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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
07 September 2018 (07.09.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/159666 A1

(51) International Patent Classification:
G06N 3/08 (2006.01) *G06N 3/04* (2006.01)

(21) International Application Number:
PCT/JP2018/007476

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2017-038492 01 March 2017 (01.03.2017) JP
2018-026134 16 February 2018 (16.02.2018) JP

(71) Applicant: **OMRON CORPORATION** [JP/JP]; 801, Minamifudodo-cho, Horikawahigashiiru, Shiokoji-dori, Shimogyo-ku, Kyoto-shi, Kyoto, 6008530 (JP).

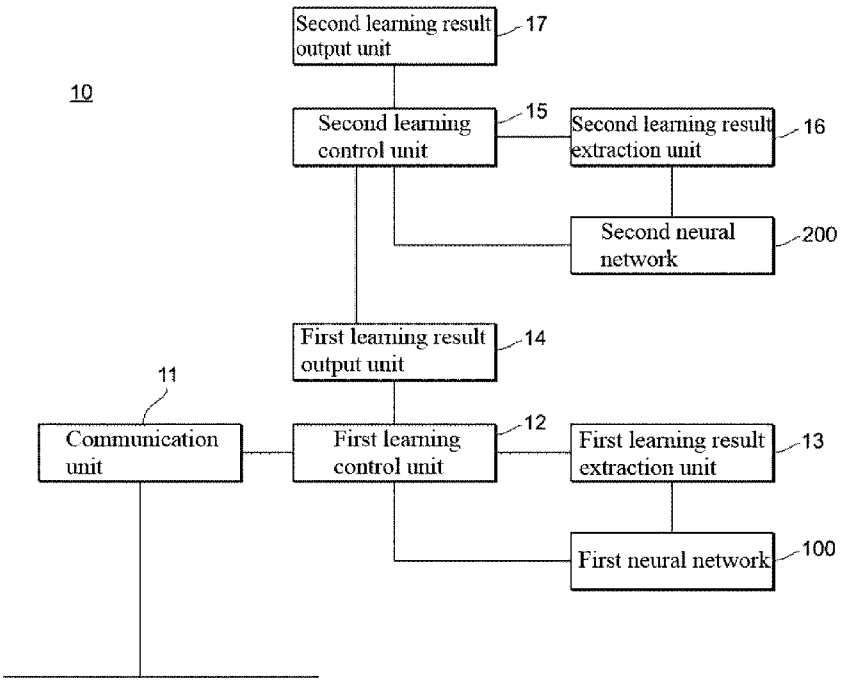
(72) Inventor: **ANDO, Tanichi**; c/o OMRON Automotive Electronics Co. Ltd., 6368, Okusa-Nenjozaka, Komaki-shi, Aichi, 4850802 (JP).

(74) Agent: **INABA, Yoshiyuki** et al.; TMI ASSOCIATES, 23rd Floor, Roppongi Hills Mori Tower, 6-10-1, Roppongi, Minato-ku, Tokyo, 1066123 (JP).

(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, 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.

WO 2018/159666 A1

(54) Title: LEARNING APPARATUS, LEARNING RESULT USING APPARATUS, LEARNING METHOD AND LEARNING PROGRAM



(57) Abstract: A technique for acquiring a desired learning result without increasing types of devices for obtaining training data is provided. A learning apparatus has a first learning control unit that trains a first learning module based on first training data and second training data associated with the first training data so as to output first output data corresponding to the features of the first training data and the second training data, and a second learning control unit that trains a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[Continued on next page]

(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))*

Description

Title of Invention: LEARNING APPARATUS, LEARNING RESULT USING APPARATUS, LEARNING METHOD AND LEARNING PROGRAM

Technical Field

[0001] The present invention relates to a learning apparatus, a learning result using apparatus, a learning method, and a learning program.

[0002] CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority to Japanese Patent Application No.2017-038492 filed March 1, 2017, the entire contents of which are incorporated herein by reference.

Background Art

[0003] In recent years, studies related to machine learning have been widely performed. In particular, due to the development of techniques called deep learning, learning modules that exhibit performance equivalent to or higher than the recognizing capability of humans have become available.

[0004] As an application example of machine learning, JP 2016-99165A describes a calculation apparatus that uses a pressure sensor that directly acquires a body weight and an image capturing apparatus that indirectly acquires a body weight and improves the accuracy of measurement by machine learning that uses measurement values of the pressure sensor and measurement values of the image capturing apparatus. JP 2016-99165A is an example of background art.

Summary of Invention

[0005] The apparatus described in JP 2016-99165A aims to acquire an accurate measurement result even if the measurement target moves by complementing the measurement value of the pressure sensor with the measurement value of the image capturing apparatus. However, with the apparatus described in JP 2016-99165A, it is necessary to install multiple types of measurement devices for machine learning in the environment in which the apparatus is used, and there are disadvantages such as the burden of installing multiple types of measurement devices, an increase in cost, and an increase in the size of the apparatus.

[0006] In view of this, it is an object of the present invention to provide a technique for acquiring a desired learning result in which multiple types of training data are incorporated without increasing the types of devices for obtaining training data.

[0007] A learning apparatus according to one aspect of the present invention has a first learning control unit that trains a first learning module based on first training data and second training data associated with the first training data so as to output first output

data corresponding to features of the first training data and the second training data, and a second learning control unit that trains a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[0008] According to this aspect, the first output data corresponding to the features of the first training data and the second training data is output by the first learning module that accepts the first training data and the second training data as input data, and the second output data is output by the second learning module that accepts the first training data as input data. The second learning module is trained by supervised learning in which supervisor data is the first output data, and thus the feature of the second training data is indirectly included in the second learning module. Therefore, the first learning module having a desired performance is generated by using both the first training data and the second training data, whereas the second learning module having the same performance as the first learning module is generated by using the first training data and the first output data from the first learning module (i.e., without using the second training data). That is, although training of the first learning module requires a device for obtaining the first training data (the first training data obtaining device) and a device for obtaining the second training data (the second training data obtaining device), training of the second learning module does not require the second training data obtaining device. As a result, it is possible to acquire a desired learning result in which multiple types of training data are incorporated without increasing types of devices for obtaining training data, and thus it is possible to reduce the number of items of hardware that are used and the processing load of the hardware processor due to a reduction in the data amount.

[0009] In the above aspect, the second learning control unit may train the second learning module after training of the first learning module.

[0010] According to this aspect, the second learning module can be trained using the first output data of the first learning module as supervisor data, after the first learning module learns the features of the first training data and the second training data, and thus the feature of the second training data is more accurately incorporated in the learning of the second learning module.

[0011] In the above aspect, the first training data may be data in the same form as input data that is input to a trained second learning module, which is acquired as a result of training performed by the second learning control unit, or a copy of the trained second learning module, and the second training data may be data temporally related to the first training data, and may be data in a form different from input data that is input to the trained second learning module or the copy of the trained second learning module.

[0012] According to this aspect, the first learning module can perform multilateral learning based on the first training data in the same form as the input data that is input to the trained second learning module and the second training data that complements or reinforces the first training data. In addition, if the first output data of the first learning module that performs multilateral learning serves as supervisor data, the second learning module can perform supervised learning that extracts a feature that is sometimes not extracted by training based only the first training data.

[0013] In the above aspect, the first learning control unit may train the first learning module by unsupervised learning based on the first training data and the second training data so as to output the first output data.

[0014] According to this aspect, the first output data corresponding to the features of the first training data and the second training data can be autonomously generated by the first learning module, making it possible to perform more objective feature extraction. In addition, it is not necessary to prepare supervisor data, and thus there is no processing load or communication load for generating and collecting supervisor data, and it is not necessary to secure the storage capacity for storing supervisor data.

[0015] In the above aspect, the first learning control unit may train the first learning module by supervised learning that uses supervisor data including attribute information of the first training data and the second training data, based on the first training data and the second training data, so as to output the first output data.

[0016] According to this aspect, it is possible to generate the first output data corresponding to the features of the first training data and the second training data in consideration of existing attribute information. In addition, it is not necessary to assign meaning to the first output data, and thus it is not necessary to perform calculation or communication in order to interpret the first output data, whereby the processing load and the communication load are suppressed.

[0017] In the above aspect, the first learning module and the second learning module may each include a neural network, and a scale of the neural network included in the second learning module may be smaller than a scale of the neural network included in the first learning module.

[0018] According to this aspect, high-load processing is performed in the learning apparatus that is relatively rich in calculation resources, and it is possible to suppress the scale of a neural network that is set in a learning result using apparatus to a small scale, and to suppress the processing load and the communication load of the learning result using apparatus.

[0019] In the above aspect, the first training data may include image data of a target, the second training data may include sensing data acquired by measuring the target using a sensor when the image data is shot, and the first output data and the second output data

may include data related to the target.

[0020] According to this aspect, the second learning module that outputs the second output data corresponding to a feature of the image data can indirectly learn a feature that is included in the sensing data and cannot be extracted from the image data, and the second learning module that outputs more accurate second output data is acquired.

[0021] In the above aspect, the first training data may include image data acquired by shooting a person, the second training data may include vital data of the person when the image data is shot, and the first output data and the second output data may be data corresponding to a human emotion.

[0022] According to this aspect, the second learning module that outputs the second output data corresponding to a feature of the person that was shot can indirectly learn a feature that is included in vital data and cannot be extracted from the image data, and the second learning module that outputs more accurate second output data is acquired.

[0023] In the above aspect, the first training data may include image data acquired by shooting a vehicle, the second training data may include sensing data acquired by performing measurement using a sensor provided in the vehicle when the image data is shot, and the first output data and the second output data may be data corresponding to an operation of the vehicle.

[0024] According to this aspect, the second learning module that outputs the second output data corresponding to a feature of the shot vehicle can indirectly learn a feature that cannot be extracted from the image data and is included in the sensing data, and the second learning module is acquired, which outputs the second output data that is more accurate.

[0025] The learning result using apparatus according to one aspect of the present invention has a learning module setting unit that acquires the trained second learning module acquired as a result of training performed by the second learning control unit provided in the learning apparatus of the above aspect or a copy of the trained second learning module, and sets the trained second learning module or the copy of the trained second learning module as a third learning module, an input unit for inputting data having the same form as the first training data to the third learning module, and an output unit for outputting output data from the third learning module.

[0026] According to this aspect, output data corresponding to a feature of input data is output by the third learning module that accepts the data having the same form as the first training data as the input data. The third learning module is set by the trained second learning module or the copy of the trained second learning module, and thus the third learning module indirectly includes a feature of the second training data. Therefore, it is possible to acquire a desired learning result in which multiple types of training data are incorporated without increasing types of devices for obtaining training

data.

[0027] A learning method according to one aspect of the present invention includes, training, by a control unit configured to control machine learning, a first learning module based on first training data and second training data associated with the first training data so as to output first output data corresponding to features of the first training data and the second training data, and training, by the control unit, a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[0028] According to this aspect, the first output data corresponding to the features of the first training data and the second training data is output by the first learning module that accepts the first training data and the second training data as input data, and the second output data is output by the second learning module that accepts the first training data as input data. The second learning module performs supervised learning in which supervisor data is the first output data, and thus data indirectly includes the feature of the second training data. Therefore, the second learning module in which the second training data is incorporated is acquired without using the second training data obtaining device.

[0029] A method for producing a trained learning module or a copy of the trained learning module according to one aspect of the present invention includes outputting a trained second learning module acquired as a result of training the second learning module by the learning method of the above aspect or a copy of the trained second learning module.

[0030] According to this aspect, the second learning module or a copy of the second learning module in which the second training data is incorporated is acquired without using the second training data obtaining device.

[0031] A trained learning module or a copy of the trained learning module according to one aspect of the present invention is acquired as a result of training the second learning module by the learning method of the above aspect.

[0032] According to this aspect, the trained learning module or the copy of the trained learning module in which the second training data is incorporated is acquired without using the second training data obtaining device.

[0033] A learning program according to one aspect of the present invention includes instructions which, when the program is executed by a computer, cause the computer to function as perform a method including training a first learning module based on first training data and second training data associated with the first training data so as to output first output data corresponding to features of the first training data and the

second training data, and training a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[0034] According to this aspect, the first output data corresponding to the features of the first training data and the second training data is output by the first learning module that accepts the first training data and the second training data as input data, and the second output data is output by the second learning module that accepts the first training data as input data. The second learning module performs supervised learning in which supervisor data is the first output data, and thus the second output data indirectly includes the feature of the second training data. Therefore, the second learning module in which the second training data is incorporated is acquired without using the second training data obtaining device.

[0035] According to the present invention, a technique is obtained that makes it possible to acquire a desired learning result in which multiple types of training data are incorporated without increasing types of devices for obtaining training data

Brief Description of Drawings

[0036] [fig.1]Fig. 1 is a diagram showing the network configuration of a learning apparatus and a learning result using apparatus according to an embodiment of the present invention.

[fig.2]Fig. 2 is a diagram showing the physical configuration of the learning apparatus according to the embodiment of the present invention.

[fig.3]Fig. 3 is a functional block diagram of the learning apparatus according to the embodiment of the present invention.

[fig.4]Fig. 4 is a functional block diagram of the learning result using apparatus according to the embodiment of the present invention.

[fig.5]Fig. 5 is a conceptual diagram showing the input/output relationship of a first neural network of the learning apparatus according to the embodiment of the present invention.

[fig.6]Fig. 6 is a conceptual diagram showing the input/output relationship of a second neural network of the learning apparatus according to the embodiment of the present invention.

[fig.7]Fig. 7 is a conceptual diagram showing the input/output relationship of a third neural network of the learning result using apparatus according to the embodiment of the present invention.

[fig.8]Fig. 8 is a flowchart of processing executed by the learning apparatus according to the embodiment of the present invention.

[fig.9]Fig. 9 is a flowchart of processing executed by the learning result using apparatus according to the embodiment of the present invention.

Description of Embodiments

[0037] Embodiments of the present invention will be described below with reference to the attached drawings. Note that in the figures, the same or similar constituent elements are denoted by the same reference numerals.

[0038] Fig. 1 is a diagram showing the network configuration of a learning apparatus 10 and a learning result using apparatus 20 according to an embodiment of the present invention. The learning apparatus 10 according to this embodiment is connected to the learning result using apparatus 20, one or more sensors 30 and a sensing data storage DB via a communication network N. The communication network N may be either a wired communication network or a wireless communication network constituted by a wired or wireless line, or may be the Internet or a LAN (Local Area Network). Note that in Fig. 1, the sensing data storage DB, the learning apparatus 10 and the learning result using apparatus 20 are configured separately, but may be configured integrally. Specifically, the sensing data storage DB, the learning apparatus 10 and the learning result using apparatus 20 may all be configured to be integrated, or two out of the sensing data storage DB, the learning apparatus 10 and the learning result using apparatus 20 may be selectively configured to be integrated. Here, in the case where the sensing data storage DB, the learning apparatus 10 and the learning result using apparatus 20 are configured to be integrated, the elements thereof are connected to each other via an internal bus.

[0039] The learning apparatus 10 trains a first learning module and a second learning module based on training data including at least one of sensing data acquired from the sensor 30 and sensing data stored in the sensing data storage DB. The learning apparatus 10 according to this embodiment is provided with the first learning module and the second learning module, but the first learning module and the second learning module may be provided in an apparatus separated from the learning apparatus 10. Note that a learning module includes a unit of dedicated or general-purpose hardware or software having a learning capability, or a combination of units of such hardware and software. Here, the learning capability refers to the ability to improve a capability of processing a certain task based on experience acquired from training data.

[0040] The learning result using apparatus 20 outputs output data corresponding to the feature of input data using a learning result of the learning apparatus 10. The learning apparatus 10 according to this embodiment acquires, from the learning apparatus 10, the trained second learning module or a copy of the trained second learning module, and sets the trained second learning module or the copy of the trained second learning

module as a third learning module. Note that a copy of a trained learning module includes a unit of dedicated or general-purpose hardware or software that can reproduce a function of the trained learning module, or a combination of units of such hardware or software. A copy of a trained learning module does not necessarily need to have a learning capability. In addition, the configuration of a trained learning module and the configuration of a copy of the trained learning module do not necessarily need to match each other. In addition, a copy of a trained learning module includes a trained learning module or a copy of the trained learning module that has completed trained and also performed additional training. In the case of the second learning module according to this embodiment, a copy of the trained second learning module includes a learning module acquired as a result of causing the trained second learning module or a copy of the trained second learning module to perform additional training based on additional training data in the same form as first training data so as to output additional output data. A copy of the trained second learning module also includes a learning module acquired as a result of causing the trained second learning module or a copy of the trained second learning module to perform additional training based on first training data so as to output additional output data. A copy of a trained learning module further includes a learning module acquired by so-called distillation. Specifically, a copy of a trained learning module includes another trained learning module that has a structure different from that of the trained learning module and has been trained so as to have a function of the trained learning module. Here, the structure of the other learning module may be simpler than the structure of the trained learning module, may be more suitable for being deployed, and output data of the trained learning module may be used for the training of the other learning module. Note that a copy of a trained learning module includes a trained learning module that is acquired by changing a method for normalization for preventing overfitting, changing a learning rate of back propagation, or changing an updating algorithm of a weight coefficient, in the training process of the learning module. In addition, acquiring the trained second learning module or a copy of the trained second learning module refers to acquiring information required to reproduce, in the learning result using apparatus 20, a function of the trained second learning module. For example, if the second learning module includes a neural network, acquiring the trained second learning module or a copy of the trained second learning module refers to acquiring at least information regarding the number of layers of the neural network, the number of nodes for each of the layers, weight parameters of links connecting nodes, bias parameters for the nodes, and the functional types of activation functions the nodes.

[0041] The sensor 30 may be either a physical quantity sensor that detects a physical quantity or an information sensor that detects information. Examples of the physical

quantity sensor include cameras that detect light and output image data or moving image data, and vital sensors such as heartbeat sensors that detect heartbeat of a person and output heartbeat data, blood pressure sensors that detect blood pressure of a person and output blood pressure data, and body temperature sensors that detect human body temperature and output body temperature data, and also include any other sensors that detect a physical amount and output an electric signal. Examples of the information sensor include sensors that detect a specific pattern from statistical data, and also include any other sensors that detect information.

[0042] The sensing data storage DB stores sensing data that has been output by the sensor 30. In the figure, the sensing data storage DB is shown as a single storage, but the sensing data storage DB may be constituted by one or more file servers.

[0043] Fig. 2 is a diagram showing the physical configuration of the learning apparatus 10 according to the embodiment of the present invention. The learning apparatus 10 has a CPU (Central Processing Unit) 10a equivalent to a hardware processor, a RAM (Random Access Memory) 10b equivalent to a memory, a ROM (Read only Memory) 10c equivalent to a memory, a communication interface 10d, an input unit 10e and a display unit 10f. These constituent elements are connected via a bus so as to be able to exchange data with each other. Note that the type of the hardware processor is not limited to a CPU. For example, a CPU, a GPU (Graphics Processing Unit), an FPGA (Field-programmable Gate Array), a DSP (Digital Signal Processor), and an ASIC (Application Specific Integrated Circuit) can be used independently or in combination as a hardware processor.

[0044] The CPU 10a performs execution of a program stored in the RAM 10b or the ROM 10c and calculation and processing of data. The CPU 10a is a calculation apparatus that executes an application for generating metadata. The CPU 10a receives various types of input data from the input unit 10e or the communication interface 10d, and displays calculation results of the input data on the display unit 10f, and stores the calculation results in the RAM 10b or the ROM 10c.

[0045] The RAM 10b is a data-rewritable storage, and is constituted by a semiconductor storage element, for example. The RAM 10b stores programs such as applications executed by the CPU 10a and data.

[0046] The ROM 10c is a data-read-only storage, and is constituted by a semiconductor storage element, for example. The ROM 10c stores programs such as firmware and data, for example.

[0047] The communication interface 10d is a hardware interface that connects the learning apparatus 10 to the communication network N.

[0048] The input unit 10e accepts input of data from the user, and is constituted by a keyboard, a mouse, or a touch panel, for example.

[0049] The display unit 10f visually displays a result of calculation performed by the CPU 10a, and is constituted by an LCD (Liquid Crystal Display), for example.

[0050] The learning apparatus 10 may be configured by a learning program according to this embodiment being executed by the CPU 10a of a general personal computer. The learning program may be stored in a computer-readable storage medium such as the RAM 10b or the ROM 10c and provided, or may be provided via the communication network N connected by the communication interface 10d.

[0051] Note that these physical configurations are examples, and do not necessarily need to be independent configurations. For example, the learning apparatus 10 may have an LSI (Large-Scale Integration) in which the CPU 10a and the RAM 10b or the ROM 10c are integrated.

[0052] Note that the learning result using apparatus 20 also has a physical configuration similar to that of the learning apparatus 10. The learning result using apparatus 20 may be configured by a learning result using program being executed by a CPU of a general personal computer. The learning result using program may be stored in a computer-readable storage medium such as a RAM or a ROM and provided, or may be provided via the communication network N connected by a communication interface.

[0053] Fig. 3 is a functional block diagram of the learning apparatus 10 according to the embodiment of the present invention. The learning apparatus 10 has a communication unit 11, a first learning control unit 12, a first learning result extraction unit 13, a first neural network 100, a first learning result output unit 14, a second learning control unit 15, a second learning result extraction unit 16, a second neural network 200 and a second learning result output unit 17. Here, the first learning control unit 12 and the second learning control unit 15 are control units that control machine learning. In addition, the first neural network 100 is an example of the first learning module, and the second neural network 200 is an example of the second learning module. The learning apparatus 10 may have a learning module other than a neural network.

[0054] The first learning control unit 12 trains the first neural network 100 based on first training data and second training data associated with the first training data so as to output first output data corresponding to the features of the first training data and the second training data. The first training data may be image data of a target, for example, and the second training data may be sensing data acquired by a sensor measuring the target or performing measurement with regard to the target when the image data was shot. In this case, the first output data is data corresponding to the features of the image data and the sensing data, and is data regarding the target that is shot. The first neural network 100 may be a CNN (Convolutional Neural Network) that is sometimes used for learning of image data, or an RNN (Recurrent Neural Network) that is sometimes used for learning of time series data. A learning result of the first neural network 100 is

extracted by the first learning result extraction unit 13, and is output to the second learning control unit 15 by the first learning result output unit 14.

[0055] The first learning control unit 12 may train the first neural network 100 by unsupervised learning based on first training data and second training data so as to output first output data. By training the first neural network 100 by unsupervised learning, the first output data that is based on the features of the first training data and the second training data can be autonomously generated by the first neural network 100, and feature extraction with higher objectivity can be performed. In addition, it is not necessary to prepare supervisor data, and thus there is no processing load or communication load for generating and collecting supervisor data, and it is not necessary to secure storage capacity for storing supervisor data.

[0056] The first learning control unit 12 may train the first neural network 100 by supervised learning that uses supervisor data including attribute information of first training data and second training data, so as to output first output data based on the first training data and the second training data. Here, attribute information of training data is information indicating a feature of the training data, and may include information regarding the type of a physical amount measured by a sensor, the type of the sensor, the type of sensing data and a target measured by the sensor. By training the first neural network 100 by supervised learning, it is possible to generate the first output data corresponding to the features of the first training data and the second training data in consideration of the existing attribute information. In addition, it is not necessary to assign a meaning such as a label and an annotation to the first output data, and thus it is not necessary to perform calculation or communication in order to interpret the first output data, and the processing load and the communication load are suppressed.

[0057] The second learning control unit 15 trains the second neural network 200 by supervised learning in which the supervisor data is the first output data that is output from the first neural network 100 in the case where the first training data is input to the first neural network 100, based on the first training data, so as to output second output data. By performing the supervised learning using the first output data as the supervisor data, the second neural network 200 shares the learning objective with the first neural network 100 and acquires the same type of capability as the first neural network 100. Specifically, both of the second output data output from the second neural network 200 and the first output data output from the first neural network 100 are data relating to the same subject and expressed in the same form. The same type of capability may include the capability for performing at least one of analysis, estimation, control with respect to the same (or substantially the same) target, state or operation, and the capability for performing determination, identification, recognition with respect to the same (or substantially the same) requirement. The data relating to the

same subject and expressed in the same form includes, for example, data indicating control values for the same variables in the same unit, and data indicating scores for the same determination (the quality of an item, presence of an object, or the like) according the the same rule. If the first training data is image data of a target, and the second training data is sensing data in the same time series as the image data, the supervisor data is the first output data that is output from the trained first neural network 100 in the case where the image data is input to the trained first neural network 100, and the second output data that is output from the second neural network 200 in the case where the image data is input to the second neural network 200 is data relating to the same subject and expressed in the same form as the first output data, that is, data corresponding to the feature of the image data, and is data regarding the target that is shot. A learning result of the second neural network 200 is extracted by the second learning result extraction unit 16, and is output to the outside via the communication unit 11 by the second learning result output unit 17.

[0058] Note that in this embodiment, the first training data used for learning performed by the first neural network 100 and the first training data used for learning performed by the second neural network 200 are the same data, but the present invention is not limited to this example.

As long as the first training data used for learning performed by the first neural network 100 and the first training data used for learning performed by the second neural network 200 have the same form (or the same type), both data may differ in contents. Specifically, the first training data used for learning performed by the first neural network 100 and the first training data used for learning performed by the second neural network 200 are data in the same form, but may be data in which part of all of the content is different. For example, a configuration may be adopted in which, in the case where image data of a first group as the first training data and sensing data as the second training data were used in learning performed by the first neural network 100, when the second neural network 200 performs learning, image data of a second group is input to the trained first neural network 100 as the first training data, and the second neural network 200 performs learning based on the image data of the second group with the first output data that is output from the trained first neural network 100 serving as supervisor data. A form of data indicates, for example, the form of images (e.g., colour images, infrared images, and range images) or the form of numerical values (e.g., binary, and continuous values). Data in the same form may include data obtained by the same type of data obtaining devices such as cameras, sensors, and measurement devices, and data in the different forms may include data obtained by the different types of data obtaining devices. In addition, data in the same form may include data obtained for the same target such as a subject of images and a sensing

target object, by the same type of data obtaining device, and data in the different form may include data obtained for the different targets. In this embodiment, the image data of the first group and the image data of the second group are both image data (i.e., the data in the same form), and the image data of the second group may or may not include the same pieces of image data as the image data of the first group.

[0059] With the learning apparatus 10 according to this embodiment, the first output data corresponding to the features of the first training data and the second training data is output by the first neural network 100 that accepts the first training data and the second training data as input data, and the second output data corresponding to the feature of the first training data is output by the second neural network 200 that accepts the first training data as input data. The second neural network 200 performs supervised learning in which supervisor data is the first output data, and thus the second output data indirectly includes the feature of the second training data. Therefore, a neural network having a desired performance without increasing types of measurement devices for obtaining training data is acquired. Specifically, with the learning apparatus 10 according to this embodiment, a neural network is acquired which provides the same performance as that in a case where a plurality of types of measurement devices that obtain the first training data and the second training data are used without using a measurement device for the second training data. A neural network in which a desired measurement result is incorporated without using a measurement device for the second training data is acquired, and thus it is possible to reduce the number of items of hardware of the learning result using apparatus 20 that uses the trained neural network, and to further reduce the processing load of the hardware processor due to a reduction in data amount.

[0060] After the first neural network 100 was trained, the second learning control unit 15 trains the second neural network 200. Accordingly, after the first neural network 100 learned the features of the first training data and the second training data, the second neural network 200 can be trained using, as supervisor data, the first output data that is output from the first neural network 100, and thus the feature of the second training data is more accurately reflected on the training of the second neural network 200.

[0061] Fig. 4 is a functional block diagram of the learning result using apparatus 20 according to the embodiment of the present invention. The learning result using apparatus 20 has a learning result input unit 231, a neural network setting unit 232, a third neural network 233, a control unit 234, an input unit 235, a communication unit 236, a data acquiring unit 237 that acquires data to be input to the third neural network 233, and an output unit 238. Here, the third neural network 233 is an example of a learning module, and the learning result using apparatus 20 may have a learning module other than a neural network, and in that case, the neural network setting unit

232 will be replaced by a constituent element that sets a learning module other than a neural network. Note that the data acquiring unit 237 may acquire data via the communication unit 236, or may acquire data via communication other than communication using the communication unit 236.

[0062] The learning result input unit 231 accepts input of a learning result. The learning result input unit 231 accepts, via the communication unit 236, a learning result that is output by the second learning result output unit 17 of the learning apparatus 10. The neural network setting unit 232 acquires the trained second neural network 200 acquired as a result of training by the second learning control unit 15 provided in the learning apparatus 10 or a copy of the trained second neural network 200, and sets the trained second neural network 200 or the copy of the trained second neural network 200 as the third neural network 233. The control unit 234 controls the data acquiring unit 237 and the input unit 235 so as to input designated input data to the third neural network 233 and to output output data. The input unit 235 inputs data having the same form as the first training data to the third neural network 233. The output unit 238 outputs the output data from the third neural network 233. The output data from the third neural network 233 is output by the output unit 238 via the communication unit 236.

[0063] With the learning result using apparatus 20 according to this embodiment, output data corresponding to the feature of input data is output by the third neural network 233 that accepts, as input data, data having the same form as the first training data. The third neural network 233 is set by the trained second neural network 200 or a copy of the trained second neural network 200, and thus the third neural network 233 indirectly includes the feature of the second training data. Therefore, a learning module having a desired performance is acquired without increasing types of measurement devices. As a result, in the environment in which the third neural network 233 is used, a desired learning result can be acquired even without using a measurement device used for obtaining sensing data (second training data), and it is possible to reduce the number of items of hardware that constitute the learning result using apparatus 20, and to further reduce the processing load of the hardware processor due to a reduction in the data amount.

[0064] In this embodiment, the first training data may be data in the same form as input data that is input to the trained second neural network 200 acquired as a result of training by the second learning control unit 15 of the learning apparatus 10 or a copy of the trained second neural network 200. In addition, the second training data may be data temporally related to the first training data. Further, the second training data may be data in a form different from that of the input data that is input to the trained second neural network 200 or the copy of the trained second neural network 200. The second

training data is data that complements or reinforces the first training data, and is data for extracting a feature that cannot be extracted through training that is based only on the first training data. Each piece of the second training data may be obtained at the same time as or in proximity to when the corresponding piece of the second training data is obtained.. The second training data temporally related to the first training data includes the second training data obtained within the predetermined period of time before or after the corresponding first training data is obtained. Accordingly, the first neural network 100 can perform multilateral learning based on the first training data in the same form as the input data that is input to the trained second neural network 200, and the second training data that complements or reinforces the first training data. In addition, if the first output data of the first neural network 100 that performed multilateral learning serves as supervisor data, the second neural network 200 can perform supervised learning that extracts a feature that is sometimes not extracted through learning that is based only on the first training data.

[0065] In addition, in the learning apparatus 10 according to this embodiment, the scale of the second neural network 200 is smaller than the scale of the first neural network 100. Here, the scale of a neural network is a scale measured based on the number of nodes, the number of links, the number of layers and the like included in the neural network. Due to the scale of the second neural network 200 being smaller than the scale of the first neural network 100, the learning apparatus 10 that is relatively rich in calculation resources performs high-load processing, and thus the scale of the third neural network 233 that is set in the learning result using apparatus 20 can be suppressed to a small scale, and the processing load and communication load of the learning result using apparatus 20 can be suppressed.

[0066] Fig. 5 is a conceptual diagram showing the input/output relationship of the first neural network 100 of the learning apparatus 10 according to the embodiment of the present invention. In the example shown in the figure, first training data is image data acquired by shooting a person, and second training data is vital data of the person at the time when the image data was shot. Note that the time when the image data was shot is a concept that includes the same time as the shooting of the image data and the temporal vicinity before and after. In addition, the first training data includes first image data 301, second image data 302 and third image data 303. Also, the second training data includes first vital data 401, second vital data 402 and third vital data 403. Here, the first vital data 401 is vital data of a subject person at the time when the first image data 301 was shot. Accordingly, the first vital data 401 is data that is the same as the first image data 301 in time series. Similarly, the second vital data 402 is vital data of the subject person at the time when the second image data 302 was shot, and the third vital data 403 is vital data of the subject person at the time when the third image

data 303 was shot. Note that vital data is any biological data such as a heart rate, a blood pressure, a body temperature, a blood component amount, a urine component amount, or a brain wave.

[0067] The learning apparatus 10 trains the first neural network 100 based on first training data and second training data so as to output first output data corresponding to the features of the first training data and the second training data. In the case of this example, the first output data includes first data 501, second data 502 and third data 503, each of which is numeric data. The first data 501 is output data that is output in the case where the first image data 301 and the first vital data 401 are input as input data to the first neural network 100, and is a three-dimensional numeric vector "(0.9, 0.05, 0.05)" in the case of this example. Similarly, the second data 502 is output data that is output in the case where the second image data 302 and the second vital data 402 are input as input data to the first neural network 100, and is a three-dimensional numeric vector "(0.05, 0.9, 0.05)". In addition, the third data 503 is output data that is output in the case where the third image data 303 and the third vital data 403 are input as input data to the first neural network 100, and is a three-dimensional numeric vector "(0.05, 0.05, 0.9)". The first output data is data corresponding to a human emotion, and each component indicates a degree of correspondence corresponding to a predetermined emotion. The larger the numeric value of the component is, the higher the reliability that is determined to indicate an emotion corresponding to the component is.

[0068] If the first learning control unit 12 trains the first neural network 100 by unsupervised learning, the user of the learning apparatus 10 compares input data and output data of the first neural network 100, and assigns meanings to the output data. In this example, a meaning "anger" is assigned to the first data 501, a meaning "relaxation" is assigned to the second data 502, and a meaning "smile/laughter" is assigned to the third data 503.

[0069] If the first learning control unit 12 trains the first neural network 100 by supervised learning that uses supervisor data including attribute information of first training data and second training data, the user of the learning apparatus 10 does not need to assign meanings to the output data. The first neural network 100 autonomously learns that a first component included in the three dimensional vector that is output as the output data is an amount indicating a degree of anger, a second component is an amount indicating a degree of relaxation, and a third component is an amount indicating a degree of smile/laughter.

[0070] The learning apparatus 10 can acquire a learning result that makes it possible to estimate a human emotion more accurately than in a case of using only image data as training data, by training the first neural network 100 using both image data and vital data as training data. Here, the image data is data that can be acquired by a camera,

which is a common sensor, and is data that can be acquired without mounting a sensor to a person to be shot. On the other hand, the vital data is data that cannot be acquired unless a dedicated sensor is used, and is data that cannot be acquired unless a sensor is mounted to the person to be shot. Generally, the learning apparatus 10 may train the first neural network 100 by combining first training data that is relatively easy to acquire and second training data that is relatively difficult to acquire, but complements or reinforces the first training data.

[0071] Fig. 6 is a conceptual diagram showing the input/output relationship of the second neural network 200 of the learning apparatus 10 according to the embodiment of the present invention. First training data shown in this figure is the same as the first training data shown in Fig. 5, and includes first image data 301, second image data 302 and third image data 303.

[0072] The learning apparatus 10 trains the second neural network 200 by supervised learning in which the supervisor data is the first output data that is output from the first neural network 100 in the case where first training data is input to the first neural network 100, based on the first training data, so as to output second output data. In the case of this example, the second output data includes fourth data 601, fifth data 602 and sixth data 603, each of which is numeric data. The fourth data 601 is output data that is output in the case where the first image data 301 is input as input data to the second neural network 200, and is a three-dimensional numeric vector "(0.96, 0.02, 0.02)" in the case of this example. Similarly, the fifth data 602 is output data that is output in the case where the second image data 302 is input as input data to the second neural network 200, and is a three-dimensional numeric vector "(0.02, 0.96, 0.02)". In addition, the sixth data 603 is output data that is output in the case where the third image data 303 is input as input data to the second neural network 200, and is a three-dimensional numeric vector "(0.02, 0.02, 0.96)". Similarly to the first output data, the second output data is data corresponding to a human emotion.

[0073] The second learning control unit 15 trains the second neural network 200 by supervised learning in which the supervisor data is the first output data that is output from the trained first neural network 100 in the case where the first training data is input to the trained first neural network 100, and thus the user of the learning apparatus 10 does not have to assign meanings to the second output data. The second neural network 200 autonomously learns that the first component included in the three dimensional vector that is output as the second output data is an amount indicating the degree of anger, the second component is an amount indicating the degree of relaxation, and the third components is an amount indicating the degree of smile/laughter.

[0074] The learning apparatus 10 trains the second neural network 200 using, as supervisor

data, output data that is output from the trained first neural network 100 in the case where the first training data is input to the trained first neural network 100, and thereby can acquire a learning result that includes vital data, using only image data as training data, and can acquire a learning result that makes it possible to estimate a human emotion more accurately. Here, the image data is data that can be acquired by a camera, which is a common sensor, and thus the trained second neural network 200 can exhibit, using only image data that is relatively easy to acquire as input data, identification performance similar to that in the case where sensing data that is relatively difficult to acquire is used for complementation.

[0075] The second neural network 200 is trained using, as supervisor data, the first output data of the first neural network 100 that was trained based on image data and sensing data, so as to output second output data, and thereby the second neural network 200 can indirectly learn a feature that cannot be extracted from only the image data, and the second neural network 200 in which the sensing data is incorporated is acquired. As a result, in an environment using the second neural network as the second learning module, a desired learning result can be acquired even without using a measurement device used for obtaining the sensing data (the second training data), and it is possible to reduce the number of items of hardware that is used, and to further reduce the processing load of the hardware processor due to a reduction in the data amount.

[0076] In addition, the second neural network 200 is trained using, as supervisor data, the first output data of the first neural network 100 that was trained based on image data and vital data of a human, so as to output second output data, and thereby the second neural network 200 can indirectly learn a feature that cannot be extracted from only the image data, and the second neural network 200 that can estimate a human emotion more accurately is acquired. As a result, in an environment using the second neural network as the second learning module, a desired learning result can be acquired without using a measurement device used for obtainingthe vital data (second training data), and it is possible to reduce the number of items of hardware that is used, and to further reduce the processing load of the hardware processor due to a reduction in the data amount.

[0077] Note that in this example, in order to simplify the description, a case has been described in which the number of types of the features of the first training data is three, but generally, a large number of features, namely, four or more features are included in first training data. For example, if thousands of types of features are included in first training data, the first neural network 100 and the second neural network 200 are trained so as to classify the thousands of types of the features of the first training data, determine which of the thousands of types of classifications input data is close to, and output output data corresponding to the features of the input data.

[0078] Note that in this example, the learning apparatus 10 has been described which has the first neural network 100 and the second neural network 200, and performs training using first training data and second training data, but the configuration of the learning apparatus 10 is not limited to this example. Accordingly, the learning apparatus 10 may have three or more neural networks, and may be configured to perform training using training data of three types or more. For example, the learning apparatus 10 may have a first neural network that is trained based on first training data, second training data and third training data so as to output first output data corresponding to the features of the first training data, the second training data and the third training data, and a second neural network that performs supervised learning in which supervisor data is the first output data, based on the first training data, so as to output second output data. In addition, for example, the learning apparatus 10 may have a first neural network that is trained based on first training data, second training data and third training data so as to output first output data corresponding to the features of the first training data, the second training data and the third training data, a second neural network that performs supervised learning in which supervisor data is the first output data, based on the first training data and the second training data, so as to output second output data, and a third neural network that performs supervised learning in which supervisor data is the second output data, based on the first training data, so as to output third output data. In addition, for example, the learning apparatus 10 may have a first neural network that is trained based on first training data and second training data so as to output first output data corresponding to the features of the first training data and the second training data, and a plurality of second neural networks that perform supervised learning in which supervisor data is the first output data, based on the first training data, so as to output second output data. Here, the plurality of second neural networks may each have a different neural network structure regarding the number of layers, the number of units and the number of links, and may each output different second output data.

[0079] Fig. 7 is a conceptual diagram showing the input/output relationship of the third neural network 233 of the learning result using apparatus 20 according to the embodiment of the present invention. Input data shown in the figure includes fourth image data 310.

[0080] The learning result using apparatus 20 acquires the trained second neural network 200 acquired as a result of training by the second learning control unit 15 provided in the learning apparatus 10 or a copy of the trained second neural network 200, and sets, as the third neural network 233, the trained second neural network 200 or the copy of the trained second neural network 200. The third neural network 233 accepts, as input data, data having the same form as first training data. In the case of this example, the

data having the same form as the first training data is image data. In addition, the third neural network 233 outputs output data corresponding to the feature of the input data. In the case of this example, the output data is seventh data 701, and the seventh data 701 is numeric data. The seventh data 701 is output data that is output in the case where the fourth image data 310 is input as input data to the third neural network 233, and is a three-dimensional numeric vector "(0.02, 0.02, 0.96)" in the case of this example. The output data of the third neural network 233 is data corresponding to a human emotion, and the output data in this example is data corresponding to "smile/ laughter".

[0081] The learning result using apparatus 20 acquires the trained second neural network 200 or a copy of the trained second neural network 200, and sets, as the third neural network 233, the trained second neural network 200 or the copy of the trained second neural network 200, and thereby a learning result including vital data can be used even if input data is image data only, and a human emotion can be estimated more accurately. Here, the image data is data that can be acquired by a camera, which is a common sensor, and thus the third neural network 233 of the learning result using apparatus 20 can exhibit an identification performance similar to that in the case where only image data that is relatively easy to acquire is used as input data, and sensing data that is relatively hard to acquire is used for complementation.

[0082] First training data and second training data are not limited to image data and vital data of a person. For example, vital data of a person may be used as first training data, and image data of the person may be used as second training data. Accordingly, the image data of a person may be used as data for complementing or reinforcing the vital data. By setting the vital data of a person as the first training data and setting the image data of the person as the second training data, a neural network is acquired which can estimate an emotion and a mental state of the person more accurately in consideration of the image data even in the case where input data is the vital data only.

[0083] In addition, for example, first training data may include image data acquired by shooting a vehicle, and second training data may include sensing data measured by a sensor provided in the vehicle at the time when the image data was shot. More specifically, image data of a second vehicle that was shot by a camera provided in a first vehicle in the state where the first vehicle was following the second vehicle may be used as the first training data, and sensing data measured by a sensor provided in the second vehicle may be used as the second training data. Here, the sensor provided in the second vehicle may be a sensor that measures an operation of the accelerator pedal of the second vehicle, an operation of the brake pedal, a steering operation, a winker operation, and the state of the driver.

[0084] In this case, the first neural network 100 is trained based on the image data of the

second vehicle that has been shot from the first vehicle and the sensing data related to a measured operation of the second vehicle, and first output data of the first neural network 100 will be data corresponding to the operation of the vehicle. Note that the data corresponding to the operation of the vehicle includes a speed, acceleration, a traveling direction vector, probability of course change, and the like. In addition, the second neural network 200 performs supervised learning in which the supervisor data is the first output data that is output from the trained first neural network 100 in the case where the image data of the second vehicle shot from the first vehicle is input to the trained first neural network 100, based on the image data of the second vehicle shot from the first vehicle, and second output data of the second neural network 200 is data corresponding to the operation of the vehicle, similar to the first output data.

[0085] Note that the second training data may include information regarding the relative distance between the first vehicle and the second vehicle. Operations of vehicles changes in a large amount according to the distance between a leading vehicle and a following vehicle. Therefore, if the second training data includes information regarding the relative distance, it is possible to improve the accuracy of operation estimation of the vehicle, which will be described later. At this time, the relative distance can be acquired by the following method. For example, on a test course on which a measurement apparatus that identifies the position of a vehicle is provided, the relative distance between the first vehicle and the second vehicle can be measured while shooting the second vehicle using a camera provided in the first vehicle. In addition, the distance between the first vehicle and the second vehicle can be acquired by attaching a focus detection apparatus (e.g., a laser radar) at the front of the first vehicle or the rear of the second vehicle. The information regarding the relative distance may be estimated based on an image from a camera provided on a general road. In addition, a configuration may be adopted in which the first vehicle and the second vehicle built as physical models run in a virtual space, and image data as the first training data, sensor data as the second training data and the relative distance are acquired from the virtual space.

[0086] The second neural network 200 is trained using, as supervisor data, the first output data of the first neural network 100 that was trained based on the image data and the sensing data of the vehicle, so as to output the second output data, in this manner, and thereby the second neural network 200 can indirectly learn a feature that cannot be extracted from only the image data of the vehicle, and the second neural network 200 that can perform operation estimation of the vehicle more accurately is acquired. In addition, it is relatively difficult for a following vehicle to acquire sensing data acquired by measuring an operation of a leading vehicle, but with the learning result using apparatus 20 according to this embodiment, by acquiring the trained second

neural network 200 or a copy of the trained second neural network 200, setting, as the third neural network 233, the trained second neural network 200 or the copy of the trained second neural network 200, and inputting image data of the leading vehicle to the third neural network 233, it is possible to perform operation estimation of the vehicle.

[0087] First training data and second training data may be data other than the above. For example, the first neural network 100 and the second neural network 200 may perform learning based on image data acquired by shooting a person and sensing data that has been output from a sensor that detects action of the person, the image data serving as first training data and the sensing data serving as second training data, so as to output data corresponding to the action of the person as first output data and second output data. In this case, the sensor that detects action of a person may be a momentum sensor or an acceleration sensor that is mounted to a person, or a sensor that is provided on a target that is operated by a person and detects an operation performed by the person. Accordingly, it is possible to output the second output data for predicting the next action of the person in the case where the image data acquired by shooting the person is input to the second neural network 200.

[0088] In addition, for example, the first neural network 100 and the second neural network 200 may be trained based on image data acquired by shooting a fruit and sensing data that has been output from a sensor that measures the degree of maturation of the fruit, the image data serving as first training data and the sensing data serving as second training data, so as to output data corresponding to the degree of the maturation of the fruit as first output data and second output data. In this case, the sensor that measures the degree of maturation of a fruit may be a weight sensor, a hardness sensor, a sugar content sensor or the like. Accordingly, it is possible to output the second output data that estimates the degree of maturation of the fruit in the case where the image data acquired by shooting the fruit is input to the second neural network 200.

[0089] In addition, for example, the first neural network 100 and the second neural network 200 may perform learning based on image data acquired by shooting the appearance of a substrate onto which electric parts are fixed by soldering and sensing data that has been output from a sensor that measures the state of the soldering (e.g., an air content of the soldering, denaturation due to overheat, and an unjoined state due to heating shortage), the image data serving as first training data and the sensing data serving as second training data, so as to output, as first output data and second output data, data corresponding to whether or not a soldering inspection criteria is met. Accordingly, it is possible to output the second output data that estimates the state of the soldering in the case where the image data acquired by shooting the appearance of the substrate is input to the second neural network 200. If the second neural network 200 that was

trained in this manner is used in a substrate inspection apparatus for checking the state of soldering between a substrate and electric parts placed on the substrate, data corresponding to whether or not the soldering inspection criteria is met can be acquired without using the sensor that measures the state of soldering, and thus it is possible to reduce the number of items of hardware of the substrate inspection apparatus, and to further reduce the processing load of the hardware processor due to a reduction in data amount.

[0090] Fig. 8 is a flowchart of processing executed by the learning apparatus 10 according to the embodiment of the present invention. The learning apparatus 10 designates first training data and second training data based on an instruction accepted from the user (step S10). After that, the learning apparatus 10 determines whether or not supervised learning is to be performed (step S11). Here, whether or not supervised learning is to be performed may be determined based on the instruction accepted from the user.

[0091] If the learning apparatus 10 determines that supervised learning is to be performed (step S11: Yes), the learning apparatus 10 designates supervisor data based on an instruction accepted from the user (step S12). The learning apparatus 10 trains the first neural network 100 by supervised learning based on the designated first training data, second training data and supervisor data (step S13).

[0092] On the other hand, if the learning apparatus 10 determines that supervised learning is not to be performed (step S11: No), the learning apparatus 10 trains the first neural network 100 by unsupervised learning based on the designated first training data and second training data (step S14).

[0093] In both cases, the learning apparatus 10 trains the second neural network 200 by supervised learning in which supervisor data is first output data that has been output from the first neural network 100, based on the designated first training data (step S15). The processing performed by the learning apparatus 10 then ends.

[0094] The trained second neural network 200 or a copy of the trained second neural network 200 can be generated by using the learning apparatus 10 according to this embodiment. Specifically, the trained second neural network 200 or a copy of the trained second neural network 200 can be generated by the first learning control unit 12 training the first neural network 100 based on first training data and second training data, so as to output first output data corresponding to the features of the first training data and the second training data, the second learning control unit 15 training the second neural network 200 by supervised learning in which the supervisor data is the first output data that is output from the first neural network 100 in the case where the first training data is input to the first neural network 100, based on the first training data, so as to output the second output data, and the second learning result output unit 17 outputting the trained second neural network 200 or the copy of the trained second

neural network 200.

[0095] Fig. 9 is a flowchart of processing executed by the learning result using apparatus 20 according to the embodiment of the present invention. The learning result using apparatus 20 acquires the trained second neural network 200 or a copy of the trained second neural network 200, using the learning apparatus 10, and sets the third neural network 233 (step S20). The learning result using apparatus 20 then designates input data that is to be input to the third neural network 233 based on an instruction accepted from the user (step S21). Here, the input data is data having the same form as the first training data.

[0096] The learning result using apparatus 20 inputs the designated input data to the third neural network 233, and outputs output data corresponding to the feature of the input data (step S22). The processing performed by the learning result using apparatus 20 then ends.

[0097] The foregoing embodiment is for the purpose of facilitating understanding of the present invention, and is not to be interpreted as limiting the present invention. Constituent elements of the embodiment and arrangement, materials, conditions, shapes and sizes thereof are not limited to those exemplified, and can be changed as appropriate. In addition, configurations described in different embodiments can be partially substituted or combined.

[0098] In addition, a portion or the entirety of the above-described embodiment can be described as Additional Remark below, but is not limited thereto.

[0099] Additional Remark 1

A learning apparatus including at least one memory and at least one item of hardware processor connected to the memory,

wherein the hardware processor trains a first learning module based on first training data and second training data associated with the first training data, so as to output first output data corresponding to the features of the first training data and the second training data, and

the hardware processor trains a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in the case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[0100] Additional Remark 2

A learning method:

wherein at least one item of hardware processor trains a first learning module based on first training data and second training data associated with the first training data, so as to output first output data corresponding to the features of the first training data and the second training data, and

the hardware processor trains a second learning module by supervised learning in which the supervisor data is the first output data that is output from the first learning module in the case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

Claims

[Claim 1] A learning apparatus comprising:
a first learning control unit configured to train a first learning module based on first training data and second training data associated with the first training data so as to output first output data corresponding to features of the first training data and the second training data; and
a second learning control unit configured to train a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[Claim 2] The learning apparatus according to claim 1,
wherein the second learning control unit trains the second learning module after training of the first learning module.

[Claim 3] The learning apparatus according to claim 1 or 2,
wherein the first training data is data in the same form as input data that is input to a trained second learning module, which is acquired as a result of training performed by the second learning control unit, or a copy of the trained second learning module, and
the second training data is data temporally related to the first training data, and is data in a form different from input data that is input to the trained second learning module or the copy of the trained second learning module.

[Claim 4] The learning apparatus according to any one of claims 1 to 3,
wherein the first learning control unit trains the first learning module by unsupervised learning based on the first training data and the second training data so as to output the first output data.

[Claim 5] The learning apparatus according to any one of claims 1 to 3,
wherein the first learning control unit trains the first learning module by supervised learning that uses supervisor data including attribute information of the first training data and the second training data, based on the first training data and the second training data, so as to output the first output data.

[Claim 6] The learning apparatus according to any one of claims 1 to 5,
wherein the first learning module and the second learning module each include a neural network, and
a scale of the neural network included in the second learning module is

smaller than a scale of the neural network included in the first learning module.

[Claim 7] The learning apparatus according to any one of claims 1 to 6, wherein the first training data includes image data of a target, the second training data includes sensing data acquired by measuring the target using a sensor when the image data is shot, and the first output data and the second output data include data related to the target.

[Claim 8] The learning apparatus according to claim 7, wherein the first training data includes image data acquired by shooting a person, the second training data includes vital data of the person when the image data is shot, and the first output data and the second output data are data corresponding to a human emotion.

[Claim 9] The learning apparatus according to claim 7, wherein the first training data includes image data acquired by shooting a vehicle, the second training data includes sensing data acquired by performing measurement using a sensor provided in the vehicle when the image data is shot, and the first output data and the second output data are data corresponding to an operation of the vehicle.

[Claim 10] A learning result using apparatus, comprising; a learning module setting unit configured to acquire a trained second learning module acquired as a result of training performed by the second learning control unit provided in the learning apparatus according to any one of claims 1 to 9 or a copy of the trained second learning module, and sets the trained second learning module or the copy of the trained second learning module as a third learning module; an input unit configured to input data having the same form as the first training data to the third learning module; and an output unit configured to output output data from the third learning module.

[Claim 11] A learning method comprising: training, by a control unit configured to control machine learning, a first learning module based on first training data and second training data associated with the first training data so as to output first output

data corresponding to features of the first training data and the second training data; and

training, by the control unit, a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

[Claim 12] A method for producing a trained learning module or a copy of the trained learning module, comprising:

outputting a trained second learning module acquired as a result of training the second learning module by the learning method according to claim 11 or a copy of the trained second learning module.

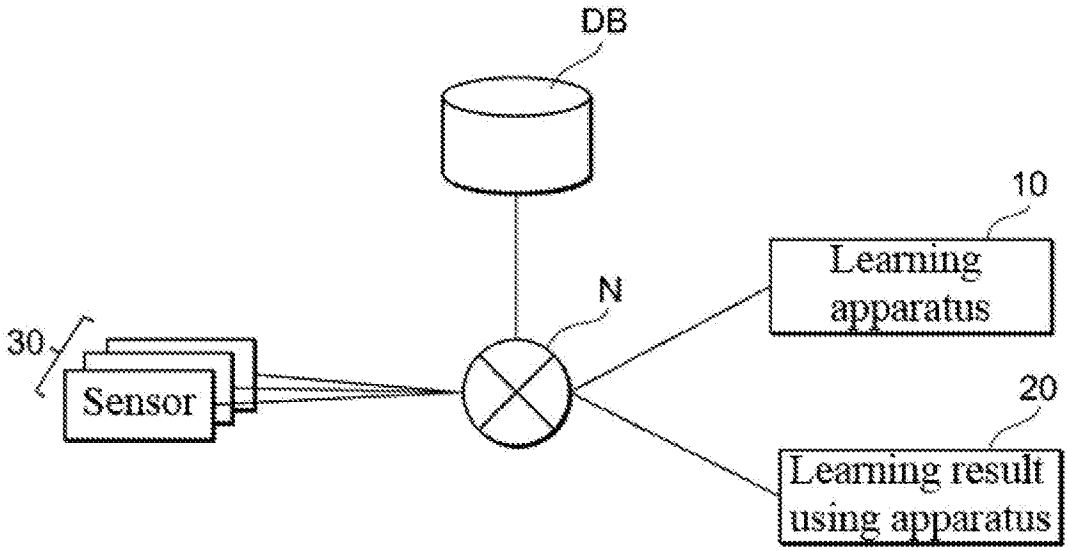
[Claim 13] A trained learning module or a copy of the trained learning module that is acquired as a result of training the second learning module by the learning method according to claim 11.

[Claim 14] A learning program comprising instructions which, when the program is executed by a computer, cause the computer to perform a method including:

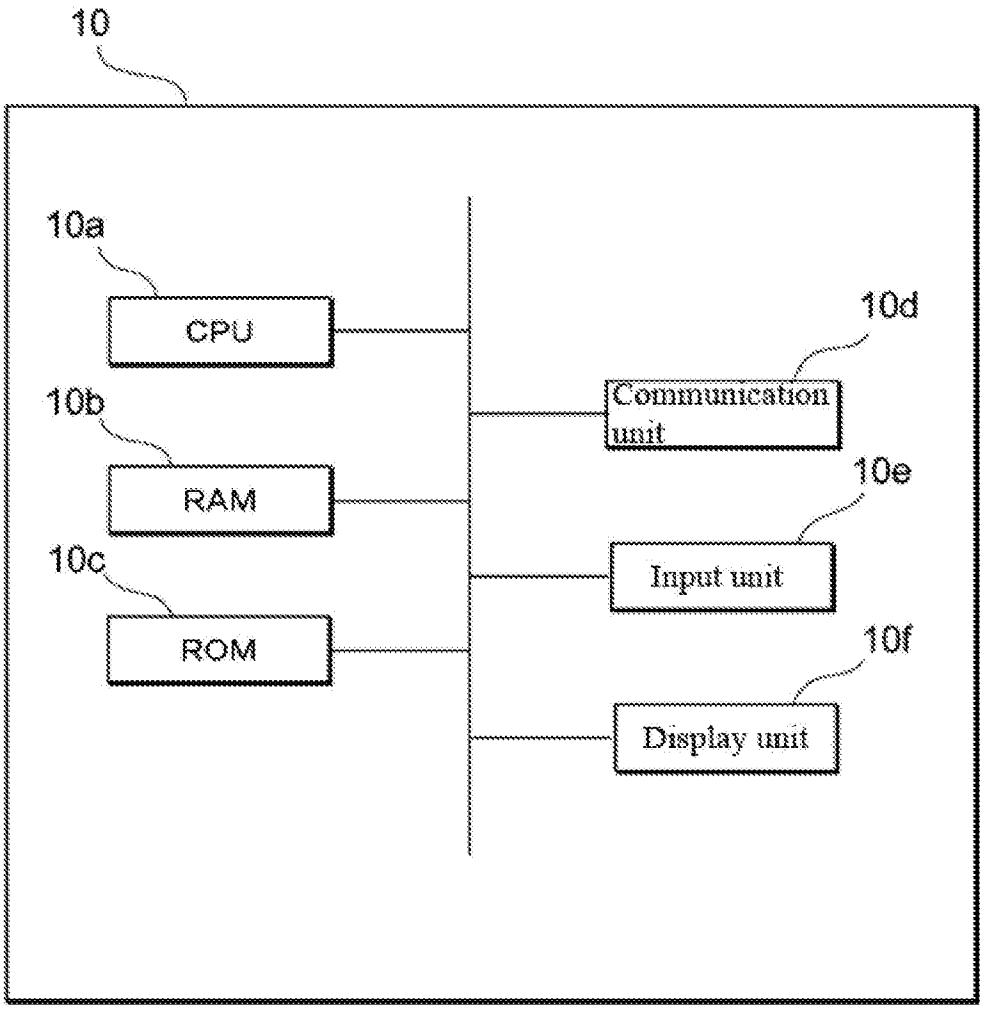
training a first learning module based on first training data and second training data associated with the first training data so as to output first output data corresponding to features of the first training data and the second training data; and

training a second learning module by supervised learning in which supervisor data is the first output data that is output from the first learning module in a case where the first training data is input to the first learning module, based on the first training data, so as to output second output data.

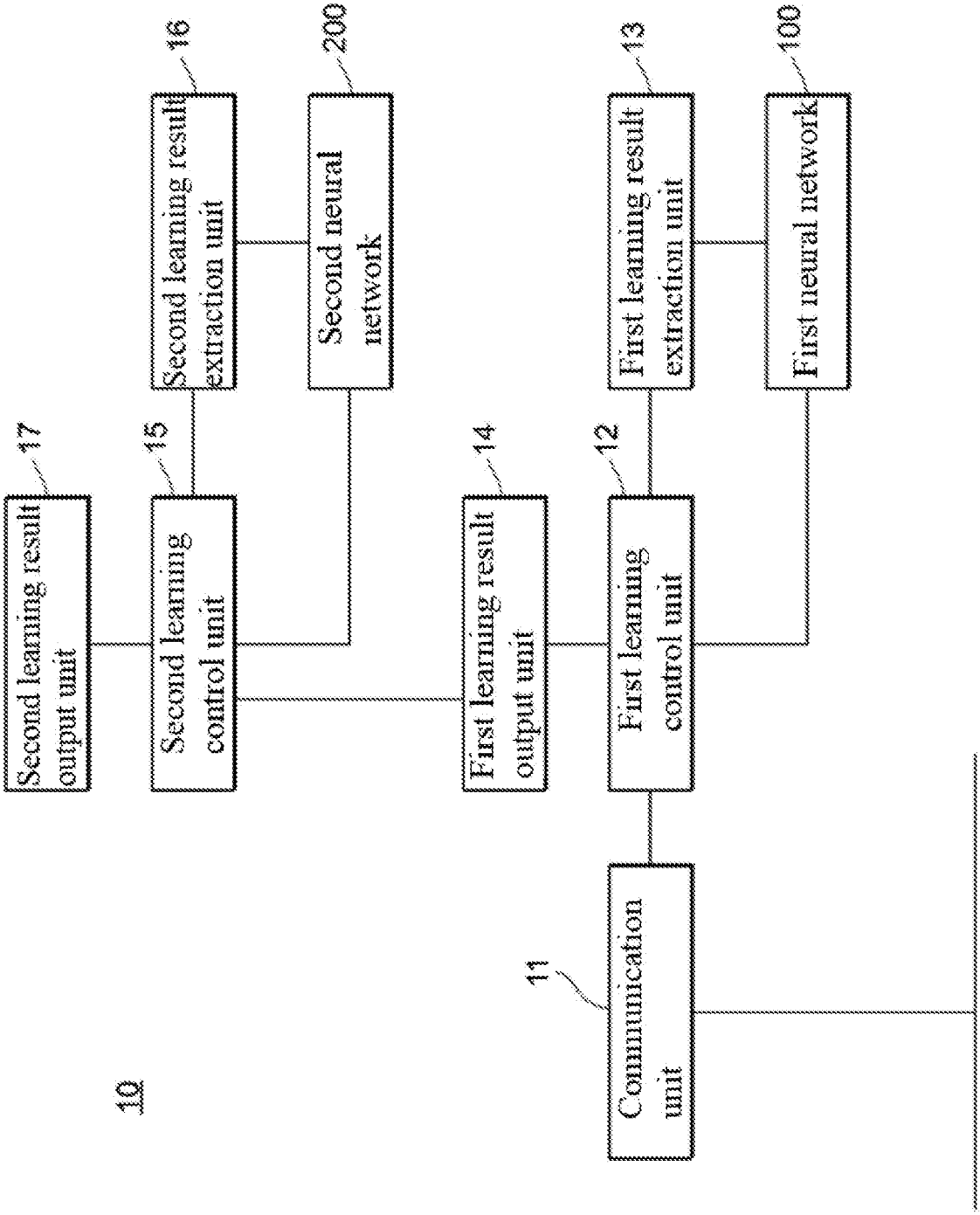
[Fig. 1]


DB
10
Learning apparatus
N
30
Sensor
20
Learning result using apparatus

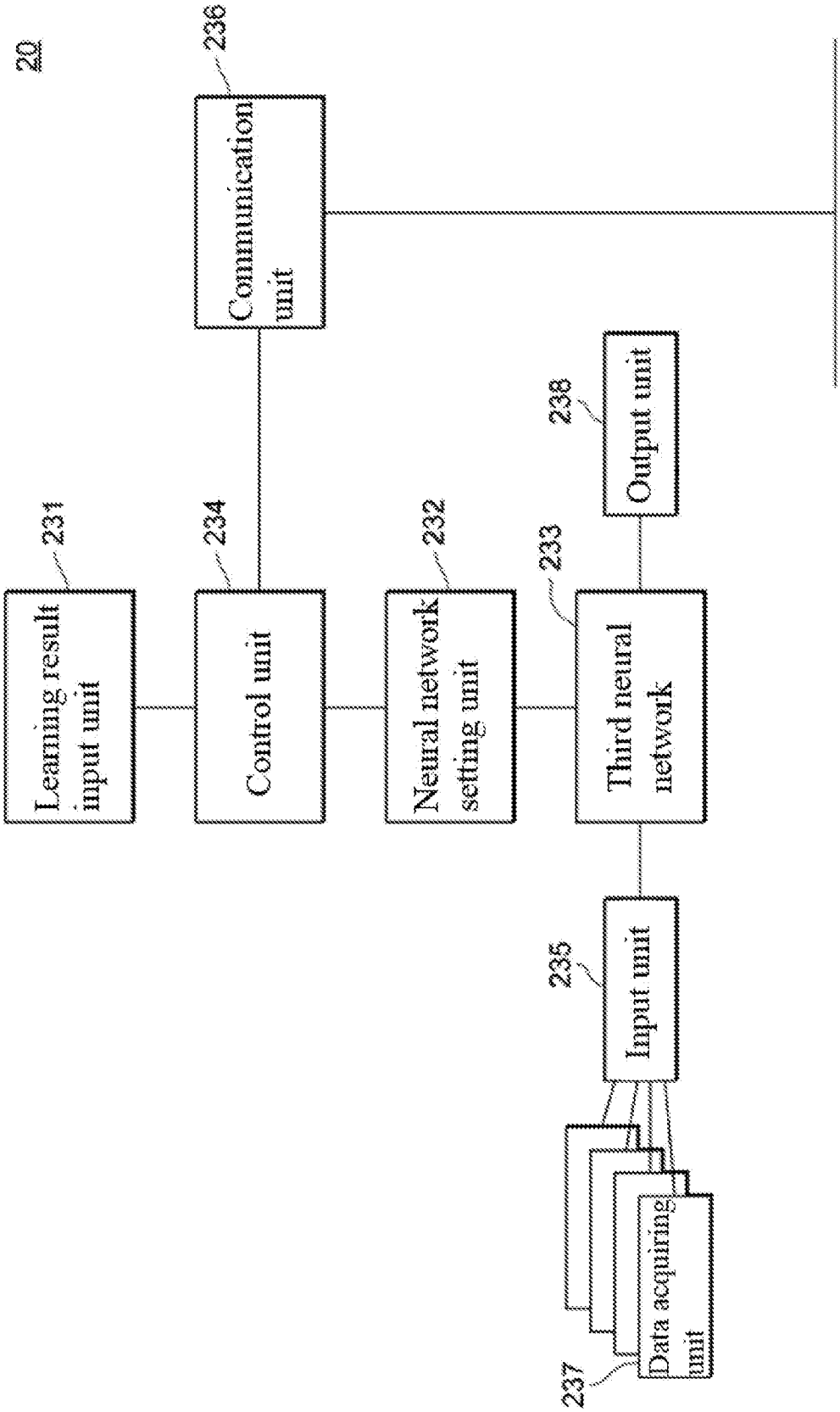
[Fig. 2]


10
10a
CPU
10d
Communication unit
10b
RAM
10e
Input unit
10c
ROM
10f
Display unit

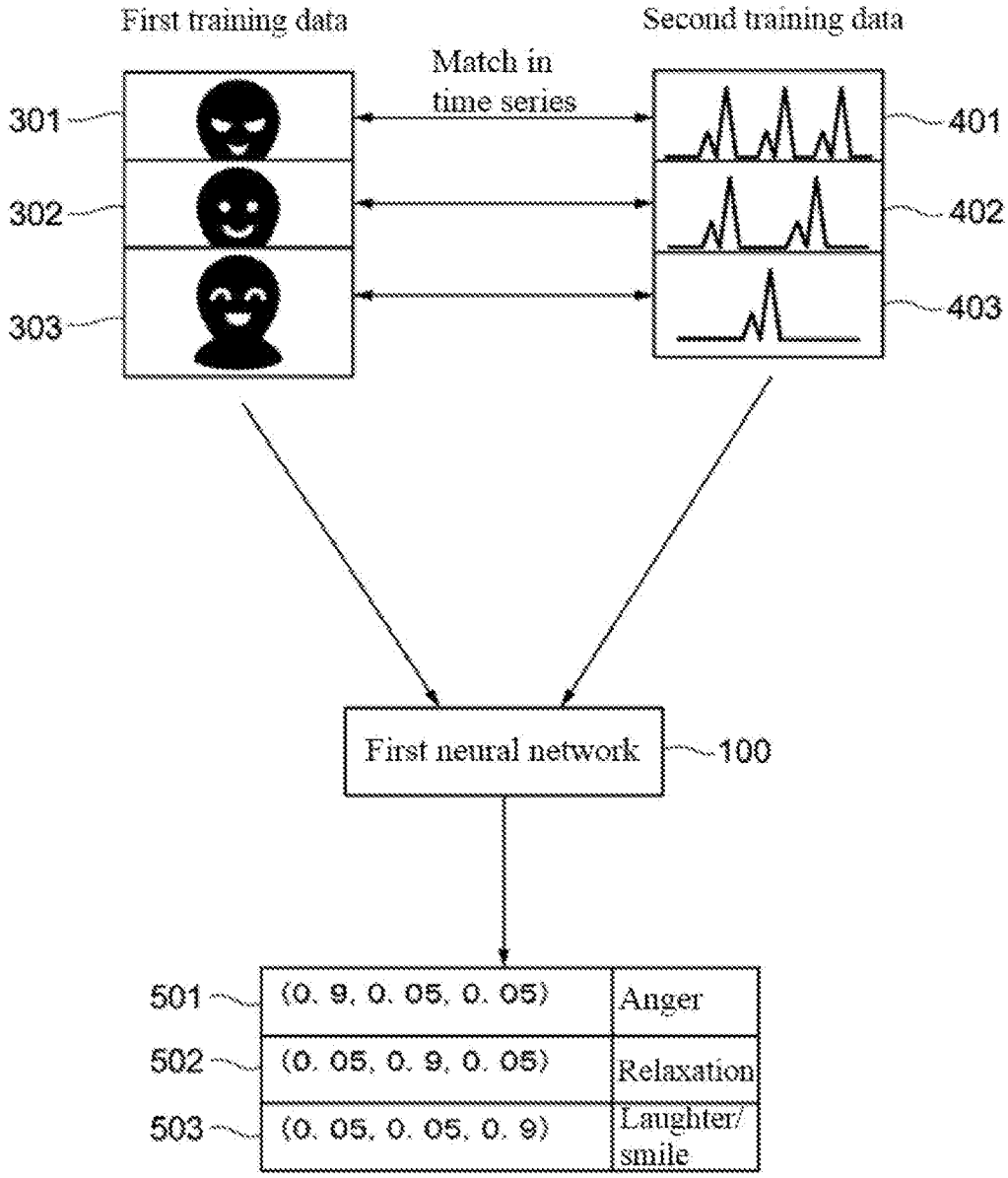
[Fig. 3]


10
11
Communication unit
12
First learning control unit
13
First learning result extraction unit
100
First neural network
14
First learning result output unit
15
Second learning control unit
16
Second learning result extraction unit
200
Second neural network
17
Second learning result output unit

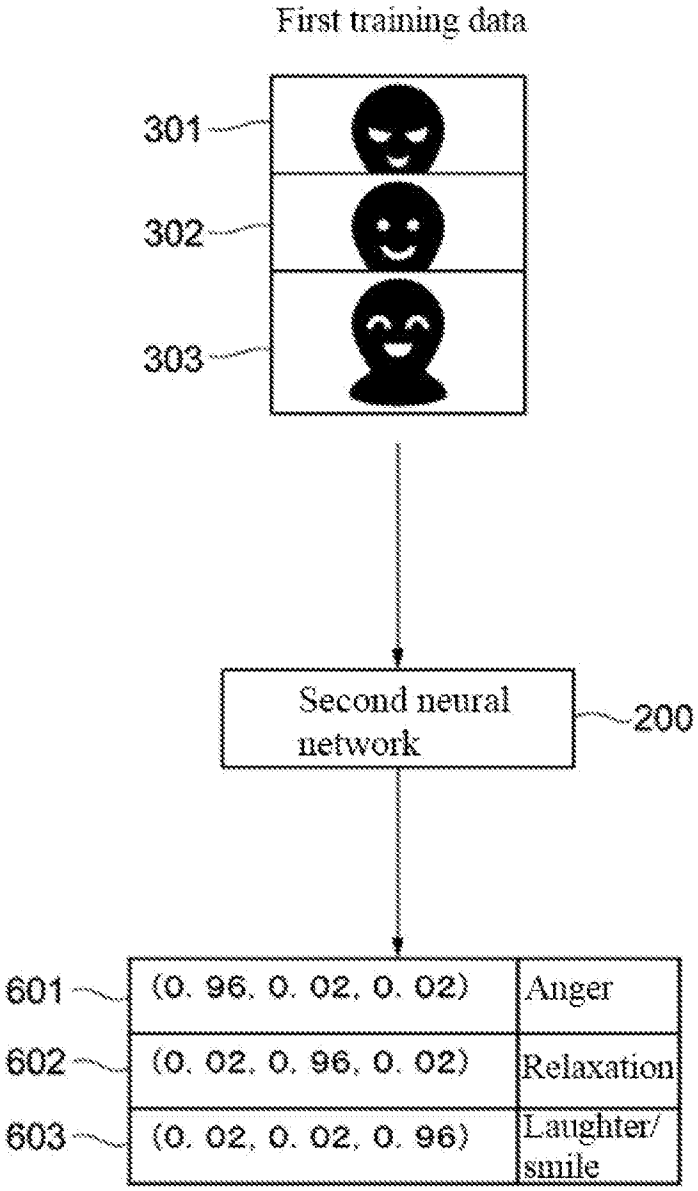
[Fig. 4]


20
236
Communication unit
231
Learning result input unit
234
Control unit
232
Neural network setting unit
233
Third neural network
238
Output unit
235
Input unit
237
Data acquiring unit

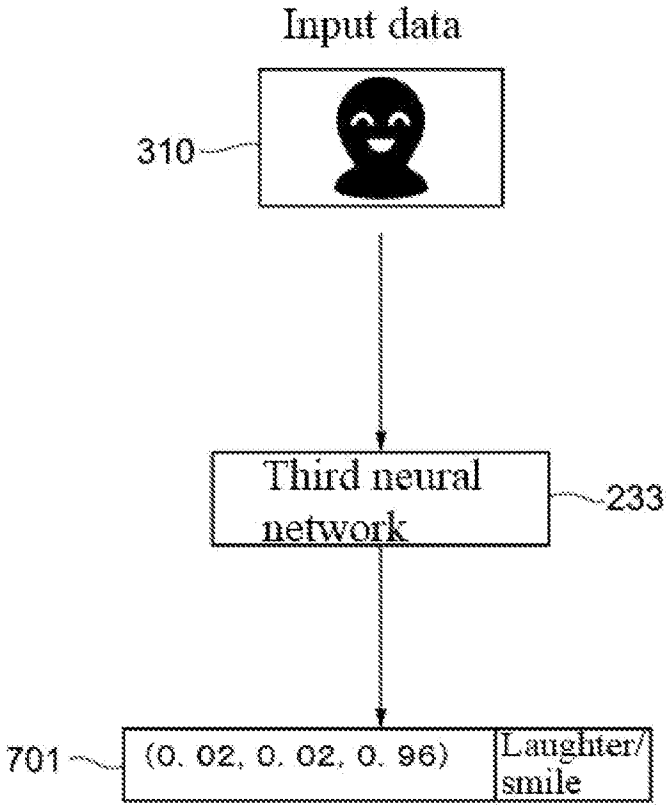
[Fig. 5]


First training data
Second training data
Match in time series
301
302
303
401
402
403
First neural network
100
501
502
503
(0. 9, 0. 05, 0. 05)
(0. 05, 0. 9, 0. 05)
(0. 05, 0. 05, 0. 9)
Anger
Relaxation
Laughter/ smile

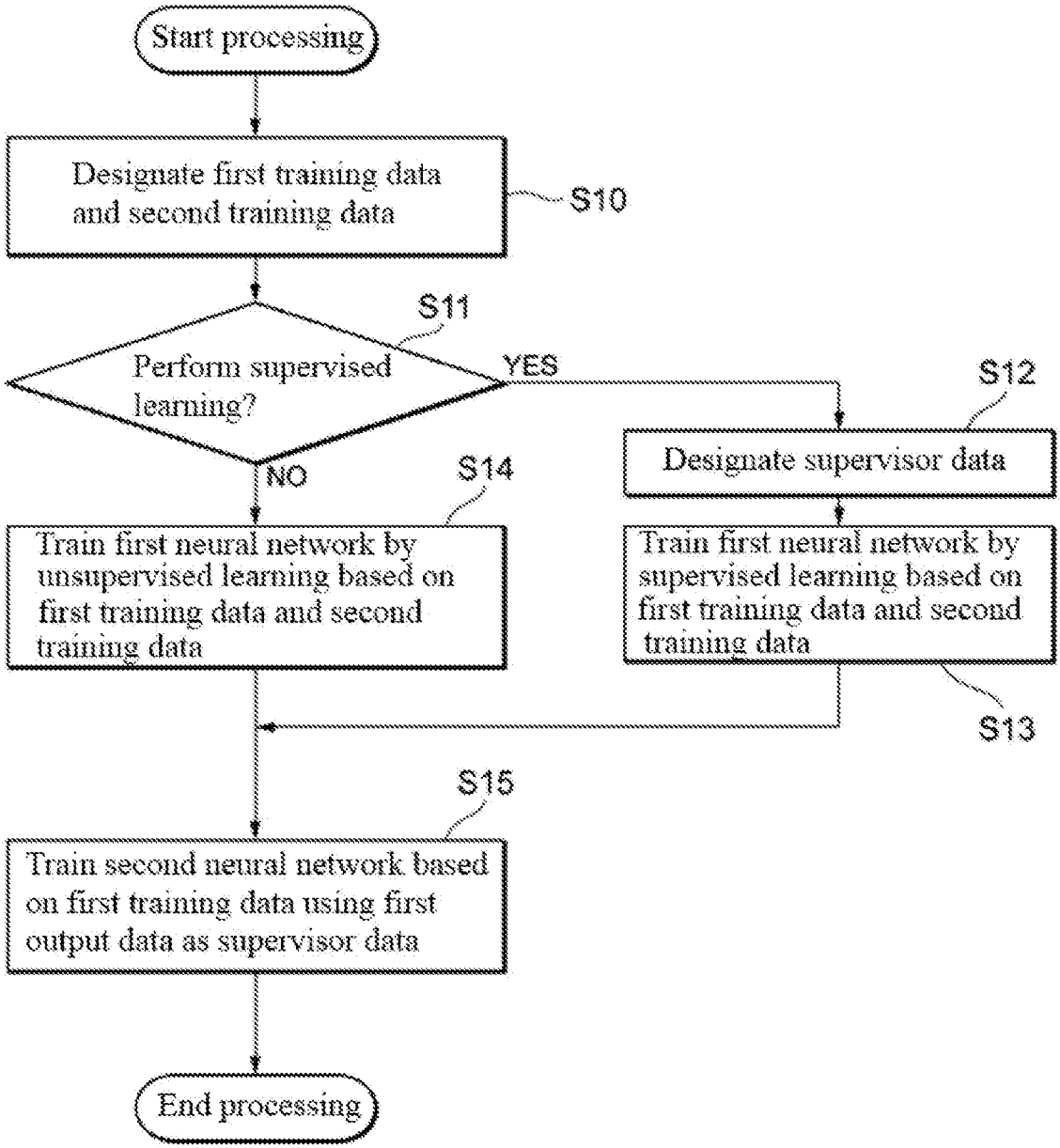
[Fig. 6]


First training data
301
302
303
Second neural network
200
601
(0. 96. 0. 02. 0. 02)
Anger
602
(0. 02. 0. 96. 0. 02)
Relaxation
603
(0. 02. 0. 02. 0. 96)
Laughter/ smile

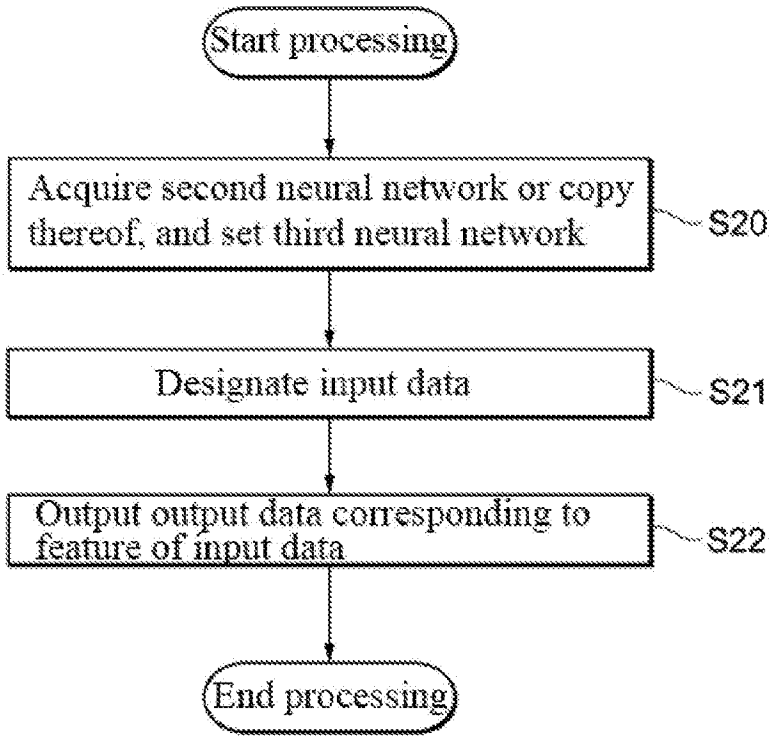
[Fig. 7]


Input data
310
Third neural network
233
701
(0. 02, 0. 02, 0. 96)
Laughter/ smile

[Fig. 8]


Start processing
Designate first training data and second training data
S10
S11
Perform supervised learning?
YES
NO
S12
Designate supervisor data
S14
Train first neural network by unsupervised learning based on first training data and second training data
Train first neural network by supervised learning based on first training data and second training data
S13
S15
Train second neural network based on first training data using first output data as supervisor data
End processing

[Fig. 9]


Start processing
Acquire second neural network or copy thereof, and set third neural network
S20
Designate input data
S21
Output output data corresponding to feature of input data
S22
End processing

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2018/007476

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

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06N

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	EP 1 939 796 A2 (FUJI XEROX CO LTD [JP]) 2 July 2008 (2008-07-02) abstract; claims 1-7; figures 2-8 paragraph [0003] - paragraph [0055] -----	1-6, 10-14 7-9
A	US 2017/024641 A1 (WIERZYNSKI CASIMIR MATTHEW [US]) 26 January 2017 (2017-01-26) abstract; claims 1-18; figures 6,8,9 paragraph [0010] - paragraph [0089] -----	1-14
A	EP 3 065 090 A2 (PANASONIC IP MAN CO LTD [JP]) 7 September 2016 (2016-09-07) abstract; claims 1-11; figures 1-16 paragraph [0003] - paragraph [0098] -----	1-14
	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 2018	11/06/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Volkmer, Markus

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HONG-WEI NG ET AL: "Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning", INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION, ACM, 2 PENN PLAZA, SUITE 701 NEW YORK NY 10121-0701 USA, 9 November 2015 (2015-11-09), pages 443-449, XP058076144, DOI: 10.1145/2818346.2830593 ISBN: 978-1-4503-3912-4 page 443 - page 447, right-hand column, paragraph 3 -----	1-14
A	CLIFFORD NASS ET AL: "Improving automotive safety by pairing driver emotion and car voice emotion", CONFERENCE PROCEEDINGS OF CHI 2005, APRIL 2-7, 2005, PORTLAND, OREGON, USA, ACM, 2 PENN PLAZA, SUITE 701 NEW YORK NY 10121-0701 USA, 2 April 2005 (2005-04-02), pages 1973-1976, XP058168602, DOI: 10.1145/1056808.1057070 ISBN: 978-1-59593-002-6 page 1973 - page 1976, right-hand column, paragraph 2 -----	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/JP2018/007476

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1939796 A2	02-07-2008	CN 101206727 A	25-06-2008
		EP 1939796 A2	02-07-2008
		JP 2008152619 A	03-07-2008
		US 2008147576 A1	19-06-2008
US 2017024641 A1	26-01-2017	CN 107851213 A	27-03-2018
		EP 3326116 A2	30-05-2018
		KR 20180034395 A	04-04-2018
		US 2017024641 A1	26-01-2017
		WO 2017052709 A2	30-03-2017
EP 3065090 A2	07-09-2016	CN 105938558 A	14-09-2016
		EP 3065090 A2	07-09-2016
		JP 2017084320 A	18-05-2017
		US 2016260014 A1	08-09-2016