

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

**PATENT OWNER'S REPLY IN SUPPORT OF
CONTINGENT MOTION TO AMEND**

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	THE SUBSTITUTE CLAIMS ARE SUFFICIENTLY SUPPORTED	1
III.	THE SUBSTITUTE CLAIMS ARE PATENTABLE OVER THE ART.....	5
A.	The Substitute Claims Are Patentable Over Grotmol-Zhu.	5
1.	Limitations [85.11]-[85.12] distinguish over the art.	5
2.	Limitations [85.13]-[85.14] distinguish over the art.	9
B.	The Substitute Claims Are Patentable Over Grotmol- Zhu-Buibas.	12
IV.	CONCLUSION.....	12

TABLE OF AUTHORITIES

	Page(s)
CASES	
<i>Arendi S.A.R.L. v. Apple Inc.</i> , 832 F.3d 1355 (Fed. Cir. 2016)	10
<i>TQ Delta, LLC v. Cisco Sys., Inc.</i> , 942 F.3d 1352 (Fed. Cir. 2019)	11
STATUTES	
35 U.S.C.	
§ 103.....	1
§ 112.....	1

EXHIBIT LIST

Exhibit	Description	Previously Submitted
2001	Declaration of Dr. Eli Saber	X
2002	Dr. Eli Saber CV	X
2003	Deposition transcript of Dr. Jason Janet dated December 19, 2024 in IPR2024-00745	X
2004	File History of U.S. Patent App. No. 15/340,991 (“the ‘991 Application”)	X
2005	Deposition Transcript of Dr. Jason, Janet dated May 16, 2025, in IPR2024-00745	X
2006	Supplemental Declaration of Dr. Eli Saber	
2007	Zeiler & Fergus, <i>Visualizing and Understanding Convolutional Networks</i> , Computer Vision – EECV, 818-833, Springer International (2014).	
2008	Duda, et al., <i>Pattern Classification</i> , 2nd ed., 282–349. Wiley-Interscience, (2001).	

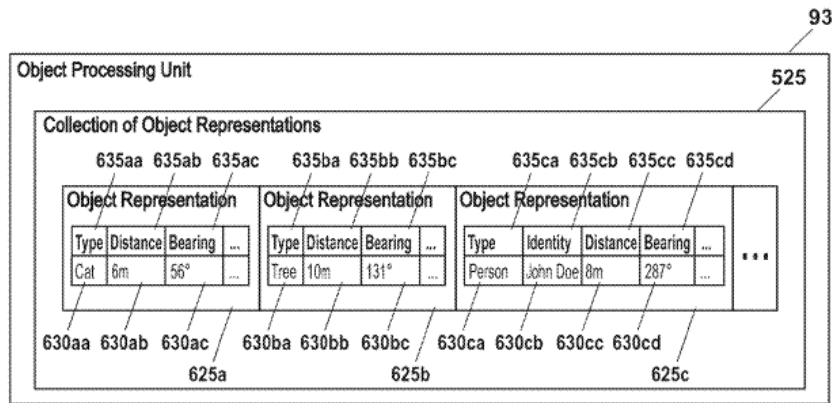
I. INTRODUCTION

PO timely submits this Reply in support of its MTA (Paper 14) pursuant to the Parties' stipulation (Paper 24). As the Board noted in its Preliminary Guidance (Paper 21, "PG"), Petitioner's Opposition (Paper 18) is incorrect regarding its § 112 challenge. Petitioner's § 103 arguments are also flawed. To the extent Petitioner attempts to inject any new arguments in its sur-reply, such attempts should be rejected, as PO would not have notice of such new arguments.

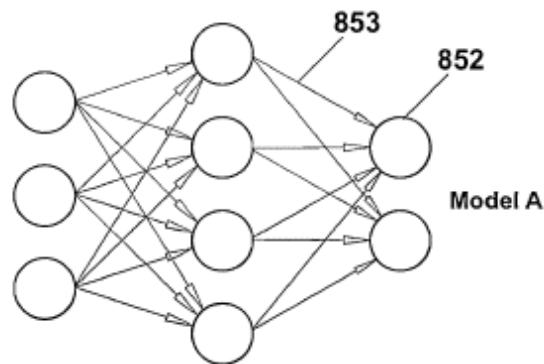
II. THE SUBSTITUTE CLAIMS ARE SUFFICIENTLY SUPPORTED

As the Board correctly noted, the specification includes adequate § 112 support for "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." PG, 5-6, 9; Opp., 23-24; Ex. 2006, ¶18. The Opposition notes the MTA's citation to Figure 17 (Opp., 23), including the neural network shown as Model A, yet Petitioner sticks its head in the sand regarding this disclosure as pertaining to inputting multiple object representation data structures into the input neuron(s).

Figure 4B shows object representation data structures:



Ex. 1001, FIG. 4B; *id.*, 56:1-3 (“resulting Collection of Object Representations 525”); Ex. 2006, ¶20. As the Board noted, the specification supports the claimed label for object type, coordinates for object location, and computer model for object size. Ex. 1002, 399:1–5; PG, 8-9; Ex. 2006, ¶20. The specification also supports *inputting* the claimed data structures *into the input neuron(s)*. For example, Figure 17 shows a neural network, where the three neurons on the left are input neurons:



Ex. 1001, FIG. 17; Ex. 2004, 398:7-9 (“[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”); Ex. 2006, ¶20. Model A’s arrows show how processing flows from input to output neurons.

Model A shows input neurons correlated using downstream connections with output neurons including instruction sets. MTA, 16-17; *see also* Ex. 2004, 356:22-26 (“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. *Id.*, 391:10-11 (“Artificial Intelligence Unit ... comprises ... Knowledgebase”), 396:14-15 (“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.*, 391:21-26. 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. *Id.*; Ex. 2006, ¶21.

Petitioner’s argument regarding “using a correlation between the one or more input neurons and the one or more instruction sets for operating the [second | third] device” is similarly meritless. Opp., 24-25; Ex. 2006, ¶22. 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.” Ex. 2004, 391:24-25. A POSA would have recognized by reading the specification that the inventor 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). Ex. 2006, ¶23. As the Board correctly noted, the specification discloses that “Device 98 may be or include Loader 98a” (Ex. 1002, 479:24, 480:3, FIG. 38), and that there may be any number of loaders 98 that utilize collective knowledge comprised in remote DCADO [] unit 100 for their autonomous operation (*id.*, 480:32-481:2). PG, 6-7. 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.” Ex. 1002, 480:32-481:2; Ex. 2004, 397:21-23 (“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.”).

For efficient use of figures, Figures 38 and 39 show examples of loader 98a and boat 98b to portray diversity of the types of devices, instead of showing multiple instances of the same loader (disclosed elsewhere in the specification). MTA, 17. As noted above, the specification describes any number of loaders 98a, which a

POSA would have understood as also disclosing any number of boats 98b (or other similar devices) without having to mix loaders and boats. Ex. 2006, ¶23. 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. *See* PG, 8; Ex. 1002, 420:18–21; Ex. 2004, 397:21-23; Ex. 2006, ¶24. As the Board noted, this would involve using correlations between input neurons in a neural network and instructions sets for operating the other Loaders 98a. *See* Ex. 1002, 427:1–2; Ex. 2006, ¶24. The specification also provides support for object representation data structures for the multiple objects detected at different times. Ex. 2004, 369:16-29 (disclosing objects with various properties and their object representations over time, e.g., at times t1 and t2).

III. THE SUBSTITUTE CLAIMS ARE PATENTABLE OVER THE ART

A. The Substitute Claims Are Patentable Over Grotmol-Zhu.

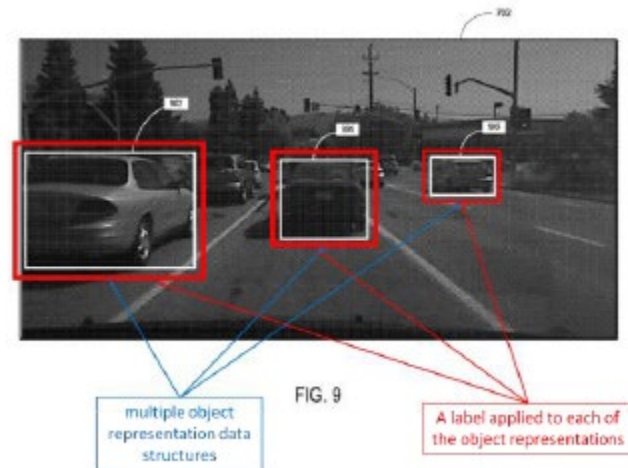
1. Limitations [85.11]-[85.12] distinguish over the art.

Grotmol-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 [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 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, 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]. Ex. 2006, ¶26. 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 ([85.pre]). 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). Ex. 2006, ¶27.

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’ 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); Ex. 2006, ¶28. Petitioner annotated the bounding boxes as follows:



Opp., 12 (annotated Zhu FIG. 9). 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 [85.11]-[85.12]. Limitations [85.11]-[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]-[85.12] also each recite a "circumstance representation that represents multiple objects detected." Thus, the circumstance *representation* (which includes object representation data structures) and the detected *objects* that the circumstance representation represents are clearly distinct. Petitioner's mapping relies on Zhu's image pixels and bounding boxes, but

those deal with *location and size of object representations within the image, not the location and size of the detected objects themselves*. Ex. 2006, ¶29.

More specifically, 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 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; Ex. 2006, ¶30. 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 contrast, for “coordinates for object location” and “model for object size” in [85.11]-[85.12], Petitioner relies on Zhu’s disclosure of pixel-based location and size of a bounding box that only pertains to an image *representation* (in the image coordinate system) *of* the underlying object. Thus, Petitioner’s mapping of “label for object type” to Zhu’s disclosure of an object type is inconsistent with Petitioner’s mapping of “coordinates for object location” and “model for object size” to Zhu’s disclosure of a pixel-based location and size of object *representations*.

Thus, Petitioner’s analysis based on Grotmol-Zhu for “coordinates for *object* location” and “model for *object* size” is flawed. Ex. 2006, ¶31. But even if the Board is inclined to side with Petitioner on those points, Petitioner’s inconsistency in its mapping of a detected “*object*” in the context of “label for object type” on the one hand, and a bounding box in pixels of an image that is clearly *not the detected object* in the context of “coordinates for object location” and “model of object size” on the other hand, is another basis for rejecting Petitioner’s prior art mapping.

2. Limitations [85.13]-[85.14] distinguish over the art.

Grotmol-Zhu does not disclose or suggest at least the following elements of [85.13]-[85.14]: “inputting the [multiple object representation data structures [of [85.11]] | another multiple object representation data structures [of [85.12]] of the stream into the one or more input neurons.” As an initial matter, the Opposition is flawed because Grotmol-Zhu does not disclose or suggest the specifically claimed data structures of [85.11]-[85.12] (*supra* § III.A.1), and thus does not disclose or suggest inputting those data structures into input neuron(s). The Opposition’s analysis of [85.13]-[85.14] is also flawed. Ex. 2006, ¶32.

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 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; Ex. 2006, ¶33. 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]. Opp., 14-16. Zhu does not disclose or suggest inputting the claimed object representation data structures into input neurons. Ex. 2006, ¶33.

Neither of the cited references actually discloses the claimed inputting of the specific claimed object representation data structures ([85.11]-[85.12]) into input neuron(s), and the Opposition merely concludes that “[i]n the combination, it would have been obvious ... to input ... the multiple object representation data structures for the objects detected at the first time (*supra*, [85.11]) into the input neurons of the neural network” and makes a similar conclusory statement for [85.12]. Opp., 15-16. The Opposition cites various background references regarding knowledge of a POSA. *Id.*, 16. But a POSA's knowledge may only be used to supply a missing claim limitation when that limitation is “unusually simple and the technology particularly straightforward.” *Arendi S.A.R.L. v. Apple Inc.*, 832 F.3d 1355, 1361 (Fed. Cir. 2016). The above portion of [85.13]-[85.14] does not satisfy those conditions. Ex.

2006, ¶34. And the testimony of Petitioner’s expert is conclusory. *TQ Delta, LLC v. Cisco Sys., Inc.*, 942 F.3d 1352, 1358 (Fed. Cir. 2019) (“Conclusory expert testimony does not qualify as substantial evidence.”).

It would have been complex, and certainly not 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 substantially different structure of the neural network to accommodate completely different inputs, training with a completely different type of inputs, and significant trial and error required to make the neural network workable and provide correct outputs. Ex. 2006, ¶34. Grotmol and Zhu do not disclose or suggest anything like this because they only contemplate using images and image features as inputs, and do not disclose or suggest inputting the specific object representation data structures of [85.11]-[85.12]. *Id.*, ¶35 (citing Exs. 2007-2008).

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. *Id.*, ¶35. That is because input neurons of Grotmol’s image-based neural network accept pixels of an image or image features (such as normalization via feature shifting or scaling) to produce any meaningful results. For example, a coordinate value or size

data cannot be inputted into an input neuron that accepts a color value for a pixel, and if attempted, this would have produced meaningless output. It would not have been obvious to just replace one type of input with a different type of input as Petitioner asserts. Ex. 2006, ¶36. Rather, it was complicated at the invention date to modify a neural network designed to accept pixels to instead accept other types of inputs such as the claimed “object representation data structures.” *Id.* A neural network for a different type of inputs was in practice not modified, but was primarily built anew. For example, 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 to be workable, which would not have been simple or predictable. *Id.*

B. The Substitute Claims Are Patentable Over Grotmol-Zhu-Buibas.

Petitioner 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 introduces Buibas as an additional secondary reference but not for the portions of those limitations discussed in § III.A.2. Thus, claims 85 and 86-91 (which inherit the above limitations via dependency) are patentable over Grotmol-Zhu-Buibas. Ex. 2006; ¶37.

IV. CONCLUSION

For the reasons above, in the MTA, and in the PG, the MTA should be granted.

Date: May 30, 2025

Respectfully submitted,

/s/ Arvind Jairam

Arvind Jairam

Reg. No. 62,759

ajairam@mckoolsmith.com

MCKOOL SMITH, P.C.

1717 K St. NW, Suite 1000

Washington, DC 20006

Telephone: (202) 370-8300

Facsimile: (202) 370-8344

Scott W. Hejny

Reg. No. 45,882

shejny@mckoolsmith.com

MCKOOL SMITH, P.C.

300 Crescent Court, Suite 1200

Dallas, TX 75201

Telephone: (214) 978-4000

George Fishback

Reg. No. 76,345

gfishback@mckoolsmith.com

MCKOOL SMITH, P.C.

303 Colorado Street, Suite 2100

Austin, TX 78701

Telephone: (512) 692-8700

Counsel for Patent Owner

Autonomous Devices, LLC

CERTIFICATION OF SERVICE

The undersigned hereby certifies that the foregoing **PATENT OWNER'S
REPLY IN SUPPORT OF CONTINGENT MOTION TO AMEND** has been
served on Petitioner via email on the following counsel of record for Petitioner:

W. Karl Renner
Nicholas W. Stephens
Usman Kham
Matthew Colvin
Fish & Richardson P.C.
60 South Sixth Street, Suite 3200
Minneapolis, MN 55402
IPR49649-0026IP1@fr.com

Ashraf Fawzy
Tesla, Inc.
800 Connecticut Ave NW
Washington, DC 20006
afawzy@tesla.com

Dated: May 30, 2025

By: /s/ Arvind Jairam

Arvind Jairam
Reg. No. 62,759

*Counsel for Patent Owner
Autonomous Devices, LLC*