

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of: Jasmin Cosic
U.S. Patent No.: 11,663,474 Attorney Docket No. 49649-0038IP1
Issue Date: May 30, 2023
Appl. Serial No.: 17/561,976
Filing Date: December 26, 2021
Title: ARTIFICIALLY INTELLIGENT SYSTEMS, DEVICES,
AND METHODS FOR LEARNING AND/OR USING A
DEVICE'S CIRCUMSTANCES FOR AUTONOMOUS
DEVICE OPERATION

SUPPLEMENTAL DECLARATION OF DR. JASON JANET

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	GROUND 1A: OBVIOUSNESS BASED ON BUIBAS-SINYAVSKIY...4	
A.	Buibas Renders Obvious “The One or More Inputs are Further Correlated with Another One or More Instruction Sets for Operating a Third Device” (Claims 56, 57).....	4
1.	Correlating the Inputs to a Knowledgebase with Multiple Instruction Sets is a Non-Critical Feature in the ’474 Patent	6
2.	Buibas’ Merge Operations Combine Fully-Defined Brain Images, Which Involves Merging More Than Weights of Brain Images	9
3.	Buibas’ Merged Brain Images Correlate the Same One or More Inputs with Instruction Sets for Operating a Second Device and Additional Instruction Sets for Operating a Third Device.....	13
4.	Buibas’ Merge Operations Would Result in Operable Networks that Produce Instruction Sets for Operating a Second Device and Additional Instruction Sets for Operating a Third Device.	18
5.	Merging Weight Vectors by Itself Would Still Produce a Brain Image Having One or More Inputs Correlated with the Claimed Instruction Sets for Operating Second and Third Devices	20
B.	Buibas Renders Obvious “At Least a Portion of the One or More Instruction Sets for Operating the Third Device ... is Learned in Another Learning Process that Includes Operating the Third Device at Least Partially by Another User” (Claim 56).....	22
III.	GROUND 2C: OBVIOUSNESS BASED ON GROTMOL-BUIBAS	25
A.	Grotmol and Buibas Render Obvious “The One or More Inputs are Further Correlated with Another One or More Instruction Sets for Operating a Third Device” (Claims 56, 57)	25

B.	Grotmol and Buibas Render Obvious “At Least a Portion of the One or More Instruction Sets for Operating the Third Device ... is Learned in Another Learning Process that Includes Operating the Third Device at Least Partially by Another User” (Claim 56).....	28
IV.	ASSESSMENT OF THE SUBSTITUTE CLAIMS	28
V.	THE SUBSTITUTE CLAIMS ARE UNPATENTABLE UNDER § 103	29
A.	Overview of Additional Prior Art	29
1.	Overview of Zhu	29
B.	Overview of the Prior Art Combinations	30
1.	Ground MTA-1A: Grotmol-Zhu.....	30
2.	Ground MTA-1B: Grotmol-Zhu-Buibas	34
VI.	OBVIOUSNESS OF THE SUBSTITUTE CLAIMS	35
1.	Elements [85.pre]-[85.7].....	35
2.	Element [85.8].....	35
3.	Element [85.9].....	39
4.	Element [85.10].....	39
5.	Element [85.11].....	41
6.	Element [85.12].....	45
7.	Elements [85.13]-[85.14]	47
8.	Claims 86-93	55
VII.	THE SUBSTITUTE CLAIMS LACK SUPPORT UNDER § 112.....	56
A.	“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 [andthe third] device”.....	56

VIII. CONCLUSION	59
-------------------------------	-----------

I, Jason Janet, Ph.D., declare as follows:

I. INTRODUCTION

1. I have been retained as a technical expert by counsel on behalf of Tesla, Inc. (“**Tesla**”/“**Petitioner**”) in Case No. IPR2024-00745 with respect to U.S. Patent No. 11,663,474 (“**the ’474 Patent**”) (TESLA-1001).

2. I understand that this declaration (“**Supplemental Declaration**”) will be submitted in support of (i) Petitioner’s Reply to Patent Owner’s Response (“**POR**”) to the Petition for *Inter Partes* Review of the ’474 Patent and (ii) Petitioner’s Opposition to Patent Owner’s Motion to Amend (“**MTA**”). This Supplemental Declaration supplements certain portions of my prior declaration in support of the Petition for this IPR Proceeding (TESLA-1003, “**my Prior Declaration**”). In my Prior Declaration, I address many topics, including (but not limited to) my background and qualifications, the level of skill in art, an overview of the ’474 Patent, a POSITA’s background knowledge, certain legal standards explained to me by counsel, legal principles applied for the purposes of my analysis and this declaration, and a detailed analysis of the prior art against the claim elements of the ’474 Patent.

3. I maintain all of my earlier analysis and continue to hold the opinions expressed in my Prior Declaration. This Supplemental Declaration does not replace my Prior Declaration. These declarations are complementary and should be read in

tandem. Accordingly, this Supplemental Declaration does not cover every topic addressed in my Prior Declaration.

4. In forming my opinions expressed in this Supplemental Declaration, I have considered, among various documents, the following: the POR, the MTA, the declaration of Patent Owner's declarant, Dr. Saber (EX-2001), the Board's Grant of Institution (Paper 9), my knowledge and experience in this field, including my research, and work experience in the field, my experience working with others involved in the field, and the third-party documents cited throughout my analysis. *See, e.g.*, TESLA-1002, ¶¶3-19 (discussing qualifications) and ¶¶21-34 (discussing legal standards and the level of skill for a "POSITA"). I have also considered the following additional references:

- U.S. Patent No. 9,221,396 to Zhu et al. ("Zhu" or TESLA-1023)
- *Neural Network Design*, Second Edition ("Hagan" or TESLA-1024)
- *Tesla, Inc. v. Autonomous Devices, Inc.*, IPR2023-01054, Paper 46 (PTAB December 12, 2024) – Final Written Decision on U.S. Patent No. 10,452,974 (TESLA-1026)
- *Tesla, Inc. v. Autonomous Devices, Inc.*, IPR2023-01055, Paper 47 (PTAB December 12, 2024) – Final Written Decision on U.S. Patent No. 11,238,344 (TESLA-1027)
- Eli Saber Deposition Transcript (April 4, 2025) (TESLA-1028)
- U.S. Patent Application Serial No. 13/623,820 ("Fisher" or TESLA-1029)
- U.S. Patent No. 10,131,052 ("Gabardos" or TESLA-1030)

- U.S. Patent No. 8,793,205 (“Fisher205” or TESLA-1031)
- Eli Saber Deposition Transcript in ’583 Patent IPR, IPR2023-00172, Exhibit 1011 (June 21, 2024) (TESLA-1032)
- *Tesla, Inc. v. Autonomous Devices, Inc.*, IPR2023-01054, Paper 26 – Patent Owner’s Response (TESLA-1033)
- *Tesla, Inc. v. Autonomous Devices, Inc.*, IPR2023-01055, Paper 26 – Patent Owner’s Response (TESLA-1034)
- U.S. Patent Pub. No. 2017/0200078 A1 (“Bichler” or TESLA-1035)
- U.S. Patent Pub. No. 2015/0363687 A1 (“Gocek” or TESLA-1036)
- U.S. Patent Pub. No. 2013/0304684 A1 (“Yufik” or TESLA-1037)
- U.S. Patent Pub. No. 2017/0010106 A1 (“Shashua” or TESLA-1038)
- File History of U.S. Patent App. No. 15/340,991 (“the ’991 Application” or EX-2004)

5. Although I cite and quote to selected portions of various references in this Declaration, the reader should understand that these citations are representative examples. A POSITA would have viewed each reference in its entirety and in combination with other references. Accordingly, I intend the references identified in this Declaration to be viewed as incorporated in their entireties.

6. I am not, and never have been, an employee of Tesla. I received no compensation for this Supplemental Declaration beyond my normal hourly compensation based on my time actually spent analyzing the ’474 Patent, the materials cited herein, and various issues related thereto. I will not receive any added compensation based on the outcome of any IPR or other proceeding involving the

'474 Patent.

7. The analysis and opinions expressed in my Prior Declaration fully explain why [1] each and every feature of the '474 Patent's claims is provided in the prior art; and [2] a POSITA would have found it obvious and would have been motivated to combine the known prior art elements in a manner that reaches the claimed invention with reasonable expectation of success. Patent Owner and its declarant, Dr. Saber, have considered my opinions and offered their own, some of which are inconsistent with my view. I will address certain points of dispute below. The fact that I may have not addressed all of Patent Owner's or Dr. Saber's opinions should not be interpreted as agreement with them.

II. GROUND 1A: OBVIOUSNESS BASED ON BUIBAS-SINYAVSKIY

A. Buibas Renders Obvious "The One or More Inputs are Further Correlated with Another One or More Instruction Sets for Operating a Third Device" (Claims 56, 57)

8. Claims 56 and 57 each depend from claims 45 and 46 and identically add to the base claims a limitation requiring that "the one or more inputs are further correlated with another one or more instruction sets for operating a third device." TESLA-1001, 177:58-178:7. Read in the context of the base limitations from claim 45, claims 56 and 57 thus require, among other things:

A first device comprising:

*a knowledgebase that includes **one or more inputs** for inputting at least a portion of a circumstance representation,
wherein the one or more inputs are correlated with one or more instruction sets for operating a **second device**; ... [and]
wherein the one or more inputs are further correlated with another one or more instruction sets for operating at **third device**.*

9. As I explained in my Prior Declaration, the phrase “instruction sets for operating a second device” in the ’474 Patent encompasses instruction sets learned at the second device by operation of the second device. TESLA-1003, ¶40; TESLA-1001, 102:30-103:39, FIG. 38. Likewise, the phrase “instruction sets for operating a third device” encompasses instruction sets learned at the third device by operation of the third device. *Id.* Taken together, the language of claims 56-57 recite features encompassing a first device having a knowledgebase with one or more inputs correlated both with (i) instruction sets learned on a second device and (ii) additional instruction sets learned on a third device—i.e., a concept that I understand Patent Owner has referred to as a “mixed” knowledgebase in its arguments on the similar claims previously found unpatentable in IPR2023-01054 and -01055. TESLA-1033, 28, 31-32, 53; TESLA-1034, 27, 32-33.

10. Based on my review, Patent Owner and Dr. Saber do not appear to dispute that the Buibas-Sinyavskiy combination (Ground 1A) provides a knowledgebase on a **first device** that includes one or more inputs correlated with

instruction sets for operating a *second device*. The POR and Dr. Saber’s declaration instead focus attention on a similar limitation that requires one or more inputs of the knowledgebase to be further correlated with additional instruction sets for operating a *third device*. I do not agree that the Buibas-Sinyavskiy combination fail to render the features recited in claims 56 and 57 obvious, as Patent Owner and Dr. Saber contend.

1. Correlating the Inputs to a Knowledgebase with Multiple Instruction Sets is a Non-Critical Feature in the ’474 Patent

11. It is important to note that Buibas discloses techniques for merging knowledge (e.g., instruction sets) learned at multiple devices into a single “mixed” knowledgebase (e.g., a brain image, robot brain, or partition thereof). Buibas therefore discloses the core concept Patent Owner touted as allegedly innovative subject matter in prior IPRs on patents related to the ’474 Patent. *See* TESLA-1033, 28, 31-32, 53; TESLA-1034, 27, 32-33. For example, Buibas describes methods for sharing and mixing trained robot behaviors learned at multiple devices, such as by “merging” brain images, as exemplified by the following disclosures:

- “ [A] user may extract the state of a neural network (or other useful training-related information) directly from the device.

Thus, the artificial mind and its traits may be copied, stored, and

later retrieved. This state information may be shared with other users.” TESLA-1005, [0070].

- “[P]artial [brain] images may be used to augment or upgrade (downgrade) a pre-loaded brain image or a stored brain image. ... [T]wo or more partial brain images may be combined to form full brain images.” TESLA-1005, [0188].

- “Users may post brain images for sharing with other users.” TESLA-1005, [0223].

- “The user may also augment the old brain image with new functionality from a partial new brain. Alternatively, the user may select multiple ones from different new functionalities of partial brain images to create a custom hybrid brain.” TESLA-1005, [0225].

12. Based on my review, Patent Owner and Dr. Saber do not dispute that Buibas discloses the underlying concept of a mixed knowledgebase that integrates instruction sets learned at multiple devices into a single knowledgebase (e.g., a brain image). Patent Owner and Dr. Saber on surrounding details of the language recited in claims 56 and 57, particularly the language directed to the *same* “one or more inputs” being *correlated* with *multiple* instruction sets: (i) instructions sets for operating a second device and (ii) additional instruction sets for operating a third

device. Notably, however, Patent Owner has identified no support in the intrinsic record of the '474 Patent that attributes any criticality or significance to this feature. Based on my review of the specification and file history, I likewise identified no such criticality or significance. Apart from the benefits of a mixed knowledgebase generally (which Buibas discloses), the '474 Patent identifies nothing unpredictable or uniquely advantageous about a *single* set of inputs of the knowledgebase being correlated with *multiple* instruction sets. On the contrary, the '474 Patent concedes that the particular model or structure of the knowledgebase employed is *independent* of the disclosed functions, and hence, any suitable model can be used to carry out the functions of the '474 Patent. TESLA-1001, 103: 44-55 ("The disclosed systems, devices, and methods are *independent* of the artificial intelligence model and/or technique used and *any model and/or technique can be utilized* to facilitate the functionalities described herein.").

13. Based on my review, the '474 Patent treats the structure of the knowledgebase as an obvious matter of design choice. A POSITA likewise would have considered the requirement for the same one or more inputs to have at least two correlations to be an obvious matter of design choice. A POSITA would have recognized this feature as reflecting just one of a number of predictable implementations of a mixed knowledgebase. A mixed knowledgebase, for example, could correlate different sets of inputs to the instruction sets for operating different

devices, or the inputs could be correlated to certain intermediate structures that effectively decouple the inputs from the instruction sets outputted by the knowledgebase, in some designs.¹ A POSITA would have chosen a suitable model depending on the needs of a particular application.

2. Buibas' Merge Operations Combine Fully-Defined Brain Images, Which Involves Merging More Than Weights of Brain Images

14. My Prior Declaration cited Buibas' "techniques for 'merging' brain images to combine behavior modules learned on two or more robots (a *second device* and a *third device*)" in support of showing that the combination in Ground 1A renders obvious "one or more inputs" of a knowledgebase further correlated with an additional one or more instruction sets for operating a third device. TESLA-1003, ¶¶99. Patent Owner and Dr. Saber argue that this mapping is deficient because "a network merge" is allegedly "a phrase Buibas uses for a merge of weight vectors of the neural networks, not for merging of neural networks." POR, 14-15; *see also id.*, 14 ("[M]erging of weight vectors results in another vector of weights, not instruction

¹ Not all intermediate structures will decouple the inputs to instruction sets at the output of a knowledgebase. A neural network, for example, may have inner layers of neurons between the input and output layers that correlate the inputs to the outputs.

sets. ... [I]t is important to realize that *a weight vector is not a neural network* since a weight vector does not include any neurons and instruction sets for operating a device.”) (emphasis in original). I disagree.

15. Patent Owner and Dr. Saber focus narrowly on Buibas’ description of FIG. 4, which explains that “*In some implementations*, the network lifecycle *may* comprise network *weight updates* comprising two or more sources, also referred herein to as ‘merges.’” TESLA-1005, [0094], FIG. 4; *generally id.*, [0090]-[0115], FIGS. 3-5. While it is true that the discussion of FIG. 4 centers on merging network weights, Buibas never suggests that a network merge operates *only* to combine the *weights* of the parent networks as opposed to the networks themselves. A POSITA would not have read the disclosure of FIG. 4 in isolation without considering Buibas’ disclosure in its entirety, which clarifies that not only weights, but “*brain images*” that fully define the networks for each task can be the subject of a network merge, thereby forming an offspring that combines the capabilities of two or more parent brain images as a whole. *See, e.g.*, TESLA-1005, [0225] (“The user may also augment the old brain image with new functionality from a partial new brain. ... *Users may also merge brain full images The merging of brain images involves the combination of one brain image with another brain image.*”), [0188] (“[T]wo or more *partial brain images may be combined to form full brain images.*”), [0189] “combining brain images”; “form full brain images from partial brain images”),

[0070] (“*the artificial mind and its traits* may be copied, stored, and later retrieved”). Buibas specifically explains that a brain image (or a partial brain image) fully defines a neural network *including not just weights but also the neurons, topology, and pathways of the network*. See, e.g., [0203] (“[T]he brain image serves as *a spatial mapping (in a simulated space) of the neurons and synapses* of an artificial neural network. The brain image is expressed as a sparse matrix (allows for empty space, growth, and connectivity mapping).”), [0205] (“The *brain image* is modified by training. Thus, the robotic learning process is achieved by adding or removing neurons, and/or generating new links between neurons.”).

16. Accounting for Buibas’ identification of “brain images” as the subject of a merge, and that brain images are comprised not only of weights but also of neurons and other information that fully define a trained neural network, a POSITA would have understood, or at least found obvious, that a network merge in Buibas would involve a merge of the neural networks themselves—including but not limited to the weights of the networks. This understanding is also consistent with Buibas’ description of the merge as a “network” merge and with Buibas’ explanation that a merged brain image incorporates the full “capabilities” of the parent images in the offspring. TESLA-1005, [0096] (“*combine capabilities* of two (or more) parents”), [0103] (“off-spring exhibiting *combined set of capabilities*”), [0092] (“may be used, for example, when updating one or more modules of a partitioned network, (e.g., a

motor-control partition, or a signal processing partition)”). Buibas further explains that a merge can be an “interleaved” merge that splices portions of the parent networks into a single network or a “non-interleaved merge” that effectively allocates each parent network to a separate partition in the offspring brain image. TESLA-1005, [0096]-[0103], FIGS. 4-5.

17. Regardless of how the merge is implemented in practice, a POSITA would have understood from Buibas’ explanation that the offspring retains the full capabilities of the parent networks that the offspring would inherit sufficient aspects of the parents that allow it to perform the same functions and thus produce equivalent outputs for a given input as the parents (e.g., output instructions sets for operating a second device and additional instructions sets for operating a third device, where those outputs are part of the respective capabilities of the parents). The inherited aspects of the parent networks would include weights, but would also include additional aspects such as additional neurons, activation functions, and the like, to the extent necessary for the offspring to implement the combined capabilities of the parents as described in Buibas.

18. This understanding of Buibas’ merge operation is further corroborated by a related patent application, filed by a co-inventor of Buibas, that includes substantially similar descriptions of network merges like those disclosed in Buibas, but which further explains how replica copies of robotic brains include “*full copies*

*that take the entire designs of one(or more) ‘parent’ robots; or partial copies, where portions of parent’s networks **may be merged to produce the offspring**.”* TESLA-1031, 8:33-45; *see also id.*, 11:34-44, 12:41-15:6, FIGS. 3-5. A POSITA would have been aware of this additional, related prior art disclosure, and would have recognized that its reference to replicating the “entire design” of parent brains images is consistent with Buibas’ explanation that the subject of a merge operation is the brain image itself, and all it defines, and the merge is not restricted only to weight vectors as alleged in the POR.

3. Buibas’ Merged Brain Images Correlate the Same One or More Inputs with Instruction Sets for Operating a Second Device and Additional Instruction Sets for Operating a Third Device

19. Based on my review, Patent Owner and Dr. Saber next err in asserting that a merged brain image must be be partitioned and that a partitioned brain image would consist of distinct networks having “separate” sets of inputs that would not meet the requirement in claims 56 and 57 for the *same* inputs to be correlated with instruction sets for operating a second device *and further* with additional instruction sets for operating a third device. POR, 20 (“partitions themselves are separate neural networks trained for different tasks[] that each includes is own separate inputs”).

20. First, the notion that Buibas’ network always results in a partitioned brain image comprised of multiple distinct networks is incorrect. While Buibas does describe this option, Buibas also describes techniques for “interleaved” merges that

interleave aspects of the parent networks to produce an offspring comprised of a single network that provides a combined set of capabilities from the parents. *See, e.g.,* TESLA-1005, ¶¶[0102]-[0104]. For example, Buibas explains that “two or more trash collector robotic apparatus, employed in a home or an office building, and comprising similar hardware (e.g., sensors and actuators) but trained under different conditions, and hence, comprising different sets of network weights, may utilize interleaved merges to produce an off-spring exhibiting combined set of capabilities.” TESLA-1005, [0103]. Here, because the offspring exhibits the “combined set of capabilities” of the parents, a POSITA would have understood the offspring to inherit the capabilities of each of the parents and to produce outputs learned by different ones of the parents, including instruction sets for operating a first parent (e.g., a second device) and additional instruction sets for operating a second parent (e.g., a third device), depending on whether the current input to the offspring corresponds to a circumstance learned by the first or second parent. And because the interleaved merge results in a single offspring network, that network would have *a single set of inputs correlated with each of its outputs* including the instruction sets for operating the first parent (e.g., a second device) and the additional instruction sets for operating the second parent (e.g., a third device). The POR fails to properly analyze this embodiment in Buibas.

21. Patent Owner’s and Dr. Saber’s further assumption that the inputs to a

partitioned brain image must be considered “separate[ly]” from each other is also illogical and inconsistent with Buibas’ express disclosure. Buibas depicts a partitioned neuromorphic apparatus in FIG. 11D, for example, that includes just a single input interface 1160 and a single output interface 1170. The input interface 1160 is correlated by way of downstream connections to the respective outputs of each partition, which are all further communicated to the output interface 1170. The outputs of each partition would include instruction sets for operating a first parent (e.g., a second device) and additional instruction sets for operating a second parent (e.g., a third device) in the scenario of a non-interleaved merge, and in this case, the input interface 1160 is naturally correlated with the instruction sets for operating both devices consistent with the features recited in claims 56 and 57:

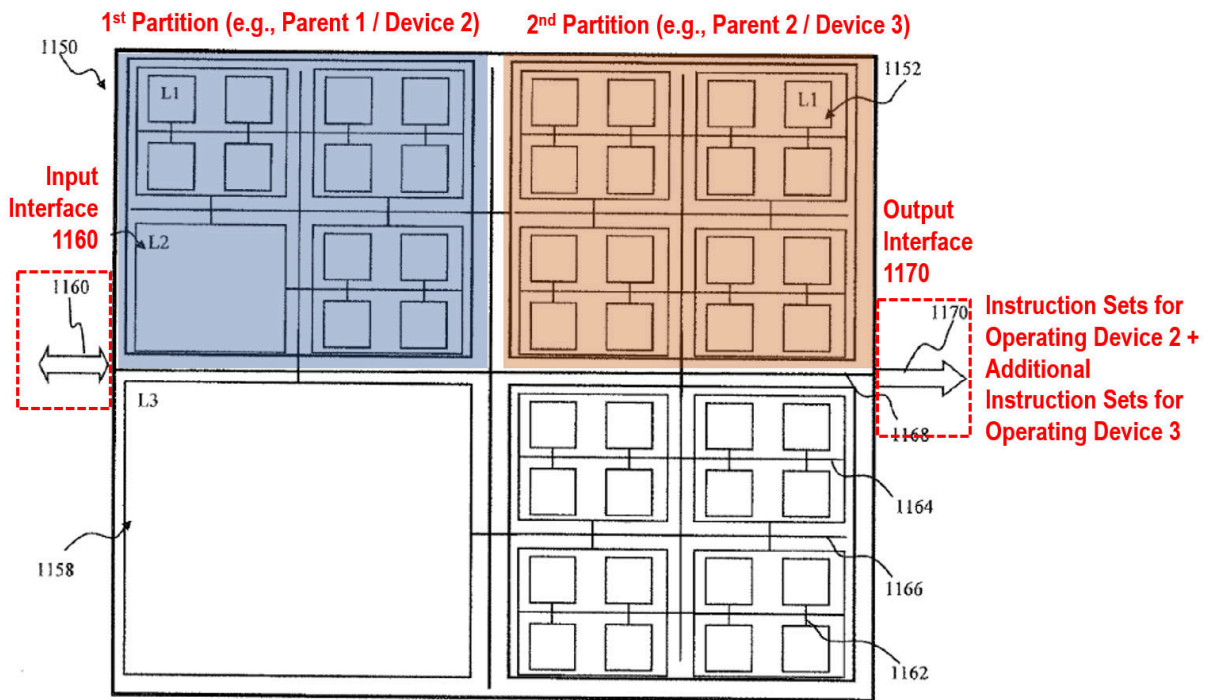


FIG. 11D

TESLA-1005, FIG. 11D (annotated); *see also id.*, [0158] (“The neuromorphic apparatus 1150 may receive *input ... via the interface 1160.*”), [0159] (“The neuromorphic apparatus 1150 may provide *output* (e.g., ... *a motor command*, e.g., to zoom/pan the image array) *via the interface 1170.*”), [0157] (“*Different L2, L3, cells may also perform different aspects of operating*, for example, a robot, with one or more L2/L3 cells processing visual data from a camera, and *other L2/L3 cells operating motor control block* for implementing lens motion what tracking an object or performing lens stabilization functions.”), [0150]-[0161], FIGS. 11A-11C.

22. Buibas further discloses that “two or more successive merges of the same or different types may be used to create new brain images (e.g., the same

‘father’ may be used in two successive merges once with the ‘mother’ and then again with the offspring).” TESLA-1005, [0225]. A single partition in a brain image after successive merges can thus itself include (sub-)partitions from a prior merge, and by the same logic described above, the input(s) to that partition would be correlated with the outputs of each (sub-)partition such as instructions sets for operating a second device and additional instruction sets for operating a third device.

23. Based on my review, the assumption of Patent Owner and Dr. Saber that the inputs to each partition must be considered “separately” from each other is also unsupported by any intrinsic evidence from the claims or specification of the ’474 Patent, let alone Buibas. The claims merely refer to “one or more inputs” of a knowledgebase (*see* [45.1]), and it logically follows that, if the knowledgebase includes multiple partitions, the “one or more inputs” to the knowledgebase would include the inputs to all the partitions in the knowledgebase. Collectively, those inputs are then correlated by downstream connections to the outputs of each partition, including instruction sets for operating a second device and additional instruction sets for operating a third device. Nothing in the claims, specification, or file history of the ’474 Patent indicates that the inputs to different partitions in a knowledgebase must or even would be considered separately and could not be grouped together as “one or more inputs” of the knowledgebase as a whole. Patent Owner identifies nothing in the specification that supports its overly restrictive

identification of inputs “correlated” with instruction sets (this language only appears in claims filed long after the earliest effective filing date). The alleged support for these features identified in Patent Owner’s Motion to Amend is generic disclosure of a feedforward neural network in Model A of FIG. 17—i.e., a model that is described in far less detail than those in Buibas. MTA, 9-11; EX2001, 79-80, 113, 127, 139, 141, 154, 164, 192, 203.

4. Buibas’ Merge Operations Would Result in Operable Networks that Produce Instruction Sets for Operating a Second Device and Additional Instruction Sets for Operating a Third Device

24. Patent Owner argues that Buibas’ “merge” operation would preclude instruction sets from multiple devices from being combined and correlated with the same inputs without re-training an existing network. POR, 20. I do not agree.

25. To start, I note that this argument does not apply to non-interleaved merges in which the system would produce a *partitioned* brain image comprised of individual parent networks integrated into *separate portions* of the offspring image. TESLA-1005, [0092]-[0098], [0225], FIG. 4; *above*, §II.A.3. For a partitioned offspring, no changes to the weights, neurons, or topology are necessary to carry each parent network into the offspring, as each partition simply corresponds to a parent network, and each partition provides a substantially identical structure and capability to its parent. Further, the offspring would provide one or more inputs

correlated with the outputted instructions sets of each partition for the reasons addressed above. *Above*, §II.A.3.

26. Even for interleaved merges that combine the parents into a single (non-partitioned) network, Buibas explains that the offspring is still operable to provide a combined set of capabilities of the parents without re-training. *See* TESLA-1005, [0103] (“utilize ***interleaved merge*** to produce an off-spring exhibiting ***combined set of capabilities***”); [0099]-[0115], [0225], FIG. 5. Patent Owner’s suggestion that re-training would be necessary add instruction sets is belied by Buibas’ detailed explanation of the interleaved network merge that makes it possible to combine the capabilities of the parents even without further training. *Id.* A POSITA would have no reason to doubt Buibas’ representations in this regard. Patent Owner’s conclusory review of conventional “backpropagation process[es]” for re-calculating/adjusting weights in a neural network fails to show that such processes are always necessary, nor does Patent Owner address why the specific neural network architectures disclosed in Buibas (e.g., artificial spiking neural networks) would require re-training to add instruction sets despite Buibas’ teachings to the contrary. TESLA-1005, [0073]-[0078]; POR, 20.

27. In any event, Patent Owner is wrong that Buibas does not disclose the option for re-training brain images. A POSITA would have known how to perform re-training if desired, and Buibas explains how this could be done (e.g., to further

improve performance of the offspring). TESLA-1005, [0085] (“a trained network or collection of networks ... may be *updated*”), [0205] (“The brain image is *modified by training*”; “the weights of various modeled synapses may be *altered* through the training process”), [0211] (“The user may create a new routine, or *continue training a previous routine*”), [0220] (“The robotic device may have its brain image replaced or *rewritten by the user at various times.*”), [0124] (“the transformation may comprise *self-transformation (with or without mutation)*”); [0098]-[0099] (“Network Evolution”), FIG. 9.

5. Merging Weight Vectors by Itself Would Still Produce a Brain Image Having One or More Inputs Correlated with the Claimed Instruction Sets for Operating Second and Third Devices

28. Even if one accepts for the sake of argument that Buibas’ merge operation is limited to merging only network weights as Patent Owner argues, Buibas would still render claims 56 and 57 obvious. Patent Owner’s arguments to the contrary are premised on a false dichotomy it creates between “weights” of the neural network and “instruction sets.” *See, e.g.*, POR, 14 (“Thus, merging of weight vectors results in another vector of weights, not instruction sets.”). A POSITA would not have conflated weights for instruction sets, but the framework of Patent Owner’s argument does not reflect how a POSITA would have analyzed the issue or Buibas’ disclosure. Although weights and instruction sets are not the same, they are nonetheless interrelated, particularly in the context of Buibas’ neural networks that

are configured to output instruction sets (e.g., motor commands) for operating a device. TESLA-1005, [0153], [0159]. Weights define the strength of connections between successive neurons, while instruction sets are produced at the output. Weights and instruction sets are thus interrelated insofar as the output (instruction sets) produced for a given input depends on the weights of connections between the input and the output. TESLA-1005, [0062], [0073]-[0078], FIGS. 1, 10. A POSITA therefore would have understood that Buibas does not specifically discuss how instruction sets are combined in a merge operation because the merging of weights itself has the effect of configuring connections between the input and output neurons that allow the offspring to produce instructions sets corresponding to the *combined* capabilities of the parents. TESLA-1005, [0103].

29. Moreover, merely because a “merge” at its lowest/most granular level is carried out by combining weight vectors does not mean that the merge does not also provide a combined set of instruction sets at the output of the merged network. All computer functions involve low-level operations that can be sterilely characterized in terms of manipulations to bits, bytes, weights, or the like, but the higher-level effect is typically much more meaningful, as it is here. In particular, as Buibas explains, the effect of a merge operation is to provide a “combined set of capabilities” in the offspring. TESLA-1005, [0092], [0103], [0225]. Accordingly, a POSITA would have understood and found obvious from Buibas that merging

weight vectors serves to combine the capabilities of the parents by weighting connections between the input and output neurons such that the output neurons would be configured to produce outputs corresponding to the capabilities of the parents, e.g., instructions sets for operating a second device and additional instruction sets for operating a third device. Additionally, the Patent Owner appears to have interpreted my testimony related to this area as limiting the merging of neural networks to only weight vectors, but this is an inaccurate representation of my opinion. EX2003, 77:4-82. Although I indicated that weight vectors would be shared in a network merge, I also clearly indicated that network merges can involve far more than simply sharing weights (consistent with Buibas' explicit disclosure – e.g., “neuron and synaptic weights, biases, activation functions, topology”). *Id.* (Q: “So you could extract a portion of a neural network and fold that into another neural network ...” A: “... you can pull a weight vector, but that weight vector is not going to tell you much about that neural network. It's just going to be a block of values.”; Q: “Okay. And so when you pull those weights out, you describe them as a means of correlating input values to output values or instruction sets. So the weights are used in that process; is that correct?” A: “If I pull weights only, it's a mystery what that would do. You can't just pull the weights.”).

B. Buibas Renders Obvious “At Least a Portion of the One or More Instruction Sets for Operating the Third Device ... is Learned in

Another Learning Process that Includes Operating the Third Device at Least Partially by Another User” (Claim 56)

30. Patent Owner’s further objections to Buibas’ disclosure of “at least a portion of the another one or more instruction sets for operating the third device or at least an information related to the another one or more instruction sets for operating the third device is learned in *another learning process* that includes *operating the third device at least partially by another user*” as recited in claim 56 lack merit. *See* POR, 27-32. Patent Owner’s arguments on this limitation repeat many of the same points that I addressed above, and they are deficient for similar reasons.

31. Moreover, my Prior Declaration cited extensive disclosure from Buibas that plainly show that, in a merged brain image, the instruction sets produced by that brain image (or information related thereto) can include both (i) instruction sets for operating a second device learned in a learning process that includes operating the second device by a first user and (2) instruction sets for operating a third device learned in another (second) learning process that includes operating the third device by another (second) user. TESLA-1003, 89-92. Indeed, Buibas specifically teaches that different users can independently train their respective robots to autonomously perform different tasks/behaviors. Brain images configured to perform the various tasks/behaviors learned on different devices by training routines carried out by

different users operating their different robots can then be shared with other users, and users can then select and incorporate different ones of the learned tasks/behaviors into a merged brain image can be loaded on a single device. *See, e.g.,* TESLA-1005, [0171]-[0180] (describing details of the training application provided to each user), [0181]-[0193] (describing details of sharing trained brain images through the cloud), [0194]-[0207], [0211]-[0213], [0222]-[0225], [0070], FIGS. 16-19.

32. Based on the teachings of Buibas, a POSITA would have recognized and found obvious that when the parents to a merged brain image are networks that have been trained by different users on different devices (i.e., a second device and a third device), the instruction sets produced by the merged brain image would include instruction sets for operating a second device learned in a learning process that includes operating the second device by a first user and additional instruction sets for operating a third device learned in another learning process that includes operating the third device by another user. Considering that the instruction sets for operating the second device and the instruction sets for operating the third device can pertain to different tasks/behaviors, and that different users can independently train the different tasks/behaviors on different devices, a POSITA would have appreciated that the training routines conducted on the two different devices constitute distinct “learning processes.”

33. Despite its critiques of Buibas, Patent Owner does not appear to identify any specific factor or criterion that would need to be met to demonstrate disclosure of separate learning processes. *See* POR, 27-32. Based on my review, the term “learning process” is never defined in the ’474 Patent and is instead used in just one paragraph of the ’474 Patent’s lengthy disclosure. TESLA-1001, 103:10-19. In a related IPR involving similar subject matter and claims that likewise referred to multiple learning processes, Patent Owner’s declarant, Dr. Saber, acknowledged that different learning processes may involve different users, different devices, and operating device(s) at different times. TESLA-1032, 38:18-39:2, 49:19-51:12, 52:11-55:21. Here, all of those factors are present in Buibas, which allows users to independently carry out different training routines on different devices at different times. A POSITA would have understood from these considerations that the instruction sets learned on a parent third device in Buibas are learned in another learning process separate from the learning process on the second device by different users.

III. GROUND 2C: OBVIOUSNESS BASED ON GROTMOL-BUIBAS

A. Grotmol and Buibas Render Obvious “The One or More Inputs are Further Correlated with Another One or More Instruction Sets for Operating a Third Device” (Claims 56, 57)

34. As explained in my Prior Declaration, claims 56 and 57 are further rendered obvious by the Grotmol-Buibas combination. TESLA-1003, ¶¶150-161.

My Prior Declaration explained that “[t]o the extent it is argued or determined that Grotmol does not expressly disclose that the inputs to an adaptive controller on a first device are correlated with control outputs (*instruction sets*) learned on at least two other devices (i.e., a *second device* and a *third device*), these features nonetheless would have been obvious when further considering the teachings of Buibas.” TESLA-1003, ¶160 (emphasis in original). Patent Owner’s arguments in the POR are based on the same flawed logic and misunderstanding of Buibas’ teachings that it earlier advanced against Buibas in Ground 1A (*see* POR, 37-43), but these arguments lack merit and fail to refute the showing in my Prior Declaration that these features were disclosed and would have been obvious to a POSITA based on Buibas for the same reasons addressed above in §II.A.

35. Patent Owner’s additional arguments against Grotmol are likewise deficient. The POR contends, for instance, that in Grotmol, “all of the instruction sets for operating the first robot where they were learned are still the same instruction sets for operating the same first robot (not a third robot) when they are loaded into a second robot,” and “nothing in Grotmol even remotely suggests that the same inputs are further correlated with another instruction for operating at third device.” POR, 39. I do not agree. Patent Owner’s arguments overlook pertinent teachings in Grotmol that underscore how the claimed features would have been obvious to a POSITA.

36. For example, my Prior Declaration cited Grotmol's teaching that "[t]he trained controller of the robot may comprise a trained configuration configured to enable autonomous operation" and "the trained configuration may be loaded to one or more other robots in order to provide learned behaviors." TESLA-1008, 33:21-35; TESLA-1003, ¶159. Grotmol discloses that the trained configuration of a controller can be updated through continuous learning, including learning in the "override" and "correct" modes. TESLA-1008, 34:59-62 ("training may comprise periods of autonomous operation and periods of learning"), 33:49-52 ("user commands ... may arrive in one or more clusters ... that may be interleaved by periods of user inactivity"), 23:9-45 ("In the override mode, the robot may execute what it has learned, unless a command is sent by the teacher. ... [T]he robot may head [sic] the teacher's command, perform the action, and try to remember it for the next time the same situation occurs. ... In the correct mode, the robot may integrate what the teacher commands with what the robot already knows."), 21:56-60 & Table 3 ("Learns the association").

37. A POSITA would have understood from these disclosures, for instance, that Grotmol's controller can first be trained on a third device to learn instruction sets for operating a third device; the trained configuration of the controller can be shared with and loaded on a second device where the trained configuration is updated through additional learning of instruction sets for operating the second device (e.g.,

in the override/correct modes); and the updated trained configuration can be shared with and loaded on a first device. This predictable situation would have been expected and obvious to a POSITA reviewing Grotmol's disclosure, as it would allow multiple devices to gain the benefit of Grotmol's capabilities for sharing and updating training configurations. The predictable result would be a trained controller on the first device whose inputs are correlated both with instruction sets for operating the second device (based on training that occurred on the second device) and with instruction sets for operating the third device (based on training that occurred on the third device).

B. Grotmol and Buibas Render Obvious “At Least a Portion of the One or More Instruction Sets for Operating the Third Device ... is Learned in Another Learning Process that Includes Operating the Third Device at Least Partially by Another User” (Claim 56)

38. Patent Owner's arguments on this limitation with respect to the Grotmol-Buibas combination are deficient for the reasons I addressed above with respect to the Buibas-Sinyavskiy combination.

IV. ASSESSMENT OF THE SUBSTITUTE CLAIMS

39. I understand that the Patent Owner has also submitted a Motion to Amend (MTA) in which (1) substitute claims 85-91 contingently amend claims 45-47, 51, 56-57, and 60, respectively, if certain claims challenged in the Petition are found unpatentable and (2) substitute claims 92-93 contingently replace claims 56-

57 if claims 56-57 are not found unpatentable.

V. THE SUBSTITUTE CLAIMS ARE UNPATENTABLE UNDER § 103

40. Based on my review, the substitute claims presented in Patent Owner's Motion to Amend would have been obvious to a POSITA before the earliest effective filing date of the '474 Patent based on the following grounds:

- **Ground MTA-1A** – Obvious based on Grotmol in view of Zhu (Claims 85-91)
- **Ground MTA-1B** – Obvious based on Grotmol in view of Zhu and Buibas (Claims 85-91)

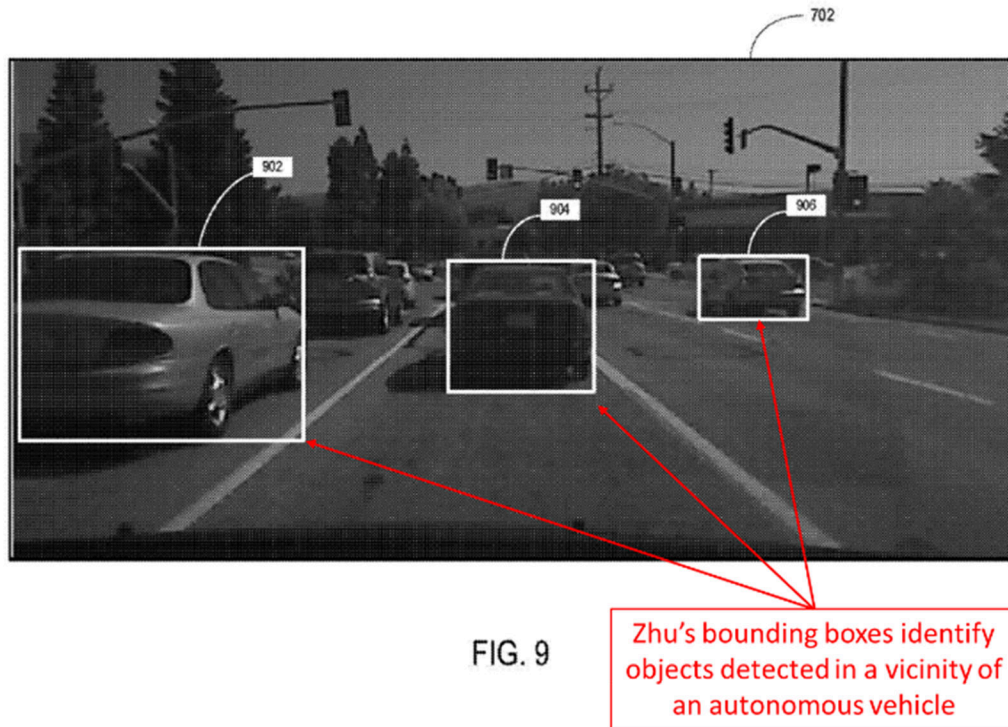
A. Overview of Additional Prior Art²

1. Overview of Zhu

41. Zhu describes techniques for, among other things, capturing images of an environment surrounding an autonomous vehicle and detecting objects including other vehicles depicted in the images. TESLA-1023, 1:48-2:4, 10:58-67, 11:16-55. Zhu discloses that autonomous vehicles include persistent data with sensor parameters that can be used to determine “when a sensor has detected an object, and the type of object detected by the sensor” (“how an object is classified”). *Id.* The

² My Prior Declaration previously provided overviews of Grotmol and Buibas. TESLA-1003, pp. 37-35.

sensor parameters are applied to the detected object with a “label” with a “bounding box” of a defined size. TESLA-1023, 14:43-17:28, FIGS. 9-10.



TESLA-1023, FIG. 9.

B. Overview of the Prior Art Combinations

1. Ground MTA-1A: Grotmol-Zhu

42. As discussed above, Zhu describes techniques for labeling objects detected by an autonomous vehicle with semantic information, including labeling objects to indicate their classification/type, size, and location with a bounding box. *Above* §V.A.1. Additionally, Zhu discloses that autonomous vehicles include sensors that detect information related to the autonomous vehicles' circumstance, including changes in that circumstance over time where the vehicle is required to

respond. TESLA-1024, 8:25-46, 10:20-28. Zhu explicitly discloses that sensors provide continuous output (e.g., as an output stream). *Id.*

43. Grotmol likewise describes techniques for training and operating autonomous vehicles (e.g., mobile robots) by acquiring images and other sensory information about the vehicles' environment, detecting objects and extracting features from the sensory information, and processing the features of the sensory information to determine control outputs/instructions sets (e.g., motor commands) for autonomously operating the vehicle. TESLA-1008, Abstract, 23:3-22, 24:1-16, 28:19-40, 33:12-35, 39:13-28, FIGS. 10A-11B; Pet., 51-54.

44. A POSITA reviewing the teachings of Grotmol and Zhu would have found it obvious to apply Zhu's suggestion for extracting additional features of detected objects—such as a classification/type of the object, and the object size and location (e.g., using a bounding box). In the combination, Grotmol's autonomous vehicle would then predictably use the additional features of the extracted objects suggested by Zhu to learn/characterize images and circumstances encountered by the vehicle during training, and to compare current circumstances obtained during autonomous operation to previously learned circumstances to determine instruction sets for operating the device consistent with Grotmol's existing methodologies for training and autonomously operating a device based on a wide range of features of detected objects and sensory inputs. TESLA-1008, 21:64-22:18, 25:30-42, 26:5-20,

27:24-40, FIGS. 10A-11B, 15; TESLA-1023, 14:43-17:28, FIGS. 9-10.

45. A POSITA would have been motivated to combine Grotmol and Zhu for multiple reasons. To start, a POSITA would have sought to classify and label the type of objects detected in the vicinity of Grotmol's autonomous vehicle based on Zhu's teachings to provide additional semantic information to Grotmol's controller that would allow it to more accurately and efficiently characterize circumstances encountered by the vehicle and to compare current circumstances to prior circumstances learned in the past (e.g., as represented in a trained neural network). TESLA-1023, 11:16-55. For example, semantic labels indicating the type of detected objects would allow the robot to readily characterize a current scene at a higher level of abstraction than the raw pixels of an image, and the semantic labels could be processed more efficiently by a neural network than the raw pixels of an image to determine the similarity of the current scene to prior learned circumstances (e.g., determine whether a tree, a ball, and an object are present in both the current scene and a prior circumstance). A POSITA likewise would have sought to apply Zhu's suggestion for labeling the coordinates and size of detected objects (e.g., using a bounding box) to Grotmol's system to more accurately and efficiently characterize circumstances encountered by the vehicle and to compare current circumstances to prior circumstances learned in the past. Knowledge of the size and location of detected objects would predictably allow Grotmol's vehicle to compare current and

past circumstances based on such information consistent with Grotmol's use of other object features that do not require comparison of raw pixels or other raw sensor information in the adaptive controller. TESLA-1008, 39:13-28, 28:19-40, FIGS. 10A-11B; TESLA-1023, 14:43-15:20, 15:36-16:4.

46. A POSITA also would have been motivated to combine Grotmol and Zhu to gain the benefit of Zhu's cross-validation techniques that allow the accuracy of different sensors on the vehicle to be evaluated. TESLA-1008, 38:43-39:12, FIGS. 20A-B; TESLA-1023, 10:29-57, 11:56-62, 14:43-15:65, 22:64-23:10.

47. Zhu also provides implementation details for sensors present on an autonomous vehicle and situations where this sensory input can be used to control an autonomous vehicle. TESLA-1024, 8:25-46, 10:20-28. Specifically, Zhu illustrates that a vehicle's sensors can be used to continuously capture data, such that evolving changes in the vehicle's circumstance can be detected expediently. *Id.*

48. Additionally, the combination is (1) merely the application of known techniques as suggested by Zhu to a known system (e.g., Grotmol's robotic vehicles) to yield predictable results (e.g., Grotmol's robots labeling objects according to Zhu), and (2) the use of known techniques as taught in Zhu to improve a similar system (e.g., an autonomous robotic device).

49. A POSITA would have reasonably expected success implementing the combination considering that Grotmol and Zhu both relate to processing sensory

information on autonomous vehicles, and Grotmol likewise describes techniques for extracting features of circumstances and objects detected in the vehicle's environment. TESLA-1008, 18:29-37, 33:49-55; TESLA-1023, 14:43-17:28, FIGS. 9-10. Further, Zhu's autonomous vehicles include sensors explicitly disclosed by Grotmol for providing input to a neural network (e.g., cameras, SONAR, LiDAR). TESLA-1008, 8:57-9:40, 10:16-22, 38:33-51; TESLA-1023, 1:18-27, 8:25-46. Additionally, Grotmol also provides a suitable mechanism (e.g., an adaptive controller) for implementing Zhu's robotic behaviors (e.g., labeling objects). TESLA-1008, 21:64-22:18, 25:30-42, 26:5-20, 27:24-40, FIGS. 10A-11B, 15. Another prior art patent (Gabardos, TESLA-1030) to the same assignee (Brain Corp.) as Grotmol likewise confirms that the use of a bounding box to indicate size and location of objects in a system substantially similar to that described in Grotmol was known before earliest effective filing date of the '474 Patent. TESLA-1030, 30:18-33, 38:56-65, FIGS. 10A-11B.

2. Ground MTA-1B: Grotmol-Zhu-Buibas

50. It would have been obvious to further combine Grotmol-Zhu with Buibas for the reasons that I explained that it would be obvious to combine Grotmol with Buibas in my Prior Declaration, e.g. such that the adaptive controllers on Grotmol's device would be implemented using brain images like those described in Buibas, where the brain images could be independently trained, shared, and merged

by different users on different devices to flexibly provide different sets of capabilities and learned behaviors to Grotmol's device. TESLA-1003, ¶¶150-157. A POSITA would have been motivated to combine Grotmol-Zhu with Buibas, and would have reasonably expected success, for the reasons discussed in my Prior Declaration.

VI. OBVIOUSNESS OF THE SUBSTITUTE CLAIMS

1. Elements [85.pre]-[85.7]

(a) All Grounds

51. The language of [85.pre]-[85.7] is identical to [45.pre]-[45.7]. Grotmol discloses and/or renders these features obvious for at least the reasons addressed in Ground 2A of my Prior Declaration. TESLA-1003, ¶¶113-131. Substitute claim 85 further recites that the claimed “knowledgebase includes a neural network” and “the one or more inputs are one or more input neurons of the neural network.” MTA, 2. Grounds MTA-1A – MTA-1C each provide these additional limitations as discussed in further detail below, and the aspects of [85.pre]-[85.7] that refer to a “knowledgebase” and “one or more inputs” are rendered obvious by these grounds for the same reasons. TESLA-1003, ¶¶113-131.

2. Element [85.8]

(a) Grotmol-Zhu

52. Grotmol discloses that the mobile robot's (“*first device*”) adaptive

controller (“*knowledgebase*”) can include a neural network. TESLA-1008, 11:28-30 (“the predictor is a neural network”), 9:41-10-15, 16:8-12, 19:10-54 (“spiking neuron network”), 21:41-46 (“multi-layer perceptron and/or a convolutional network”), 21:64-68, 26:5-20 (“‘Adaptive controller’ of type ‘2’ (e.g., perceptron classifier)”), 34:3-17, 36:19-31. A POSITA would have known that neural networks conventionally included input neurons (as evidenced by Buibas), and it would have been obvious for a neural network in Grotmol’s adaptive controller to include input neurons that would allow it to receive information characterizing a sensory input from which a control output could be generated. TESLA-1005, ¶¶[0068], [0073]-[0076], [0136]-[0138], [0144]-[0154], FIGS. 1, 10, 11. For example, Fisher, incorporated by reference in Grotmol (TESLA-1008, 9:53-60) discloses a spiking neural network 1040 having input neurons 1042. TESLA-1029, [0227]-[0228], FIG. 10A. Any of a variety of neural network architectures would be obvious and suitable to apply in the combination (e.g., as discussed in Bichler and Gocek below). TESLA-1035, ¶¶[0025]-[0040], [0119]-[0122], FIG. 7; TESLA-1036, ¶¶[0044]-[0047], FIG. 5.

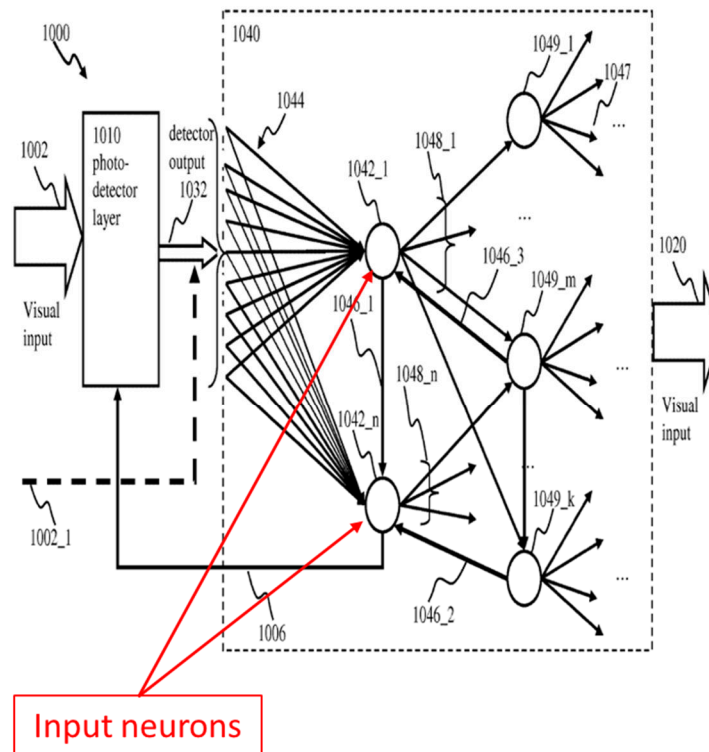


FIG. 10A

TESLA-1029, FIG. 10A.

53. As another example, Hagan (an academic textbook on neural network architecture) explicitly discloses that sensory input is provided to input neurons, which use the input to determine an output based on a transfer function (also referred to as an “activation function”). TESLA-1024, 36-43. Hagan also discloses that neurons are “activated” by particular input stimuli to provide an output. TESLA-1024, 197, 578, 581, 726; *cf.*; TESLA-1001, 94:1-3. Hagan thus exemplifies that neural networks like those disclosed in the ’474 Patent were well-known to a POSITA, and indeed, a POSITA would have been familiar with the same types of neural networks those disclosed in the ’474 Patent (e.g., feedforward, convolutional,

recurrent neural networks) before the earliest effective filing date. TESLA-1001, 103:40-55. Thus, regardless of which architecture of a neural network is deemed suitable for a particular application, the additional teachings of Hagan and the background knowledge of a POSITA confirm that [85.13]-[85.14] would have been obvious for substantially the same reasons addressed above in connection with the Grotmol-Zhu combination.

54. Finally, input neurons were commonplace in various other types of neural network architectures at the time of the Critical Date. For example, Bichler discloses a convolutional neural network where each convolutional layer includes a set of input neurons. TESLA-1035, ¶¶[0025]-[0040], [0119]-[0122], FIG. 7. Gocsek discloses that both feed-forward and recurrent neural networks included input neurons. TESLA-1036, ¶¶[0044]-[0047], FIG. 5.

(b) Grotmol-Zhu-Buibas

55. In addition to Grotmol's disclosures of neural networks, Buibas further describes a spiking neural network that would have been obvious to provide in Grotmol's adaptive controller in some implementations of the combination, which network includes input neurons for receiving information characterizing sensory input(s). TESLA-1005, ¶¶[0068], [0073]-[0076], [0136]-[0138], [0144]-[0154], FIGS. 1, 10, 11.

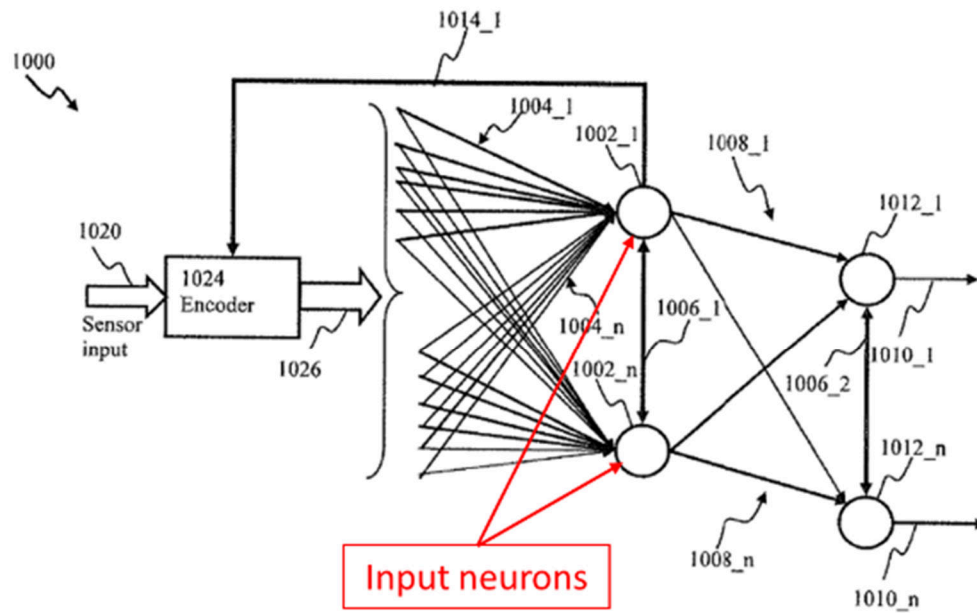


FIG. 10

TESLA-1005, FIG. 10.

3. Element [85.9]

(a) All Grounds

56. Grotmol discloses that its mobile robot includes machine-readable code that, when executed by one or more processors, causes the one or more processors to perform operations. TESLA-1008, 2:15-18, 3:27-36, 7:66-8:4, 38:33-42, FIG. 8; Pet., 62-63 (Element [45.4]).

4. Element [85.10]

(a) Grotmol-Zhu

57. Grotmol's mobile robot includes a sensor component that "provide[s] information associated with objects within the field-of-view" and that includes a "a

stream of pixel values associated with one or more digital images.” TESLA-1008, 8:57-9:3, 44:46-55 (describing a “camera stream”). Grotmol further discloses that the robot, during autonomous operation, generates a stream of observations of the environment (e.g., images 2000 and 2010) that represent circumstances encountered by the robot over time. TESLA-1008, 38:52-39:28, 38:38-41 (“sequence of images”), 42:46-43-11 (describing autonomous operation including capturing sensory inputs to be correlated with motor commands at each “given time step”), 44:56 (“At individual time steps[] ...”), 44:34-46:10, FIGS. 8, 20A-B; Pet., 63-64 at [45.5] (“generating a circumstance representation”). In the combination, Zhu likewise discloses generating a stream of circumstance representations representative of the circumstances of an autonomous vehicle. TESLA-1023, 10:20-28 (“[T]he sensors 146 may continuously update their output to reflect the environment being sensed at or over a range of time, and continuously or as demanded.”), 8:25-46, 13:46-58, FIGS. 1, 7-10.

(b) Grotmol-Zhu-Buibas

58. In addition to the disclosures from Grotmol and Zhu provided above, Buibas discloses generating a stream of sensory input including “a sequence of images” and “streams of input stimuli.” TESLA-1005, ¶¶[0136], [0227], [0236]. Like Grotmol and Zhu, Buibas describes a controller that “responds to input” detected from the environment (e.g., “increasing/decreasing synaptic weights ...

related to the behavior being executed at the time of input”). TESLA-1005, ¶¶[0172], [0228].

5. Element [85.11]

(c) All Grounds

59. As discussed above with respect to [85.10], Grotmol’s mobile robot generates a stream of circumstance representations that enables the robot to autonomously perform various tasks/behaviors. *Above*, [85.10]; Pet., 63-64 at [45.5]; *see also* TESLA-1023, 10:20-28, 8:25-46, 13:46-58, FIGS. 1, 7-10. For example, the robot may capture an image 2000 of its environment that includes representations of multiple objects (e.g., a tree, a box, and a ball) at a first time. TESLA-1008, 38:43-39:12 (“object representations”). According to Grotmol, the robot can detect objects from sensory signals such as image 2000 and can further extract features associated with the detected objects that enable comparison to stored features. TESLA-1008, 39:13-28 (“images (e.g., raw and/or pre-processed) ... may be stored in a memory buffer” and the “pre-processing operations may comprise ... feature extraction”), 28:19-40 (“the term feature may be used to describe ... values characterizing the input, e.g., the presence or absence of ... [an] **object**, at particular **locations** in the image” including “properties of the mentioned features such as **size**, ... **position**”), 21:41-43, 9:41-2 (“object recognition and/or classification”), 10:23-34 (“The tasks of the robot may be configured based on a context. ... [T]he context

may comprise ... state of the environment (e.g., *object size, location*)[] ...”).

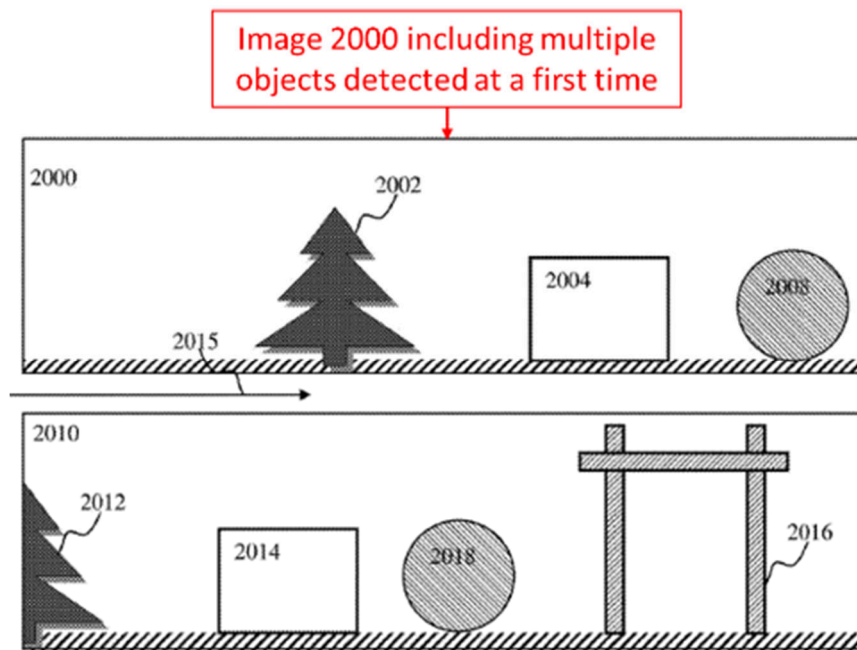
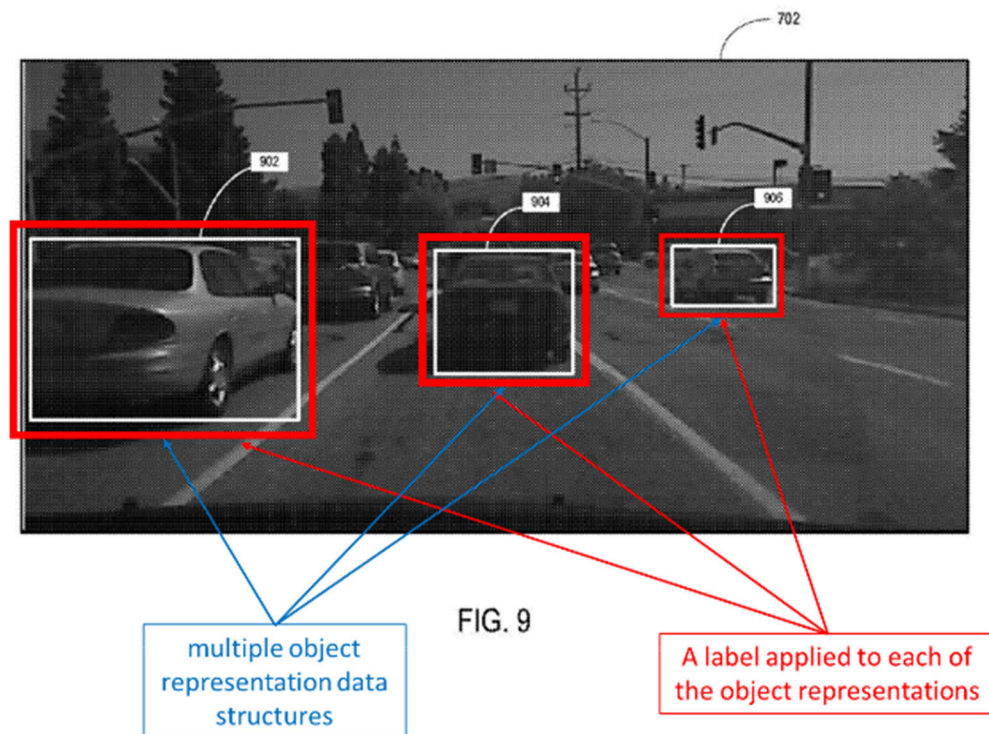


FIG. 20A

TESLA-1008, FIG. 20A.

60. Zhu further describes techniques for classifying objects detected in a vicinity of an autonomous vehicle at a given time and generating, for each object, a label for object type, coordinates for object location, and a computer model for object size. For example, 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.” TESLA-1024, 11:16-36, 14:43-52 (“may apply an object label to the detected object”), 10:58-64 (data 112 defining parameters for “the *type* of object detected by the sensor”). Zhu’s system is also configured to “define[] ... a bounding

box encompassing the object,” where “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).” TESLA-1024, 14:43-67; *see also id.*, 15:1-20, FIGS. 9-10. Zhu thus defines coordinates for the location of each object and a computer model (e.g., a bounding box and associated width/height parameters) for each object’s size.



TESLA-1024, FIG. 9.

61. Zhu teaches that the parameters associated with each object can be stored in respective object representation data structures such as the “label” or “bounding box” for each object. TESLA-1024, 14:54-15:20 (“An object label may have one or more object label parameters that define the shape of the object label.”). Zhu further explains that such data can be stored in any suitable “data structure” such as “a relational database,” “a table having a plurality of different fields and records,” or “XML documents.” *Id.*, 6:8-25; *cf.*, TESLA-1001, 65:63-66:14 (’474 Patent disclosing wide range of data structures).

62. Zhu also discloses that object bounding boxes are defined by “coordinates” that define the horizontal and vertical placement of the bounding box (“*coordinates for object location*” – shown above in FIG. 9). TESLA-1024, 14:43-15:20, FIGS. 9-10. A POSITA would have recognized and found obvious that these bounding boxes would have additionally been assigned a coordinate location in space, because as Zhu discloses, the location of the object is relative to the location of the autonomous vehicle, and as discussed in my Prior Declaration, coordinate systems were commonly used before the Critical Date to define such relationships. TESLA-1024, 7:14-19, 8:56-63.

63. It would have been obvious to a POSITA to apply Zhu’s suggestion for labeling the type of detected objects and defining bounding boxes for detected objects that provides a computer model for object size coordinates for object location

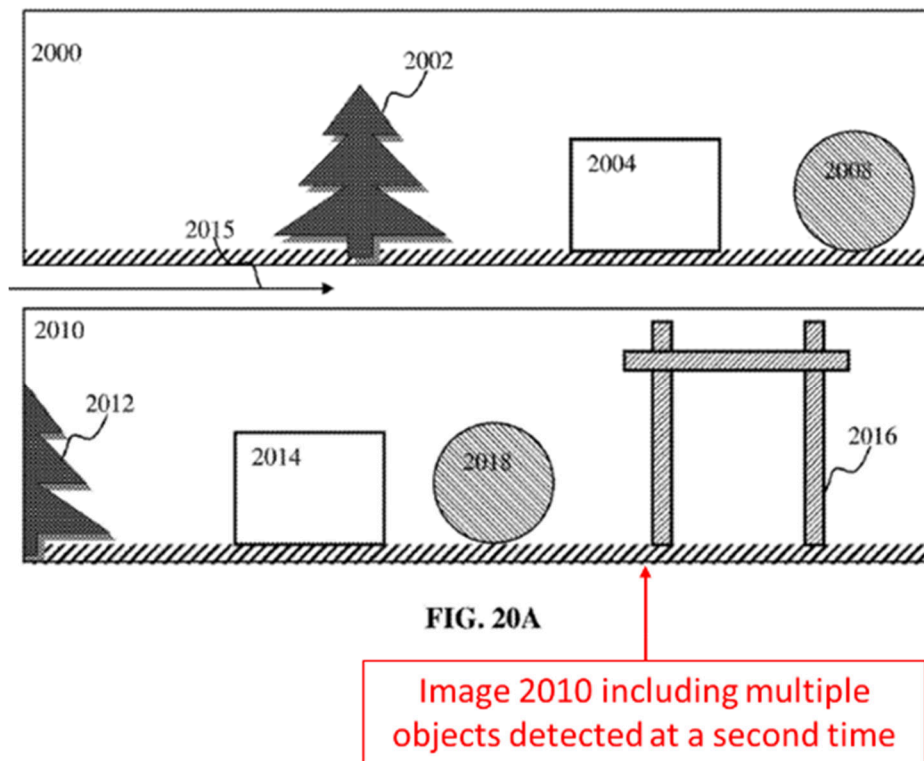
to Grotmol for the reasons described above in the combination overview. *Above*, §II.B.1; Pet., 45-51, 73-77. Zhu's bounding boxes would have beneficially allowed the combined system to accurately reflect not just the presence of an object, but also the object's geometry, as Zhu discloses that bounding boxes can be tailored to the object's shape (e.g., "a bounding oval or bounding shape"). TESLA-1024, 14:43-52. Grotmol and Zhu both indicate that multiple object types can be present in any given circumstance, and thus tailoring the object's spatial representation would have improved spatial awareness. TESLA-1008, 38:43-39:12; TESLA-1024, 11:16-36. Additionally, adding a semantic label to the bounding shape, as disclosed in Zhu in the form of a classification, would have allowed even further refinement of the object's bounding shape (e.g., based on an expected geometry associated with the object type). TESLA-1024, 11:16-36.

6. Element [85.12]

(a) All Grounds

64. As discussed above with respect to [85.10]-[85.11], the teachings of Grotmol and Zhu render obvious processors of a mobile robot (*first device*) that generate a stream of circumstance representations, each circumstance representation representing a circumstance detected at least in part by one or more sensors of the robot at different times. *Above*, [85.10]-[85.11]; *see also* Pet., 63-64 at [45.5]. For example, the robot may capture an image 2000 of its environment that includes

representations of multiple objects (e.g., a tree, a box, and a ball) at a first time and shortly thereafter capture an image 2010 that includes representations of the multiple objects (e.g., the tree, the box, and the ball) at a second time. TESLA-1008, 38:43-39:12 (“object representations”), FIG. 20A; TESLA-1023, 10:20-28 (“[T]he sensors 146 may continuously update their output to reflect the environment being sensed at or over a range of time, and continuously or as demanded.”), 8:25-46, 13:46-58, FIGS. 1, 7-10. In the combination, and for the reasons addressed in [85.11], the circumstance representation generated at the later (second) time would provide 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.



TESLA-1008, FIG. 20A.

7. Elements [85.13]-[85.14]

(a) Grotmol-Zhu

65. Grotmol discloses that the adaptive controllers determine control outputs (*instruction sets*) for operating a robot using features of the sensory input that are inputted into the controller. TESLA-1008, 29:37-39, 29:54-30:20, FIGS. 10A-11B. As explained in my Prior Declaration, Grotmol's system uses a correlation between the one or more inputs of the adaptive controller and instruction sets learned by the adaptive controller to determine instruction sets for operating the robot. TESLA-1008, 8:57-9:3, 44:46-55.

66. Moreover, as discussed above in connection with [85.8], Grotmol discloses that the adaptive controller can be implemented using a neural network having one or more input neurons. *Above*, [85.8]. In such embodiments, features of the sensory input are inputted into the input neurons and an output layer of neurons produces control outputs (*instructions sets*) for operating a robot. TESLA-1004, 21:42 (“[C]onfigured to detect targets and/or obstacles based on sensory input” in which “detection may be configured based on an operation of a multi-layer perceptron and/or a convolutional network”), 19:10-42 (control signals for a “spiking controller output” include positive/negative “spikes indicative of a control command”), FIGS. 10A-11B (depicting feature inputs and control commands produced by adaptive controllers); TESLA-1029 (incorporated by reference in Grotmol at 9:53-60), [0227] (“neural spiking network configured to encode the input 1032 into output 1020”), [0226]-[0234], [0252] (“output” includes “motor command”), [0257], FIG. 10A.

67. In the combination, it would have been obvious to a POSITA to input (i) the multiple object representation data structures for the objects detected at the first time (*above*, [85.11]) into the input neurons of the neural network for the benefit of obtaining a first control output conditioned on the first input and (ii) the another object representation data structures for the objects detected at the second time (*above*, [85.12]) into the input neurons of the neural network for the benefit of

obtaining a second control output conditioned on the second input. TESLA-1008, 25:30-47-42, 17:28-44, 28:41-44, FIGS. 10-11. Inputting such object representation data structures into the input neurons would have been obvious especially because it involves no more than using the controllers/neural networks in Grotmol according to their ordinary and intended purpose—namely, to process information characterizing sensory inputs including features of detected objects to produce controls for autonomously operating a robot. TESLA-1008, FIGS. 10A-11B. It is also consistent with the conventional use of input neurons in a neural network to receive inputs. *See, e.g.*, TESLA-1029, [0227]-[0234], FIG. 10A. A POSITA would have reasonably expected success inputting object representation data structures into the input neurons since neural networks were known to be capable of processing a wide range of data including features of detected objects as of the Critical Date. For example, Buibas corroborates that neurons can be used to detect “complex features and objects.” TESLA-1005, ¶¶[0140]-[0142], [0228]. Likewise, Yufik also discloses that neurons can receive and process semantic understanding of objects and circumstances. TESLA-1037, ¶¶[0004]-[0006], [0043]-[0048]. The ’474 Patent likewise acknowledges that a POSITA would have possessed the skill necessary to implement such features on a neural network. TESLA-1001, 103:40-60 (“The disclosed systems, devices, and methods are independent of the artificial intelligence model and/or technique used and any model and/or technique can be utilized to

facilitate the functionalities described herein.”), 104:2-27, 169:30-54, FIG. 17.

68. Grotmol-Zhu also renders obvious (i) using a correlation between the one or more input neurons and one or more instruction sets for operating a second device to determine the instruction sets for operating the second device and (ii) using another correlation between the one or more input neurons and another one or more instruction sets for operating a third device to determine the instruction sets for operating the third device. For example, Grotmol discloses that “[t]he trained controller of the robot may comprise a trained configuration configured to enable autonomous operation” and “the trained configuration may be loaded to one or more other robots in order to provide learned behaviors.” TESLA-1008, 33:21-35. Grotmol also discloses that the trained configuration of a controller can be updated through continuous learning, including learning in the “override” and “correct” modes. TESLA-1008, 34:59-62 (“training may comprise periods of autonomous operation and periods of learning”), 33:49-52 (“user commands ... may arrive in one or more clusters ... that may be interleaved by periods of user inactivity”), 23:9-45 (“In the override mode, the robot may execute what it has learned, unless a command is sent by the teacher. ... [T]he robot may head [sic] the teacher’s command, perform the action, and try to remember it for the next time the same situation occurs. ... In the correct mode, the robot may integrate what the teacher commands with what the robot already knows.”), 21:56-60 & Table 3 (“Learns the association”). A POSITA

would have understood from these disclosures, for instance, that the neural network for a controller can first be trained on a third device to learn instruction sets for operating a third device; the trained configuration of the neural network can be shared with and loaded on a second device where the trained configuration is updated through additional learning of instruction sets for operating the second device (e.g., in the override/correct modes); and the updated trained configuration can further be shared with and loaded on a first device from the second device. This predictable situation would have been expected and obvious to a POSITA reviewing Grotmol's disclosure, as it would allow multiple devices to gain the benefit of Grotmol's capabilities for sharing and updating training configurations. *Id.* The predictable result would be a trained neural network on the first device whose input neurons are correlated both with instruction sets for operating the second device (based on training that occurred on the second device) and with instruction sets for operating the third device (based on training that occurred on the third device). *Id.*

69. The Patent Owner and its declarant, Dr. Saber, have explained that even a basic neural network like the feedforward network Model A depicted in FIG. 17 of the '474 Patent has input neurons correlated with instruction sets for operating a device merely because the input neurons have "downstream connections with output nodes/neurons that include/produce instruction sets for operating a device." MTA, 15-16 (identifying written description support in Model A); EX2001, pp. 79-80, 113,

127, 139, 141-142, 151, 154-155, 164-165, 192, 203; TESLA-1001, FIG. 17. The input neurons of the neural network in the Grotmol-Zhu combination likewise have downstream connections to output neurons that produce instruction sets for operating the second and third devices (based on the training that occurred on each device). TESLA-1008, FIGS. 10A-11B; TESLA-1029, FIG. 10A. A POSITA thus would have recognized that the system in the Grotmol-Zhu combination uses correlations like that those recited in [85.13]-[85.15] to determine instruction sets for operating the second and third devices. It also would have been obvious that differences in the inputs provided to the Grotmol-Zhu neural network that correspond to differences in circumstances learned on the second and third devices would cause the neural network to determine/output different instruction sets for operating either the second or third devices. *Id.* For instance, differences in the locations or detected sizes of the multiple objects between the two sets of object representation data structures (and/or differences in other inputted information) could lead to the determination of different instruction sets. *Id.*

70. As an example of multiple autonomous vehicles interacting with the same object representation, it was known before the Critical Date that autonomous vehicle systems would “crowd-source” data related to their operating environments to improve the operation of individual vehicles (e.g., by refining a shared “road model”). TESLA-1038, Cover, ¶¶[0008], [0414]-[0437], [0533]-[0581], FIGS. 22,

28. Shashua discloses one example where the position of a “landmark” (e.g., a common object representation) is updated as information is received from different autonomous vehicles 2201 (e.g., a “second device”) and 2202 (a “third device”). *Id.* In Shashua’s system, any other given autonomous vehicle (e.g., a collection of “first devices”) would receive changes to their road map based on the interactions of one or both of vehicles 2201 and 2202. *Id.*

(b) Grotmol-Zhu-Buibas

71. The Grotmol-Zhu-Buibas combination renders [85.13]-[85.14] for substantially the same reasons addressed above in connection with the Grotmol-Zhu combination, except that Buibas further demonstrates that it would be obvious (with or without Grotmol’s teachings of sharing trained controller configurations and re-training/updating controllers) for the neural network in Grotmol’s adaptive controller to use correlations between the one or more input neurons and both (i) instruction sets for operating the second device (ii) instruction sets for operating the third device. Buibas’ teachings to these claim features, including in combination with Grotmol, have been addressed in detail in my Prior Declaration and above in this Supplemental Declaration with respect to original claims 56 and 57.³ TESLA-

³ Although claims 56 and 57 do not specifically require the knowledgebase to include a neural network, the mappings of Buibas for these claims are based on

1003, ¶¶99-100, 150-161.

72. For example, Buibas discloses techniques by which many different users can independent train their respective robotic devices to autonomously perform different tasks/behaviors. Users can then make the trained configurations of their robots available to others (e.g., by uploading trained brain images to the cloud). TESLA-1005, [0171]-[0180] (describing details of the training application provided to each user), [0181]-[0193] (describing details of sharing trained brain images through the cloud), [0194]-[0207], [0211]-[0213], [0222]-[0225], [0070], FIGS. 16-19. Users can then select one or more brain images to download and load on their own robots to take advantage of the training performed by others, and can even merge multiple brain images learned on other devices to obtain a combined set of capabilities. TESLA-1005, [0090]-[0127], [0188]-[0189], [0225], FIGS. 3-9. The merged brain image, whether by an interleaved merge or a non-interleaved merge, produces an offspring that has one or more input neurons correlated with instruction sets learned at each of the parents (e.g., at a second device and a third device), which correlations are used to determined instruction sets during autonomous operation. Specifically, in the context of non-interleaved merges, Buibas discloses that input from parents A and B are provided to “one of the parent network” (possessing its

Buibas’ teachings of brain images that implement neural networks.

own input neurons. TESLA-1005, ¶¶[0094]-[0098], FIG. 4.

8. Claims 86-93

(a) All Grounds (Unless Otherwise Indicated)

73. **Claim 86** – The additional features recited in Claim 86 are obvious based on Grotmol’s teachings for the reasons my Prior Declaration described in connection with Claim 46. Pet., 68-70; TESLA-1003, ¶¶132-133. Further, the “one or more inputs” would be input neurons for inputting at least a portion of the one or more object representations as described for [85.8] and [85.13]-[85.14].

74. **Claim 87** – The additional features recited in Claim 87 are obvious based on Grotmol’s teachings for the reasons my Prior Declaration described in connection with Claim 47. TESLA-1003, ¶¶134-135.

75. **Claim 88** – The additional features recited in Claim 88 are obvious based on Grotmol’s teachings for the reasons my Prior Declaration described in connection with Claim 51. Pet., 71; TESLA-1003, ¶136. Further, in the Grotmol-Zhu-Buibas combination, Buibas describes an additional learning process for operating a second device at least partially by a user. TESLA-1005, [0171]-[0193], [0070]; TESLA-1003, ¶¶99-100, 150-161.

76. **Claims 89-90** – In the Grotmol-Zhu-Buibas combination, the additional features recited in Claims 89-90 are obvious based on Grotmol’s and Buibas’ teachings for the reasons this and my Prior Declaration describe in

connection with Claims 56 and 57, and further above for [85.13]-[85.14]. TESLA-1003, ¶¶99-100, 150-161.

77. **Claim 91** – The limitation in claim 91 that the one or more object representations include one or more “coordinates” of one or more objects is obvious based on Grotmol’s and Zhu’s teachings for the reasons described above in connection with [85.11]-[85.12]. *Above*, [85.11]-[85.12].

78. **Claims 92-93** – Claims 92 and 93 merely re-write dependent claims 56 and 57 in independent form, respectively, but these claims are unpatentable based on Grounds 1A and 2C for the reasons described in this Supplemental Declaration and my Prior Declaration. Pet., 41-44, 77-82; TESLA-1003, ¶¶99-100, 150-161.

VII. THE SUBSTITUTE CLAIMS LACK SUPPORT UNDER § 112

A. “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”

79. Substitute claim 85 requires inputting a very specifically-recited set of multiple object representation data structures into input neuron(s) of a neural network and using correlations between the input neurons and instruction sets for operating a second device and a third device to determine instruction sets from the inputted data structures. MTA, 1-3. Based on my review, Patent Owner has not shown, however, that the inventor was in possession of these concepts, and hence

has not shown that the substitute claims are supported by the '474 Patent or the '991 Priority Application. The MTA includes lengthy string citations to various portions of the lengthy specification of the '991 Priority Application, but the citations to alleged support for processing the recited data structures with a neural network are quite limited.

80. For example, the MTA cites to FIG. 17 and the feedforward neural network shown in Model A as alleged support for these features. MTA, 15-16. But none of the disclosures cited in the MTA describe inputting the multiple object representation data structures into the input neurons of a neural network for processing as claimed—and particularly not multiple object representation data structures that include the recited label for object type, coordinates for object location, and computer model for object size. There is simply no evidence that the inventor possessed the use of a neural network for this specific scenario. Patent Owner instead improperly treats the '991 Application's boilerplate that "[t]he disclosed systems, devices, and methods are independent of the artificial intelligence model and/or technique used and any model and/or technique can be utilized to facilitate the functionalities described herein," as license for asserting long after the filing date that the inventor possessed the specifically claimed combination of features recited in claim 85, in the context of a neural network. EX2004, 93:24-32.

81. The MTA likewise fails to demonstrate support for a number of specific

nuances recited in Substitute Claim 85, including that the instruction sets are determined using correlations between input neurons and two *different* devices (i.e., the second and third devices) for object representation data structures pertaining to the *same* “multiple objects” detected at different times. MTA, [85.11]-[85.14]. For example, Patent Owner points to FIG. 38 as depicting a “second device” (“Loader 98a”), even though this disclosure does not address the claimed “third device,” let alone demonstrate that instruction sets can be determined for operating a third device based on the representations of the same objects that were used to determine instructions for operating the second device (“Loader 98a”). MTA, 16-17; TESLA-1001, FIG. 38. Citing to an entirely different figure, Patent Owner then alleges that “Boat 98b” is the claimed third device. *Id.*; TESLA-1001, FIG. 39. The Patent Owner’s theory would appear to require that (1) the Loader 98a and Boat 98b learned instructions for operating their respective devices based on representations of the same objects, and (2) the first device would need to be capable of executing instructions from both Loader 98a (second device) and Boat 98b (third device). *Id.* There is no support in the ’474 Patent, or explanation in Patent Owner’s motion, that bridges the gap between Figures 38 and 39, or explains how the two devices in these two separate figures would operate as claimed. The Patent Owner has failed to demonstrate written description support for the Substitute Claims. *Id.*

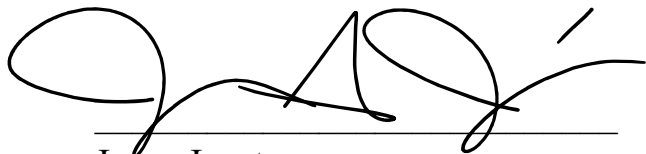
VIII. CONCLUSION

82. After reviewing the POR and Dr. Saber's declaration, I was not persuaded by any of the arguments therein. As such, I still conclude that the Remaining Challenged Claims (claims 56 and 57) are obvious.

83. Moreover, the substitute claims are also obvious over prior art and lack written description support in the '474 Patent or in its Priority Application, as I noted above.

84. I declare that all statements made in this Declaration of my own knowledge are true and that all statements made on information and belief are believed to be true. I further declare that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both (under Section 1001 of Title 18 of the United States Code).

Executed this 11th day of April, 2025.



Jason Janet