoding or decoding computation based on the extracted object information.

[0013] The training may include training the parameter information by using a correlation between a degree of de-identification defining similarity between the third image and the first image and prediction accuracy of the neural

network for the object information.

[0014] A size of p-th dimensions may be determined based on the degree of the de-identification.

[0015] N-th dimensions and q-th dimensions may have the same size.

[0016] The method may further include storing the second image encoded in the determined p-th dimensions, wherein the decoding may include decoding the stored second image into the third image when the neural network performs the training.

[0017] The decoding may include decoding the second image to have a data value different from that of the first image when decoding the second image in the q-th dimensions.

[0018] In another general aspect, a neural network training server using a de-identified image includes: a de-identification unit encoding a first image represented by an n-th dimensional vector into a predetermined p-th dimensional second image, and then decoding the second image into a q-th dimensional third image; and a training unit inputting the third image to the neural network and extracting an object information included in the third image to train at least one parameter information used for a computation in the neural network.

[0019] The training unit may train the parameter information used for the encoding or decoding computation based on the extracted object information.

[0020] The training unit may train the parameter information by using a correlation between a degree of de-identification defining similarity between the third image and the first image and prediction accuracy of the neural network for the object information.

[0021] The server may further include a database storing the second image encoded in determined p -th dimensions, wherein the de-identification unit decodes the stored second image into the third image when the neural network performs the training.

[0022] The de-identification unit may decode the second image to have a data value different from that of the first image when decoding the second image in the q -th dimensions.

【Advantageous Effects】

[0023] According to the disclosure, it is possible for the neural network to perform the training without using the personal information included in the image because the de-identified image is used for the training of the neural network.

[0024] According to the disclosure, it is also possible to manage more efficiently the training data because the image is de-identified and stored for the de-identified image not to be restored to the original image.

[0025] According to the disclosure, it is also possible to process the image more efficiently by adjusting the dimension of the information included in the image based on

the prediction performance of the neural network and encoding the same.

【Description of Drawings】

[0026] FIG. 1 is an exemplary diagram showing a system providing a service based on de-identified image recognition according to an embodiment of the disclosure.

[0027] FIG. 2 is an exemplary diagram showing object information detection in a neural network training method performed by a service operation server according to an embodiment of the disclosure.

[0028] FIG. 3 is a flowchart showing a training process of the service operation server according to an embodiment of the disclosure.

[0029] FIG. 4 is an exemplary diagram showing an image encoding method of the service operation server according to an embodiment of the disclosure.

[0030] FIG. 5 is an exemplary diagram showing an image decoding method of the service operation server according to an embodiment of the disclosure.

[0031] FIG. 6 is an exemplary diagram showing a configuration of the neural network of the service operation server according to an embodiment of the disclosure.

[0032] FIG. 7 is an exemplary diagram showing the neural network training method performed by the service operation server according to an embodiment of the disclosure.

[0033] FIG. 8 is a block diagram showing a configuration of the service operation server according to another embodiment of the disclosure.

【Best Mode】

[0034] The following description illustrates only a principle of the disclosure. Therefore, those skilled in the art may implement the principle of the disclosure and invent various devices included in the spirit and scope of the disclosure although not clearly described or shown in the specification. In addition, it is to be understood that all conditional terms and embodiments mentioned in the specification are obviously intended only to allow those skilled in the art to understand a concept of the disclosure in principle, and the disclosure is not limited to the embodiments and states particularly mentioned as such.

[0035] The above-mentioned objects, features and advantages will become more obvious from the following detailed description provided in relation to the accompanying drawings. Therefore, those skilled in the art to which the disclosure pertains may easily practice a technical idea of the disclosure.

[0036] Further, in describing the disclosure, in the case in which it is judged that a detailed description of a well-known technology associated with the disclosure may unnecessarily make the gist of the disclosure unclear, it

will be omitted. Hereinafter, the embodiments of the disclosure will be described in detail with reference to the accompanying drawings.

[0037] FIG. 1 is an exemplary diagram showing a system providing a service based on de-identified image recognition according to an embodiment of the disclosure.

[0038] Referring to FIG. 1, a user 5 may capture the user's own face by using a user terminal 10.

[0039] A service operation server 100 may receive a captured face image 20 from the user 5, compare an identity image with the captured facial image 20, perform user authentication, and provide a service based thereon.

[0040] Here, the service operation server 100 may perform the user authentication by using a neural network and perform a faster and more accurate user verification procedure than a conventional image processing method by using the neural network.

[0041] Iterative training may be required in order to the neural network used for the user authentication to determine the identity of the user through the image, and the service operation server 100 may thus store various training data in a separate database 200.

[0042] However, as described above, the neural network used for the user authentication may be required to training images including facial information of the user 5 as the training data and may thus be required to directly use

personal information.

[0043] Therefore, in this embodiment, the service operation server 100 may protect the personal information by de-identifying image including the personal information such as the facial information used as the training data and making it impossible to restore the de-identified image to an original image.

[0044] The service operation server 100 according to this embodiment may de-identify the image received from the user terminal 10 and store the de-identified image as the training data in the database 200 to be used for the training of the neural network.

[0045] In general, the neural network used for detecting the image may use data values in the images represented as multi-dimensional vectors in a color space.

[0046] The data values in the image may be defined as a height, a width and a depth, which define its size as a three-dimensional vector. Here, the depth may be defined as a size of a channel, which is a method of representing information on each pixel.

[0047] For example, the depth of the image may be represented as a red-green-blue (RGB) code as information that defines its color. When the image is represented by the RGB code, a size of the channel may be three (3) and each channel may store a RGB value of the corresponding pixel.

[0048] Alternatively, it is possible to represent information in the image as the multi-dimensional vector by allowing the channel to include grayscale or hue, saturation and value (HSV).

[0049] In addition, the plurality of images may be used for the training of the neural network, and the training data may also include a four-dimensional vector by adding the number of these images.

[0050] The neural network may apply a filter defined for each channel to the data values in the image represented by the multi-dimensional vector to perform a convolutional computation and extract feature information in the image.

[0051] The convolutional computation may be sequentially performed in units of layers included in the neural network, and the neural network may include a network in which the convolutional layers are stacked.

[0052] For example, when an object is defined by a distribution or pattern of a specific color, the neural network may be trained by adjusting a weight of a filter that emphasizes the specific color.

[0053] In detail, in this embodiment, the neural network may be trained to extract object information 30 from the facial image including the facial information of the user 5 for the user verification. Here, the object information 30 may be the feature information having the highest importance for distinguishing a large number of data, and may be, for

example, unique information having the highest identification power in the facial information. For example, the neural network may extract eyes, nose, mouth and the like, included in the facial information, from the image, and may be trained to extract, as the object information 30, information determined to be necessary to identify whose facial image these extracted images belong to, such as their shapes, relative positional relationship and the like.

[0054] The object information 30 may include one or more feature information, and a weight may be given to each feature information based importance required for the identification.

[0055] However, the facial image here may include the personal information as it is, and the training data itself used for the training may be the personal information. In addition, the neural network may use an actual image as the training data to have a robust performance on images under various capturing conditions, and the actual image may include more diverse personal information.

[0056] Therefore, an additional processing process for separately protecting the personal information may be required for using the facial image for the training of the neural network, and the training method according to the disclosure may de-identify the image including the personal information and train the neural network.

[0057] Hereinafter, the training method of the service operation server 100 according to this embodiment is described in more detail.

[0058] Referring to FIG. 2, the service operation server 100 according to this embodiment may include a de-identification unit 110 and a training unit 120.

[0059] The de-identification unit 110 may de-identify the face image captured by and transmitted from the user terminal 10.

[0060] In detail, the de-identification unit 110 may include an encoder 112 and a decoder 114.

[0061] The de-identification unit 110 may encode the received face image by using the encoder 112 and convert the image through decoding again to generate training data.

[0062] Here, the image converted by the decoder 114 may include information represented in a format different from that of the original face image.

[0063] In detail, the converted image may be converted for the identity of the object included in the image not to be recognized with naked eyes and may have a format which is impossible to be restored to the original face image.

[0064] The converted image may be used as input data of a neural network model and may thus be information making it impossible to recognize the identity with the naked eyes and represented in a more efficient format for the neural network model to extract the object information 30.

[0065] The training unit 120 may train the neural network which extracts the object information 30 by using the de-identified image received from the de-identification unit 110.

[0066] When the de-identified image is input to the neural network, the neural network may output the object information 30. The training unit 120 may train and update weights of various parameters used for extracting the feature information in the neural network by using the output object information 30.

[0067] In addition, the training unit 120 may also train various parameters of the encoder 112 or decoder 114, used to de-identify the image together with the neural network.

[0068] Hereinafter, the training method is sequentially described in more detail with reference to FIG. 3.

[0069] First, the first image 20 represented by an n-th dimensional vector may be encoded into a predetermined p-th dimensional second image 21 (S10).

[0070] N-th dimensions may define the number of features representing information included in the first image 20 captured by the user terminal 10 and transmitted to the service operation server 100. For example, the n-th dimensions may be a size of the image and defined as three dimensions of height, width and depth of the image.

[0071] That is, the n-th dimensional first image 20 may be compressed using the encoder 112 and converted into the p-

th dimensional second image 21.

[0072] Referring to FIG. 4, the received first image 20 may be input to the encoder 112, and the encoder 112 may generate the p-th dimensional second image 21 according to a predetermined compression rule.

[0073] The second image 21 may be encoded into the p-th dimensions, which are reduced dimensions different from dimensions representing the first image 20.

[0074] For example, data values of the n-th dimensions representing the first image 20 may be compressed according to an encoding standard, and the second image 21 may thus be represented in the p-th dimensions having a reduced data size compared to the n-th dimensions.

[0075] Alternatively, it is possible to reduce the number of dimensions by projecting the data values based on a specific dimension among the n-th dimensions, and to generate the reduced p-th dimensional second image 21.

[0076] As described above, the encoder 112 may encode the first image 20 in the p-th dimensions in which the number of dimensions itself is reduced compared to the n-th dimensions or the data size representing the data value of each dimension is reduced.

[0077] Next, the p-th dimensional second image 21 generated by the encoder 112 may be decoded into a q-th dimensional third image 22 (S20).

[0078] Q-th dimensions may re-expand the p-th dimensions

through the decoder 114 and redefine the data values in the input second image 21.

[0079] Referring to FIG. 5, the decoder 114 may expand the data values of the second image 21 defined in the p -th dimensions back to data values in the q -th dimensions. In detail, the decoder 114 may expand the data values of the second image 21 into q -th dimensional information by using weight parameters.

[0080] Here, the q -th dimensions decoded in this embodiment may be defined as the same dimensions as the n -th dimensions of the first image 20 before being encoded. However, the third image 22 decoded in the q -th dimensions may be generated from the encoded second image 21 and may have different data values from the original first image 20 on the same dimensions.

[0081] That is, the first image 20 may be an image captured in a general way and it is thus possible to identify the object included therein with the naked eyes. However, the third image 22 may include information redefined after being encoded, and include data values in a form in which the object is impossible to be identified with the naked eyes.

[0082] Through the above process, the training method according to this embodiment may generate the second image 21 and the third image 22 by de-identifying the first image 20.

[0083] Next, the de-identified third image 22 may be input to the neural network, and the object information 30 included in the third image 22 may be extracted (S30).

[0084] In addition, the neural network may be trained using the extracted object information 30 (S40).

[0085] The neural network may update parameters in the network to have a more robust performance on the de-identified images, which are the input data.

[0086] Referring to FIG. 6, the neural network may include a plurality of hidden layers 124 and 126, which perform the convolutional computation in an input layer 122 and an output layer 128. The hidden layers 124 and 126 may be nodes performing the weight computation in response to the input data.

[0087] In this embodiment, the neural network may include the convolutional layers performing the convolutional computation, and the nodes in the hidden layers 124 and 126 may be filters performing the convolutional computation on the data values in the de-identified image for each unit region.

[0088] That is, the data values for each channel in the input image may be sequentially computed while passing through the plurality of convolutional layers, the characteristic distribution or pattern of the data values emphasized may be found through this computation, and the object information 30 in the image may be extracted.

[0089] The object information 30 extracted as an output signal for the training may be compared with a target value of the actual image to calculate an error, and the calculated error may be propagated in the neural network in a reverse direction according to a backpropagation algorithm. For example, the neural network may be trained to find a cause of the error by using the backpropagation algorithm and change a connection weight between the nodes of the neural network.

[0090] In addition, referring to FIG. 7, the training method according to this embodiment may be performed by training a weight parameter $W2$ between each node performing the convolutional computation in the neural network, together with a de-identification parameter $W1$ used for the encoding or the decoding.

[0091] That is, it is possible to update the parameter $W2$ defining the connection weight between the nodes in the layer according to the backpropagation algorithm by using the extracted object information 30, and also possible to update the parameter $W1$ used for computing the encoding or the decoding of the de-identification unit 110.

[0092] For example, it is possible to determine encoding parameters for reducing the n -th dimensional first image 20 to the p -th dimension "z" through the training when a size of the p -th dimensions "z" is determined in consideration of a compression rate as a de-identification degree of the

de-identification unit 110.

[0093] For example, the training may be performed to update the parameters used for the computation as the size and sampling rate of a unit block that becomes a data sampling standard when the dimensions of the encoder 112 are reduced.

[0094] Conversely, the training may also be performed to update the relevant weight parameters when the dimensions used for the decoding are expanded.

[0095] Through the above process, the training method according to this embodiment may update the parameters in the neural network by using the extracted object information 30 and simultaneously update the parameters of the encoder 112 and decoder 114, which perform the de-identification.

[0096] In addition, the degree of the de-identification may be determined based on the compression rate of the encoder 112, and it may thus be advantageous to protect the personal information included in the image when the degree of the de-identification is increased. However, prediction accuracy of the neural network by using the decoded third image 22 may be lower.

[0097] Therefore, the training unit 120 may also relatively adjust the parameters used for the de-identification in consideration of the degree of the de-identification and the prediction accuracy. For example, the training unit 120 may update the encoding or decoding parameters to

weaken information necessary to recognize the identity with the naked eyes and to strengthen information advantageous for extracting the feature information of the neural network.

[0098] The training unit may further train the parameter information by using a correlation between the degree of the de-identification defining similarity between the third image 22 and the first image 20 and the prediction accuracy of the neural network for the object information 30.

[0099] As described above, the training method according to this embodiment may train the encoder 112, the decoder 114 and the neural network together, change the image to specify a portion that may be recognized well by a machine, not a human, and simultaneously allow the neural network to easily adapt thereto.

[00100] That is, the neural network may be trained to emphasize a portion necessary to mechanically extract the object from the image and weaken the rest.

[00101] Hereinafter, the configuration of the service operation server 100 according to this embodiment is described with reference to FIG. 8.

[00102] In this embodiment, the service operation server 100 may include the de-identification unit 110 and the training unit 120.

[00103] The de-identification unit 110 may include the encoder 112 and the decoder 114.

[00104] The encoder 112 may compress the n-th dimensional first image 20 captured by the user terminal 10 and generate the compressed p-th dimensional second image 21.

[00105] The encoder 112 may compress the data values of the n-th dimensions in the first image 20 according to the encoding standard, and thus may generate the p-th dimensional second image 21 having the reduced dimensions compared to the n-th dimensions.

[00106] The decoder 114 may change the second image 21 having the compressed information compared to the first image 20 to the q-th dimensional third image 22.

[00107] Here, the q-th dimensions may be defined as the same dimensions as the n-th dimensions of the first image 20.

[00108] However, the third image 22 changed to the q-th dimensions may be decoded from the second image 21 and may have different data values from the original first image 20.

[00109] Ultimately, the third image 22 may include the information redefined after being encoded, and may include the data values in the form in which the object is impossible to be identified with the naked eyes.

[00110] That is, the de-identification unit 110 may use the encoder 112 and the decoder 114 to use the second image 21 as an intermediate image and finally generate the third image 22 as the training data.

[00111] The training unit 120 may use the third image 22 as the input data and extract the object information 30 by using the neural network.

[00112] The neural network may include the plurality of layers sequentially performing the convolutional computation on the input data, and each layer may include the plurality of nodes including the data for the convolutional computation.

[00113] In this embodiment, the neural network may perform the convolutional computation on the data values in the q -th dimensions of the third image 22. That is, the node in the convolution layer may perform the convolutional computation on the data values through the filter.

[00114] Here, each node may define the weight when transmitting the data value as the parameter, and in this embodiment, the training may be performed to update the weight by using the object information 30 output to the neural network.

[00115] In addition, the training unit 120 may also train the parameters of the de-identification unit 110 in addition to the neural network.

[00116] The de-identification unit 110 may encode the first image 20 as the second image 21 and encode the same back into the q -th dimensions, and the number of the p -th dimensions defining the dimensions of the second image 21 may thus be determined through the training.

[00117] In addition, other parameters used in the encoder 112 and the decoder 114 may also be updated through the training.

[00118] As described above, the training unit 120 may train the neural network to output an optimal result for the de-identified image in consideration of both the degree of the de-identification and the degree of the prediction of the neural network.

[00119] Furthermore, the service operation server 100 may include a separate database 200 storing the de-identified image.

[00120] The iterative training may be required to be performed through a large amount of the training data for the training of the neural network, and it is thus also possible to store the de-identified image in the database 200 and train the neural network through the same.

[00121] That is, the service operation server 100 may de-identify the received first images 20 and then delete the same, and store only the de-identified second image 21 or third image 22 in the database 200 to prevent damage caused by leakage of the personal information.

[00122] In addition, when the second image 21 is stored in the database 200, it is possible to further increase efficiency in the training of the neural network by accumulating a determined amount of the second images 21, and bulk-decoding the accumulated second images 21 to

generate the third image 22.

[00123] According to the disclosure above, it is possible for the neural network to perform the training without directly using the personal information included in the image because the de-identified image is used for the training of the network. In addition, according to the disclosure, it is possible to increase efficiency in protecting the personal information and managing the data by de-identifying the image and storing the same for the image not to be restored to the original image.

[00124] In addition, according to the disclosure, it is possible to increase efficiency in processing the image by adjusting the dimensions of the information included in the image based on the prediction performance of the neural network and encoding the same.

[00125] Meanwhile, in the disclosure and the claims, terms such as "first", "second", "third", "fourth" and the like, if any, will be used to distinguish similar components from each other and used to describe a specific sequence or a generation sequence, and is not necessarily limited thereto. It will be understood that these terms are compatible with each other under an appropriate environment so that the exemplary embodiments of the disclosure set forth herein may be operated in a sequence different from a sequence illustrated or described herein. Likewise, in the case in which it is described herein that

a method includes a series of steps, a sequence of these steps suggested herein is not necessarily a sequence in which these steps may be executed. That is, any described step may be omitted and/or any other step that is not described herein may be added to the method.

[00126] As described above, the various embodiments described herein may be implemented in a computer-readable recording medium by using, for example, software, hardware or a combination thereof.

[00127] According to a hardware implementation, the embodiments described herein may be implemented using at least one of application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), processors, controllers, micro-controllers, microprocessors or electric units for performing other functions. In some cases, the embodiments described in the disclosure may be implemented by the processor itself.

[00128] According to a software implementation, the embodiments such as procedures and functions described in the specification may be implemented by separate software modules. Each of the software modules may perform one or more functions and operations described in the specification. A software code may be implemented as a software application written in a suitable programming

language. The software code may be stored in a memory module and executed by a control module.

[00129] The spirit of the disclosure has been illustratively described hereinabove. It will be appreciated by those skilled in the art to which the disclosure pertains those various modifications, alterations and substitutions may be made without departing from the essential features of the disclosure.

[00130] Accordingly, the embodiments and the accompanying drawings disclosed in the disclosure are provided not to limit the spirit of the disclosure, but to fully describe the disclosure, and the scope of the disclosure is not limited by the embodiments or the accompanying drawings. The scope of the disclosure should be interpreted by the following claims, and all the spirit equivalent to the following claims should be interpreted to fall within the scope of the disclosure.

【CLAIMS】

【Claim 1】

A neural network training method using a de-identified image, performed by a neural network training server, the method comprising:

encoding a first image represented by a vector of n -th dimensions into a second image of predetermined p -th dimensions;

decoding the second image into a third image of q -th dimensions;

inputting the third image to the neural network and extracting object information included in the third image; and

training at least one parameter information used for a computation in the neural network by using the extracted object information, and a de-identification parameter information used for the encoding or decoding computation based on the extracted object information,

wherein the training comprises updating the de-identification parameter information determining a compression rate of the encoding computation after updating the at least one parameter information used for the computation in the neural network by backpropagating an error of the extracted object information compared with a target value of the first image.

【Claim 2】

The method of claim 1, wherein a size of p-th dimensions is determined based on the compression rate of the encoding computation.

【Claim 3】

The method of claim 2, wherein n-th dimensions and q-th dimensions have the same size.

【Claim 4】

The method of claim 1, further comprising storing the second image encoded in the determined p-th dimensions,

wherein the decoding comprises decoding the stored second image into the third image when the neural network performs the training.

【Claim 5】

The method of claim 1, wherein the decoding comprises decoding the second image to have a data value different from that of the first image when decoding the second image in the q-th dimensions.

【Claim 6】

A neural network training server using a de-identified image, the server comprising:

a de-identification unit configured to encode a first

image represented by a vector of n -th dimensional vector into a second image of predetermined p -th dimension, and then decode the second image into a third image of q -th dimension; and

a training unit configured to input the third image to the neural network and extract an object information included in the third image to train at least one parameter information used for a computation in the neural network, and a de-identification parameter information used for the encoding or decoding computation based on the extracted object information,

wherein the training unit is configured to update, the de-identification parameter information determining a compression rate of the encoding computation after updating the at least one parameter information used for the computation in the neural network by backpropagating an error of the extracted object information compared with a target value of the first image.

【Claim 7】

The server of claim 6, further comprising a database storing the second image encoded in determined p -th dimensions,

wherein the de-identification unit is configured to decode the stored second image into the third image when the neural network performs the training.

【Claim 8】

The server of claim 6, wherein the de-identification unit is configured to decode the second image to have a data value different from that of the first image when decoding the second image in the q -th dimensions.

【Claim 9】

One or more computers and a program stored in a non-transitory computer-readable recording medium to allow the one or more computers to perform operations of each method of any one of claims 1 to 5 when executed by the one or more computers.