

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.,
Petitioner
v.

AUTONOMOUS DEVICES, LLC,
Patent Owner

Case: IPR2024-00745
U.S. Patent No. 11,663,474

SUPPLEMENTAL DECLARATION OF DR. ELI SABER, PH.D.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	MATERIALS REVIEWED AND TIME OF ALLEGED INVENTION.....	2
III.	PERSON OF ORDINARY SKILL IN THE ART (POSA)	4
IV.	OVERVIEW OF ALLEGED PRIOR ART REFERENCES	4
1.	Grotmol.....	4
2.	Buibas	6
3.	Zhu.....	8
V.	THE ORIGINAL APPLICATION FOR THE ‘474 PATENT INCLUDES SUPPORT FOR THE SUBSTITUTE CLAIMS.....	10
VI.	THE SUBSTITUTE CLAIMS ARE PATENTABLE OVER THE CITED PRIOR ART.....	15
1.	Limitations [85.11]-[85.12] are not disclosed or suggested by Grotmol in view of Zhu.....	17
2.	Limitations [85.13]-[85.14] are not disclosed or suggested by Grotmol in view of Zhu.....	20
3.	Grotmol in view of Zhu and Buibas does not disclose or suggest all the elements of the substitute claims.....	25
VII.	CONCLUSION.....	26

I, Dr. Eli Saber, declare as follows:

I. INTRODUCTION

1. As I mentioned in my earlier Declaration (Ex. 2001) in this case, I have been retained as an independent expert consultant on behalf of Autonomous Devices, LLC (“Patent Owner” or “Autonomous Devices”) in this *Inter Partes* Review proceeding (“IPR”) before the Patent Trial and Appeal Board (“Board” or PTAB”) of the United States Patent and Trademark Office (“PTO”) regarding U.S. Patent No. 11,663,474 (“the ’474 patent” or “the Challenged Patent”) (Ex. 1001).¹

2. My earlier declaration in this IPR contains various opinions of mine, including opinions that I understand Patent Owner cited in support of its Motion to Amend (“MTA”; I understand it is Paper 14 in this IPR). Subsequent to my earlier declaration in this IPR, I understand that Tesla, Inc. (“Petitioner” or “Tesla”) submitted its Opposition to Patent Owner’s MTA (“Opposition brief,” “the Opposition,” or “Opp.”; Paper 18 in this IPR), which cites the second expert declaration of Petitioner’s expert Dr. Jason Janet (Ex. 1022). I provide herein

¹ In this Declaration, I reference documents by exhibit numbers that, as I understand, are attached as exhibits in the *Inter Partes* Reviews of the ’474 patent (IPR2024-00745).

various opinions regarding written description support and patentability of Patent Owner's proposed substitute claims (set forth in the MTA and discussed in my earlier declaration) over certain prior art references cited in the Opposition brief and/or Dr. Janet's second declaration.

3. This declaration is based on the information currently available to me. To the extent that additional information becomes available, I reserve the right to continue my investigation and study, which may include review of documents and information that may be produced, as well as testimony from depositions that have not yet taken place. My opinions are set forth below.

II. MATERIALS REVIEWED AND TIME OF ALLEGED INVENTION

3. The opinions contained in this declaration are based on the documents that I reviewed, my professional judgment, as well as my education, experience, and knowledge regarding technologies related to the '474 patent.

4. In forming my opinions expressed in this declaration, I reviewed:

- the '474 patent (Ex. 1001);
- the file history of the '474 patent (Ex. 1002);
- the first declaration of Dr. Jason Janet (Ex. 1003);
- the second declaration Dr. Janet (Ex. 1022) and references cited therein;
- U.S. Patent No. 9,604,359 to Grotmol, et al. ("Grotmol" – Ex.

1008);

- U.S. Patent Application Publication No. 2014/0089232 to Buibas, et al. (“Buibas” – Ex. 1005);
- U.S. Patent No. 9,221,396 to Zhu, et al. (“Zhu” – Ex. 1023);
- Patent Owner’s Motion to Amend (“MTA” – Paper 14);
- Petitioner’s Opposition brief (Paper 18) and references cited therein;
- The Board’s Preliminary Guidance regarding PO’s MTA (Paper 21);
- any other materials that I refer to in this declaration in support of my opinions, including the Zeiler and Fergus paper (Ex. 2007) and the chapter from the Duda, et al. textbook (Ex. 2008).

5. I have been asked to consider the time of the alleged invention to be around November 2, 2016, which I understand is the filing date of the application (U.S. Patent Application No. 15/340,991) to which the ’474 patent claims priority (Ex. 1001, Cover). My opinions reflect how a person of ordinary skill in the art (as I define in Section V below) would have understood the claims and the specification of the ’474 patent, the prior art to the patent, and the state of the art, at the time of the alleged invention.

6. As I discuss in detail below, it is my opinion that the proposed substitute claims listed in the MTA have adequate written description support and are not disclosed or suggested by the prior art references cited by Petitioner in its Opposition brief.

III. PERSON OF ORDINARY SKILL IN THE ART (POSA)

7. For my analysis regarding the substitute claims, I used the same characterization of a person of ordinary skill in the art (POSA) that I set forth in my earlier declaration. Ex. 2001, ¶¶46-47. All of my opinions in this declaration are from the perspective of a person of ordinary skill in the art, as I have defined it in my earlier declaration and during the relevant time frame (e.g., 2016). During this time frame, I possessed at least the qualifications of a POSA.

IV. OVERVIEW OF ALLEGED PRIOR ART REFERENCES

8. In its Petition, Tesla relies on several references for its arguments, including Buibas, Sinyavskiy, and Grotmol. I describe these references briefly below.

1. Grotmol

9. Grotmol (discussed in my earlier declaration) discloses a method for training and operating a robotic device to follow a simple trajectory based on a simple sequence (i.e., ordered buffer) of images. Ex. 1008, Abstract. If any circumstance occurs for which the simple trajectory is not trained (i.e., different

objects, etc.) the robot in Grotmol would not know what to do or would collide with an obstacle. The robotic device “may utilize a camera” to store a “sequence of images of a visual scene seen when following a trajectory during training in an ordered buffer.” *Id.* During training, “[m]otor commands issued by a trainer corresponding to time instances when the images are acquired may be stored.” *Id.*, 39:17-19. During autonomous operation, the robotic device may be configured to compare an acquired image “with one or more images from the training buffer in order to determine the most likely match.” *Id.*, Abstract. After an acquired image is matched to an image stored in the training buffer, the best match may be used “to predict motor command for controlling operation of a robot.” *Id.*, 27:57-58.

10. Figure 1 from Grotmol, shown below, discloses that robot 160 includes sensor 166. *Id.*, 8:57-61. Sensor 166 provides “information associated with objects within the field-of-view 168.” *Id.*, 8:64-66. Grotmol discloses that its device also includes a memory that stores “training sets,” which may include “an instance of the video frame, an instance of the binary signal, and an instance of a motor control indication configured to cause the robot to execute an action.” *Id.*, 3:35-39.

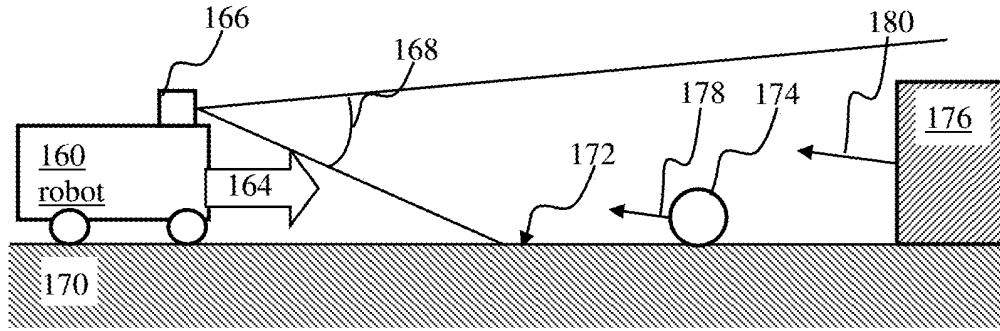


FIG. 1

Id., Fig. 1.

2. Buibas

11. Buibas (discussed in my earlier declaration) is titled “Neural Network Learning and Collaboration Apparatus and Methods” and discloses an “[a]pparatus and methods for learning and training in neural network-based devices.” Ex. 1005, Abstract. Buibas discloses a “robotic brain” that “may comprise one or more neural networks.” *Id.*, ¶[0079]. Figure 10 of Buibas, shown below, discloses that a processing apparatus 1000 receives sensor input 1020. *Id.*, ¶[0136]. This input is transformed into an encoded signal 1026, which is sent to detectors 1002_n, which are used for object detection. *Id.*, ¶¶[0137]-[0139].

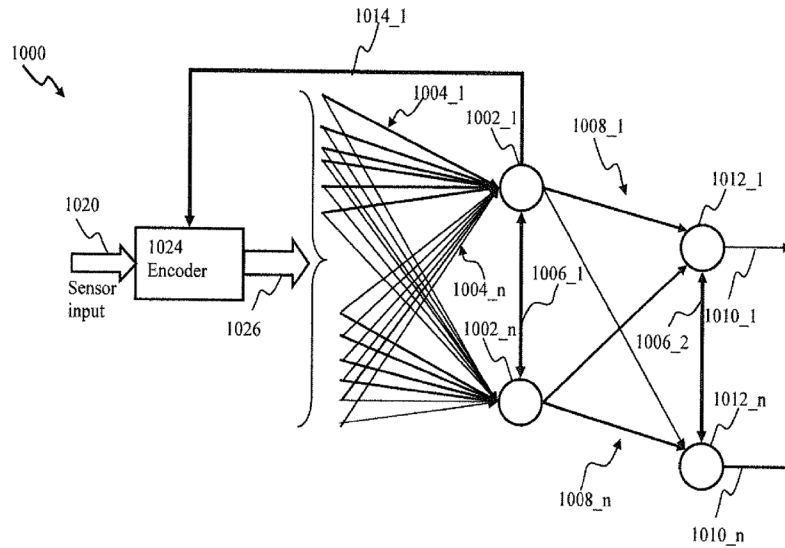


FIG. 10

Id., Fig. 10.

12. Figure 4 from Buibas, shown below, discloses that weight vectors, which are vectors that comprise weight values, can be merged in merge 400, “where the parent[] A provides the portion 432 and the parent B provides the portion 434 to the off-spring 436.” *Id.*, ¶[0096]. Figure 4 merely shows that “offspring network configuration (e.g., the weight vector W) 408 may be obtained using two parent weight vectors,” *id.*, ¶[0095], and never discloses or suggests instruction sets for operating a device.

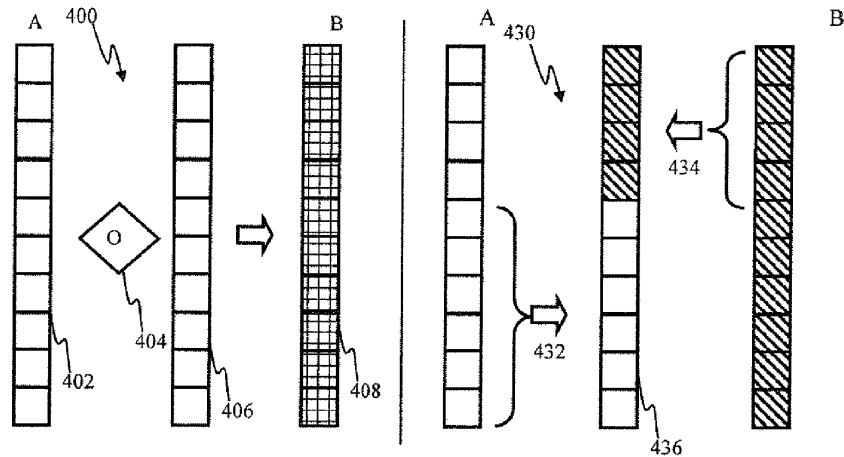


FIG. 4

Id., FIG. 4.

3. Zhu

13. Zhu is titled “Cross-Validating Sensors of an Autonomous Vehicle,” and describes methods and systems for cross-validating a second sensor with a first sensor. Ex. 1023, Abstract. Zhu discloses that such cross-validation may include obtaining sensor readings from the first sensor and comparing them with sensor readings from the second sensor. Such comparison may include obtaining images from each sensor and then comparing various characteristics of the images. *Id.*

14. Zhu discloses detecting vehicles in “various driving environments.” *Id.*, 11:16-18. Zhu describes classifications of detected vehicles, e.g., “such classifications as ‘passenger car,’ ‘bicycle,’ ‘motorcycle,’ and other such classifications.” *Id.*, 11:17-20.

15. Zhu describes that, in the example of a sensor that captures images, a bounding box is used:

For example, where the shape of the object label is a bounding box, and the sensor that captured the image is a [sic] one or more of the cameras, the object label parameters may include a height parameter that defines the height of the bounding box (e.g., in pixels), a width parameter that defines the width of the bounding box (e.g., in pixels), a first pixel coordinate that defines the latitudinal placement of the bounding box (e.g., an X-coordinate), and a second pixel coordinate that defines the longitudinal placement of the bounding box (e.g., a Y-coordinate).

Id., 14:58-67.

16. Zhu shows rectangular bounding boxes 902/904/906 in Figure 9:

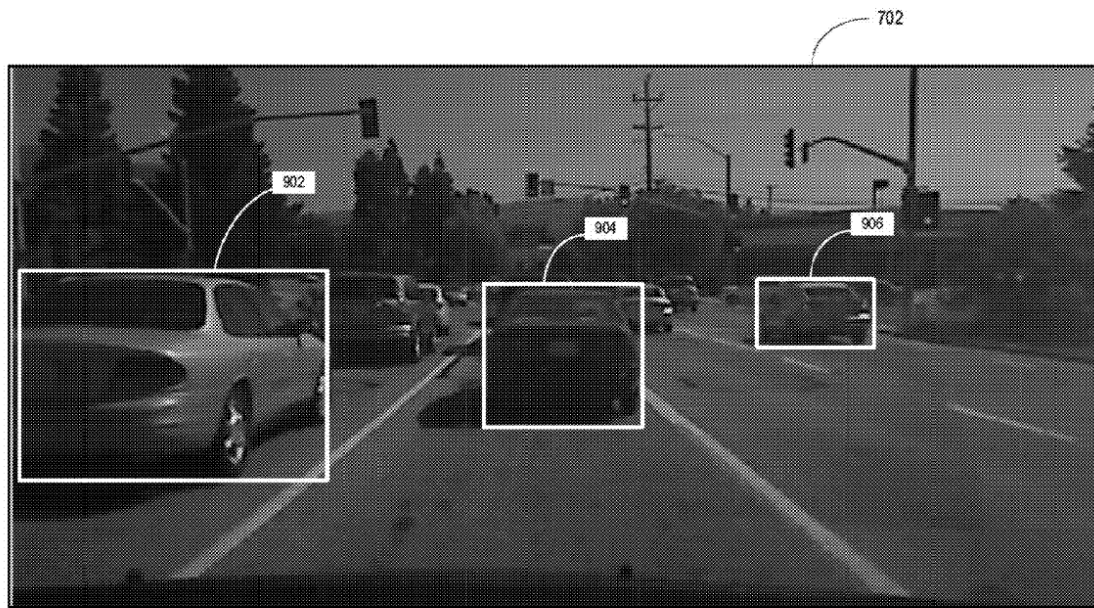


FIG. 9

Id., FIG. 9.

V. THE ORIGINAL APPLICATION FOR THE '474 PATENT INCLUDES SUPPORT FOR THE SUBSTITUTE CLAIMS

17. I noted in footnote 4 (at paragraph 134) of my previous declaration that the '474 patent is a continuation of U.S. Patent Application No. 16/540,972 (“the '972 Application”), which issued as U.S. Patent No. 11,238,344 (“the '344 patent”), and the '972 Application is a continuation of U.S. Patent Application No. 15/340,991 (“the '991 Application”) (Ex. 2004), which issued as U.S. Patent No. 10,452,974. In that footnote, I noted that in my previous declaration citations were provided to the originally-filed '976 and '991 Applications—e.g., '976 Application at [1:25-30] refers to Ex. 1007, 334:25-30, and '991 Application at [1:15-20] refers to Ex. 2004, 305:15-20. I follow that same citation convention here, for citations to the '991 Application.

18. I see that the Opposition brief and Dr. Janet argue that the following claim phrases in substitute claim 85 (at the portions of the claim designated as limitations [85.13] and [85.14] in the MTA) are not sufficiently supported by the written description: “inputting the multiple [and another] object representation data structures of the stream into the one or more input neurons, and using a correlation between the one or more input neurons and the one or more instruction sets for operating the second [and the third] device.” Opp., 23-25; Ex. 1022, ¶¶79-81. I disagree with Petitioner and Dr. Janet, as I believe there is adequate written description support. For convenience, I discuss in this present declaration various

arguments and citations found in the Opposition brief, but they are also generally found in Dr. Janet's second declaration (Ex. 1022) and so my opinions here are equally applicable to Dr. Janet's second declaration.

19. Figure 4B shows object representation data structures that are disclosed in the '474 patent:

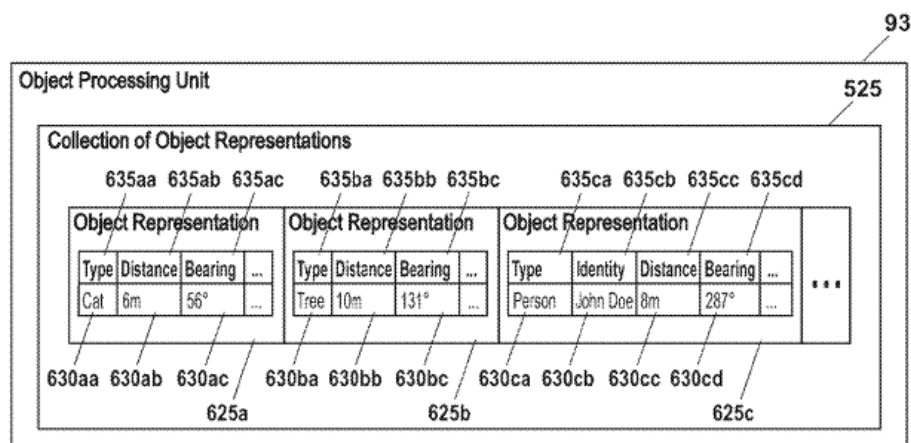
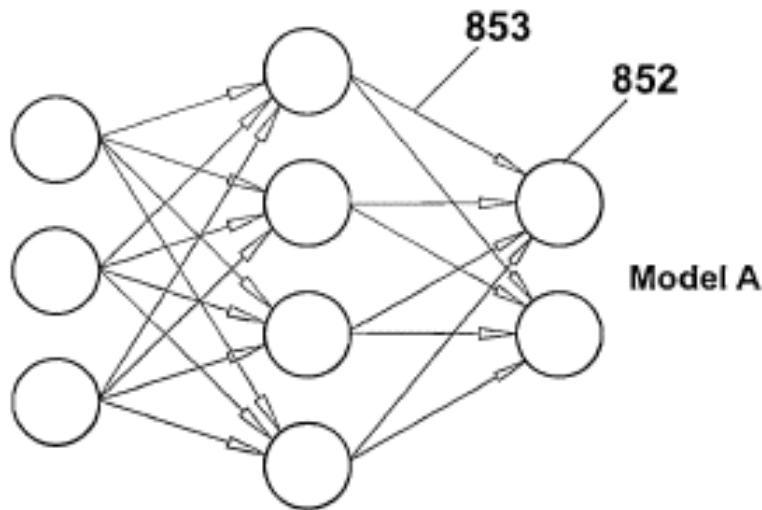


FIG. 4B

Ex. 1001, FIG. 4B; '991 Application at FIG. 4B.

20. The specification supports the claimed label for object type, coordinates for object location, and computer model for object size. '991 Application at [136:7-12]. The specification also supports inputting the claimed data structures *into the input neuron(s)* as claimed in substitute claim 85. For example, Figure 17 shows a neural network, where the three neurons on the left are input neurons:



Ex. 1001, FIG. 17; '991 Application at FIG. 17; *see also id.* at [94:7-9] (“[A] neural network may include ... input neurons ... activated by inputs. Activations of these neurons can then be passed on, weighted, and transformed by a function to other neurons”). The arrows shown above in Figure 17 demonstrate how processing flows from input neurons (at left) to output neurons (at right), via an intermediate layer of neurons (the four neurons in the middle).

21. Model A (in Figure 17) shows input neurons correlated using downstream connections with output neurons including instruction sets. MTA, 16-17; *see also* '991 Application at [52:22-26] (“correlated instruction sets for operating the device”). The specification explains that the neural network is a knowledgebase, which in turn is part of an AI unit. *See* '991 Application at [87:10-11] (“Artificial Intelligence Unit ... comprises ... Knowledgebase”), [92:14-15] (“Knowledgebase ... may be or include Neural Network”). The

specification further discloses that the “Artificial Intelligence Unit ... comprises the functionality for *anticipating one or more Instruction Sets* ... based on one or more *incoming Collections of Object Representations*” (e.g., collections of object representations 525 from Figure 4B). *Id.* at [87:21-26]. Thus, the collections of object representations, and hence the multiple object representation data structures ([85.13]) and the another multiple object representation data structures ([85.14]), are *inputted to the input neuron(s)* because they are incoming from the perspective of the input neurons of the neural network which is part of the AI unit.

22. I also disagree with the argument Petitioner and Dr. Janet have made regarding alleged lack of support for “using a correlation between the one or more input neurons and the one or more instruction sets for operating the [second | third] device” in limitations [85.13] and [85.14].

23. The neural network of Model A (discussed above) has output neurons (at right in figure) which are used for “anticipating one or more Instruction Sets ... based on one or more incoming Collections of Object Representations.” ’991 Application at [87:24-25]. In my opinion, a POSA would have recognized by reading the specification that the inventor of the ’474 patent had possession of, for example, the top-right neuron in Figure 17 providing instruction set(s) for operating one device (second device) and the bottom-right

neuron providing instruction set(s) for operating another device (third device). The specification discloses that “Device 98 may be or include Loader 98a” (*Id.* at [146:29], FIG. 38), and that there may be “[a]ny number of loaders 98a [that] utilize such collective knowledge comprised in remote DCADO Unit 100 for their autonomous operation.” *Id.* at [148:6-8], [147:8]. The specification explains that “remote DCADO Unit 100 enables learning of the operators’ collective knowledge of operating Loaders 98 in circumstances including objects with various properties.” ’991 Application at [147:38-148:6], [93:21-23] (“For example, a Device 98 including or interfaced with DCADO Unit 100 having access to a combined Knowledgebase 530 can use the collective knowledge learned from multiple Devices 98 for the Device’s 98 autonomous operation.”). Figures 38 and 39 show examples of different types of devices, e.g., loader 98a and boat 98b, instead of showing multiple instances of the same loader (which is disclosed elsewhere in the specification). MTA, 17. As noted above, the specification describes any number of loaders 98, which a POSA would have understood as also disclosing any number of boats 98 (or other similar devices) without having to mix loaders and boats.

24. Therefore, a POSA would have recognized by reading the specification that the inventor had possession of the fact that when there are, for example, multiple loaders 98a in communication with a remote DCADO Unit

100, one loader 98a can operate autonomously based on prior operation of multiple other loaders 98a that, at other times, contributed to collective knowledge of DCADO unit 100. '991 Application at [87:23-26], [93:21-23]. This would involve using correlations between input neurons in a neural network and instructions sets for operating the other Loaders 98a. *Id.* at [94:7-8]. The specification also provides support for object representation data structures for the multiple objects detected at different times. For example, the '991 Application discloses that a "Collection of Object Representations 525 may, therefore, include knowledge (i.e., unit of knowledge, etc.) of Device's 98 circumstances including objects with various properties at a particular time," and that "one Collection of Object Representations 525 may be associated with time stamp t1, another Collection of Object Representations 525 may be associated with time stamp t2, and so on," where "Time stamps t1, t2, etc. may indicate the times of generating Collections of Object Representations 525, for instance" and "a stream of Collections of Object Representations 525 may ... include knowledge (i.e. unit of knowledge, etc.) of Device's 98 circumstances including objects with various properties over time." *Id.* at [65:16-29].

VI. THE SUBSTITUTE CLAIMS ARE PATENTABLE OVER THE CITED PRIOR ART

25. In my opinion, the alleged prior art fails to disclose or suggest to a POSA all of the elements of the substitute claims listed in the Motion to Amend

(MTA) under the two grounds presented for these claims in the Opposition brief. Each ground is addressed individually below. I focus on certain limitations ([85.11], [85.12], [85.13], and [85.14]) of substitute claim 85, because various other substitute claims (86-91) are dependent claims that inherit those limitations from claim 85. For convenience, I have recited limitations [85.11]-[85.14] as they are listed in the Motion to Amend (underlining that was present in the MTA is omitted here for clarity):

[85.11] the generated circumstance representation that (i) includes multiple object representation data structures that each comprises: a label for object type, coordinates for object location, and a computer model for object size, and (ii) represents multiple objects detected at a first time, and

[85.12] another circumstance representation that (i) includes another multiple object representation data structures that each includes: a label for object type, coordinates for object location, and a computer model for object size, and (ii) represents the multiple objects detected at a second time,

[85.13] wherein the determining the one or more instruction sets for operating the second device includes: inputting the multiple object representation data structures of the stream into the one or more input neurons, and using a correlation between the one or more input neurons and the one or more instruction sets for operating the second device; and

[85.14] determining another one or more instruction sets for operating a third device at least by: inputting the another multiple object representation data structures of the stream into the one or more input neurons, and using a

correlation between the one or more input neurons and the another one or more instruction sets for operating the third device.

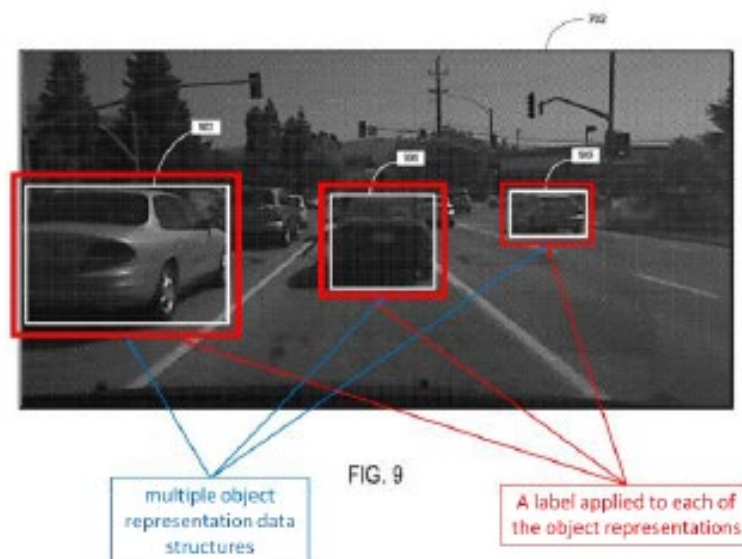
1. Limitations [85.11]-[85.12] are not disclosed or suggested by Grotmol in view of Zhu.

26. In my opinion, Grotmol in view of Zhu does not disclose or suggest “multiple object representation data structures that each comprises: ... coordinates for object location, and a computer model for object size” as recited in limitations [85.11]-[85.12]. Although the Opposition mentions certain disclosures from Grotmol (Opp., 10-11), it relies on obviousness in view of Zhu for [85.11]-[85.12]. Opp., 13 (“It would have been obvious to a POSITA to apply Zhu’s suggestion for ... defining bounding boxes for detected objects that provides a computer model for object size coordinates for object location [sic] to Grotmol...”); *see also id.*, 11-12. Limitations [85.11] and [85.12] recite “object representation data structures” that comprise the “coordinates for object location” and “model for “object size.” The object representation data structures are part of a circumstance representation, which represents a circumstance detected by one or more actual, real-world sensors (*see* Ex. 1001, Figs. 3A-3E, showing Camera 92a, Microphone 92b, Lidar 92c, Radar 92d, and Sonar 92e) for detecting objects and their properties, as recited in [85.2] and [85.5].

27. In my opinion, a POSA would have understood that the object (and its properties) that is detected by the sensor(s) is an actual, real-world object in

the surroundings of the first device recited in [85.pre] that itself exists in the real world. Ex. 1001, 56:1-3 (“FIGS. 4A-4B, illustrate an exemplary embodiment of Objects 615 detected in Device's 98 surrounding, and resulting Collection of Object Representations 525.”), FIGS. 4A (showing an object 615a that is a cat at a distance of 6m from device 98), 4B (showing properties of the detected objects).

28. In contrast, the Opposition relies on disclosures showing a bounding box in an image, the height and width of the box in pixels for the size of the bounding box (not the object's actual size), and first and second pixel coordinates that define location of the box in the image (not the object's actual coordinates). Opp., 11-12 (citing Ex. 1023, 14:43-67, 15:1-20, FIGS. 9-10). Petitioner annotated the bounding boxes as follows:



Opp., 12 (annotated Zhu FIG. 9).

29. These portions of Zhu just show representations of objects, as

indicated by bounding boxes in the image coordinate system, not coordinates for location and model for size of objects as claimed in limitations [85.11]-[85.12]. The bounding boxes of Zhu do not provide “coordinates for object location” and certainly do not constitute a “computer model for object size” as claimed in limitations [85.11]-[85.12]. Bounding boxes are typically utilized to provide a rough indication of the location of the object in the imagery. Limitations [85.11] and [85.12] do not recite coordinates for object representation location and model for object representation size like disclosed by Zhu’s bounding boxes, but instead recite “coordinates for object location” and “computer model for object size” of the object. Limitations [85.11] and [85.12] also each recite a “circumstance representation that represents multiple objects detected,” so it is apparent that the circumstance representation (and its object representation data structures for which Opposition relies on Zhu’s pixels or bounding boxes in an image representing depicted objects) is distinct from the detected objects that the circumstance representation represents.

30. The height and width of Zhu’s bounding box “in pixels” and the “first pixel coordinate” and “second pixel coordinate” that define the latitudinal/longitudinal placement of the bounding box (id.) are distinct from, and do not disclose or suggest, the claimed “coordinates for object location, and a computer model for object size” for detected objects of [85.11]-[85.12]. Opp., 11-

12. Petitioner argues that “Zhu discloses labeling the object type for each object by generating ‘such classifications as passenger car, bicycle, motorcycle, and other such classifications’ that indicate ‘the type of vehicle detected by a given sensor.’” Opp., 11. Thus, for “label for object type” in [85.11]-[85.12], Petitioner relies on Zhu’s disclosure of “the type of vehicle detected by a given sensor.” In other words, Petitioner’s mapping of “label for object type” to Zhu’s disclosure of an object type is at odds with Petitioner’s mapping of “coordinates for object location” and “computer model for object size” to pixel-based properties of a rectangular bounding box that only pertains to a representation (in the image coordinate system) of the underlying object.

31. Therefore, it is my opinion that the Opposition brief’s analysis based on Grotmol-Zhu for “coordinates for object location” and “model for object size” is flawed.

2. Limitations [85.13]-[85.14] are not disclosed or suggested by Grotmol in view of Zhu.

32. In my opinion, Grotmol in view of Zhu does not disclose or suggest at least the following elements of limitations [85.13]-[85.14]: “inputting the [multiple object representation data structures [of limitation [85.11]] | another multiple object representation data structures [of limitation [85.12]] of the stream into the one or more input neurons.” As an initial matter, the Opposition’s analysis is flawed because Grotmol-Zhu does not disclose or suggest the

specifically claimed data structures of [85.11]-[85.12] (*see* above at Section VI.1), and thus does not disclose or suggest inputting those data structures into input neuron(s). The Opposition's analysis of limitations [85.13]-[85.14] is also flawed.

33. The Opposition first mentions Grotmol's disclosures. Opp, 14-15. Grotmol discloses that a video camera obtains images, which are fed as input into a neural network. Ex. 1008, 30:40-42. Grotmol describes image processing techniques (e.g., normalization such as feature shifting and/or scaling) for processing the pixels before providing the image features (i.e., edges, corners, etc.) as inputs to the neural network. *Id.*, 7:63-65, 30:54-31:35. Thus, image pixels or image features are used to provide inputs to Grotmol's neural network. Grotmol does not disclose the claimed object representation data structures of [85.11]-[85.12] being inputted into input neurons. Opp., 14-15. For the claimed inputting of the object representation data structures into input neuron(s), the Opposition does not cite to Zhu or point to anything in Zhu, and does not even mention or cite Zhu for [85.13]-[85.14]. *Id.*, 14-16. Nevertheless, Zhu does not disclose or suggest inputting the claimed object representation data structures into input neurons. Thus, neither of the cited references (Grotmol and Zhu) actually discloses the claimed inputting of the specific claimed object representation data structures (recited in limitations [85.11]-[85.12]) into input neuron(s).

34. In my opinion, it would have been complex, and certainly not unusually simple or straightforward, to replace one type of input (Grotmol's images or image features) into a neural network with another type of input (the object representation data structures of [85.11]-[85.12]) due to the complexities associated with a completely different structure of the neural network to accommodate completely different inputs, training with a different type of inputs, and the significant iterative nature of the training process that typically utilizes backpropagation type algorithms that are often combined with regularization techniques to appropriately train the network. This is so because, generally speaking during retraining, proper weights for correlations/connections are calculated/adjusted through the backpropagation process (*see* Ex. 1001, 108:64-67) for specific instruction sets associated with specific same inputs in each training set. Backpropagation, generally speaking, typically involves an iterative process of training a neural network where a training input is provided to a given network and the corresponding calculated output is computed via feed forward operations through the network. The error between the calculated and expected output is then backpropagated throughout the network using algorithms such as gradient descent in order to properly update the weights throughout the layers. This process is repeated multiple times with various inputs using an appropriately chosen training dataset and proper regularization techniques (to ensure that the

weights do not diverge and cause the overall network to become unstable) until convergence is achieved. The optimized network is then tested with hyperparameters (e.g., learning rates, line search, batch selection, etc.) potentially further optimized using a validation dataset and then further tested using a test dataset to measure/quantify its performance.

35. Grotmol and Zhu do not disclose or suggest anything like this because they only contemplate using images and image features as inputs, and in any case, do not disclose or suggest inputting the specific object representation data structures of [85.11]-[85.12]. Petitioner's proposed modification of Grotmol's system would have resulted in an inoperable system, with incorrect and unusable results, if object representation data structures as claimed were to be inputted into Grotmol's input neurons. That is because input neurons of Grotmol's image-based neural network can only accept pixels of an image or potentially simple image features (such as normalization via feature shifting or scaling) to produce any meaningful results (e.g., a coordinate value or size data cannot be inputted into an input neuron that expects a color value for a pixel, and if attempted, this would have produced meaningless output). For example, convolutional neural networks (CNN) are typically designed and widely accepted as well suited to accept images as input and perform convolutions² followed by

² CNN typically perform cross-correlation rather than convolution even though

pooling in a series of layers to extract features from structured data such as imagery (pixels of an image typically have a well-defined spatial relationship to each other because the input variables are arranged in a two-dimensional grid, and nearby pixels have highly correlated values). *See generally* Ex. 2007. These features are then provided as input to fully connected neural networks that are designed and trained to perform various tasks such as classification, recognition, etc. Inputting images of practical size (e.g., 512 x 512 grayscale image) directly into a fully connected neural network would enormously increase the size of the input layer causing undue complexities in training given the enormous number of weights that would need to be handled and often yield erroneous, meaningless, and/or unstable results. Hence, neural networks are designed and trained for specific tasks with specific inputs and outputs. It is not straightforward or unusually simple to simply swap inputs (e.g., input object representations instead of image pixels for instance). Introducing different style of inputs (e.g., object representations instead of pixels) into a neural network that is already trained for a specific task would potentially require a complete re-design of its layers, structure, redefinition of appropriate loss functions to quantify the differences between desired and actual outcomes, selection of a properly diversified training dataset and many more followed by a complete retraining with appropriate use of

they are recognized as convolutional neural networks.

techniques such as regularization, normalization, boosting, etc., and proper tuning of hyperparameters such as learning rates, line search, batch selection, etc., in order to provide a final product that generalizes well to the problem at hand as demonstrated by proper testing on appropriately selected validation and test datasets. In essence, neural networks are highly specialized systems that are specifically designed for specific tasks. *See* Ex. 2008.

36. It would not have been predictable to just replace one type of input with a different type of input as Petitioner and Dr. Janet assert. Rather, it was complicated at the invention date to modify a neural network designed to accept pixels, instead to accept other types of inputs (such as object representation data structures) for the reasons described above. The neural network would have to be structured with different numbers of neurons, layers and connections, trained on the completely different type of inputs, and tested and optimized through significant trial and error as described above to prove it is workable, which would not have been simple or predictable. Grotmol in view of Zhu and Buibas does not disclose or suggest all the elements of the substitute claims.

3. Grotmol in view of Zhu and Buibas does not disclose or suggest all the elements of the substitute claims.

37. The Opposition brief does not present separate arguments regarding the Grotmol-Zhu-Buibas combination for [85.11]-[85.12]. Opp., 10-14. For [85.13]-[85.14], Petitioner introduced Buibas as an additional secondary

reference but not for the portions of those limitations that I discussed above in Section VI.2. Thus, in my opinion Grotmol in view of Zhu and Buibas does not disclose or suggest all the features of claims 85 and 86-91 (which inherit the above limitations via dependency).

VII. CONCLUSION

38. This declaration represents the opinions that I have formed up to date based on the material currently available to me. I reserve the right to revise, supplement, and/or amend my opinions stated in this declaration based on new information that becomes available to me and on my continuing analysis of the materials available to me.

39. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine, imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: May 22, 2025



Dr. Eli Saber