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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 February 2024 (22.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/038653 A1

(51) International Patent Classification:
G06Q 50/02 (2012.01) *A01K 61/95* (2017.01)
G01S 15/96 (2006.01)

(21) International Application Number:
PCT/JP2023/019775

(22) International Filing Date:
26 May 2023 (26.05.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2022-129707 16 August 2022 (16.08.2022) JP

(71) Applicant: **FURUNO ELECTRIC CO., LTD.** [JP/JP]; 9-52, Ashihara-Cho, Nishinomiya-City, Hyogo, 6628580 (JP).

(72) Inventors: **KASAI Akinori**; 9-52 Ashihara-cho, Nishinomiya-City, Hyogo, 6628580 (JP). **HIRABAYASHI Yuta**; 9-52 Ashihara-cho, Nishinomiya-City, Hyogo, 6628580 (JP). **MURAGAKI Masashi**; 9-52 Ashihara-cho, Nishinomiya-City, Hyogo, 6628580 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

(54) **Title:** DEVICE, METHOD, AND PROGRAM FOR LEARNING FISH SPECIES

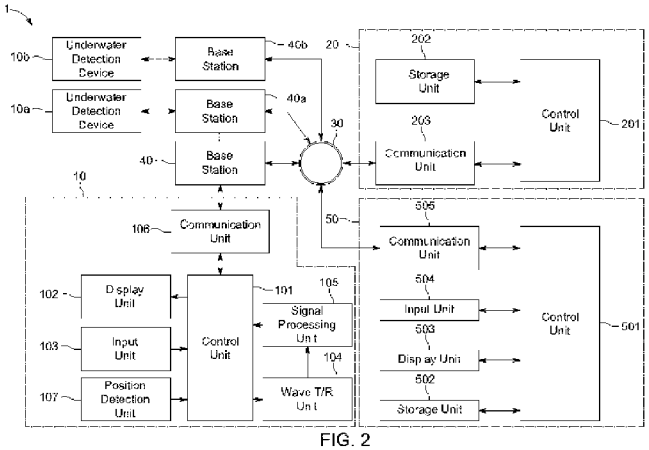


FIG. 2

(57) **Abstract:** To provide a fish species learning device, a fish species learning method and a program that can perform machine learning for fish species identification with better accuracy. To solve this problem, a server 20 (fish species learning device) includes a storage unit (202) that stores annotation data for each unit echo image (Pn, Pn+1), and a control unit (201). The control unit (201) extracts a first annotation data and a second annotation data of a fish school extending over a first unit echo image (Pn) and a second unit echo image (Pn+1) from the annotation data of the first and second unit echo images, respectively, the first and second unit echo images (Pn, Pn+1) being temporally consecutive in time and generates the annotation data of the entire fish school that extends over the first and second unit echo images by integrating the extracted first and second annotation data.

WO 2024/038653 A1

Description

Title of Invention: DEVICE, METHOD, AND PROGRAM FOR LEARNING FISH SPECIES

Technical Field

[0001] The present invention relates to a fish species learning device, a fish species learning system, a fish species learning method and a program for performing machine learning for fish species discrimination using training data (annotation data).

Background

[0002] Fish finders have been known to detect fish schools in water. Typically, fish finders send ultrasonic waves underwater and receive reflected waves. Echo data is generated according to an intensity of the reflected waves received, and an echo image is displayed based on the echo data generated. A user can confirm the fish school from the echo image, and a capture of the fish school can proceed smoothly.

[0003] In this case, it is preferable that a fish species of the fish school on the echo image is further discriminated and displayed. This enables the user to efficiently catch fish of the fish species as desired.

[0004] Such fish species discrimination can be performed using, for example, machine learning models (machine learning algorithms). For the machine learning models, learning is performed using a large number of training data (annotation data). Each annotation data includes echo data of a school of fish, range (depth, time) of the school of fish, and the species of fish in the school of fish. A Japanese Patent, Publication No. JP2019-200175, describes a configuration when such annotation data is generated by a user.

[0005] PTL 1: The Patent publication is JP2019-200175.

Summary

[0006] Generally, generation of the annotation data is performed for each unit echo image. That is, a series of echo images for a given time is divided into multiple units with a time width corresponding to one frame, and multiple unit echo images targeted for generation of annotation data are generated. For each unit echo image, an operator such as a specialist designates the range of the fish school, and additionally assigns a fish species label to the designated fish school. This generates annotation data in which the range of the fish school, the fish species in the range of the fish school and the image data (echo data) included in the range of the fish school are associated.

[0007] With such generation method, the unit echo image does not necessarily include the fish school in its entirety. That is, one school of fish may extend over two temporally consecutive unit echo images. In such a case, two fish school parts separated from one

another will be included in each of the two echo images. In the above generation method, since processing is performed in unit echo image, the annotation data is generated individually for each separated fish school part.

[0008] However, in the above machine learning models, features arising from the fish school in its entirety, such as tailing and distribution of echoes that occur uniquely in the fish species, can affect an accuracy of the machine learning. Therefore, when the annotation data generated for a part of the fish school is used for machine learning, as described above, the accuracy of the machine learning can be degraded.

[0009] In view of such issues, the present invention aims to provide a fish species learning device, a fish species learning method and a program that can perform machine learning for fish species discrimination with better accuracy.

[0010] A first aspect of the present invention relates to a fish species learning device. The fish species learning device, according to this aspect, is provided with a storage unit configured to store annotation data for each unit echo image and a control unit. The control unit is configured to extract a first annotation data and a second annotation data of a fish school extending over a first unit echo image and a second unit echo image from the annotation data of the first unit echo image and the second unit echo image, respectively. The first unit echo image and the second unit echo image are temporally consecutive in time. Further, the control unit is configured to generate the annotation data of the fish school in its entirety by integrating the extracted first and second annotation data.

[0011] According to the fish species learning device, the first and second annotation data of the fish school extending over the first unit echo image and the second unit echo image are integrated to generate the annotation data for the entire fish school. Therefore, since the annotation data for the entire fish school can be used for machine learning, the machine learning for fish species discrimination can be performed with higher accuracy.

[0012] In the fish species learning device, the control unit may be configured to identify in the first unit echo image a first fish school extending to a vicinity of a first boundary on a side of the second unit echo image and identify in the second unit echo image a second fish school extending to a vicinity of a second boundary on a side of the first unit echo image, and extract the annotation data of the first fish school and the second fish school as the first annotation data and the second annotation data, respectively, when a depth range of the first fish school and a depth range of the second fish school substantially match.

[0013] According to this configuration, the first fish school and the second fish school, which have a high probability of constituting one fish school, are smoothly identified in the first unit echo image and the second unit echo image. Therefore, by integrating

the annotation data of these fish schools, the annotation data of one entire fish school may be generated with high accuracy.

[0014] In this configuration, the control unit may be configured to determine whether the depth range of the first fish school and the depth range of the second fish school substantially match based on a match rate between the depth range of the first fish school near the first boundary and the depth range of the second fish school near the second boundary.

[0015] According to this configuration, it can be determined whether the depth range of the first fish school roughly matches the depth range of the second fish school near the first boundary and the second boundary, which are the boundaries between the first unit echo image and the second unit echo image, so that it can be accurately determined whether these first and second fish schools are temporally consecutive with each other. Therefore, the annotation data of the entire fish school can be generated with high accuracy from the annotation data of the first and second fish schools.

[0016] Also, in this configuration, the control unit may be configured to extract, as the first fish school, a fish school at least partially included in a given time range from the first boundary, and extract, as the second fish school, a fish school at least partially included in a given time range from the second boundary.

[0017] According to this configuration, the first fish school and the second fish school, which constitute the fish school extending over the first unit echo image and the second unit echo image, are identified in the first unit echo image and the second unit echo image without omission even when the distribution of fish is sparse near the boundaries. Therefore, the annotation data of one entire fish school may be properly generated.

[0018] In the fish species learning device, according to this aspect, when a fish species included in the first annotation data and a fish species included in the second annotation data are different, the control unit may be configured to set a fish species of the annotation data of the fish school based on a given setting condition.

[0019] In this case, the given setting condition may include, for example, setting the fish species of the annotation data to the newer of the first annotation data and the second annotation data.

[0020] According to these configurations, when the fish species included in the first and second annotation data to be integrated are different from each other, a highly reliable fish species may be set in the annotation data after integration. Therefore, the accuracy of machine learning may be enhanced.

[0021] A second aspect of the invention relates to a fish species learning method. The fish species learning method, according to this aspect, comprises extracting a first annotation data and a second annotation data of a fish school extending over a first unit

echo image and a second unit echo image from annotation data of the first unit echo image and the second unit echo image, respectively, the first unit echo image and the second unit echo image being temporally consecutive in time. Further, the method comprises generating the annotation data of the fish school in its entirety by integrating the extracted first annotation data and second annotation data.

[0022] A third aspect of the invention relates to a program that makes a computer perform given functions. The program, according to this aspect, includes a function for extracting a first annotation data and a second annotation data of a fish school extending over a first unit echo image and a second unit echo image from annotation data of the first unit echo image and the second unit echo image, respectively, the first unit echo image and the second unit echo image being temporally consecutive in time and a function for generating the annotation data of the fish school in its entirety by integrating the extracted first annotation data and second annotation data.

[0023] According to the second and third aspects, the same effect as the first aspect is achieved.

[0024] As described above, according to the present invention, a fish species learning device, a fish species learning method, and a program that can perform machine learning for fish species discrimination with higher accuracy can be provided.

[0025] The effect or significance of the present invention will be further clarified by the following description of the embodiment. However, the following embodiment is only one example when implementing the present invention, and the present invention is not in any way limited to those described in the following embodiment.

Brief Description of Drawings

[0026] [Fig.1]Fig. 1 is a diagram showing a configuration of a fish species discrimination system according to an embodiment.

[Fig.2]Fig. 2 is a block diagram showing a configuration of a fish species discrimination system according to an embodiment.

[Fig.3]Fig. 3 is a diagram schematically showing fish species discrimination processing by a neural network according to an embodiment.

[Fig.4]Fig. 4 is a diagram schematically showing a display example of an echo image including fish species discrimination results according to an embodiment.

[Fig.5]Fig. 5 is a diagram schematically showing how annotation data is generated in a terminal device according to an embodiment.

[Fig.6]Fig. 6 is a diagram showing how annotation data is stored in a storage unit of a server according to an embodiment.

[Fig.7]Fig. 7 is a flowchart showing an annotation data integration process executed by a control unit of a server according to an embodiment.

[Fig.8]Fig. 8 is a flowchart showing a processing of extracting annotation data of a fish school extending over two-unit echo images according to an embodiment.

[Fig.9]Fig. 9(a) and Fig. 9(b) are diagrams schematically showing a state of a fish school near a boundary between a first unit echo image and a second unit echo image according to an embodiment, respectively.

[Fig.10]Fig. 10(a) and Fig. 10(b) are diagrams showing an integration method of annotation data of first and second fish schools according to an embodiment, respectively.

[Fig.11]Fig. 11(a) is a diagram schematically showing a state of a fish school near a boundary between a first unit echo image and a second unit echo image according to a first modification, and Fig. 11(b) is a diagram showing an integration method of annotation data of first and second fish schools according to the first modification.

[Fig.12]Fig. 12 is a flowchart showing a fish species setting process executed in an annotation data integration process according to a second modification.

Detailed description

[0027] Embodiments of the present invention will be described below with reference to the drawings.

[0028] In the following embodiment, server (20) corresponds to "fish species learning device" described in the claims. However, the "fish species learning device" according to the present invention is not necessarily limited to the server (20), and for example, other devices such as terminal devices (50) may perform the functions of the "fish species learning device" according to the present invention.

[0029] Fig. 1 is a diagram showing a configuration of a fish species discrimination system 1.

[0030] The fish species discrimination system (1) is equipped with an underwater detection device (10) and the server (20). The underwater detection device (10) is a fish finder installed in a vessel (2). The underwater detection device (10) can communicate with the server (20) via an external communication network (30) (for example, the Internet) and a base station (40). The underwater detection device (10) and the server (20) each hold address information for communicating with each other. The respective address information is set in the underwater detection device (10) and the server (20) at the initial setting.

[0031] The underwater detection device (10) is equipped with a transmitter/receiver (11) and a control unit (12). The transmitter/receiver (11) is installed on the bottom of the vessel (2), and the control unit (12) is installed in the wheelhouse or the like of the vessel (2). The transmitter/receiver (11) and the control unit (12) are connected by a signal cable (not shown). The transmitter/receiver (11) is equipped with an ultrasonic transducer (not shown) for transmitting and receiving waves. The transmitter/receiver (11) transmits an ultrasonic wave (3) (transmitted wave) toward a seabed (4) and receives

its reflected wave by the ultrasonic transducer in response to control from the control unit (12). The transmitter/receiver (11) transmits a received signal based on the received reflected wave to the control unit (12).

[0032] The control unit (12) processes the received signal to generate an echo data indicating echo intensity at each depth. The control unit (12) arranges the echo intensity at each depth based on the echo data in time series to generate an echo image for one screen. The control unit (12) displays the generated echo image on a display unit. The control unit (12) updates the echo image for each ultrasonic wave transmitted and received. A user can grasp presence and location of a fish school (5) by referring to the echo image.

[0033] Furthermore, the control unit (12) transmits the generated echo data to the server (20) at any time. The server (20) stores the received echo data and generates an echo image similar to that of the control unit (12). The server (20) uses a machine learning model (machine learning algorithm) to calculate a prediction probability (probability of being a fish species) for each fish species included in the echo image.

[0034] The server (20) acquires a discrimination result of the fish species for the fish school based on the prediction probabilities for each fish species calculated by the machine learning model. The server (20) transmits the discrimination result of the fish species thus acquired along with a range (depth, time) of the fish school to be discriminated to the underwater detection device (10) that received the echo data.

[0035] Based on the received discrimination result and the range (depth, time) of the fish school, the underwater detection device (10) superimposes the discrimination result of the fish species on the corresponding range on the echo image. Thus, the user can confirm the fish species of each fish school on the echo image and smoothly go ahead with capture of the desired fish.

[0036] Furthermore, the server (20) acquires annotation data (training data) for training the machine learning model from a plurality of terminal devices (50) via an external communication network (30). That is, the server (20) distributes the echo data received from the underwater detection device (10) to any of the plurality of terminal devices (50). The terminal devices (50) are used to generate annotation data from the received echo data.

[0037] That is, the terminal devices (50) are owned by an operator, such as an expert, who generates annotation data. The operator causes the terminal devices (50) to display an echo image based on the received echo data, and sets the fish species for each fish school included in these echo images. The set fish species is associated with the range (depth, time) of the fish school and the echo data within the range. The associated fish species, the range of the fish school and the echo data constitute the annotation data for the fish school. The annotation data is transmitted to the server (20) together with an

identification information of the echo image on which the annotation data is generated.

[0038] The server (20) uses the received annotation data to teach the machine learning model. This improves accuracy of the machine learning model for fish species discrimination.

[0039] Here, annotation data are generated for the echo data provided by the server (20) to the terminal devices (50). The generated annotation data are returned to the server (20). It can be noted that the method of providing the annotation data to the server (20) is not limited to this. For example, the echo data may be provided to the terminal devices (50) from a device other than the server (20) to generate the annotation data, and the generated annotation data may be provided to the server (20).

[0040] In addition, only one underwater detection device 10 is shown in Fig. 1, but in fact, a plurality of underwater detection devices may communicate with the server (20) via the external communication network (30) and the nearest base station. In addition, the underwater detection devices (10) that communicate with the server (20) may include several types of underwater detection devices with different fishing methods, such as underwater detection devices installed in fixed nets, in addition to those installed on the vessel (2), as shown in Fig. 1.

[0041] Fig. 2 is a block diagram showing the configuration of the fish species discrimination system 1.

[0042] The underwater detection device (10) includes a control unit (101), a display unit (102), an input unit (103), a wave transmit/receive unit (104), a signal processing unit (105), a communication unit (106), and a position detection unit (107).

[0043] The control unit (101) is composed of a microcomputer, a memory, etc. The control unit (101) controls each part of the underwater detection device (10) according to a program stored in the memory.

[0044] The display unit (102) is equipped with a monitor and displays an image by control from the control unit (101). The input unit (103) is equipped with a trackball for moving a cursor on the image displayed on the display unit (102), an operation key, etc., and outputs a signal corresponding to an operation from the user to the control unit (101). The display unit (102) and the input unit (103) may be integrally constituted by a liquid crystal touch panel, etc.

[0045] The wave transmit/receive unit (104) includes the transmitter/receiver (11), as shown in Fig. 1, a transmission circuit (not shown) for supplying a transmission signal to the transmitter/receiver (11), and a reception circuit (not shown) for processing the received signal output from the transmitter/receiver (11) and outputting it to the signal processing unit (105). The transmission and reception circuits are included in the control unit (12), as shown in Fig. 1.

[0046] The wave transmit/receive unit (104) transmits a transmission wave (ultrasonic

wave) of a given frequency according to the control from the control unit (101). The wave transmit/receive unit (104) receives the reflected wave of the transmission wave and outputs the received signal. The reception circuit extracts the received signal of the frequency of the transmission wave and outputs it to the signal processing unit (105).

[0047] The signal processing unit (105) generates echo data indicating the intensity of the reflected wave according to the depth from the received signal input from the wave transmit/receive unit (104), and outputs the generated echo data to the control unit (101). An elapsed time from a timing of transmitting the transmission wave corresponds to the depth. Here, the intensity of the reflected wave decreases as the depth increases. Therefore, the signal processing unit (105) corrects the intensity of the reflected wave that decreases according to the elapsed time and outputs the corrected intensity echo data to the control unit (101) so that echo data can be quantitatively handled regardless of depth difference.

[0048] The control unit (101) generates the echo image based on the received echo data and causes the display unit (102) to display it. From the echo data, the control unit (101) generates one row of image in the depth direction in which the echo intensity at each depth is expressed in gradation by a color scale. The control unit (101) integrates the images of each row from the present time to a given time in the past in the time direction to generate the echo image for one screen.

[0049] The communication unit (106) is a communication module capable of wireless communication with the base station (40). The position detection unit (107) is equipped with GPS and detects a position of the underwater detection device (10). The position detection unit (107) outputs the detected position information to the control unit (101).

[0050] As described with reference to Fig. 1, the control unit (101) transmits echo data to the server (20) via the communication unit (106) at any time. The control unit (101) also receives the fish species discrimination result from the server (20) via the communication unit (106). The control unit (101) further transmits the position information detected by the position detection unit (107) to the server (20).

[0051] As shown in Fig. 2, in addition to the underwater detection device (10), a plurality of underwater detection devices (10a, 10b ...) may communicate with the server (20) via the external communication network (30) and the nearest base stations (40a, 40b ....). As shown above, the underwater detection device (10) that communicates with the server (20) includes the one installed on the vessel (2), as shown in Fig. 1, as well as several types of underwater detection devices with different fishing methods, such as underwater detection devices installed in fixed nets. The basic configuration of the other underwater detection devices is similar to that of the underwater detection device (10), as shown in Fig. 2.

[0052] The server (20) includes a control unit (201), a storage unit (202), and a commu-

nication unit (203). The control unit (201) is composed of a CPU or the like. The storage unit (202) is composed of a ROM, a RAM, a hard disk or the like. The storage unit (202) stores a program for fish species discrimination and a program for machine learning. The control unit (201) controls each unit according to the program stored in the storage unit (202). The communication unit (203) communicates with the underwater detection device (10) via the external communication network (30) and the base station (40) under control from the control unit (201). The server (20) also communicates with the plurality of terminal devices (50) via the external communication network (30).

[0053] The terminal devices (50) are devices such as personal computers and tablet-type computers that can input and output information. The terminal devices (50) are provided with a control unit (501), a storage unit (502), a display unit (503), an input unit (504), and a communication unit (505).

[0054] The control unit (501) is composed of a CPU or the like. The storage unit (502) is composed of a ROM, a RAM, a hard disk or the like. The storage unit (502) stores a program for generating annotation data. The control unit (501) controls each unit according to the program stored in the storage unit (502). The display unit (503) is composed of a liquid crystal monitor or the like, and displays an image under control from the control unit (501). The input unit (504) is provided with input means such as a mouse and a keyboard. The communication unit (505) communicates with the server (20) by control from the control unit (501).

[0055] Fig. 3 is a diagram schematically showing the fish species discrimination processing by a neural network.

[0056] In this embodiment, machine learning using a neural network is applied as the machine learning. For example, a neural network by deep learning combining the neurons in multiple stages is applied. However, the machine learning applied is not limited to this, and other machine learning such as support vector machines and decision trees may be applied.

[0057] The control unit (201) of the server (20) extracts the range (depth, time) of the fish school from the echo data of one screen to be processed. On the echo image, an area where the echo intensity is above a given threshold and where the echo intensity is connected is extracted as the fish school, and a rectangular range composed of the maximum time width and maximum depth width of the fish school is extracted as the range of the fish school. For the method of extracting the fish school, the description in the applicant's earlier application, International Publication No. 2019/003759, is incorporated by reference.

[0058] The control unit (201) applies the echo data of the range of the extracted fish school to inputs (301a) of a machine learning model (machine learning algorithm by neural

network) (301), as shown in Fig. 3.

[0059] Items of fish species such as sardine, horse mackerel and mackerel are assigned to the outputs (301b) of the machine learning model (301). When the echo data of the range of the fish school is applied to the inputs (301a) of the machine learning model (301), probability that the fish species of the fish school is the fish species of each item (prediction probability) is outputted from each item of outputs (301b) of the machine learning model (301). In an example, as shown in Fig. 3, the prediction probability of 85% is outputted from the sardine item, the prediction probability of 70% is outputted from the mackerel item, and the prediction probability of 10% is outputted from the horse mackerel item.

[0060] The prediction probability of each item is checked against an output condition (302). For the output condition (302), for example, a condition is applied in which the fish species of the item whose prediction probability is equal to or greater than a given lower limit and which has the highest rank is outputted as a discrimination result (303). The lower limit is set in order to prevent fish species with low probability from being outputted as the discrimination result. In the example of Fig. 3, sardine with a prediction probability of 85% is outputted as the fish species discrimination result (303).

[0061] The machine learning for the machine learning model (301) is performed by sequentially applying a series of annotation data (training data) to the inputs (301a) and the outputs (301b) of the machine learning model (301). That is, the echo data of the fish school having an annotation data is inputted to inputs (301a) of the machine learning model (301), and the item corresponding to the fish species included in the annotation data is set to 100% in the outputs (301b) of the machine learning model (301), and the other items are set to 0% to perform machine learning.

[0062] In addition to the echo data of the fish school, other information that can be used for fish species discrimination, such as the position where the echo data is obtained and oceanographic data of the position, may be inputted to inputs (301a) of the machine learning model (301). The annotation data may further include the other information.

[0063] Fig. 4 is a diagram schematically showing a display example of an echo image (P1) including fish species discrimination results. For convenience, in Fig. 4, a line in the depth direction is attached only to the part where the echo intensity is high.

[0064] When the control unit (101) of the underwater detection device (10) receives the discrimination result and the range (depth, time) of the fish school from the server (20), a frame-shaped marker (M0) indicating the range of the fish school is displayed in an area on the echo image (P1) corresponding to the depth width and time width corresponding to the received range of the fish school. Furthermore, the control unit (101) further displays a label (L0) indicating the received discrimination result of the fish

school around the marker (M0).

[0065] In the example of Fig. 4, based on the discrimination result and the range (depth, time) of the fish school received from the server (20), the marker M0 is displayed for fish schools F1 to F8. Further, the label (L0) indicating the discrimination result of the fish species is displayed around these markers (M0). The current date and time is displayed near the upper left corner of the echo image (P1).

[0066] In the example of Fig. 4, the marker and label are not displayed for the fish school (F9) because the result of the fish species discrimination was not outputted according to the machine learning model (301) and the output condition (302 of Fig. 3.). This can happen, for example, if the prediction probability of the fish school (F9) by the machine learning model (301) does not satisfy the output condition (302). In such a case, because the discrimination result is not transmitted from the server (20) to the underwater detection device (10) for this fish school, the discrimination result of the fish species is not displayed as in the case of the fish school (F9 in Fig. 4).

[0067] Fig. 5 is a diagram schematically showing a generation method of annotation data performed in the terminal device (50).

[0068] The generation of annotation data is performed in the terminal device (50) by displaying the echo image based on the echo data to be processed on the display unit (503).

[0069] An operator such as a specialist designates the range of the fish school on the displayed echo image. For example, the range of the fish school is specified by a rectangular range. The operator manipulates the input unit (504) to specify the range of the fish school by specifying apexes of the diagonal of the rectangle. Furthermore, the operator manipulates the input unit (504) to set the fish species for each fish school. For example, according to the specification of the range of the fish school, a list of fish species selection candidates is displayed in the display unit (503), and the operator selects the fish species to be set from this list. Thus, the fish species is set in the range of the fish school.

[0070] Here, the generation of annotation data is performed for each unit echo image as shown in Fig. 5. That is, a series of echo images of a given time is divided into multiple units with a time width corresponding to one frame, and a plurality of unit echo images targeted for the generation of annotation data are generated. Of the boundaries (B1) and (B2) of the unit echo images in the time direction, the boundary (B1) on a side with the new time coincides with the boundary (B2) on a side with the old time of the next unit echo image.

[0071] For each unit echo image, the operator, such as a specialist, specifies the range of the fish school and further sets the fish species for the specified fish school. This generates annotation data in which the range of the fish school, the fish species in the range of

the fish school and the image data (echo data) included in the range of the fish school are associated.

[0072] In the example of Fig. 5, the range of five fish schools or the range of rectangular dashed lines is specified by annotation processing (An) on the unit echo image (Pn), and the fish species is set in the range of each fish school. In addition, the range of five fish schools or the range of rectangular dashed lines is specified by annotation processing (An + 1) on the unit echo image (Pn + 1), and the fish species is set in the range of each fish school. Therefore, from the unit echo image (Pn), for each of the five fish schools, annotation data in which the range of the fish school, the fish species in the range of the fish school, and the image data (echo data) included in the range of the fish school are associated is generated. Also, from the unit echo image (Pn + 1), for each of the five fish schools, annotation data in which the range of the fish school, the fish species in the range of the fish school, and the image data (echo data) included in the range of the fish school are associated is generated.

[0073] The annotation data thus generated is transmitted from the terminal device (50) to the server (20) and stored in the storage unit (202) of the server (20).

[0074] Fig. 6 shows a storage form of the annotation data in the storage unit (202) of the server (20).

[0075] Further, unit echo data and annotation data are stored in the storage unit (202) in association with an echo ID and an image ID.

[0076] The echo ID is identification information for identifying the above-described series of echo images (echo data) of the given time. The echo ID is set uniquely by the server (20). The image ID is identification information for identifying the unit echo image. The image ID is, for example, a number assigned to each unit echo image in order from the oldest one. The unit echo data is echo data for one frame corresponding to each unit echo image. The annotation data is annotation data generated from each unit echo image.

[0077] When the control unit (201) of the server (20) newly receives echo data for a given time to be subjected to annotation processing from the underwater detection device (10) or the like, it attaches the echo ID to the echo data. Furthermore, the control unit (201) divides the echo data by a time width of one frame from the beginning of the side with the old time to generate unit echo data corresponding to each unit echo image. Then, the control unit (201) attaches the image ID to each unit echo data and stores the unit echo data in the storage unit (202) in association with the image ID. In this way, the echo ID, the image ID and the unit echo data are configured in the data structure of Fig. 6.

[0078] Then, the control unit (201) sends each unit echo data together with the echo ID and the image ID to the terminal device (50). The control unit (501) of the terminal device

(50) stores these received data in the storage unit (502) for annotation processing by the operator. When the operation of annotation processing is performed by the operator, the control unit (501) causes the display unit (503) to display the unit echo image based on the unit echo data to be processed. The operator performs annotation processing on the unit echo image to be processed by the processing, as shown in Fig. 5. This generates annotation data for the unit echo image.

[0079] Thus, when the annotation processing for all the unit echo data received from the server (20) is finished, the control unit (501) transmits the annotation data generated for each unit echo image to the server (20) along with the echo ID in association with the image ID of each unit echo image. The control unit (201) of the server (20) stores the received annotation data in the storage unit (202) in association with the image ID of Fig. 6. As shown in Fig. 5, the annotation data corresponding to the number of fish schools included in the unit echo image is associated with each image ID. Thus, by storing the annotation data in the storage unit (202), all the data structure shown in Fig. 6 are constructed.

[0080] By the way, the unit echo image shown in Fig. 5 does not necessarily include a fish school in its entirety, and one fish school may extend over two temporally consecutive unit echo images.

[0081] Referring to Fig. 5, for example, the sardine fish school in the upper right of the unit echo image (Pn) is thought to be connected to the sardine fish school in the upper left of the next unit echo image (Pn + 1). However, in the above generation method, since the fish school is specified for each unit echo image, the sardine fish school extending over the unit echo image (Pn) and the unit echo image (Pn + 1) is divided and specified into the unit echo image (Pn) and the unit echo image (Pn + 1). Annotation data would then be generated individually for each separated fish school part.

[0082] However, in the machine learning described above, features arising from the fish school in its entirety, such as tailing and distribution of echoes that occur uniquely in the fish species, can affect the accuracy of the machine learning. Therefore, when the annotation data generated for a part of the fish school is used for machine learning as described above, the accuracy of the machine learning decreases.

[0083] The present embodiment enhances accuracy of the machine learning, from the annotation data of two temporally consecutive unit echo images. The annotation data of the fish school extending over these unit echo images are extracted, respectively, and the extracted two annotation data are integrated to generate the annotation data of the fish school. As discussed in the example of Fig. 5, the annotation data of the sardine fish school at the upper right of the unit echo image (Pn) and the annotation data of the sardine fish school at the upper left of the unit echo image (Pn + 1) are integrated to generate the annotation data for the entire sardine fish school.

[0084] Fig. 7 is a flowchart showing an annotation data integration process executed by the control unit (201) of the server (20).

[0085] At first, the control unit (201) refers to the annotation data of two temporally consecutive unit echo images among the unit echo images (unit echo data) of the same echo ID (S11). Next, the control unit (201) extracts the annotation data of the fish school extending over these two-unit echo images from the referred annotation data (S12). Then, the control unit (201) integrates the extracted annotation data to generate the annotation data of the entire fish school extending over the two-unit echo images, and stores the generated annotation data in the storage unit (202) (S13). The control unit (201) performs the processing of Fig. 7 for all unit echo images (unit echo data) of the same echo ID.

[0086] When training the machine learning model (301 of Fig. 3), the annotation data of Fig. 6 generated by the terminal device (50) and the annotation data generated by the annotation processing of Fig. 7, are used. At that time, the two annotation data integrated by the processing of Fig. 7 may be excluded from the annotation data used for machine learning. For example, the two annotation data integrated by the processing of Fig. 7 may be erased from the data structure of Fig. 6 constructed in the storage unit (202).

[0087] Fig. 8 is a flowchart showing an example of the processing in step (S12) of Fig. 7.

[0088] For convenience, Fig. 8 includes steps (S11) and (S13) of Fig. 7. Steps (S101) to (S105) of Fig. 8 correspond to step (S12) of Fig. 7.

[0089] At first, the control unit (201) refers to the annotation data of two consecutive unit echo images, at step (S11), and grasps the range of the fish school in each annotation data. Next, in the first unit echo image corresponding to the older of the two unit echo images, the control unit (201) determines whether a fish school exists near the right boundary (near the boundary B1 in Fig. 5) (S101). In one case, if no fish school exists near the boundary (B1) on the right side of the first unit echo image (S101: NO), the control unit (201) terminates the processing in Fig. 8. In another case, if a fish school exists near the boundary (B1) on the right side of the first unit echo image (S101: YES), the control unit (201) identifies this fish school as a first fish school to be integrated, at step (S102).

[0090] Further, the control unit (201) determines whether a fish school exists near the boundary on the left side (that is, near the boundary (B2 in Fig. 5)) of the second unit echo image, at step (S103) corresponding to the newer of these two unit echo images. In one case, when there is no fish school near the left boundary (B2) of the second unit echo image (S103: NO), the control unit (201) terminates the processing of Fig. 8. In another case, when there is a fish school near the left boundary (B2) of the second unit echo image (S103: YES), the control unit (201) identifies this fish school as a second fish school to be integrated, at step (S104).

[0091] Next, the control unit (201) determines whether a depth range of the first fish school roughly matches that of the second fish school, at step (S105). More specifically, the control unit (201) determines whether a match rate between the depth range of the first fish school and that of the second fish school is more than a given threshold (Th1) or not. In one case, when the determination at step (S105) is NO, the control unit (201) terminates the processing in Fig. 8 without performing the processing at step (S13).

[0092] In another case, when the determination at step (S105) is YES, the control unit (201) integrates the annotation data of the first and second fish schools, at step (S13) and generates the annotation data of the entire fish school, assuming that the first and second fish schools constitute a fish school extending over the first unit echo image and the second unit echo image (S13). With this, the control unit (201) terminates the processing for the annotation data of these two-unit echo images.

[0093] The control unit (201) performs the processing of Fig. 8 for the annotation data of the next two temporally consecutive unit echo images. Thus, the control unit (201) performs the processing of Fig. 8 for the annotation data of all the unit echo images with the same echo ID.

[0094] Fig. 9(a) and Fig. 9(b) schematically show a state of the fish school near the boundary between the first unit echo image and the second unit echo image.

[0095] Fig. 9(a) shows the upper right portion of the unit echo image (Pn) (first unit echo image) and the upper left portion of the unit echo image (Pn + 1) (second unit echo image), as shown in Fig. 5. (Fa) and (Fb) are sardine fish schools, and (Ra) and (Rb) are the range (time range, depth range) of fish schools included in the annotation data. As described above, the ranges (Ra, Rb) of the fish schools are set to be rectangular.

[0096] As discussed, at step (S101) of Fig. 8, it is determined whether a fish school exists in a given time range (Ta) from the boundary (B1) of the first unit echo image (Pn), and at step (S103), it is determined whether a fish school exists in a given time range (Tb) from the boundary (B2) of the second unit echo image (Pn + 1). The time ranges (Ta) and (Tb) are set to the order of the maximum value of a temporal gap on the echo image that can occur in a fish school. The time ranges (Ta) and (Tb) may be set to the same time width.

[0097] In the example of Fig. 9(a), since the fish school (Fa) (fish school range Ra) exists within the time range (Ta), the determination at step (S101) of Fig. 8 is YES and the fish school (Fa) is identified as the first fish school. In the example of Fig. 9(a), since the fish school (Fb) (fish school range Rb) exists within the time range (Tb), the determination at step (S103) of Fig. 8 is YES and the fish school (Fb) is identified as the second fish school.

[0098] In comparison, in the example of Fig. 9(b), since the fish school (Fa) (fish school range Ra) does not exist within the time range (Ta), the determination at step (S101) of

Fig. 8 is NO and the fish school (Fa) is not identified as the first fish school. In another example of Fig. 9(b), since the fish school (Fb) (fish school range Rb) does not exist within the time range (Tb), the determination in step (S103) of Fig. 8 is NO and the fish school (Fb) is not identified as the second fish school.

[0099] Further, at step (S105) of Fig. 8, the determination is made according to the following equation (1).

[0100] (min?(Amax,Bmax)-max?(Amin,Bmin))/(max?(Amax,Bmax)-min?(Amin,Bmin))?Th1 … (1).

[0101] As shown in Fig. 9(a), (Amax) and (Amin) are the maximum and minimum depths of the first fish school (Fa) (fish school range Ra) in the first unit echo image (Pn), and (Bmax) and (Bmin) are the maximum and minimum depths of the second fish school (Fb) (fish school range Rb) in the second unit echo image (Pn + 1), respectively.

[0102] The denominator on the left side of equation (1) is the depth range from the smaller (shallower) of (Amin) and (Bmin) to the larger (deeper) of (Amax) and (Bmax), and in the example of Fig. 9(a), it is the depth range from (Amin) to (Bmax). Also, the numerator on the left side of equation (1) is the depth range from the larger (deeper) of (Amin) and (Bmin) to the smaller (shallower) of (Amax) and (Bmax), and in the example of Fig. 9(a), it is the depth range from (Bmin) to (Amax). The left side of equation (1) calculates the match rate between the depth range of the first fish school (Fa) and the depth range of the second fish school (Fb). At step (S105) of Fig. 8, whether or not the depth range of the first fish school (Fa) roughly matches the depth range of the second fish school (Fb) is determined depending on whether or not the match rate is above the threshold (Th1). The threshold (Th1) may be a statistically set default value or may be set arbitrarily by the administrator of the server (20).

[0103] Fig. 10(a) and Fig. 10(b) are diagrams showing a method of integrating the annotation data of the first and second fish schools at step (S13) of Fig. 8.

[0104] Further, at step (S13) of Fig. 8, the control unit (201) matches the depth range of the fish school range (Ra) of the first fish school (Fa) and the fish school range (Rb) of the second fish school (Fb). That is, as shown in Fig. 10(a), the control unit (201) modifies the depth range of the fish school ranges (Ra and Rb) so that the smaller of the maximum depths (Amax) and (Bmax) (Amax) matches the larger (Bmax) and the larger of the minimum depths (Amin) and (Bmin) (Bmin) matches the smaller (Amin). Then, as shown in Fig. 10(b), the control unit (201) integrates the annotation data of the modified fish school ranges (Ra) and (Rb) to generate the annotation data of one fish school composed of the fish schools (Fa) and (Fb).

[0105] The generated annotation data consists of the modified fish school range (Rab) integrating the modified fish school ranges (Ra) and (Rb), the echo data contained in the fish school range (Rab), and the fish species (sardine in this case) of the fish school.

Like the fish school ranges (Ra) and (Rb), the fish school range (Rab) is also rectangular, and this rectangular range contains the fish schools (Fa) and (Fb). As described above, the control unit (201) of the server (20) uses the annotation data after integration for training the machine learning model (301).

[0106] Although the fish schools in the upper right part of the first unit echo image (Pn) and the upper left part of the second unit echo image (Pn + 1) are illustrated in Fig. 9(a) to Fig. 10(b), when there are multiple fish schools extending over the first unit echo image and the second unit echo image, the same processing as described above is performed for each fish school to generate multiple sets of annotation data after integration.

[0107] When one fish school extends over three or more-unit echo images, the above processing is performed sequentially for adjacent unit echo images among these unit echo images. As a result, the annotation data of the parts of the fish school divided into the three or more-unit echo images are integrated to generate annotation data for one fish school.

Effect of the embodiment

[0108] According to the embodiment, the following effects can be achieved.

[0109] As shown in Fig. 7 to Fig. 10(b), the annotation data (first and second annotation data) of a fish school extending over the first unit echo image (Pn) and the second unit echo image (Pn + 1) are integrated to generate annotation data for the fish school in its entirety. Therefore, since the annotation data for the entire fish school can be used for machine learning, the accuracy of machine learning for fish species discrimination can be enhanced.

[0110] As shown in Fig. 9(a), the control unit (201) identifies in the first unit echo image (Pn) the fish school (Fa) extending to the vicinity of the boundary (B1) (first boundary) on the side of the second unit echo image (Pn + 1) as the first fish school, identifies in the second unit echo image (Pn + 1) the fish school (Fb) extending to the vicinity of the boundary (B2) (second boundary) on the side of the first unit echo image (Pn) as the second fish school, and when the range Ra of the fish school Fa (first fish school) and the range (Rb) of the fish school (Fb) (second fish school) roughly match, the annotation data of these fish schools (Fa) and (Fb) (first and second fish schools) are extracted as the annotation data (first and second annotation data) to be integrated, respectively. This allows the fish schools (Fa) and (Fb) (first and second fish schools), which have a high probability of constituting one fish school, to be smoothly identified in the first unit echo image (Pn) and the second unit echo image (Pn + 1). Therefore, by integrating the annotation data of these fish schools, the annotation data of one entire fish school can be generated with high accuracy.

[0111] As shown in Fig. 9(a), the control unit (201) extracts the fish school (Fa) at least

partially included in the given time range (Ta) from the boundary (B1) (first boundary), as the fish school to be integrated (first fish school), and extracts the fish school (Fb) at least partially included in the given time range (Tb) from the boundary (B2) (second boundary), as the fish school to be integrated (second fish school). This enables the first and second fish schools, which constitute the fish school extending over the first unit echo image (Pn) and the second unit echo image (Pn + 1), to be identified without omission in the first unit echo image (Pn) and the second unit echo image (Pn + 1) even when the distribution of fish is sparse near the boundaries (B1) and (B2). Therefore, the annotation data of one entire fish school can be properly generated.

First Modification

[0112] The present invention is not limited to the above embodiment, and the embodiment of the present invention can be modified in various ways other than the above configuration.

[0113] For example, in the above embodiment, the range (time range, depth range) of the fish school constituting the annotation data is rectangular, but the range of the fish school constituting the annotation data may be any other shape other than rectangular.

[0114] For example, as shown in Fig. 11(a), the ranges (Ra, Rb) of the fish schools may be shaped along the outer edge of the fish schools (Fa, Fb). In this case, as described above, on the echo image, the area where the echo intensity is above a given threshold and where the echo intensity is connected is extracted as the fish schools (Fa, Fb), and the range along the outer edge of the extracted fish schools (Fa, Fb) is extracted as the fish school ranges (Ra, Rb). When the operator indicates the position of the fish school on the echo image, the control unit (501) of the terminal device (50) superimposes the range of the fish school extracted by the above processing on the echo image for the fish school including the specified position.

[0115] In this case, as in the first embodiment, in step (S101) of Fig. 8, it is determined whether a fish school exists in the time range (Ta) from the boundary (B1), and in step (S103), it is determined whether a fish school exists in the time range (Tb) from the boundary (B2). With this, as shown in Fig. 11(a), fish schools (Fa) and (Fb) are identified as the first fish school and the second fish school, respectively.

[0116] In this case, in step (S105) of Fig. 8, whether or not the depth range of fish school (Fa) (first fish school) roughly matches that of fish school (Fb) (second fish school) is determined based on the match rate between the depth range of fish school (Fa) (first fish school) near boundary (B1) (first boundary) and the depth range of fish school (Fb) (second fish school) near boundary (B2) (second boundary).

[0117] Further, for fish school (Fa), the maximum depth (Amax) and the minimum depth (Amin) are obtained for the portion of the fish school range (Ra) included in the time

range (Ta), and for fish school (Fb), the maximum depth (Bmax) and the minimum depth (Bmin) are obtained for the portion of the fish school range (Rb) included in the time range (Tb).

[0118] The control unit (201) applies the obtained maximum depths Amax and Bmax and minimum depths Amin and Bmin to equation (1) above to determine step S105. Thus, by determining whether the depth range of fish school (Fa) (first fish school) roughly matches that of fish school (Fb) (second fish school) in the vicinity of boundaries (B1) (first boundary) and (B2) (second boundary), it can be accurately determined whether these fish schools (Fa) and (Fb) (first and second fish schools) are consecutive with each other. Therefore, the annotation data of an entire fish school can be generated with high accuracy from the annotation data of fish schools (Fa) and (Fb) (first and second fish schools).

[0119] In this case, at step (S13), of Fig. 8, the range (Rab) of the entire fish school can be set by connecting the positions of the acquired maximum depths (Amax) and (Bmax) from the ranges (Ra) and (Rb) of the fish schools to each other and by connecting the positions of the acquired minimum depths (Amin) and (Bmin) from the ranges (Ra) and (Rb) of the fish schools to each other as in Fig. 11(b). Thus, the range (Rab) of the fish school constituting the annotation data of the entire fish school can be set smoothly and appropriately.

[0120] However, the method of setting the range (Rab) of the fish school is not limited to this. For example, the range (Rab) of the fish school may be set by connecting the positions of the ranges (Ra) and (Rb) of the fish school slightly closer to the boundaries (B1) and (B2) than the positions of the acquired maximum depths (Amax) and (Bmax) from the ranges (Ra) and (Rb) of the fish schools, and by connecting the positions of the ranges (Ra) and (Rb) of the fish school slightly closer to the boundaries (B1) and (B2) than the positions of the acquired minimum depths (Amin) and (Bmin) from the ranges (Ra) and (Rb) of the fish schools. Again, the range (Rab) of the fish school constituting the annotation data of the entire fish school can be set smoothly and appropriately.

Second Modification

[0121] In the above embodiment, it is assumed that the fish species constituting the annotation data of the first fish school and the fish species constituting the annotation data of the second fish school are the same. However, judging fish species based on images of fish schools can be difficult for the operator. For this reason, the fish species in the annotation data set for the first and second fish schools, respectively, identified as constituting one fish school may differ from each other.

[0122] In the second modification, in such a case, the fish species is set for the annotation data after integration based on a given condition.

[0123] Fig. 12 is a flowchart showing the fish species setting process executed in the annotation data integration process in step (S13) of Fig. 8.

[0124] At first, the control unit (201) determines whether the fish species of the annotation data of the first and second fish schools to be integrated match, at step (S201). In one case, if both match (S201: YES), the control unit (201) sets the matching fish species as the fish species of the fish school after integration, at step (S202). On the other hand, if both do not match, (S201: NO), the control unit (201) sets the fish species of the fish school after integration according to a given setting condition, at step (S203).

[0125] The setting condition of step (S203) may include, for example, setting the fish species of the annotation data to the newer of the annotation data of the first and second fish schools (first annotation data, second annotation data).

[0126] For example, in the example of Fig. 11(a), when the fish species constituting the annotation data of the fish school (Fa) (the first fish school) and the fish school (Fb) (the second fish school) are different, the control unit (201) sets the fish species of the annotation data of the fish school Fb acquired from the newer unit echo image (Pn + 1) as the fish species for the range (Rab) of the integrated fish school.

[0127] It should be noted that the setting condition at step (S203) is not limited to the condition (first condition) of setting the fish species to the newer annotation data as described above, and may be another condition as long as a highly reliable fish species can be set for the annotation data after integration.

[0128] For example, the setting condition at step (S203) may include the condition (second condition) of setting the fish species of the annotation data of the integrated fish school to the fish species of the annotation data of the fish school with the larger range among the annotation data (first and second annotation data) of the first and second fish schools to be integrated.

[0129] Alternatively, the setting condition at step (S203) may include the condition (third condition) of setting the fish species of the annotation data of the integrated fish school to the fish species obtained by processing with the fish species discrimination machine learning model (301).

[0130] In this case, for example, if the fish species cannot be set by the second condition (for example, if the range of the first fish school is approximately the same as the range of the second fish school), the first or third condition may be applied to set the fish species. Alternatively, among the fish species acquired by the first to third conditions respectively, the fish species with the highest number of fish species matching each other may be set as the fish species in the annotation data of the integrated fish school. The setting conditions are not limited to the first to third conditions but may further include other conditions.

[0131] In this way, when the fish species constituting the annotation data of the fish school

(Fa) (first fish school) and the fish school (Fb) (second fish school) are different, a highly reliable fish species can be set as the annotation data after the integration by performing the processing at step (S203). Therefore, the accuracy of the machine learning can be enhanced.

[0132] If the determination at step (S201) is NO, the process of integrating the annotation data of the first fish school and the second fish school may be stopped. In this case, the control unit (201) may exclude both the annotation data of the first fish school and the second fish school from the annotation data used for machine learning of the machine learning model (301), or one of these annotation data (for example, the older or the fish school with the smaller range) may be excluded from the annotation data used for machine learning of the machine learning model (301).

Other Modifications

[0133] In the above embodiment and first and second modifications, the processes of Figs. 7 and 8 are performed in the control unit (201) of the server (20), but these processes may be performed in other devices other than the server (20), and the annotation data after the integration may be provided to the server (20). For example, the processes of Figs. 7 and 8 may be performed by the control unit (501) of the terminal device (50) after the generation of the annotation data by the operator, and the annotation data generated by the operator and the annotation data generated by the processes of Figs. 7 and 8 may be transmitted by the control unit (501) to the server (20). In this case, the terminal device (50) corresponds to the fish species learning device described in the claims.

[0134] In the above embodiment and first and second modifications, it is determined whether or not the depth ranges of the first and second fish schools roughly match according to the above equation (1), but this determination method is not limited to this. For example, in Fig. 9(a), if the conditions are satisfied that a difference between a first depth width between the maximum depth (Amax) and the minimum depth (Amin) and a second depth width between the maximum depth (Bmax) and the minimum depth (Bmin) is less than a given threshold and that a difference between a center depth of the first depth width and a center depth of the second depth width is less than a given threshold, it may be determined that the depth ranges of the first and second fish schools roughly match, and if these conditions are not satisfied, it may be determined that the depth ranges of these fish schools do not match. In the case of Fig. 11(a), the same determination method may be used to determine whether the depth ranges of the first and second fish schools roughly match.

[0135] In addition, the method for extracting the annotation data of the fish school extending over the first unit echo image and the second unit echo image is not limited to the method shown at steps (S101) to (S105) in Fig. 8. For example, based on a continuity

of a distribution of the echo data of the first fish school and the echo data of the second fish school, it may be determined whether the first and second fish schools constitute one fish school.

[0136] In the above embodiment and first and second modifications, it is assumed that the frequency of the transmission wave transmitted by the wave transmit/receive unit (104) is one frequency, but a transmission wave with two different frequencies may be transmitted from the wave transmit/receive unit (104). In this case, the echo data based on the received signal of each frequency may be applied to the machine learning model (301) for each fish school to determine the fish species of each fish school. By transmitting and receiving waves at two different frequencies in this way, the fish species determination by the machine learning model (301) can be performed with higher accuracy. For example, the presence or absence of a swim bladder causes a difference in the echo intensity at each frequency. Therefore, by referring to the difference in the echo intensity from a fish school, the fish species in that fish school can be accurately identified.

[0137] In this case, the annotation data generated by the unit echo image may be generated for each fish school at each frequency. In addition, the integration processing of the annotation data shown in Figs. 7 and 8 need only be performed individually for the echo data (unit echo image) of each frequency.

[0138] In the above embodiment, the determination of the fish species using the machine learning model (301) is performed on the server (20) side, but this determination processing may be performed on the underwater detection device (10) side. In this case, the machine learning model (301) updated by the annotation data is transmitted to the underwater detection device (10) at any time and stored in the underwater detection device (10). Based on the echo data acquired by the wave transmit/receive unit (104) and the signal processing unit (105), the control unit (101) of the underwater detection device (10) executes fish species discrimination using the machine learning model (301), as in the control unit (201) of the above server (20), and displays the result of the discrimination on the echo image.

[0139] Also, in the above embodiment, the annotation data is generated in the terminal device (50), but the device generating the annotation data is not limited to this. For example, a user of the underwater detection device (10) may input a fish school and its fish species on the echo image from his/her own catch result, and the echo data of this fish school and the fish species in the fish school may be transmitted as annotation data to the server (20).

[0140] In the above embodiment, the underwater detection device (10) is a fish finder, but the underwater detection device (10) may be a device other than a fish finder such as sonar.

[0141] In addition, the embodiments of the present invention may be modified as appropriate to the extent stated in the claims.

[0142] List of Reference Numerals

20 server (machine learning device)

201 control unit

202 storage unit

301 machine learning model

Pn, Pn + 1 unit echo images (first unit echo image, second unit echo image)

Fa, Fb fish schools (first and second fish schools)

B1, B2 boundaries (first boundary, second boundary)

Ta, Tb time ranges

Claims

[Claim 1] A fish species learning device (20), comprising:
a storage unit (202) configured to store annotation data for each unit echo image (Pn, Pn+1); and
a control unit (201), wherein
the control unit is configured to:
extract a first annotation data and a second annotation data of a fish school (5) extending over a first unit echo image (Pn) and a second unit echo image (Pn+1) from the annotation data of the first unit echo image and the second unit echo image, respectively, the first unit echo image (Pn) and the second unit echo image (Pn+1) being temporally consecutive in time, and
generate the annotation data of the fish school (5) in its entirety by integrating the extracted first annotation data and second annotation data.

[Claim 2] The fish species learning device (20) of claim 1, wherein
the control unit (201) is configured to:
identify in the first unit echo image (Pn) a first fish school (Fa) extending to a vicinity of a first boundary (B1) on a side of the second unit echo image (Pn+1) and identify in the second unit echo image (Pn+1) a second fish school (Fb) extending to a vicinity of a second boundary (B2) on a side of the first unit echo image (Pn), and
extract the annotation data of the first fish school (Fa) and the second fish school (Fb) as the first annotation data and the second annotation data, respectively, when a depth range of the first fish school (Fa) and a depth range of the second fish school (Fb) substantially match.

[Claim 3] The fish species learning device (20) of claim 2, wherein
the control unit (201) is configured to:
determine whether the depth range of the first fish school (Fa) and the depth range of the second fish school (Fb) substantially match based on a match rate between the depth range of the first fish school (Fa) near the first boundary (B1) and the depth range of the second fish school (Fb) near the second boundary (B2).

[Claim 4] The fish species learning device (20) of claim 2, wherein
the control unit (201) is configured to:
extract, as the first fish school (Fa), a fish school at least partially included in a given time range from the first boundary (B1), and
extract, as the second fish school (Fb), a fish school at least partially

included in a given time range from the second boundary (B2).

[Claim 5] The fish species learning device (20) of claim 1, wherein when a fish species included in the first annotation data is different from a fish species included in the second annotation data, the control unit (201) is configured to set a fish species of the annotation data of the fish school based on a given setting condition.

[Claim 6] The fish species learning device (20) of claim 5, wherein the given setting condition includes setting the fish species of the annotation data to the newer of the first annotation data and the second annotation data.

[Claim 7] A fish species learning method, comprising:
extracting a first annotation data and a second annotation data of a fish school extending over a first unit echo image (Pn) and a second unit echo image (Pn+1) from annotation data of the first unit echo image (Pn) and the second unit echo image (Pn+1), respectively, the first unit echo image (Pn) and the second unit echo image (Pn+1) being temporally consecutive in time, and
generating the annotation data of the fish school in its entirety by integrating the extracted first annotation data and second annotation data.

[Claim 8] A program comprising functions which, when executed by a computer, cause the computer to:
extract a first annotation data and a second annotation data of a fish school extending over a first unit echo image (Pn) and a second unit echo image (Pn+1) from annotation data of the first unit echo image (Pn) and the second unit echo image (Pn+1), respectively, the first unit echo image (Pn) and the second unit echo image (Pn+1) being temporally consecutive in time, and
generate the annotation data of the fish school in its entirety by integrating the extracted first annotation data and second annotation data.

[Fig. 1]

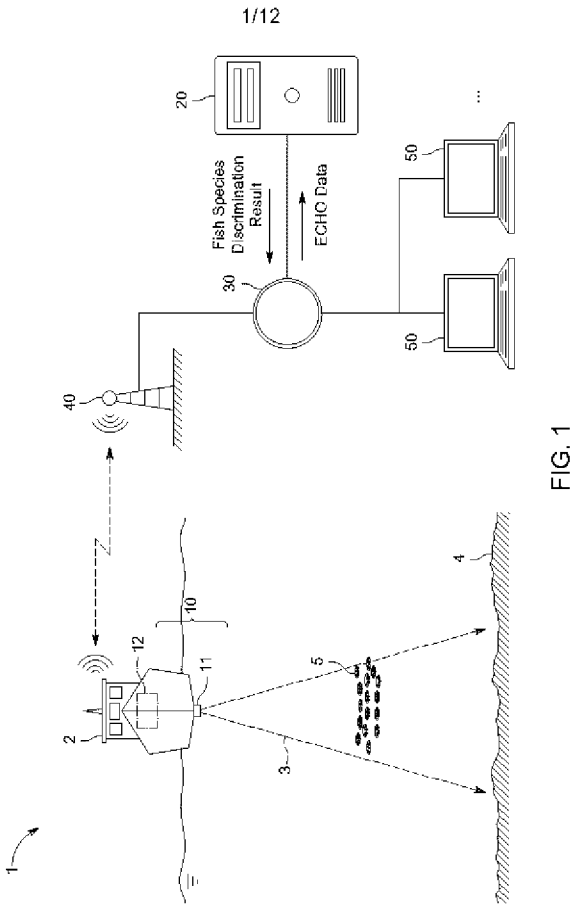

1/12
20
50
50
Fish Species Discrimination Result
ECHO Data
30
40
Z
10
12
11
2
5
3
4
1

FIG. 1

[Fig. 2]

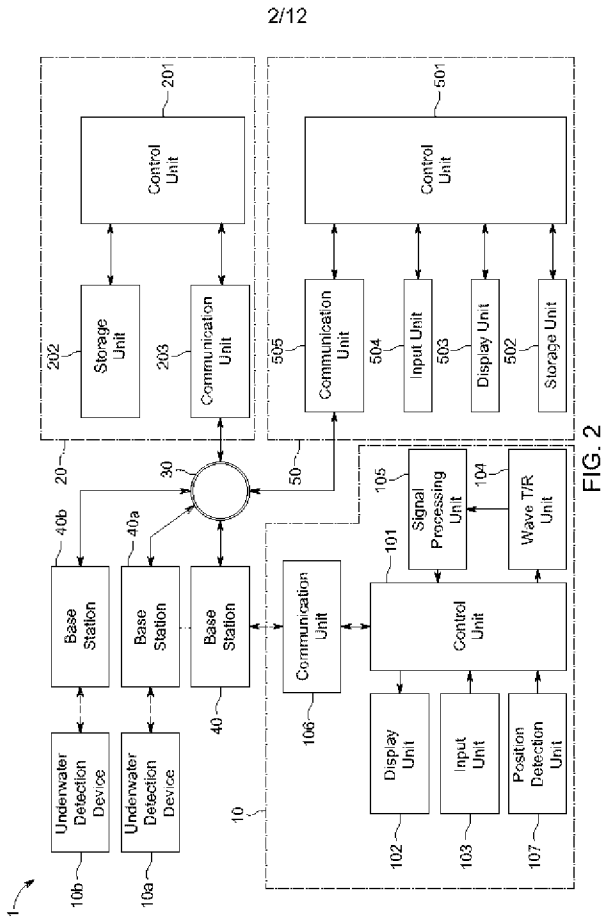

2/12
Control Unit
201
Storage Unit
202
Communication Unit
203
Control Unit
501
Communication Unit
505
Input Unit
504
Display Unit
503
Storage Unit
502
20
30
50
Base Station
40b
Base Station
40a
Base Station
40
Underwater Detection Device
10b
Underwater Detection Device
10a
1
10
Communication Unit
106
Signal Processing Unit
105
Wave T/R Unit
104
Control Unit
101
Display Unit
102
Input Unit
103
Position Detection Unit
107

FIG. 2

[Fig. 3]

3/12

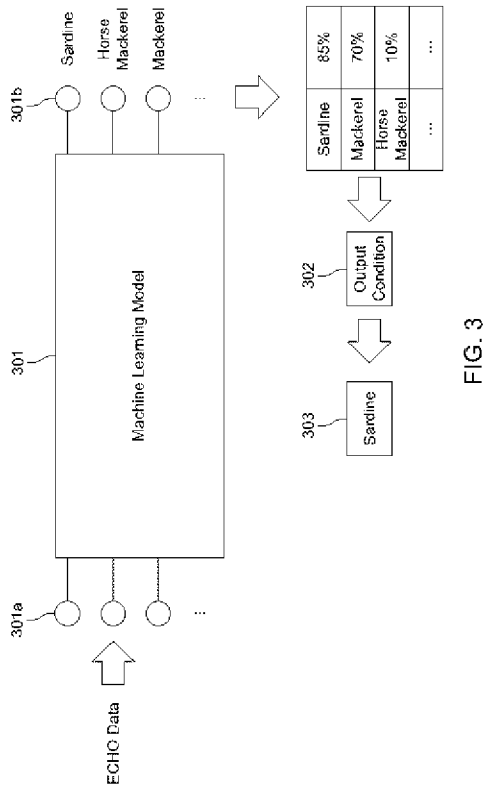


FIG. 3

[Fig. 4]

4/12

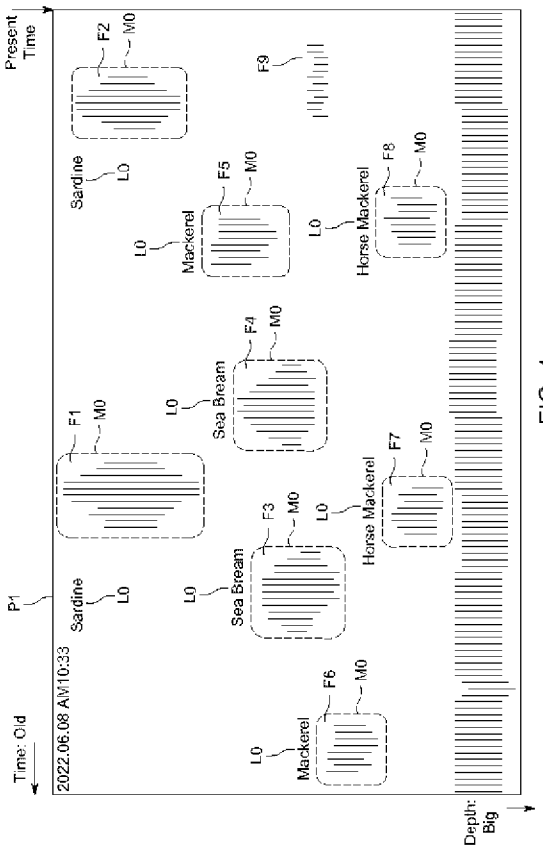


FIG. 4

[Fig. 5]

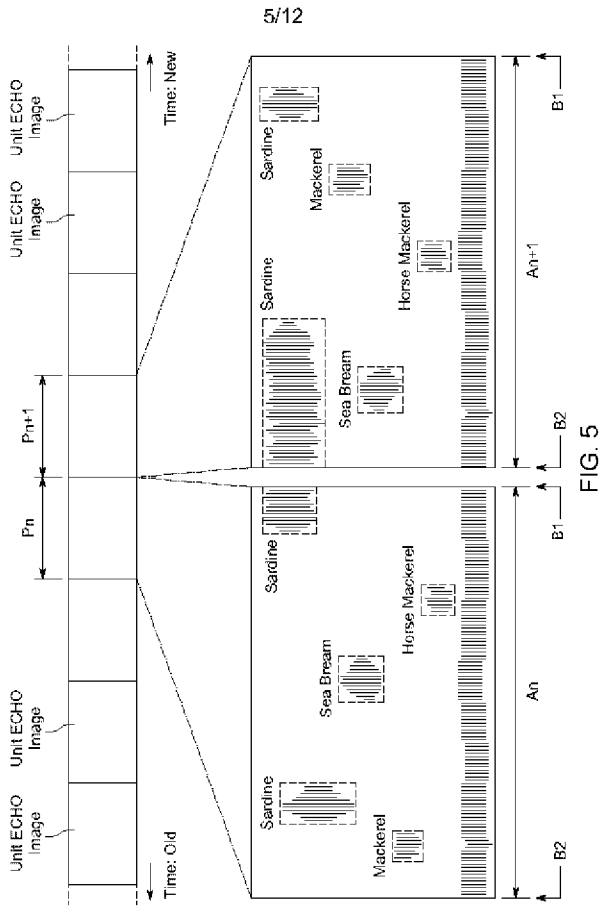

5/12
Unit ECHO Image
Unit ECHO Image
Unit ECHO Image
Unit ECHO Image
Time: New
Time: Old
Pn+1
Pn
Sardine
Mackerel
Sardine
Horse Mackerel
Sea Bream
Sardine
Sea Bream
Horse Mackerel
Sardine
Mackerel
An+1
An
B1
B2
B1
B2

FIG. 5

[Fig. 6]

6/12

Echo ID	Image ID	Unit Echo Data	Annotation Data
E1	P1	E11	A11
	P2	E12	A12
	...	...	...
...	...	...	...

FIG. 6

[Fig. 7]

7/12

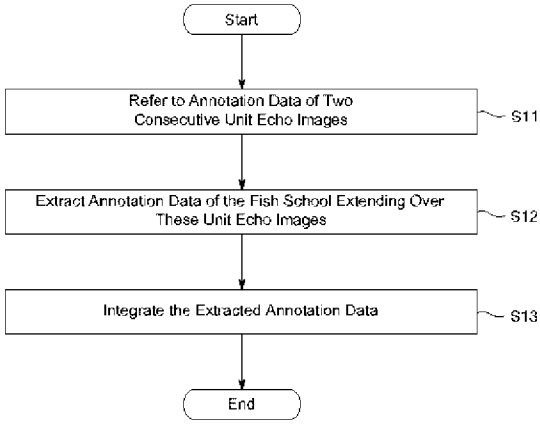


FIG. 7

[Fig. 8]

8/12

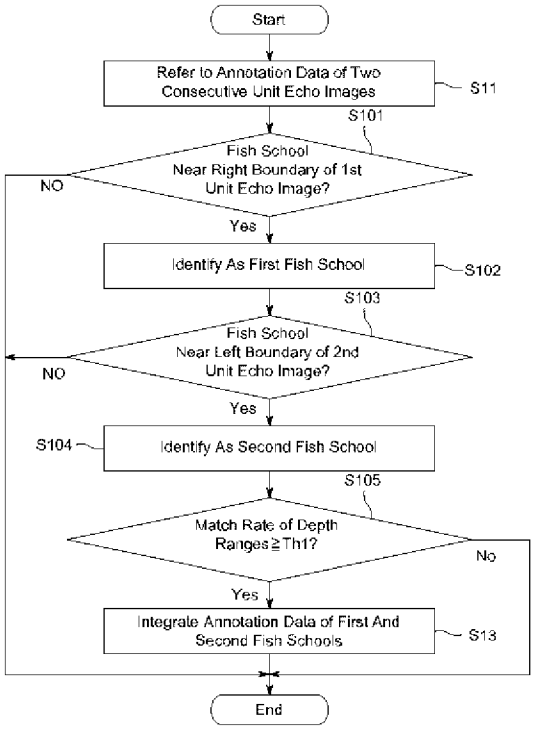


FIG. 8

[Fig. 9]

9/12

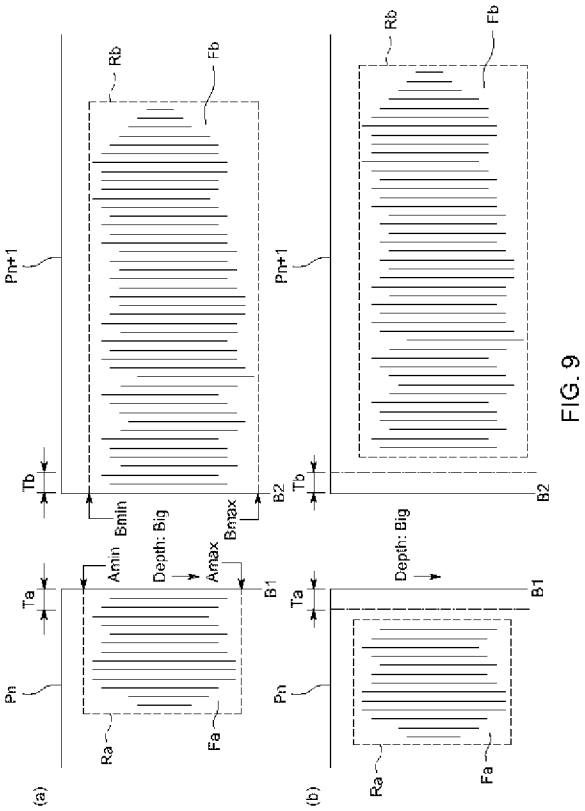


FIG. 9

[Fig. 10]

10/12

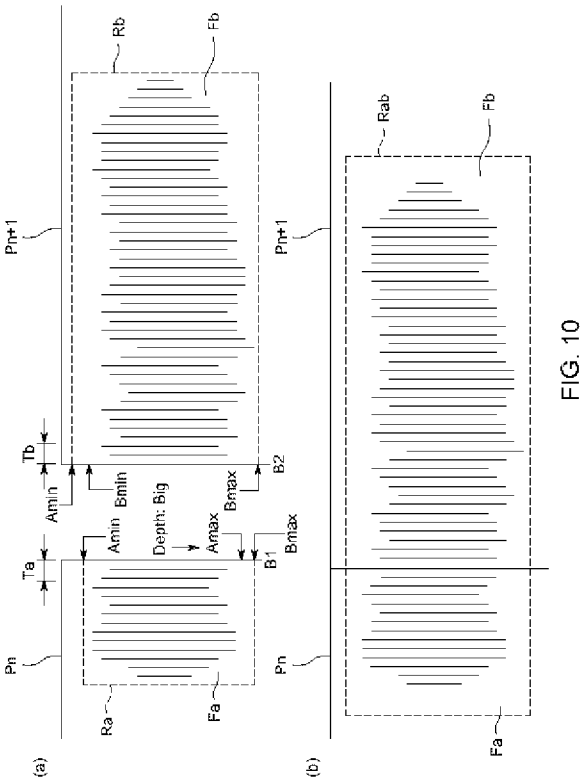


FIG. 10

[Fig. 11]

11/12

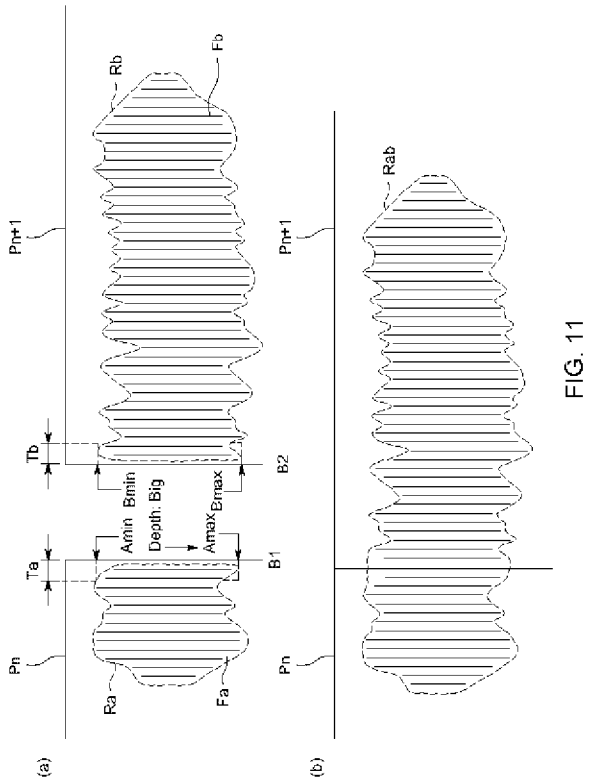


FIG. 11

[Fig. 12]

12/12

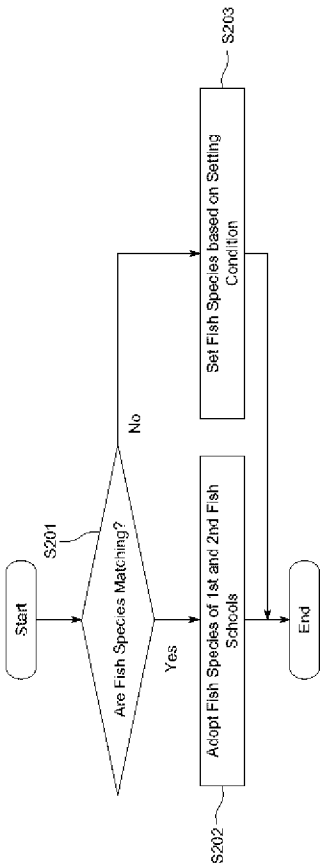


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2023/019775

A. CLASSIFICATION OF SUBJECT MATTER
G06Q50/02,G01S15/96,A01K61/95 Version=2023.01

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)

G06Q, G01S, A01K

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

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

Databases: PatSeer, IPO Internal Database, Google Patents
Keywords: Training data,fish species,school,sonar,echo,ML,discrimination

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP3761059A1(FURUNO ELECTRIC CO), 06.01.2021 (6 January 2021) [whole document]	1-8
Y	JP2014190787A(ROYAL KOGYO KK), 06.10.2014 (6 October 2014) [whole document]	1-8

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"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
06-09-2023	06-09-2023

Name and mailing address of the ISA/	Authorized officer
Indian Patent Office Plot No.32, Sector 14,Dwarka,New Delhi-110075 Facsimile No.	Rakesh Kumar Telephone No. +91-1125300200

Citation	Pub.Date	Family	Pub.Date
EP 3761059 A1	06-01-2021	JP 2021012193 A	04-02-2021