

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.

Petitioner

v.

PERCEPTIVE AUTOMATA LLC

Patent Owner

IPR2025-01573

U.S. Patent No.: 10,614,344 B2

**PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-22 OF U.S.
PATENT NO. 10,614,344**

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
United States Patent and Trademark Office
Submitted Electronically

TABLE OF CONTENTS

	Page
PETITIONER’S EXHIBIT LIST	VI
I. MANDATORY NOTICES UNDER 37 C.F.R. 42.8(A)(1)	1
A. Real Party-in-Interest under 37 C.F.R. 42.8(b)(1)	1
B. Related Matters under 37 C.F.R. 42.8(b)(2)	1
C. Lead and Back-Up Counsel under 37 C.F.R. 42.8(b)(3)	1
D. Service Information under 37 C.F.R. 42.8(b)(4).....	2
II. CLAIM LISTING	3
III. INTRODUCTION	9
IV. PAYMENT OF FEES UNDER 37 C.F.R. § 42.103.....	9
V. CERTIFICATION OF GROUNDS FOR STANDING	9
VI. OVERVIEW OF CHALLENGE AND RELIEF REQUESTED	9
A. Prior Art Printed Publications	9
B. Identification of Challenge and Statement of Precise Relief Requested	11
VII. SUMMARY OF THE ’344 PATENT	11
A. Priority Date	11
B. Prosecution History	14
C. Technology Background	15
D. The Alleged Invention.....	15

VIII. LEVEL OF ORDINARY SKILL IN THE ART	17
IX. CLAIM CONSTRUCTION.....	17
X. THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	18
A. Ground 1: Claims 1-22 are unpatentable under 35 U.S.C. § 103 over Ross and Cox.....	18
1. Summary of Prior Art	18
a. Ross–EX1004.....	18
b. Cox–EX1005.....	19
2. Motivation to Combine Ross and Cox.....	20
3. Claim 1	24
a. 1[Pre].....	24
b. 1[a]	25
c. 1[b]	28
d. 1[c]	34
e. 1[d]	36
4. Claim 2	37
5. Claim 3	38
6. Claim 4	38
7. Claim 5	39
8. Claim 6	40
9. Claim 7	41
10. Claim 8	43
11. Claim 9	43
12. Claim 10	44
a. 10[Pre].....	44

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

b.	10[a]	44
c.	10[b]	44
d.	10[c]	44
e.	10[d]	44
13.	Claim 11	45
14.	Claim 12	45
15.	Claim 13	45
16.	Claim 14	45
17.	Claim 15	45
18.	Claim 16	45
19.	Claim 17	45
20.	Claim 18	46
21.	Claim 19	46
a.	19[Pre]	46
b.	19[a]	46
c.	19[b]	46
d.	19[c]	46
e.	19[d]	46
22.	Claim 20	46
23.	Claim 21	46
24.	Claim 22	47
B.	Ground 2: Claims 1-22 are unpatentable under 35 U.S.C. § 103 over Ross, Cox, and Ogale	47
1.	Summary of Prior Art	47
a.	Ogale–EX1010	47

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

2.	Motivation to Combine Ross, Cox, and Ogale	48
3.	Claim 1	51
a.	1[Pre] and 1[a].....	51
b.	1[b]	51
c.	1[c]	53
d.	1[d].....	56
4.	Claims 10 and 19.....	56
5.	Claims 2-9, 11-18, and 20-22	56
C.	Ground 3: Claims 1-4, 7-13, and 16-22 are unpatentable under 35 U.S.C. § 103 over Zhang.....	57
1.	Summary of the Prior Art	57
a.	Zhang–EX1006	57
2.	Claim 1	58
a.	1[Pre].....	58
b.	1[a]	58
c.	1[b]	60
d.	1[c]	67
e.	1[d]	69
3.	Claim 2	70
4.	Claim 3	70
5.	Claim 4	70
6.	Claim 7	71
7.	Claim 8	73
8.	Claim 9	73
9.	Claim 10	73

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

a.	10[Pre].....	73
b.	10[a]	73
c.	10[b]	74
d.	10[c]	74
e.	10[d]	74
10.	Claim 11	74
11.	Claim 12	74
12.	Claim 13	74
13.	Claim 16	74
14.	Claim 17	74
15.	Claim 18	75
16.	Claim 19	75
a.	19[Pre].....	75
b.	19[a]	75
c.	19[b]	75
d.	19[c]	75
e.	19[d]	75
17.	Claim 20	75
18.	Claim 21	75
19.	Claim 22	76
D.	Printed Matter	76
XI.	CONCLUSION	78

PETITIONER’S EXHIBIT LIST

Ex[No.]	Description of Documents
1001	U.S. Patent No. 10,614,344 (“’344 Patent”)
1002	File history of U.S. Patent No. 10,614,344
1003	Declaration of Dr. Jason Janét
1004	U.S. Patent No. 10,496,091 (“Ross”)
1005	International Publication No. WO2014/210334 (“Cox”)
1006	U.S. Patent Application Publication No. 2019/0072966 (“Zhang”)
1007	U.S. Provisional Application 62/528,771
1008	File history of U.S. Patent No. 10,402,687
1009	U.S. Patent Application Publication No. 2016/0162456 (“Munro”)
1010	U.S. Patent No. 10,733,506 (“Ogale”)
1011	Complaint in <i>Perceptive Automata LLC v. Tesla, Inc.</i> , Case No. 2:25-cv-00742 (E.D. Tex. filed July 23, 2025)
1012	CV of Dr. Jason Janét

I. MANDATORY NOTICES UNDER 37 C.F.R. 42.8(a)(1)

A. Real Party-in-Interest under 37 C.F.R. 42.8(b)(1)

The real party-in-interest is Tesla, Inc. No other parties exercised or could have exercised control over this Petition; no other parties funded or directed this Petition.

B. Related Matters under 37 C.F.R. 42.8(b)(2)

As of the filing date of this Petition, and to the best knowledge of Petitioner, U.S. Patent No. 10,614,344 (the “’344 Patent”) is involved in *Perceptive Automata LLC v. Tesla, Inc.*, Case No. 2:25-cv-00742 (E.D. Tex. filed July 23, 2025) (“EDTX Litigation”). EX1011.

C. Lead and Back-Up Counsel under 37 C.F.R. 42.8(b)(3)

Petitioner provides the following designation of counsel:

Lead Counsel	Backup Counsel
Roger Fulghum, Reg. No. 39,678 Baker Botts L.L.P. One Shell Plaza 910 Louisiana Street Houston, TX 77002 Phone: (713) 229-1234 roger.fulghum@bakerbotts.com	Mark Speegle, Reg. No. 77,512 Baker Botts L.L.P. 401 S. 1 st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2536 mark.speegle@bakerbotts.com Ellyar Y. Barazesh, Reg. No. 74,096 Baker Botts L.L.P. 401 S. 1 st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2507 ellyar.barazesh@bakerbotts.com William Gaines, Reg. No. 83,138 Baker Botts L.L.P.

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

	401 S. 1st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2550 william.gaines@bakerbotts.com Lance J. Goodman, Reg. No. 77,989 Baker Botts L.L.P. 401 S. 1st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2629 lance.goodman@bakerbotts.com Ashraf Fawzy, Reg. No. 67,914 Tesla, Inc. 800 Connecticut Ave. NW Washington, DC 20006 Phone: (202) 905-9221 afawzy@tesla.com
--	--

D. Service Information under 37 C.F.R. 42.8(b)(4)

A copy of this entire Petition, including all Exhibits and a power of attorney, is being served by FEDERAL EXPRESS, costs prepaid, to the correspondence address of record for the '344 Patent at the USPTO:

758 - FENWICK & WEST LLP
c/o Rajendra B. Panwar (rpanwar@fenwick.com)
Silicon Valley Center
801 California Steet
Mountain View, CA 94041

and by email to the attorney or agent of record for Patent Owner in the EDTX Litigation:

Patrick J. Conroy (pat@nelbum.com)
Andrea L. Fair (andrea@millerfairhenry.com)

Please address all correspondence to lead and back-up counsel. Petitioner consents to service at lead counsel's address provided above. Petitioner consents to electronic service, provided it is made to all of the following e-mail address:

Roger Fulghum (roger.fulghum@bakerbotts.com)
Mark Speegle (mark.speegle@bakerbotts.com)
Ellyar Y. Barazesh (ellyar.barazesh@bakerbotts.com)
William Gaines (william.gaines@bakerbotts.com)
Lance J. Goodman (lance.goodman@bakerbotts.com)
Ashraf Fawzy (afawzy@tesla.com)
DLTeslavPerceptiveAutomataIPRs@bakerbotts.com

A Power of Attorney is filed concurrently herewith under 37 C.F.R. §42.10(b).

II. CLAIM LISTING

Claim 1

1[Pre]	A computer-implemented method comprising:
1[a]	receiving, by a computing device associated with an autonomous vehicle operating on a road, sensor data captured by a sensor installed on the autonomous vehicle, the sensor data displaying an object on the road;
1[b]	providing the sensor data as input to a supervised learning based model, the supervised learning based model configured to receive an input sensor data displaying a particular object on a road and predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object;
1[c]	executing the trained supervised learning based model to generate a statistical summary data characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users; and
1[d]	controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data.

Claim 2

2	The computer-implemented method of claim 1, wherein the supervised learning based model is one of: a random forest regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, or a long short-term memory (LSTM) neural network.
----------	--

Claim 3

3	The computer-implemented method of claim 1, wherein the object on the road is one of: a pedestrian, a cyclist, or a motor vehicle.
----------	--

Claim 4

4	The computer-implemented method of claim 1, wherein the user responses describe an action associated with the object, the action including one of the object staying in place, changing lanes, or crossing a street.
----------	--

Claim 5

5	The computer-implemented method of claim 1, wherein the statistical data is associated with a parameter of the subset of the plurality of response data, the parameter including at least one of a content of a response, a time associated with entering a response, and a position of an eye of a human observer associated with the response, the position being measured with respect to a display associated with the user interface.
----------	--

Claim 6

6	The computer-implemented method of claim 5, wherein the parameter is further associated with at least one of a central tendency, a variance, a skew, a kurtosis, a scale, and a histogram.
----------	--

Claim 7

7	The computer-implemented method of claim 1, wherein the supervised
----------	--

	learning based model is configured to predict user responses describing a state of mind of a road user associated with the object.
--	--

Claim 8

8	The computer-implemented method of claim 1, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing an intention of the road user to move into the path of the autonomous vehicle.
----------	---

Claim 9

9	The computer-implemented method of claim 1, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing awareness of the road user of the autonomous vehicle.
----------	--

Claim 10

10[Pre]	A non-transitory computer readable storage medium storing instructions that when executed by one or more processors, cause the one or more processors to perform the steps comprising:
10[a]	receiving, by a computing device associated with an autonomous vehicle operating on a road, sensor data captured by a sensor installed on the autonomous vehicle, the sensor data displaying an object on the road;
10[b]	providing the sensor data as input to a supervised learning based model, the supervised learning based model configured to receive an input sensor data displaying a particular object on a road and predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object;
10[c]	executing the trained supervised learning based model to generate a statistical summary data characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users; and
10[d]	controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data.

Claim 11

11	The non-transitory computer readable storage medium of claim 10, wherein the supervised learning based model is one of: a random forest regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, or a long short-term memory (LSTM) neural network.
-----------	---

Claim 12

12	The non-transitory computer readable storage medium of claim 10, wherein the object on the road is one of: a pedestrian, a cyclist, or a motor vehicle.
-----------	---

Claim 13

13	The non-transitory computer readable storage medium of claim 10, wherein the user responses describe an action associated with the object, the action including one of the object staying in place, changing lanes, or crossing a street.
-----------	---

Claim 14

14	The non-transitory computer readable storage medium of claim 10, wherein the statistical data is associated with a parameter of the subset of the plurality of response data, the parameter including at least one of a content of a response, a time associated with entering a response, and a position of an eye of a human observer associated with the response, the position being measured with respect to a display associated with the user interface.
-----------	---

Claim 15

15	The non-transitory computer readable storage medium of claim 14, wherein the parameter is further associated with at least one of a central tendency, a variance, a skew, a kurtosis, a scale, and a histogram.
-----------	---

Claim 16

16	The non-transitory computer readable storage medium of claim 10, wherein the supervised learning based model is configured to predict user responses describing a state of mind of a road user associated with the object.
-----------	--

Claim 17

17	The non-transitory computer readable storage medium of claim 10, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing an intention of the road user to move into the path of the autonomous vehicle.
-----------	--

Claim 18

18	The non-transitory computer readable storage medium of claim 10, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing awareness of the road user of the autonomous vehicle.
-----------	---

Claim 19

19[Pre]	A computing system comprising: one or more processors; and a non-transitory computer readable storage medium, storing instructions that when executed by the one or more processors, cause the one or more processors to perform steps comprising:
19[a]	receiving, by a computing device associated with an autonomous vehicle operating on a road, sensor data captured by a sensor installed on the autonomous vehicle, the sensor data displaying an object on the road;
19[b]	providing the sensor data as input to a supervised learning based model, the supervised learning based model configured to receive an input sensor data displaying a particular object on a road and predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object;
19[c]	executing the trained supervised learning based model to generate a

	statistical summary data characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users; and
19[d]	controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data.

Claim 20

20	The computing system of claim 19, wherein the supervised learning based model is configured to predict user responses describing a state of mind of a road user associated with the object.
-----------	---

Claim 21

21	The computing system of claim 19, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing an intention of the road user to move into the path of the autonomous vehicle.
-----------	---

Claim 22

22	The computing system of claim 19, wherein the object is associated with a road user and the supervised learning based model is configured to predict user responses describing awareness of the road user of the autonomous vehicle.
-----------	--

III. INTRODUCTION

Tesla, Inc. (“Petitioner”) requests *inter partes* review (“IPR”) of claims 1-22 (the “Challenged Claims”) of U.S. Patent No. 10,614,344 (EX1001) (the “’344 Patent”). This Petition demonstrates that the Challenged Claims are unpatentable.

IV. PAYMENT OF FEES UNDER 37 C.F.R. § 42.103

The undersigned authorizes the Office to charge the fee set forth in 37 C.F.R. §42.15(a)(1) for this Petition to Deposit Account No. 02-0384. The undersigned further authorizes payment for any additional fees that may be due in connection with this Petition.

V. CERTIFICATION OF GROUNDS FOR STANDING

Petitioner certifies under 37 C.F.R. § 42.104 that the ’344 Patent is available for IPR and that Petitioner is not barred or estopped from requesting IPR of the Challenged Claims based on the grounds identified in this Petition.

VI. OVERVIEW OF CHALLENGE AND RELIEF REQUESTED

A. Prior Art Printed Publications

The ’344 Patent was filed on July 16, 2019, as a continuation of U.S. Pat. App. No. 15/830,549 (issued as U.S. Pat. No. 10,402,687) (the “’687 Patent”) filed on December 4, 2017, which claims priority to U.S. Provisional Application 62/528,771, filed on July 5, 2017 (the “’771 Provisional”). EX1001, Cover Page. However, the ’344 Patent is not entitled to the July 5, 2017 priority date. *See* Section VII.A.

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

Regardless, the Ross, Cox, and Ogale references are prior art to the '344 Patent before July 5, 2017.

- U.S. Patent No. 10,496,091 (“Ross”) (EX1004); filed on August 17, 2016; prior art under 35 U.S.C. § 102(a)(2).
- International Publication No. WO2014/210334 (“Cox”) (EX1005); published December 31, 2014 and filed June 26, 2014; prior art under 35 U.S.C. § 102(a)(1) and § 102(a)(2).¹
- U.S. Patent No. 10,733,506 (“Ogale”) (EX1010); filed December 14, 2016; prior art under 35 U.S.C. § 102(a)(2).

Zhang is prior art to the '344 Patent before the claimed December 4, 2017 priority date, after priority to the provisional application's July 5, 2017 date is broken.

- U.S. Patent Application Publication No. 2019/0072966 (“Zhang”) (EX1006); filed November 7, 2017; prior art under 35 U.S.C. § 102(a)(2).

None of these references were cited or discussed during prosecution.

¹ Cox lists Harvard as applicant and Samuel Anthony as an inventor. EX1005, Cover Page. This inventor appears to be the same Samuel Anthony listed as an inventor of the '344 Patent. EX1001, Cover Page.

B. Identification of Challenge and Statement of Precise Relief Requested

Petitioner requests cancellation of the Challenged Claims based on the specific grounds of this Petition, which are set forth below and are supported by the declaration of Dr. Janét (EX1003).

Ground	Challenged Claims	Basis	Reference(s)
1	1-22	§ 103	Ross and Cox
2	1-22	§ 103	Ross, Cox, and Ogale
3	1-4, 7-13, and 16-22	§ 103	Zhang

VII. SUMMARY OF THE '344 PATENT

A. Priority Date

The '344 Patent was filed on July 16, 2019, as a continuation of the '687 Patent, which was filed on December 4, 2017 and claims priority to the '771 Provisional filed July 5, 2017. EX1001, Cover Page.²

The priority date of the '344 Patent is no earlier than December 4, 2017, because it is not entitled to the priority date of the '771 Provisional. First, there is no presumption that the '344 Patent is entitled to the July 5, 2017 provisional filing date.

² Because the application leading to the '344 Patent was filed after March 16, 2013, patentability is governed by the America Invents Act ("AIA") version of 35 U.S.C. §§102 and 103. Pub L. No. 112-29, 125 Stat. 284 (2011).

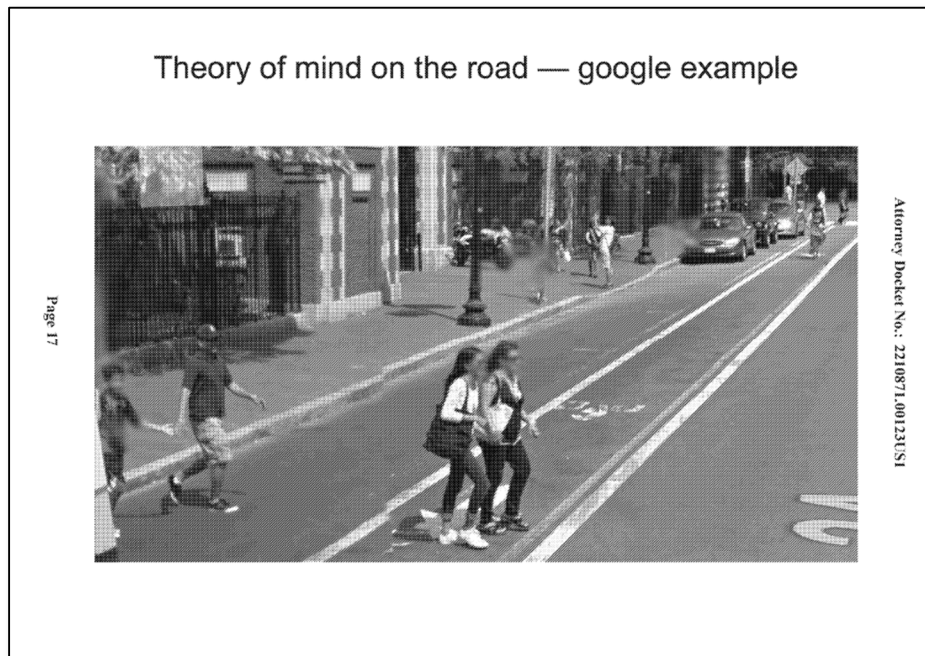
Dynamic Drinkware, LLC v. National Graphics, Inc., 800 F.3d 1375, 1380 (Fed. Circ. 2015). Patent Owner bears the burden to establish that the '771 Provisional provides written description support under 35 U.S.C. §112(a) for each Challenged Claim. *Chiron Corp. v. Genentech, Inc.*, 363 F.3d 1247, 1253 (Fed. Cir. 2004); *New Railhead Mfg., L.L.C. v. Vermeer Mfg. Co.*, 298 F.3d 1290, 1294-95 (Fed. Cir. 2002). This means the specification “must describe the invention sufficiently to convey to a person of skill in the art that the patentee had possession of the claimed invention at the time of the application[.]” *Lizardtech, Inc. v. Earth Resource Mapping, Inc.*, 424 F.3d 1336, 1345 (Fed. Cir. 2005).

The '771 Provisional does not provide written description support for at least the following limitations of claim 1 (and similarly in claims 10 and 19):

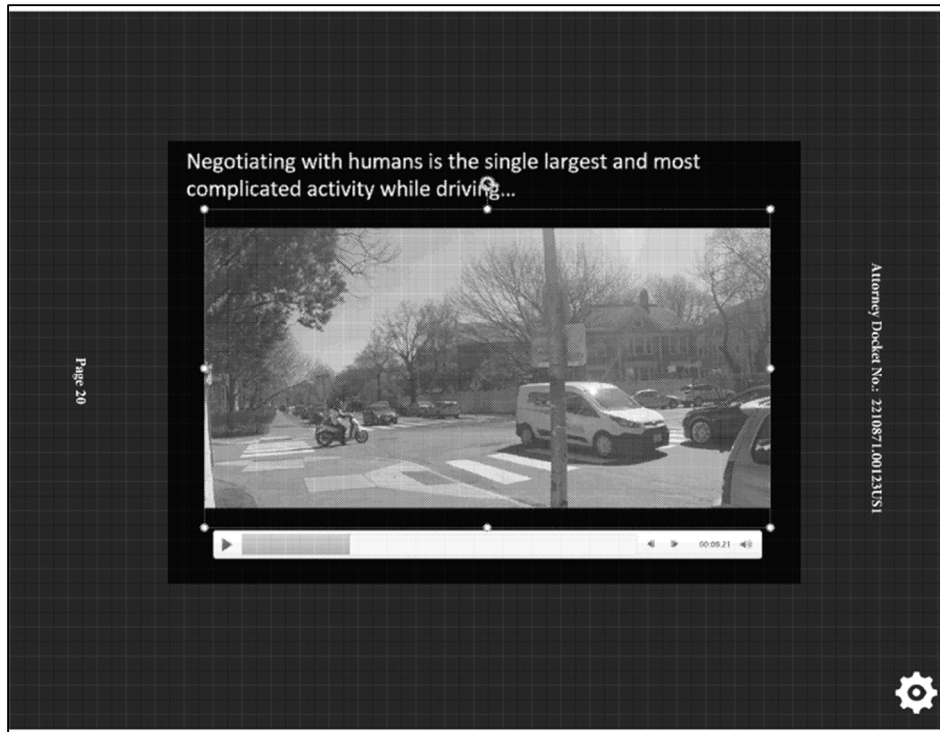
- “*the supervised learning based model configured to receive an input sensor data displaying a particular object on a road and predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object,*”
- “*executing the trained supervised learning based model to generate a statistical summary data characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users,*” and

- “controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data.”

EX1003, ¶58. The '771 Provisional includes only PowerPoint slides—the majority of which only show low resolution apparent screenshots of a video without any descriptive content—that are devoid of any reference to at least these claim limitations. *See* EX1007, 9-38. Two exemplary slides showing the lack of detail in the provisional are below. EX1003, ¶59.



EX1007, 25



EX1007, 28

Notably, the slides in the provisional are all conceptual. None describe the claimed “*supervised learning based model*,” much less the specific claimed features such as “*predicting an output statistical summary characterizing a distribution of user responses*.” Absent any description of these features, conceptual slides do not show that the patentee was in possession of the claimed inventions, and thus the Challenged Claims are not entitled to the ’771 Provisional’s filing date. *SAP America, Inc. v. Pi-Net International, Inc.*, IPR2014-00414, Paper 11 at 11-16 (P.T.A.B. Aug. 18, 2014); EX1003, ¶60.

B. Prosecution History

The Office issued a first non-final office action including only a non-statutory double patenting rejection over the application which later issued into the ’687

Patent. EX1002, 91-97. In response, the applicant filed a terminal disclaimer, and the Office issued a notice of allowance without identifying allowable claim limitations. *Id.*, 114, 133-139.

C. Technology Background

The '344 Patent generally relates to supervised machine learning models, which are computer programs that analyze input data and generate an output based on the data. EX1003, ¶62. Such models are “supervised” because they are trained using labeled training data that is known to represent a correct output. *Id.* Human annotators (also referred to as “labelers” or “reviewers”) can label such training data. *Id.* After training, the model can be applied to new, non-training data, to provide an output for the new data that reflects the human annotators’ decision making when labeling training data. *Id.*

D. The Alleged Invention

The '344 Patent’s alleged invention relates to (1) training a machine learning model using human labeled training data and (2) applying the trained model to new input data to generate an output that replicates how humans would have labeled the new input data. EX1001, Abstract, 6:1-7:39. But this alleged invention was well-known, as demonstrated herein. Section X; EX1003, ¶64; *see also, e.g.*, EX1010, 13:30-14:9 (machine learning model trained by human labeled training data “predicts how a human would have labeled the [new] input sensor data”). Indeed, the *inventor’s own prior art* shows that the '344 Patent should not have issued. The

Cox reference—which also lists Samuel Anthony as an inventor—discloses machine learning model predictions that are “more consistent with the decisions of the human annotators” and “better mimic human performance.” EX1005, Cover, ¶¶0004, 0032; EX1003, ¶64. Neither the patentee nor Mr. Anthony brought the Cox reference to the patent office’s attention during prosecution of the ’344 Patent. *See* EX1001; EX1002.

The ’344 Patent also says that after training, the model is used to make a prediction based on “live” input data for operation of an autonomous vehicle. EX1001, 7:5-39. For example, the predictions reflect “predicted summary statistics for the ‘real world’ or ‘live data’ image” where the “predicted summary statistics are the model's best estimation of what the summary statistics would be on the image if the image had human annotations collected.” *Id.*, 12:44-49, Fig. 9. Such predictions were well known. EX1003, ¶¶65-66. For example, the Ogale reference explains that a machine learning model trained by human labeled training data “predicts how a human would have labeled the [new] input sensor data.” EX1010, 13:30-14:9.

And the ’344 Patent’s application of these known concepts to autonomous driving does not save it. EX1001, 1:53-56, 3:44-46; EX1003, ¶67. Controlling an autonomous vehicle based on the predictions of a machine learning model was well known as demonstrated by at least Ross and Zhang. EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 6:33-59, 7:57-8:31, 9:13-10:30, 15:1-16:52, Fig. 1; EX1006, ¶¶0054, 0057-0061, Fig. 8; EX1003, ¶67. Ross explains how an “intent model” trained using

human observation is used so that an autonomous vehicle can determine how vehicles, bicyclists, and pedestrians around the autonomous vehicle will act—for example, whether a vehicle wishes to change lanes or a pedestrian wishes to cross a road. EX1004, 1:30-2:53, 4:33-5:11, 9:13-10:10, 15:1-16:52. Zhang similarly discloses predicting the behavior and intent of road users using “well-known” neural networks that are trained using human labeled training data that “associate[s] a label having behavior and direction information with each instance of a vehicle in the training data.” EX1006, ¶¶0049-0052, Abstract.

VIII. LEVEL OF ORDINARY SKILL IN THE ART

A person having ordinary skill in the art (“POSITA”) relevant to the ’344 Patent as of July 5, 2017 would have had at least: (1) a bachelor’s degree in electrical engineering, computer engineering, computer science, or equivalent course work, with three years of work experience in computer vision, autonomous vehicles, and/or machine learning; or (2) a master’s degree in electrical engineering, computer engineering, computer science, or equivalent course work, with a focus in computer vision, autonomous vehicles, and/or machine learning. EX1003, ¶69.

IX. CLAIM CONSTRUCTION

Claim terms in an IPR are construed according to their “ordinary and customary meaning” to those of skill in the art. *Phillips v. AWH Corp.*, 415 F.3d 91303 (Fed. Cir. 2005) (en banc); 37 C.F.R. §42.100(b). Because there is no need for claim construction to determine that the cited prior art renders all Challenged Claims

obvious, Petitioner does not present any constructions. EX1003, ¶70.

X. THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. Ground 1: Claims 1-22 are unpatentable under 35 U.S.C. § 103 over Ross and Cox.

Ross alone renders obvious most limitations of the Challenged Claims, and is further combined with Cox to render obvious the remaining limitations, which relate to a *supervised learning model* and predicting outputs that reflect “*user responses expected to be received.*” As discussed below, these limitations are explicitly taught by Cox, and it would have been obvious to combine Cox’s teachings with Ross. EX1003, ¶71.

1. Summary of Prior Art

a. Ross–EX1004

Ross—a Waymo patent—discloses techniques for “maneuvering” an autonomous vehicle. EX1004, 1:30-31, 3:39-42, 3:64-5:5. Ross’s system generates human influenced training data and uses the data to train an “intent model” for “predicting [the] intent of an object” (e.g., another vehicle, a pedestrian) in an autonomous vehicle’s environment. EX1004, 1:30-2:53, 4:33-5:11, 9:13-10:10, 15:1-16:52. The human influenced training data is formed by “human operators” who manually generate a “list of predetermined actions” performed by road users. *Id.*, 9:13-10:10.

Ross teaches that an autonomous vehicle uses the trained intent model for

navigation. EX1004, 7:57-8:10, 9:13-10:10, 15:1-16:52. Cameras of the autonomous vehicle capture “raw” camera data around the vehicle. *Id.* The trained model is executed on this data to predict a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories)” for objects around the vehicle. *Id.*, 3:64-4:46, 7:57-8:10, 9:13-10:10, 15:1-16:52 (parenthetical in original). The model also generates a set of likelihood values associated with the different possible actions that “indicate which of the predictions are more likely to occur (relative to one another).” EX1004, 9:13-10:10 (parenthetical in original).

Ross is analogous art to the ’344 Patent, from the same field of endeavor as the patent (e.g., predictive model techniques), and is reasonably pertinent to the particular problem that the patent was trying to solve (e.g., accounting for moving objects when navigating an autonomous vehicle in an environment). EX1004, Abstract, 3:64-5:1; 9:27-47; EX1001, 1:15-17, 1:36-56; EX1003, ¶76.

b. Cox–EX1005

Cox lists a common inventor—Samuel Anthony—with the ’344 Patent. EX1001, Cover; EX1005, Cover.

Cox is directed toward “machine learning augmented by measurements acquired from human populations.” EX1005, Abstract, ¶¶0001-0004. Cox teaches crowd-sourcing training data used to train supervised machine learning models. *Id.*, Abstract, ¶¶0001-0004, 0015-0016, 0025-0034, 0044, 0060. After transmitting images to a plurality of human annotators/labelers, Cox compiles statistics about the

human responses, including psychometric data associated with the responses, and trains a supervised learning model on this information. EX1005, ¶¶0015-19, 0029-0032, Figs. 1A-B, 3. The goal of Cox’s machine learning techniques is to “better mimic human performance.” EX1005, ¶0004. Cox touts that its “supervised learning” techniques may be “directly applied to several important domains where machine learning is found” and expressly suggests applying the invention to “driverless” automobile applications. EX1005, ¶¶0056, 0060.

Cox is analogous art to the ’344 Patent, from the same field of endeavor as the patent (e.g., predictive model techniques), and is reasonably pertinent to the particular problem that the patent was trying to solve (e.g., accounting for moving objects when navigating an autonomous vehicle in an environment). EX1005, ¶¶0029-0032, Figs. 1A-B, 3; EX1001, 1:15-17, 1:53-56, 5:39-46; 6:1-42; Fig. 2A; EX1003, ¶79

2. Motivation to Combine Ross and Cox

A POSITA would have been motivated to combine Ross and Cox such that Ross’s intent model is a supervised learning model that is trained with human annotated image data, as taught by Cox (providing the “Ross-Cox model”). EX1004, 4:33-65, 9:27-10:10, 15:1-16:22; EX1005, ¶¶0007-0010, 0015-0019, 0025-0035, 0051, 0060; EX1003, ¶80. Based on this training approach, a POSITA would have further configured the Ross-Cox model to generate predictions for objects shown in camera data that are “more consistent with the decisions of the

human annotators” who annotated the training data. EX1004, 1:55-2:13, 4:33-46, 9:13-10:10, 15:1-21; EX1005, ¶¶0004, 0007-0010, 0015-0019, 0025-0035, 0051, 0060; EX1003, ¶80. A POSITA would have been motivated to combine Ross and Cox in this way for several reasons. EX1003, ¶80.

First, Cox simply expands on Ross’s disclosure of a predictive intent model for an autonomous vehicle that is trained using human observations. EX1004, 9:13-10:10; Section X.A.1.a; EX1003, ¶81. Ross discloses generating training data, using the training data to train an intent model, and executing the model on input raw camera data from cameras of an autonomous vehicle to predict trajectories for vehicles, pedestrians, and bicyclists around the autonomous vehicle. EX1004, 3:39-51; 3:64-5:11, 7:57-8:10, 9:13-10:10, 13:29-33, 15:1-16:52. Cox expands on this by identifying such a model as a supervised learning model, and teaching that the model is trained and optimized using human observations, so that model predictions better reflect human decision making. EX1005, ¶¶0007-0010, 0015-0019, 0025-0038, 0051, 0060, Fig. 4B; EX1003, ¶81. The goal of Cox’s machine learning techniques is to “better mimic human performance,” and Cox explains that its machine learning “may be applied to any form of supervised learning, including neural networks.” EX1005, ¶¶0004, 0060; *see also id.*, ¶¶0008-0010, 0015-0019, 0025-0035, 0051. Cox thus provides a natural expansion and improvement on Ross’s base disclosure. EX1003, ¶81.

A POSITA seeking to implement Ross’s system would also have looked to

Cox to provide more instructions for training a machine learning model to predict human behavior. EX1003, ¶82. Both references are directed to machine learning technologies and involve (1) gathering training data, (2) training a model with the training data, and (3) executing the model, and as noted above, both are explicitly applicable to autonomous vehicle applications. *Id.* Cox even acknowledges as much, explaining that its disclosure is relevant to many “visual-recognition problems” including those related to detecting “objects” such as a “car,” as well as “driverless” vehicle applications. EX1005, ¶¶0016, 0056. Because Cox suggests that its system is relevant for such applications, Cox therefore provides an explicit teaching, suggestion, or motivation to combine with a system for controlling driverless (or autonomous) vehicles, like Ross. EX1003, ¶82.

The stated goal in Ross to provide an autonomous vehicle that “function[s] in a more ‘human-like’ or ‘polite’ way” would also have motivated a POSITA to combine Ross with Cox. EX1004, 5:7-12; EX1003, ¶83. Consistent with Ross’s goal, Cox provides a system for training a model “to better mimic human performance” and to generate predictions that are “more consistent with the decisions of the human annotators.” EX1005, ¶¶0004, 0029-0032; EX1003, ¶83. A POSITA would have recognized that using Cox’s teachings would help ensure that Ross’s predictive intent model is more “human-like.” *Id.*; EX1004, 5:7-12

A POSITA would have also been motivated to combine the teachings of Ross and Cox because the combination would have benefitted Ross’s system. EX1003,

¶84. Enhancing Ross’s system with Cox’s human annotated training data that considers the state of mind of training subjects, as well as Cox’s associated model training and optimization processes, would have further refined Ross’s prediction model so it better considers the state of mind of such road users and what they might be thinking when present in autonomous vehicle’s environment. *Id.*; EX1005, ¶¶0007-0010, 0015-0019, 0025-0035, 0051, 0060. This would have improved the accuracy of Ross’s intent predictions, and thus improved safety by helping to reduce collisions between the autonomous vehicle and human drivers or pedestrians. EX1003, ¶84.

All the reasons explained above establish that a POSITA would have been motivated to combine Ross and Cox as described. This combination would have been the predictable combination of well-known prior art elements (e.g., Ross’s prediction model; Cox’s model training and optimization techniques) according to known methods to yield predictable results (Ross’s intent model being a supervised learning model; the Ross-Cox model generating predictions for objects shown in camera data that are “more consistent with the decisions of the human annotators” who annotated the training data). *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007); EX1003, ¶85. The combination would have further used known techniques (known machine learning model training and optimization techniques as taught by Cox) to improve similar devices (Ross’s model for predicting intents) in the same way (using Cox’s aforementioned techniques). *Id.* Lastly, in addition to the

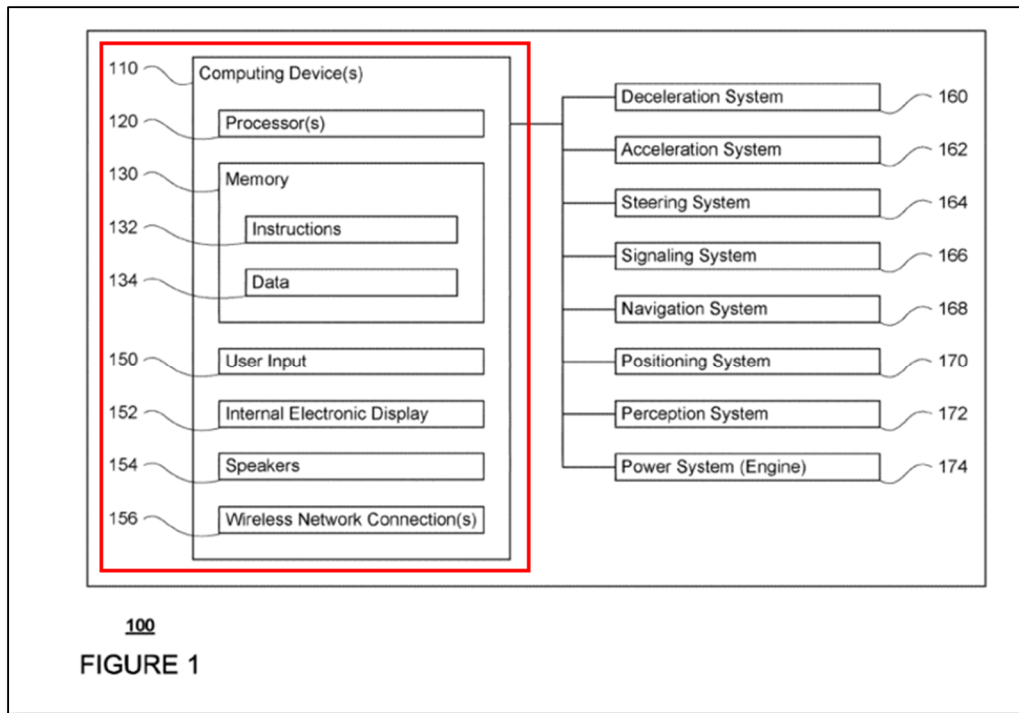
above reasons, because Ross uses human-labeled data to train its model, combining Ross with Cox’s teachings about generating and using the same type of data in a supervised learning model would have required minimal changes. *Id.* This would have been well within the skill of a POSITA, and a POSITA would have had a reasonable expectation of success. *Id.*

3. Claim 1

Ross and Cox render obvious claim 1. EX1003, ¶86.

a. 1[Pre]

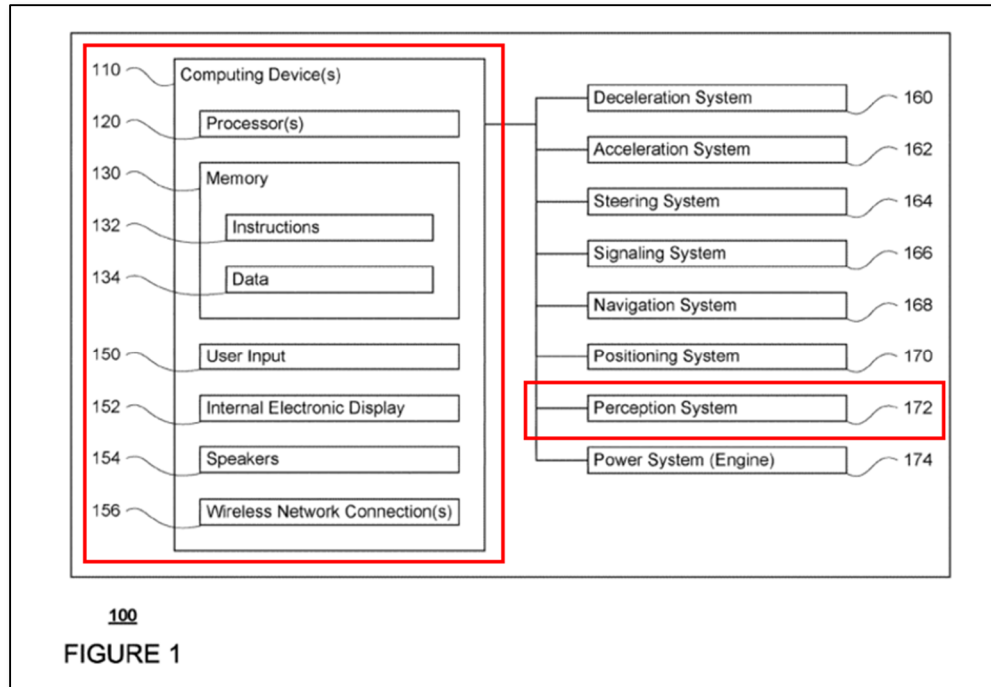
To the extent the preamble is limiting, Ross teaches it. EX1003, ¶87. Ross describes systems and methods for “maneuvering a vehicle” using a “computing device 110;” computing device 110 is an “autonomous driving computing system” incorporated into a vehicle 100 that controls the vehicle (*[a] computer-implemented method*). EX1004, 1:30-2:53, 3:39-42, 5:13-22, 6:33-59, Fig. 1; EX1003, ¶87.



EX1004, Fig. 1 (annotated)

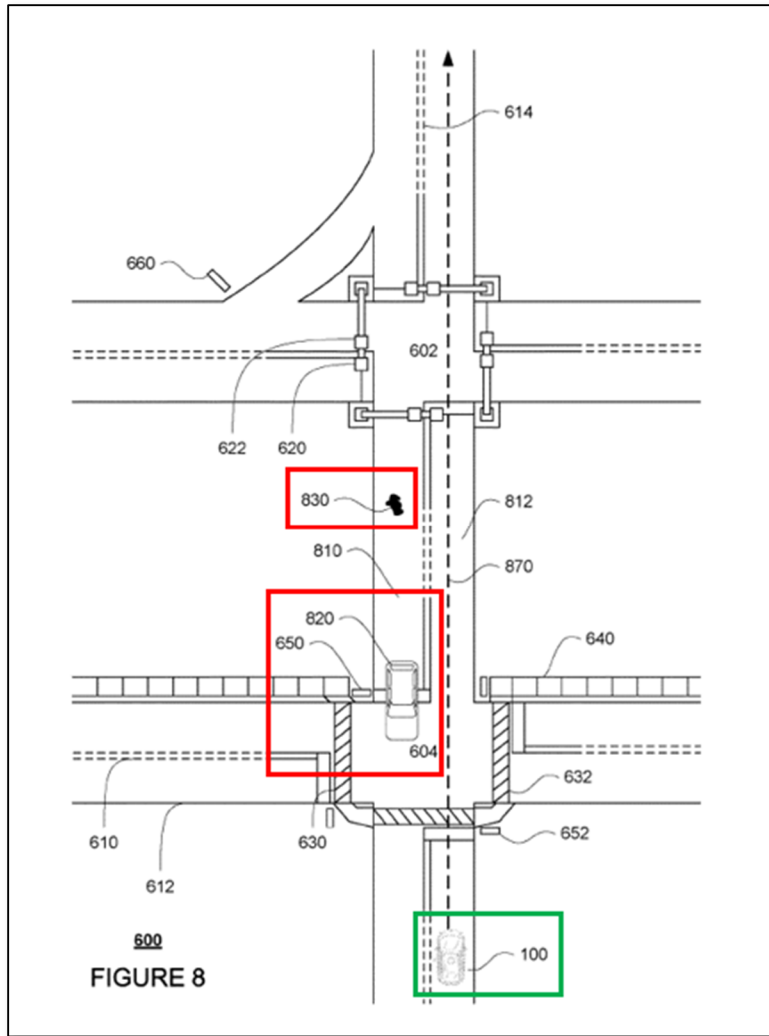
b. 1[a]

Ross teaches element 1[a] because Ross’s vehicle 100 includes a “perception system 172” that includes “cameras” for “detecting objects external to the vehicle” such as other vehicles, pedestrians, and bicyclists, as well as “obstacles in the roadway, traffic signals, signs, tress, etc.” EX1004, 4:66-5:11, 7:57-8:31, 8:32-52, 9:41-65, 15:50-16:41, Claims 2 and 3, Figs. 1, 3A-3D; EX1003, ¶88.



EX1004, Fig. 1 (annotated)

As vehicle 100 travels on a road, the cameras of perception system 172 generate “raw data” that displays objects corresponding to other “road user[s]” such as other vehicles, bicycles, or pedestrians—for example, as vehicle 100 travels on a road, the data can display a vehicle 820 and pedestrian 830 in the road as shown by Figure 8. EX1004, 3:39-47, 4:66-5:11, 7:57-8:52, 9:41-65, 13:29-14:58, 15:50-16:41, Fig. 8, Claims 1, 2, and 3; EX1003, ¶89.



EX1004, Fig. 8 (annotated)

The raw camera image data captured by the cameras of perception system 172 are transmitted to computing device 110 of autonomous vehicle 100 operating on a road. EX1004, 7:57-8:31, 9:4-10:10, 13:29-15:21, Claim 1, Fig. 8. Computing device 110 of the vehicle 100 receives and processes this image data captured by the cameras (*receiving, by a computing device associated with an autonomous vehicle operating on a road, sensor data captured by a sensor installed on the autonomous vehicle*), where the image data displays the various objects on the road in the vehicle's

environment (e.g., other vehicles, pedestrians, bicyclists) (*the sensor data displaying an object on the road*). *Id.*; EX1003, ¶90.

c. 1[b]

Ross and Cox teach element 1[b] as described below. EX1003, ¶91.

- ***supervised learning based model***

The combination of Ross and Cox teach a *supervised learning based model* as claimed. EX1003, ¶92. Ross’s “intent model” is a prediction model trained using human observations to generate likelihoods for various road user intents. EX1004, 4:33-65, 9:27-10:10, 15:1-16:22; EX1003, ¶92. Cox similarly teaches a machine learning model that generates predictions and is trained with human labeled image data but specifies that its model is “**any form of supervised learning**, including neural networks” and “random forests” and various other model types. EX1005, ¶¶0027-0032, 0060 (emphasis added). Accordingly, combining the teachings of Cox with Ross provides that Ross’s intent model is a supervised learning model, such as a neural network or one of Cox’s other explicitly described supervised learning models, that is trained with human labeled image data (the “Ross-Cox model”), thus teaching a *supervised learning model* as claimed. *Id.*; EX1004, 4:33-65, 9:27-10:10, 15:1-16:22; Section X.A.2; EX1003, ¶92.

- ***providing the sensor data as input to a supervised learning based model,***

Ross describes providing the particular received “raw data” from the cameras of vehicle 100, referred to in Section X.A.3.b (element 1[a]) (*providing the sensor*

data), as the input to the Ross-Cox model (*as input to a supervised learning based model*) to “generate a set of possible intents for a given road user by looking at all the possible actions that the road user could execute.” EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, Claims 1 and 7; EX1005, ¶¶0026-0032; EX1003, ¶93.

- ***the supervised learning based model configured to receive an input sensor data displaying a particular object on a road***

This recitation refers to functionality that the '344 Patent's model is configured to perform, which may be interpreted as occurring during “live” non-training use (referred to as “operational use” hereafter) as well as during training. EX1003, ¶94; EX1001, 11:5-12:60, Figs. 8, 9. In either situation, Ross and Cox teach this limitation. EX1003, ¶94.

The Ross-Cox model (*the supervised learning based model*) is configured to receive input “raw data” from cameras of vehicle 100, including the specific raw camera data discussed in Section X.A.3.b (element [1a]) as well as other raw camera data captured at other times by the vehicle (*configured to receive an input sensor data*). EX1004, 1:30-2:53, 3:39-47, 3:64-4:5, 4:33-5:11, 7:57-8:31, 8:32-52, 9:13-10:10, 15:1-16:52, Fig. 8, Claims 1, 2, 3, and 7; EX1003, ¶95. The camera data pertains to objects corresponding to other road users, such as vehicles, bicycles, and pedestrians on the road (*displaying a particular object on a road*). *Id.* Input raw camera data displaying road objects is provided to the Ross-Cox model in this

manner during training of the model as well as during operation. EX1004, 4:33-65, 9:27-10:10, 15:1-16:22; EX1003, ¶95.

- ***and predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object***

This limitation refers to functionality that the '344 Patent's model is configured to perform, which may be interpreted as occurring during operational use or during training. EX1003, ¶96; EX1001, 11:5-12:60, Figs. 8, 9. In either situation, Ross and Cox teach this limitation. EX1003, ¶96.

Model Functionality During Operational Use

After receiving input raw camera data from cameras of autonomous vehicle 100 (including the specific raw camera data discussed in Section X.A.3.b (element [1a])), the Ross-Cox model predicts for an object (e.g., vehicle, pedestrian, bicyclist) a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” (*predicting an output statistical summary*). EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, 17:43-18:2, Fig. 10, Claims 1 and 7 (parenthetical in original); EX1003, ¶97. This model output (referred to hereinafter as the “Ross predicted data”) is *an output statistical summary* because it is a summary of the various predicted intents and likelihood values for a particular object in input raw camera image data. EX1004, 4:33-5:11, 9:13-10:10,

15:1-16:52, Fig. 10. The likelihood values summarize how likely each of the intents and trajectories of the object is to occur, based on the training data that was used to train the intent model. *Id.* Because these values are a collection of numerical measurements, a POSITA would have recognized that such values form statistics. *Id.*; EX1003, ¶97. Moreover, the likelihood values indicate the relative likelihoods of predictions, where the greatest likelihood value corresponds to the prediction most likely to occur. EX1004, 9:27-40, 15:50-16:52; EX1003, ¶97. The predicted intents and associated likelihood values thus reflect statistics summarizing how an object within the environment of Ross’s autonomous vehicle may behave. EX1003, ¶97.

Cox teaches that the Ross predicted data should reflect *user responses expected to be received responsive to presenting the sensor data to a plurality of users*. EX1003, ¶98. Cox’s purported novelty is using a crowdsourced, “large population” of human annotators to annotate a training image, generating “statistical information” based on the responses received from the human annotators, and training a supervised machine learning model using the statistical information. EX1005, ¶¶0007-0010, 0015-0019, 0025-0035, 0049, 0051, Figs. 1A, 1B. Annotator responses describe a state of mind of subjects displayed in training images. *Id.* The purpose of this training in Cox is to allow the model “to make predictions (based on various query objects) **that are more consistent with the decisions of the human annotators.**” *Id.*, ¶¶0004, 0027-0032 (parenthetical in original, bolding added), 0051. Thus, as applied to Ross in this regard, Cox teaches that the output of the Ross-

Cox model should reflect the likely responses from human annotators, including with respect to the state of mind of subjects in analyzed images. EX1003, ¶98.

The Ross-Cox combination therefore teaches that the Ross predicted data (i.e., the predicted “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” for an object discussed earlier in this section) predicted by the Ross-Cox model (*output statistical summary*), is consistent with the decisions of human annotators during training and thus reflects the distribution of the various human annotator responses received from human annotators (*characterizing a distribution of user responses*). EX1004, 1:55-2:13, 4:33-46, 9:13-10:10, 15:1-21; EX1005, ¶¶0004, 0007-0010, 0015-0019, 0025-0035, 0051; Section X.A.2; EX1003, ¶99. These various human annotator responses thus indicate possible next actions and state of mind of the object (e.g., vehicle, pedestrian, bicyclist) that human annotators would have identified for the object in the raw camera data and responded with when observing the raw camera data input into the Ross-Cox model (*expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object*). *Id.*

Model Functionality During Training

Ross and Cox further teach *predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated*

with the particular object if this limitation is interpreted to refer to functionality the model is configured to perform during model training. EX1003, ¶100.

In Cox, training data in the form of annotation responses and psychometric data from human annotators is utilized for training the model, “resulting in a human-weighted loss function.” EX1005, ¶¶0031-0032, Fig. 4B; EX1003, ¶101. The loss function penalizes misclassification by the model in relation to human annotated training data and thus reduces model error. EX1005, ¶¶0032, 0034. Cox’s model is trained by predicting classifications of training data and optimizing the model such that the difference between the predicted classifications output by the model and the training data is minimized. EX1005, ¶¶0008, 0019, 0029-0038, 0060; Fig. 4B. Thus, as applied to Ross, Cox teaches model training that would have included comparing Ross’s predictions to the human labeled training data to identify prediction errors and optimizing the Ross-Cox model such that the errors are minimized. *Id.*; EX1004, 4:47-65, 7:57-8:31, 9:13-10:10, 15:1-16:52, 17:43-18:2, Fig. 10; Section X.A.2; EX1003, ¶101. This is consistent with the ’344 Patent, which discusses training as similarly calculating (and minimizing) differences between training data and model predictions. EX1001, 12:15-20; EX1003, ¶101.

In view of the above, the Ross-Cox combination teaches the claimed “*predicting an output statistical summary...*” during training, just like during operational use. EX1003, ¶102. In order to optimize the model as taught by Cox, the same type of predictions would be made during training as during operational use.

Id.

d. 1[c]

Ross and Cox teach element 1[c]. EX1003, ¶103. The particular raw camera data from the cameras of vehicle 100 (*sensor data captured by the sensor installed on the autonomous vehicle*) is provided as the input to the trained Ross-Cox model (see Sections X.A.3.b, X.A.3.c (elements 1[a], 1[b])), where the model was trained via the human observations, human annotations, and model optimization as discussed for element 1[b]. *Id.* The trained model is then executed on this input data (*executing the trained supervised learning based model*) to generate a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” (*to generate a statistical summary data*), which is the Ross predicted data, for objects in the particular raw camera data. EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, Claims 1 and 7 (parenthetical in original); Section X.A.3.c (element 1[b]); EX1003, ¶103. Specifically, the trained model is executed on the particular raw camera data from the cameras of vehicle 100 input into the model as discussed in element 1[b] in the “Model Functionality During Operational Use” example discussed for element 1[b].

Id.

The Ross predicted data generated by the trained model forms *a statistical summary data* with respect to the particular raw camera data for the same reasons

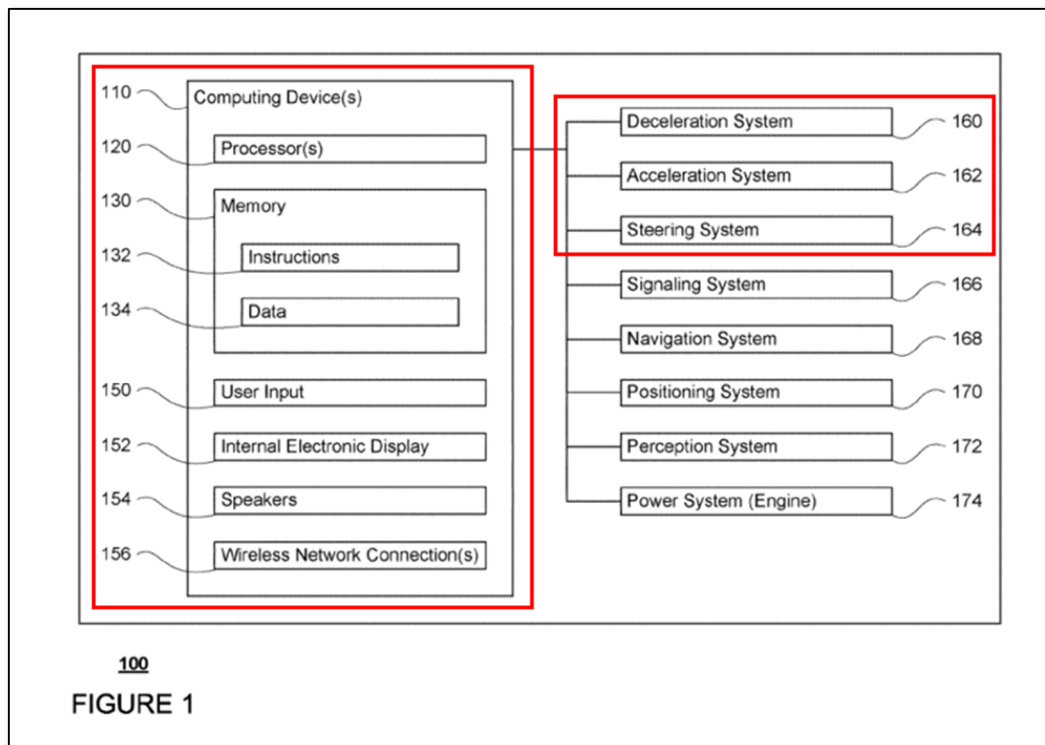
discussed regarding the *statistical summary* recitation in element 1[b]. See Section X.A.3.c (element 1[b]); EX1003, ¶104. The Ross predicted data output in element 1[c] is data specifically tied to the live operation of the autonomous vehicle (*statistical summary data*), which would have the same general characteristics as the model output described generally in element 1[b] (*statistical summary*). *Id.* Moreover, the Ross predicted data for the particular raw camera data from the cameras of vehicle 100 (*statistical summary data*) is *characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users* with respect to the particular raw camera data for the same reasons discussed regarding the *characterizing a distribution* recitation in element 1[b] for both Ross and Cox. *Id.*

The Ross predicted data for the particular raw camera data (*output statistical summary*) would have been consistent with the decisions of human annotators during training and thus reflect the distribution of the various human annotator responses received from human annotators (*characterizing a distribution of user responses*). Section X.A.3.c (element 1[b]); EX1003, ¶105. This is because such human annotator responses would have trained the model and the model would have been optimized with respect to the responses, with errors between the model predictions and annotator responses minimized. *Id.* Moreover, the various human annotator responses indicate possible next actions (and thus state of mind) of the object that human annotators would have identified for the object in the particular raw camera

data and responded with when observing the particular raw camera data (*expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users*). EX1004, 4:47-65, 7:57-8:31, 9:13-10:10, 15:1-16:52, 17:43-18:2, Fig. 10; EX1005, ¶¶0019, 0029-0038, 0060; Fig. 4B; Section X.A.2; Section X.A.3.c (element 1[b]); EX1003, ¶105.

e. 1[d]

Ross and Cox teach element 1[d]. Autonomous vehicle 100's computing device 110 controls the vehicle's deceleration system, acceleration system, and steering system, which controls speed and direction of the vehicle. EX1004, 6:33-59, 10:11-30, Fig. 1; EX1003, ¶106.



EX1004, Fig. 1 (annotated)

Moreover, executing the Ross-Cox model for the particular input raw camera data predicts a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” for an object vehicle, pedestrian, or bicyclist displayed in the camera data (*the generated statistical summary data*). EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, Claims 1 and 7 (parenthetical in original); Sections X.A.3.c, X.A.3.d (elements 1[b], 1[c]). One or more of the predictions with the highest likelihoods are identified and used by the autonomous vehicle’s computing device 110 to “maneuver the vehicle in order to promote the other road user’s intent.” EX1004, 15:50-16:52, 10:11-30, 4:66-5:11. Ross’s “maneuvering” or control of the vehicle 100 is thus based on the predicted “set of possible intents...” (*the generated statistical summary data*) generated by the model (*controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data*). *Id.*, 9:13-10:30, 15:50-16:52; EX1003, ¶107.

4. Claim 2

Ross and Cox render obvious claim 2. EX1003, ¶108. The Ross-Cox model is, for example, a neural network or one of Cox’s other described supervised learning models, e.g., “any form of supervised learning, including neural networks, boosting, bagging, random forests, nearest neighbor algorithms, naive bays classifiers, density estimators, and other forms of statistical regression” (*wherein the supervised*

learning based model is one of: a random forest regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, or a long short-term memory (LSTM) neural network). See Section X.A.3.c (element 1[b]).

5. Claim 3

Ross and Cox render obvious claim 3. EX1003, ¶109. The raw camera image data displays objects on the road in the vehicle's environment including other vehicles, pedestrians, and bicyclists (*wherein the object on the road is one of: a pedestrian, a cyclist, or a motor vehicle*). See Section X.A.3.b (element 1[a]).

6. Claim 4

Ross and Cox render obvious claim 4. EX1003, ¶110. Ross and Cox teach that the various human annotator responses (*user responses*) indicating possible next actions of the object (vehicle, pedestrian, bicyclist) that would have been received responsive to humans observing the raw camera data including the object include actions associated with the object (*wherein the user responses describe an action associated with the object*). EX1004, 4:47-65, 9:41-10:10; see Sections X.A.3.c, X.A.3.d (elements [1b], [1c]). When the object is a vehicle, the actions include the vehicle “changing lanes” as well as “turning at intersections, making a U-turn[], entering or exiting driveways, pulling over, [and] pulling out from a parking spot or parked position” (*the action including one of the object staying in place, changing lanes, or crossing a street*). EX1004, 4:47-65, 9:41-10:10. For an object bicyclist,

the actions include “all of the actions for vehicles as well as crossing crosswalks and ‘jaywalking’ (for instance, crossing a roadway at a location other than an intersection or crosswalk).” *Id.* For an object pedestrian, the actions include “crossing a roadway” and “crossing a roadway in a crosswalk,” for example. *Id.*; EX1003, ¶110.

7. Claim 5

Ross and Cox render obvious claim 5. EX1003, ¶111. Cox teaches that its training involves displaying training images to human annotators for manual annotations/labels. EX1005, ¶0017. Cox describes that “[d]uring and/or after the annotation, psychometric data is also acquired that characterizes the annotation of the training objects by the annotators.” *Id.*, ¶¶0018, 0029, Fig. 3. The psychometric data includes “response times for classifying one or more features, the accuracy of feature classification, and/or the presentation time (i.e., the amount of time presented to each annotator) of one or more training objects.” *Id.*, ¶¶0018-0029 (parenthetical in original); EX1003, ¶111. Cox further describes that “[e]ye-tracking hardware may be used to collect measures suitable for deep annotation such as saccade-to-target accuracy, saccade-to-target latency, number of saccade hops to target, and total number of saccades.” EX1005, ¶0063.

Thus, applying Cox’s teachings to Ross would have provided that the predicted outputs of the Ross-Cox model during training would have been generated according to training using human annotation/label responses, and would have thus been associated with psychometric data for at least some of the human

annotation/label responses (*wherein the statistical data is associated with a parameter of the subset of the plurality of response data*). EX1005, ¶¶0017-0018, 0029; *see* Section X.A.3.c (element 1[b]); EX1003, ¶112. As taught by Cox, the psychometric data would have included response times for annotations, accuracy of the annotation, as well as the amount of time a training image is presented for annotation (*the parameter including at least one of a content of a response, a time associated with entering a response*). *Id.* The psychometric data would have further included information on human annotator eye movements when providing annotation responses, which would have been measured with respect to the display screen displaying the training image that the annotator is viewing (*and a position of an eye of a human observer associated with the response, the position being measured with respect to a display associated with the user interface*). *Id.*; EX1005, ¶0063; EX1003, ¶112. A POSITA would have been motivated to combine Ross and Cox for the reasons discussed in Section X.A.2.

8. Claim 6

Ross and Cox render obvious claim 6. EX1003, ¶113. The combination of Ross and Cox teach *the parameter* in the form of psychometric data including response times for annotations, accuracy of the annotation, as well as the amount of time a training image is presented for annotation. Section X.A.7 (claim 5). Cox further teaches that collected “accuracy and reaction time may be measured on a per-image, per-subject, or per-condition level.” EX1005, ¶¶0025-0029. Cox describes

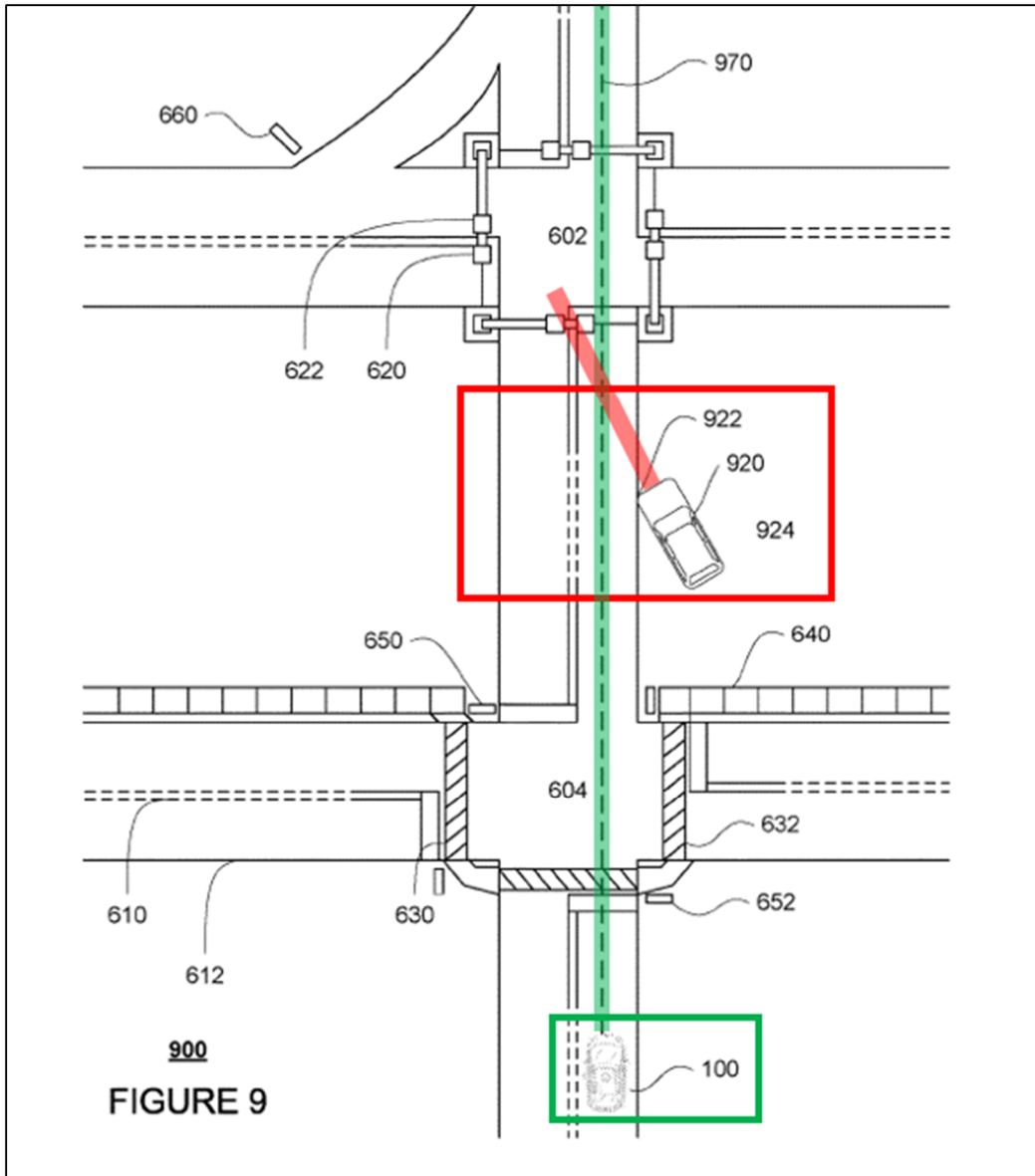
that its model may then use statistical information related to these measurements, such as “mean accuracy” or “mean response time” to refine the training data used to train the supervised machine learning model. *Id.* Thus, by determining a “mean accuracy” and “mean response time” of responses from multiple human annotators, Cox teaches to Ross the association of accuracy and response time parameters with central tendencies reflecting average (mean) accuracy and response time (*the parameter is further associated with at least one of a central tendency . . .*). *Id.*; Section X.A.7 (claim 5); EX1003, ¶113.

9. Claim 7

Ross and Cox render obvious claim 7. EX1003, ¶114. The Ross-Cox model is configured to predict a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values.” EX1004, 1:55-2:13, 4:33-46, 9:13-10:10, 15:1-16:52; EX1005, ¶¶0016, 0056; Sections X.A.3.c, X.A.3.d (elements [1b], [1c]). As discussed with respect to elements 1[b] and 1[c], Cox teaches model training using human annotated training data that provides model predictions are more consistent with human annotation decisions and reflect state of mind of a subject. Sections X.A.3.c, X.A.3.d (elements [1b], [1c]). Thus, due to such training, the predictions by the Ross-Cox model reflect predicted intents that human annotators would have responded with when annotating the image data under analysis by the model, teaching *wherein the supervised learning based model is*

configured to predict user responses. Id. Moreover, the predicted intents are, for an object vehicle or object bicycle, the intent of a vehicle driver or bicyclist riding the bicycle to turn the vehicle/bicycle, change lanes, drive through an intersection, or cross and move into the path of a roadway in which the autonomous vehicle is traveling. *Id.*; EX1004, 2:4-13, 9:13-10:10, 15:1-16:52; EX1003, ¶114. The predicted intents therefore describe *a state of mind of a road user associated with the object* because they describe how a road user would like to operate their vehicle. *Id.*

For example, as shown in annotated Figure 9 below, an intent may be that the driver of vehicle 920 intends that vehicle 920 will “exit from parking lot 924 and cross edge 922,” and thereby move (shown in red) into the path of autonomous vehicle 100 (shown in green) when it is “clear.” EX1004, 16:5-41, Fig. 9. In such a case, the predictions of human responses to viewing the input camera sensor data describe how an annotator would have identified that the driver of vehicle 920 has an intention to move vehicle 920 into the pathway of vehicle 100 when it is clear, and that the driver of vehicle 920 thus has an awareness of the vehicle 100 driving in the roadway they wish to enter so they can enter the roadway safely. *Id.*; EX1003, ¶115.



EX1004, Fig. 9 (in part, annotated)

10. Claim 8

Ross and Cox render obvious claim 8. EX1003, ¶116. *See* Section X.A.9 (Claim 7).

11. Claim 9

Ross and Cox render obvious claim 9. EX1003, ¶117. *See* Section X.A.9

(Claim 7).

12. Claim 10

Ross and Cox render obvious claim 10 for the same reasons discussed with respect to claim 1. EX1003, ¶118.

a. 10[Pre]

See Section X.A.3.a (element 1[Pre]). Moreover, Ross describes that its computing device 110 contains “one or more processors 120” and “memory 130,” where “the memory 130 stores information accessible by the one or more processors 120, including instructions 132 and data 134 that may be executed or otherwise used by the processor 120” to effectuate the operation of computing device 110 (*[a] non-transitory computer readable storage medium storing instructions that when executed by one or more processors, cause the one or more processors to perform the steps*). EX1004, 5:13-6:5, 6:33-59, Fig. 1, Claim 11; EX1003, ¶119.

b. 10[a]

See Section X.A.3.b (element 1[a]).

c. 10[b]

See Section X.A.3.c (element 1[b]).

d. 10[c]

See Section X.A.3.d (element 1[c]).

e. 10[d]

See Section X.A.3.e (element 1[d]).

13. Claim 11

Ross and Cox render obvious claim 11 for the same reasons discussed above in Section X.A.4 (claim 2). EX1003, ¶124.

14. Claim 12

Ross and Cox render obvious claim 12 for the same reasons discussed above in Section X.A.5 (claim 3). EX1003, ¶125.

15. Claim 13

Ross and Cox render obvious claim 13 for the same reasons discussed above in Section X.A.6 (claim 4). EX1003, ¶126.

16. Claim 14

Ross and Cox render obvious claim 14 for the same reasons discussed above in Section X.A.7 (claim 5). EX1003, ¶127.

17. Claim 15

Ross and Cox render obvious claim 15 for the same reasons discussed above in Section X.A.8 (claim 6). EX1003, ¶128.

18. Claim 16

Ross and Cox render obvious claim 16 for the same reasons discussed above in Section X.A.9 (claim 7). EX1003, ¶129.

19. Claim 17

Ross renders obvious claim 17 for the same reasons discussed above in Section X.A.10 (claim 8). EX1003, ¶130.

20. Claim 18

Ross renders obvious Claim 18 for the same reasons discussed above in Section X.A.11 (claim 9). EX1003, ¶131.

21. Claim 19

Ross and Cox render obvious claim 19 for the same reasons discussed with respect to claims 1 and 10. EX1003, ¶132.

a. 19[Pre]

See Sections X.A.3.a (element 1[Pre]), X.A.12.a (element 10[Pre]).

b. 19[a]

See Sections X.A.3.b (element 1[a]).

c. 19[b]

See Sections X.A.3.c (element 1[b]).

d. 19[c]

See Sections X.A.3.d (element 1[c]).

e. 19[d]

See Sections X.A.3.e (element 1[d]).

22. Claim 20

Ross renders obvious Claim 20 for the same reasons discussed above in Section X.A.9 (claim 7). EX1003, ¶138.

23. Claim 21

Ross renders obvious Claim 21 for the same reasons discussed above in Section X.A.10 (claim 8). EX1003, ¶139.

24. Claim 22

Ross renders obvious Claim 22 for the same reasons discussed above in Section X.A.11 (claim 9). EX1003, ¶140.

B. Ground 2: Claims 1-22 are unpatentable under 35 U.S.C. § 103 over Ross, Cox, and Ogale

Ross and Cox render obvious the Challenged Claims for the reasons discussed in Section X.A. However, to the extent that the Patent Owner argues, or the Board finds, that the predictions generated by the Ross-Cox model do not correspond to (1) *user responses expected to be received responsive to presenting the sensor data to a plurality of users* or (2) *user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users* (as recited by elements 1[b] and 1[c], *infra*), the combination of Ross, Cox, and Ogale teaches these elements. EX1003, ¶141.

1. Summary of Prior Art

a. Ogale–EX1010

Ogale—another Waymo patent—is generally directed to “autonomous vehicles” and “generating object predictions using a neural network.” EX1010, Abstract, 1:5. Ogale explains that for a neural network trained using human labeled training data prepared by “human participants,” the prediction generated by the neural network based on input sensor data “predicts how a human would have labeled the input sensor data.” *Id.*, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9. Ogale’s

disclosure aligns with the '344 Patent's purported invention, which includes a model predicting information "characterizing a distribution of user responses expected to be received responsive to presenting [] sensor data to a plurality of users." EX1001, Claim 1. Ogale is analogous art to the '344 Patent, from the same field of endeavor as the patent (e.g., predictive model techniques), and reasonably pertinent to the particular problem that the patent was trying to solve (e.g., accounting for moving objects when navigating an autonomous vehicle in an environment). EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9; EX1001, 1:15-17, 1:53-56; EX1003, ¶142.

2. Motivation to Combine Ross, Cox, and Ogale

A POSITA would have been motivated to combine Ross and Cox as discussed in Section X.A.2. A POSITA would have been further motivated to combined Ross and Cox with Ogale. EX1003, ¶143. The combination would have provided that the Ross-Cox model, trained using human annotated training data, would have provided predictions that "predict[] how a human would have labeled the input sensor data." *See* Section X.A.2; EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9.

The Ross-Cox combination already describes model training using "human observations" that are manually marked as correct and human labeled training data that describes state of mind of a training object. *See* Section X.A.2. The Ross-Cox combination further provides that the Ross-Cox model "function[s] in a more 'human-like' or 'polite' way" (EX1004, 5:7-12) and "mimic[s] human performance" to ensure that model predictions that are "more consistent with the decisions of the

human annotators” (EX1005, ¶¶0004, 0029-0032). *See* Section X.A.2. Ogale further elaborates that such a model, trained using human participants, would have generated predictions based on input sensor data that “predicts how a human would have labeled the input sensor data.” EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9; EX1003, ¶144. Ogale thus merely provides implementation details for the Ross-Cox combination. *Id.* This explicitly provides that the result of the combination (i.e., model predictions that “predict[] how a human would have labeled the input sensor data.”) would have been predictable to a POSITA, and also had a reasonable expectation of success for at least these same reasons. *Id.*

Moreover, a POSITA would have looked to Ogale. EX1003, ¶145. Ross, Cox, and Ogale are all directed to machine learning technologies and involve gathering training data, training a model with the training data, and executing the model, providing that a POSITA seeking to implement the Ross-Cox combination would have looked to Cox. *Id.*; *see* Sections X.A.1.a, X.A.1.b, X.B.1.a. Ross, Cox, and Ogale overlap in that they are all related to machine learning models within the context of autonomous vehicles and human annotated training data. EX1004, 1:30-31, 3:39-42, 3:64-5:5, 9:30-10:10; EX1005, ¶¶0027-0032, 0060; EX1010, Abstract, 1:5. And as noted above, each of Ross, Cox, and Ogale are directed to techniques that provide for a predictive model that outputs more “human” predictions, further motivating the combination. EX1003, ¶145. Combining Ogale with Ross and Cox would also have improved how “‘human-like’ or ‘polite’” the model behaves—

which is a stated goal of Ross—because, as noted above, Ogale teaches model predictions that reflect “how a human would have labeled the input sensor data” and are thus similarly human-like. *Id.*; EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9; EX1004, 5:7-12. This would have improved the Ross-Cox model’s predictions, and in turn, the autonomous vehicle’s safe navigation around other humans/vehicles near the autonomous vehicle. EX1003, ¶145.

Therefore, for the reasons discussed above, a POSITA would have been motivated to combine Ross, Cox, and Ogale. EX1003, ¶146. In addition, as demonstrated above, the combination would have merely involved combining prior art elements (e.g., Ross’s model for predicting intents; Cox’s model training and optimization techniques; Ogale’s prediction model trained using human labels) according to known methods (known software techniques with respect to machine learning model training) to yield the predictable results of model predictions that “predict[] how a human would have labeled the input sensor data.” *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007); EX1003, ¶146; *see* above analysis in this Section X.B.2; EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9. The combination would have further been use of a known technique (known machine learning model training using human labeling as taught by Ogale) to improve similar devices (Ross-Cox model for predicting intents) in the same way (using Ogale’s aforementioned techniques). EX1003, ¶146. And, because Ross, Cox, and Ogale all use the same type of human-labeled data to train models to be more human-like (as

discussed above), combining Ogale with Ross and Cox would have required only minimal modifications that were well within the skill of a POSITA, and for which a POSITA would have had a reasonable expectation of success. *Id.*

3. Claim 1

Ross, Cox, and Ogale render obvious claim 1. EX1003, ¶147.

a. 1[Pre] and 1[a]

Ross, Cox, and Ogale teach elements 1[Pre] and 1[a] for the same reasons discussed in Section X.A.3.a and X.A.3.b (elements 1[Pre] and 1[a]). EX1003, ¶148.

b. 1[b]

Ross, Cox, and Ogale teach element 1[b]. EX1003, ¶149.

The Ross-Cox combination teaches the Ross-Cox model (*supervised learning based model*). See Section X.A.3.c (element 1[b]). Moreover, Ross describes that raw camera data including an object (e.g., another vehicle, bicyclist, pedestrian) on a road is provided as an input to the model (*providing the sensor data as input to a supervised learning based model*), and Ross further explains that the model is configured to receive this specific raw camera data (discussed in Sections X.A.3.a (elements 1[Pre]), X.A.3.b (element 1[a])) as well as other raw camera data captured at other times by the vehicle, for example during training of the model and when the model is in operational use (*the supervised learning based model*

configured to receive an input sensor data displaying a particular object on a road).

See X.A.3.c (element 1[b]); EX1003, ¶149.

Ross and Cox teach *predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object* in the operational and training contexts for the reasons discussed in Section X.A.3.c (element 1[b]). But to the extent the Patent Owner argues or the Board finds that Ross and Cox do not teach this recitation, the combination of Ross, Cox, and Ogale does. EX1003, ¶150.

Like Ross and Cox, Ogale describes that its prediction model for an autonomous vehicle that is trained using human labeled training data and in particular, labeled training data provided by multiple “human participants.” EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9; EX1003, ¶151. By training Ogale’s model with such human labeled data, the model outputs predictions based on input sensor data that “predicts how a human would have labeled the input sensor data.” *Id.*

Thus, as applied to the Ross-Cox combination, in both the non-training and training contexts discussed in Section X.A.3.c (element 1[b] of Ground 1), Ogale teaches training the Ross-Cox intent model with human observation data and human annotator responses, expanding on the disclosure taught by Ross individually and the combination of Ross and Cox (*See id.*), such that the Ross-Cox model predicts how a human would have labeled the input sensor data. EX1010, 3:29-37, 6:51-7:12,

13:7-11, 13:30-14:9 (bracketing added); EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, Claims 1 and 7 (parenthetical in original); EX1005, ¶¶0004, 0007-0010, 0015-0019, 0025-0035, 0051; Section X.A.3.c (element 1[b] of Ground 1); EX1003, ¶152. The Ross-Cox model’s predicted “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” (*predicting an output statistical summary*) based on input raw camera data characterizes how “multiple human participants” observing training data and labeling the training data would have responded to the input raw camera data and predicted possible next actions (and trajectories) (*id.*); indeed, the predictions would have reflected “how a human would have labeled the input [raw camera] data” as taught by Ogale (*characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object*). *Id.* This is because Ogale makes clear through its disclosure that human annotated training data used to train a predictive model provides that trained model outputs reflect how human annotators would have annotated the input data. EX1010, 3:29-37, 6:51-7:12, 13:7-11, 13:30-14:9; EX1003, ¶152.

c. 1[c]

Ross, Cox, and Ogale teach element 1[c]. EX1003, ¶153. Ross discloses that the particular raw camera data from the cameras of vehicle 100 (*sensor data*

captured by the sensor installed on the autonomous vehicle) is provided as the input to the trained Ross-Cox model (*see* Sections X.A.3.b, X.A.3.c (elements 1[a], 1[b]) of Ground 1), which was trained via the human observations, human annotations, and model optimization as discussed for element 1[b] in Ground 1 and element 1[b] in Ground 2 with respect to Ross, Cox, and Ogale. *Id.*; *see also* Sections X.A.3.a-X.A.3.c (elements 1[Pre], 1[a], and 1[b] of Ground 1), X.B.3.a-X.B.3.c. The trained model is then executed on this input data (*executing the trained supervised learning based model*) to generate a “set of possible intents or hypotheses identifying possible next actions (and predicted trajectories), a given point in time when the action is likely to occur, and associated likelihood values” (*to generate a statistical summary data*), which is the Ross predicted data, for objects in the particular raw camera data. EX1004, 1:30-2:53, 3:64-4:5, 4:33-5:11, 7:57-8:31, 9:13-10:10, 15:1-16:52, Claims 1 and 7 (parenthetical in original); Section X.A.3.c (element 1[b] of Ground 1); EX1003, ¶153. Specifically, the trained model is executed on the particular raw camera data from the cameras of vehicle 100 input into the model as in the “Model Functionality During Operational Use” example for element 1[b] of Ground 1. *See* Section X.A.3.c (element 1[b]); EX1003, ¶153. As discussed in Section X.B.3.b (element 1[b] of Ground 2), the Ross, Cox, and Ogale combination further teaches that the model output characterizes how “multiple human participants” observing training data and labeling the training data would have responded to the input raw camera data and predicted possible next actions (and trajectories), and in turn “how

a human would have labeled the input [raw camera] data” as taught by Ogale. Section X.B.3.b (element 1[b] of Ground 2).

The Ross predicted data generated by the trained model forms *a statistical summary data* with respect to the particular raw camera data from the cameras of vehicle 100 for the same reasons discussed regarding the *statistical summary* recitation in element 1[b] of Ground 2. *See* Sections X.B.3.b (element 1[b]), X.A.3.c. Moreover, the Ross predicted data for the particular raw camera data from the cameras of vehicle 100 (*statistical summary data*) is *characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users* with respect to the particular raw camera data for the same reasons discussed regarding the *characterizing a distribution* recitation in element 1[b] of Ground 2 for Ross, Cox, and Ogale. *Id.* The Ross predicted data generated for the particular raw camera data (*output statistical summary*) characterizes how “multiple human participants” observing training data and labeling the training data would have responded to the particular raw camera data and predicted possible next actions (and trajectories and thus state of mind) for the object (e.g., a vehicle, pedestrian, bicyclist) in the data (*id.*); the model would have been trained using human annotated data and thus generated predictions that reflect “how a human would have labeled the input [raw camera] data” as taught by Ogale (*characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the*

sensor installed on the autonomous vehicle to users). Sections X.B.3.a-X.B.3.b (elements 1[Pre]-1[b] of Ground 2), X.A.3.c (element 1[b] of Ground 1); EX1003, ¶154.

d. 1[d]

Ross, Cox, and Ogale teach element 1[d] for the same reasons discussed in Sections X.A.3.e (element [1d]), X.B.3.b (element 1[b]), and X.B.3.c (element 1[c]). EX1003, ¶155.

4. Claims 10 and 19

Ross, Cox, and Ogale render obvious Claims 10 and 19. EX1003, ¶156. Ross, Cox, and Ogale teach elements 10[b], 10[c], and 10[d], and 19[b], 19[c], and 19[d] for the same reasons discussed above in Sections X.B.3.b (element 1[b]), X.B.3.c (element 1[c]), and X.B.3.d (element 1[d]). EX1003, ¶156.

Ross, Cox, and Ogale teach the remaining elements of claims 10 and 19 for the same reasons discussed in Sections X.A.12 (claim 10) and X.A.21 (claim 19). EX1003, ¶157.

5. Claims 2-9, 11-18, and 20-22

Ross, Cox, and Ogale render obvious claims 2-9, 11-18, and 20-22 for the same reasons discussed in Sections X.A.4-X.A.11, X.A.13-X.A.20, and X.A.22-X.A.24. EX1003, ¶158.

C. Ground 3: Claims 1-4, 7-13, and 16-22 are unpatentable under 35 U.S.C. § 103 over Zhang

1. Summary of the Prior Art

a. Zhang–EX1006

Zhang is generally directed to a “prediction-based system and method of trajectory planning of autonomous vehicles.” EX1006, Abstract. Zhang describes an autonomous vehicle 105 that implements a trained neural network to analyze image data of the environment around the vehicle, captured by cameras of the autonomous vehicle, and determine a “set of predicted trajectories and confidence levels” of moving vehicles and other “dynamic objects” in the image data. *Id.*, ¶¶0019, 0025, 0029-0030, 0047, 0049-0065, Figs. 1, 3, 4. The neural network is trained using “real world, human behavior data” that represents “human driving behaviors,” and the training data includes images of driving behavior that are labeled by human labelers. EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 8, 9, 10.

Zhang is analogous art to the '344 Patent, from the same field of endeavor as the patent (e.g., predictive model techniques), and reasonably pertinent to the particular problem that the patent was trying to solve (e.g., accounting for moving objects when navigating an autonomous vehicle in an environment). EX1006, ¶¶0019, 0025, 0029-0030, 0047, 0049-0065, Figs. 1, 3, 4; EX1001, 1:15-17, 1:53-56; EX1003, ¶160.

2. Claim 1

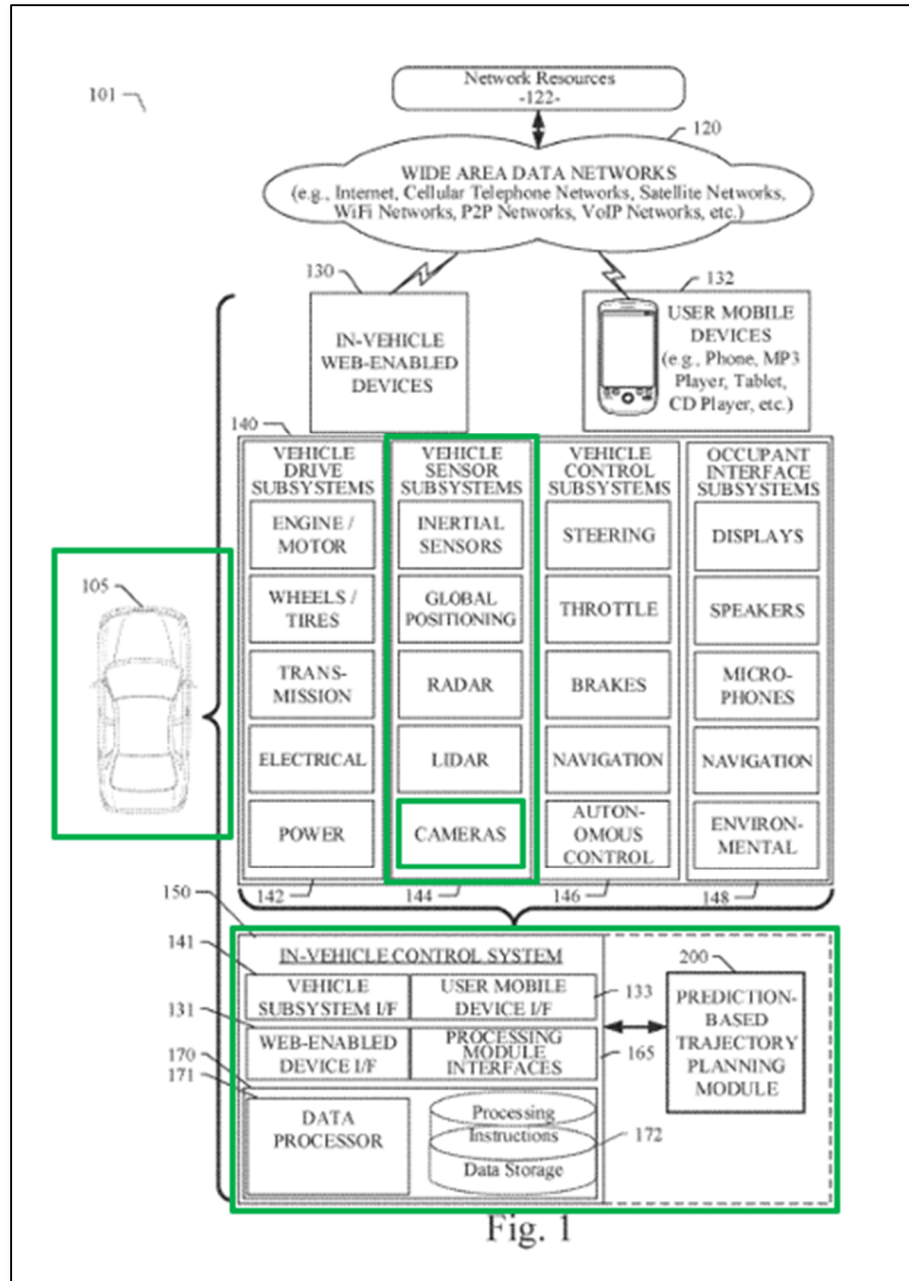
Zhang renders obvious Claim 1. EX1003, ¶161.

a. 1[Pre]

To the extent the preamble is limiting, Zhang teaches it. EX1003, ¶162. Zhang describes techniques for “trajectory planning, vehicle control systems, and autonomous driving systems,” including use of “a prediction-based method” (*[a] computer-implemented method*). EX1006, ¶¶0003, 0006, 0018-0032, 0036, 0052-0061, 0073, Figs. 1, 5, 6, 9, 11; EX1003, ¶162.

b. 1[a]

Zhang teaches element 1[a]. Zhang describes a vehicle 105, (a car) that operates in a fully autonomous mode on a “roadway.” EX1006, ¶¶0006, 0018-0026, 0050, 0073, Figs. 1, 3, 11, claim 1, claim 19; EX1003, ¶163. The vehicle includes an “in-vehicle control system 150” having a “prediction-based trajectory planning module 200” that is implemented by a “computing system” (*a computing device associated with an autonomous vehicle operating on a road*). *Id.* “[O]ne or more cameras or image capture devices” are attached to vehicle 105 (*a sensor installed on the autonomous vehicle*) and “sense information about an environment” of vehicle 105 to detect “objects” in the environment. EX1006, ¶¶0019, 0029-0030, 0049-0051, 0059, Figs. 1, 3, 4; EX1003, ¶163.



EX1006, Fig. 1 (annotated)

The in-vehicle control system 150 receives the image data captured by the cameras (receiving, by a computing device associated with an autonomous vehicle operating on a road, sensor data captured by a sensor installed on the autonomous vehicle). EX1006, ¶¶0019, 0029-0030, 0049-0051, 0058-0059, Figs. 1, 3, 4; EX1003,

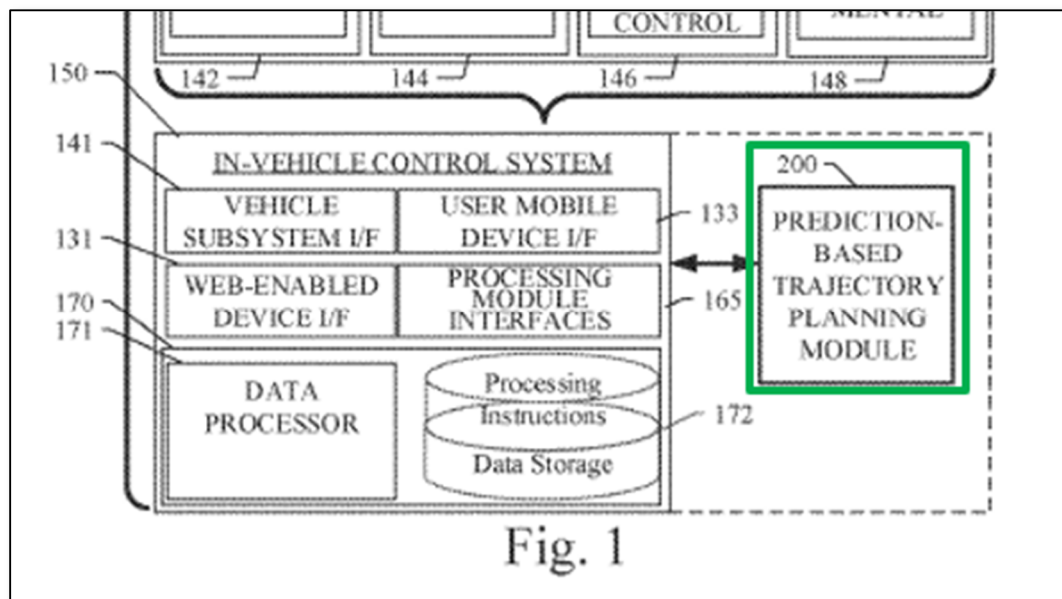
¶164. The image data displays objects such as “moving vehicles, dynamic objects, or other objects in the proximate vicinity of the vehicle 105” that are on the roadway (*the sensor data displaying an object on the road*). *Id.* This reception of image data occurs during both a “training phase” and an “operational phase” in Zhang’s system. EX1006, ¶¶0019, 0049-0054, 0058-0065, Figs. 1, 2, 3, 7; EX1003, ¶164.

c. 1[b]

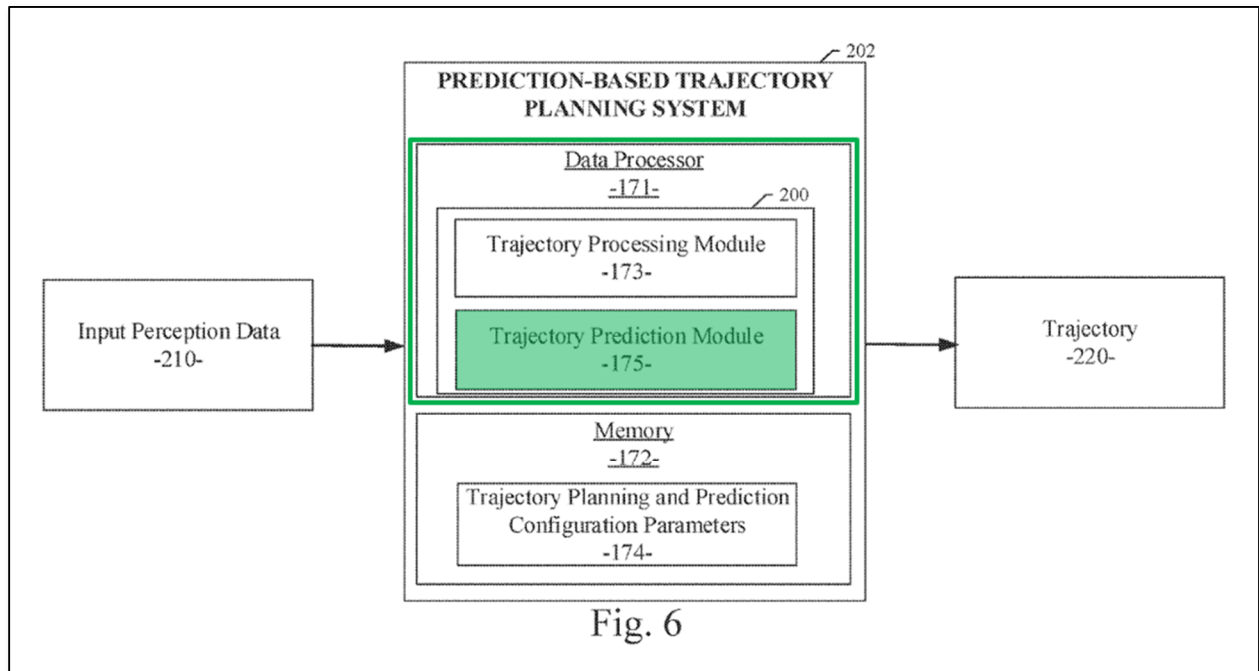
Zhang teaches element 1[b]. EX1003, ¶165.

- ***supervised learning based model***

Zhang discloses prediction-based trajectory planning module 200 includes a “trajectory prediction module 175” implemented as a trained “neural network” (*supervised learning based model*). EX1006, ¶¶0045-0059, 0064, Figs. 1, 3, 6, claims 4, 5; EX1003, ¶166.



EX1006, Fig. 1 (in part, annotated)



EX1006, Fig. 6 (annotated)

Zhang’s neural network model is trained by human labeled training data “to produce [a] configured output based on training data provided to the neural network or other machine learning system in a training phase.” EX1006, ¶¶0045-0059, 0064, Figs. 1, 3, 6, claims 4, 5. Moreover, a “loss function” is used to “examine and correct the results of the training provided to the trajectory prediction module 175.” *Id.* Loss functions are well known mathematical functions used in model training that measure how well a model’s predicted output matches a known correct output, thereby quantifying model error. *Id.*; EX1003, ¶167. These model training techniques disclosed by Zhang thus demonstrate that its neural network is supervised to ensure that its outputs are accurate in view of training data. *Id.* Zhang therefore

teaches a *supervised learning based model*. *Id.*

Zhang's neural network forming a *supervised learning based model* also aligns with claim 2 of the '344 Patent itself, which recites that the *supervised learning based model* is a *simple neural network*. EX1006, ¶¶0045-0059, 0064, Figs. 1, 3, claims 4, 5; EX1003, ¶168; EX1001, claim 2.

- ***providing the sensor data as input to a supervised learning based model***

Image data from the cameras of Zhang's autonomous vehicle is part of "perception data 210" that is provided as an input to the trajectory prediction module 175 implemented as a neural network (*providing the sensor data as input to a supervised learning based model*) EX1006, ¶¶0045-0059, 0064, Figs. 1, 3, claims 4, 5; Section X.C.2.b (element 1[a]); EX1003, ¶169. Image data is provided to Zhang's neural network during both a "training phase" of the model and an "operational phase" where the trained model is deployed to navigate the autonomous vehicle. *Id.*

- ***the supervised learning based model configured to receive an input sensor data displaying a particular object on a road***

The '344 Patent limitation of *the supervised learning based model configured to receive an input sensor data displaying a particular object on a road* may be interpreted to refer to functionality the model is configured to perform during "live" non-training use (referred to as "operational use" hereafter), as well as during a training. EX1003, ¶170. In either situation, Zhang teaches the limitation. *Id.*

As noted above, Zhang discloses that its neural network receives image data

from the cameras of Zhang’s autonomous vehicle, where the image data displays objects such as “moving vehicles, dynamic objects, or other objects in the proximate vicinity of the vehicle 105” (*the supervised learning based model configured to receive an input sensor data displaying a particular object on a road*). EX1006, ¶¶0045-0059, 0064, Figs. 1, 3, claims 4, 5; Section X.C.2.b (element 1[a]); EX1003, ¶171. The data is received during both training and operational use. *Id.*

- ***predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object***

The ’344 Patent limitation *predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object* may be interpreted to refer to functionality the model is configured to perform during operational use as well as during training. EX1003, ¶172. In either situation, Zhang teaches the limitation. *Id.*

Model Functionality Operational Use

Zhang’s neural network of trajectory prediction module 175 is executed to generate a “set of predicted trajectories and confidence levels” for each object (e.g., a moving vehicle) in the image data, where the confidence levels “correspond[] to a likelihood” of the object to “perform the predicted behavior.” EX1006, ¶¶0025, 0058. Zhang’s “set of predicted trajectories and confidence levels” teaches *an output*

statistical summary. Id. The set of predicted trajectories and confidence levels summarize the various trajectories than an object in image data might take and the confidences associated with each prediction. *Id.* This provides a statistical distribution and organization reflecting how an object is predicted to travel within the environment of Zhang’s autonomous vehicle, and how likely each prediction is. *Id.*; EX1003, ¶173. Because these values are a collection of numerical measurements, a POSITA would have recognized that such values form statistics. *Id.*

Zhang further teaches that the “set of predicted trajectories and confidence levels” (*output statistical summary*) is *characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object* as claimed. *Id.*, ¶0058; EX1003, ¶174. Zhang describes that its neural network of trajectory prediction module 175 is trained by training data that is “real world, human behavior data” that represents “human driving behaviors.” EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10. Zhang explains that “[b]ecause the trajectory prediction module 175 is trained in the training phase using real world, human behavior data, the predicted behavior and trajectories of vehicles or objects produced by the trajectory prediction module 175 are closely correlated to the actual behavior and trajectories of vehicles in real world environments with human drivers.” *Id.* Zhang’s training data includes images that are labeled by human labelers to “associate a label having behavior and direction information with each instance of a

vehicle in the training data.” *Id.* The labeled training represents different driving scenarios and includes “human driver behavior data related to how drivers typically react to actions performed by proximate vehicles.” *Id.* A “loss function” is used to “examine and correct the results of the training.” *Id.*, ¶0049. A POSITA would have understood that loss functions are used in machine learning to quantify the error between the correct human labels of training images and a desired model output, and thus allows for adjustment of the model to reduce the error. EX1003, ¶174.

Thus, when Zhang’s neural network predicts a “set of predicted trajectories and confidence levels” for each object (e.g., moving vehicle) in the image data, the model is predicting the driving behaviors (e.g., trajectories) that human labelers would have identified for the object in the image data and responded with by labeling in the data when presented with the data. EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10; EX1003, ¶175. The set of confidence values output by the model are the likelihoods of the behaviors/trajectories occurring based on the input camera data of the object. EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10; EX1003, ¶175. The model predicts these behaviors by being trained on the training data of human labels from humans observing training images of human driving behavior, with the training (and use of a loss function) reducing the error between human labels and the predicted behaviors/trajectories. *Id.* The “set of predicted trajectories and confidence levels” (*output statistical summary*) therefore reflects the distribution of the various human

labeler responses (*characterizing a distribution of user responses*) indicating possible trajectories/behaviors of the object (e.g., a moving vehicle) that human labelers would have identified for the object in the image data and responded with by labeling in the data when presented with the data (*expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object*). EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10; EX1003, ¶175.

Model Functionality During Training

In the training context, Zhang performs the same functionality as discussed above in the operational use example. *See* prior analysis in this Section X.C.2.c (element 1[b]), *supra*. As noted above, Zhang’s neural network of trajectory prediction module 175 is trained during a training phase by training data, including camera images, that is “real world, human behavior data” that represents “human driving behaviors” and is human labeled, and a “loss function” is used to “examine and correct the results of the training.” *Id.*; EX1006, Abstract, ¶¶0025, 0043-0049, 0051-0065, claims 1, 4, 5, Figs. 2, 4, 5, 9, 10; EX1003, ¶176. Moreover, Zhang also teaches that its model is refined and corrected during training. *Id.* Refining the model allows the model predictions to accurately reflect the distribution of the various human labeler responses indicating possible trajectories/behaviors of the object (e.g., a moving vehicle) that human labelers would have identified for the object in the image data and responded with by labeling that data when presented with it. *Id.*

The foregoing confirms that Zhang’s “set of predicted trajectories and confidence levels” described above for operational use are also generated during training. EX1006, ¶0058; EX1003, ¶177. As noted above, a POSITA would have understood that loss functions are used in machine learning to quantify the error between the correct human labels of training images and a desired model output, and thus allows for adjustment of the model to reduce the error. *Id.* Zhang therefore teaches *predicting an output statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object* in the training context for the same reasons discussed above in operational use. See prior analysis of this Section X.C.2.c (element 1[b]), *supra*; EX1003, ¶177.

d. 1[c]

Zhang teaches, element 1[c]. EX1003, ¶178. Zhang discloses that the particular camera data from the cameras of vehicle 105 (*sensor data captured by the sensor installed on the autonomous vehicle*) is provided as the input to the Zhang’s trained neural network (*see* Sections X.C.2.b, X.C.2.c (elements 1[a], 1[b])), which was trained using the human annotated behavior training data and loss function as discussed for element 1[b]. *Id.* The trained model is then executed on this input data (*executing the trained supervised learning based model*) to generate a “set of predicted trajectories and confidence levels” for each object (e.g., moving vehicle) in the image data (*to generate a statistical summary data*). EX1006, ¶¶0025, 0058;

Section X.C.2.c (element 1[b]); EX1003, ¶178. For example, the trained model is executed on the particular image data from the cameras of vehicle 105 input into the neural network as discussed in element 1[b] in the “Model Functionality During Operational Use” example discussed for element 1[b]. *See* Section X.C.2.c (element 1[b]); EX1003, ¶178.

Zhang’s “set of predicted trajectories and confidence levels” forms *a statistical summary data* with respect to the particular input image data for the same reasons discussed regarding the *statistical summary* recitation in element 1[b]. *See* Section X.C.2.c (element 1[b]); EX1003, ¶179. Moreover, the “set of predicted trajectories and confidence levels” for the particular input image data from the cameras of vehicle 105 (*statistical summary data*) is *characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users* with respect to the particular input image data for the same reasons discussed regarding the *characterizing a distribution* recitation in element 1[b] for Zhang. *Id.* Zhang’s generated “set of predicted trajectories and confidence levels” for the particular input image data (*output statistical summary*) reflects the distribution of the various human labeler responses (*characterizing a distribution of user responses*) indicating possible trajectories/behaviors of the object (e.g., a moving vehicle) that human labelers would have identified for the object in the particular input image data and responded with by labeling in the data when presented with the data (*expected to be*

received responsive to presenting the sensor data to a plurality of users, the user responses associated with the particular object). Section X.C.2.c (element 1[b]); EX1003, ¶179.

e. 1[d]

Zhang teaches element 1[d]. EX1003, ¶180. Autonomous vehicle 105 generates a “proposed trajectory” for autonomous vehicle 105 itself to follow on the road. EX1006, ¶¶0057-0061, Fig. 8; *id.*, 0019, 0044, 0047, 0050-0051, 0054, Figs. 3, 4. The proposed trajectory of vehicle 105 is analyzed to determine if it will “conflict with any of the predicted trajectories” of the “set of predicted trajectories and confidence levels” (*generated statistical summary data*) for objects in the particular input image data, and if there is a conflict, the proposed trajectory of autonomous vehicle 105 is changed to eliminate the conflict. *Id.*, ¶0060, Fig. 8; *see* Section X.C.2.d (element 1[c]). Thereafter, autonomous vehicle 105 is “control[led]” such that the “vehicle is directed along the proposed trajectory, which will not be in conflict with any of the proximate vehicles.” EX1006, ¶0060, Fig. 8. Thus, because the control of vehicle 105 is based on a determination that the proposed trajectory does not conflict with any of the “set of predicted trajectories and confidence levels” (*generated statistical summary data*), the control of vehicle 105 on the road to follow the proposed trajectory is based on the “set of predicted trajectories and confidence levels” (*controlling the operation of the autonomous vehicle on the road based on the generated statistical summary data*). EX1006, ¶¶0057-0061, Fig. 8; *see also id.*,

¶¶0054 (explaining autonomous vehicle 105 is controlled based on a predicted trajectory “to more efficiently and safely control the host vehicle 105”), 0019, 0044, 0047, 0050-0051, Figs. 3, 4; EX1003, ¶180.

3. Claim 2

Zhang renders obvious claim 2. EX1003, ¶181. Zhang discloses that its model includes neural networks (*wherein the supervised learning based model is one of: a random forest regressor, a support vector regressor, a simple neural network, a deep convolutional neural network, a recurrent neural network, or a long short-term memory (LSTM) neural network*). See Section X.C.2.c (element 1[b]).

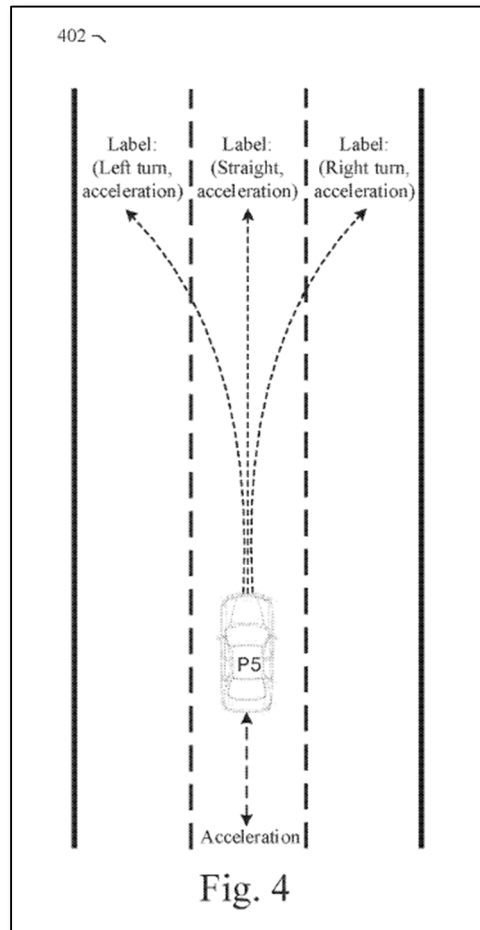
4. Claim 3

Zhang renders obvious claim 3. EX1003, ¶182. The input perception data 210 from cameras of Zhang’s vehicle displays objects on the road in the vehicle’s environment including other “moving vehicles and dynamic objects” (*wherein the object on the road is one of: a pedestrian, a cyclist, or a motor vehicle*). See Section X.C.2.b (element 1[a]).

5. Claim 4

Zhang renders obvious claim 4. EX1003, ¶183. The possible trajectories/behaviors of the object (e.g., a moving vehicle) that human labelers would have identified for the object in the input image data when presented with the image include a “left turn” of an object vehicle into a left lane of a road or a “right

turn” of an object vehicle into a right lane of a road, as illustrated by Figure 4 (wherein the user responses describe an action associated with the object, the action including one of the object...changing lanes....). EX1006, ¶0051, Fig. 4; Sections X.C.2.c, X.C.2.d (elements 1[b], 1[c]); EX1003, ¶183.



EX1006, Fig. 4

6. Claim 7

Zhang renders obvious claim 7. EX1003, ¶184. Zhang’s neural network of trajectory prediction module 175 is executed to generate a “set of predicted trajectories and confidence levels” for each object (e.g., moving vehicle and its

associated driver road user) in the input image data, where the confidence levels “correspond[] to a likelihood” of the object to “perform the predicted behavior” indicated by the trajectories. EX1006, ¶¶0025, 0058; EX1003, ¶184. Human labeler responses of “behavior” in training images are used to train the neural network. EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10; Section X.C.2.c (element 1[b]). Therefore, when Zhang’s neural network (*the supervised learning based model*), trained by the human labeled data noted above, predicts the “set of predicted trajectories and confidence levels” for a moving vehicle and its associated road user driver with each trajectory representing behavior of the moving vehicle and driver, the model is predicting the human labeler responses describing this behavior that would have been received when the labeler is presented with the input image data (*predict user responses describing a state of mind of a road user associated with the object*). EX1006, Abstract, ¶¶0025, 0047, 0049, 0051-0065, claims 1, 4, 5, Figs. 4, 5, 9, 10; Section X.C.2.c (element 1[b]); EX1003, ¶184.

The behavior (to which predicted human labeler responses relate to) indicates, for example, an intent of a vehicle driver to drive their vehicle into the path of autonomous vehicle 105, such as by making a “left turn” or “right turn” into the vehicle 105’s path, because Zhang’s system identifies predicted trajectories of such vehicles in autonomous vehicle 105’s environment and determines whether their predicted trajectories conflict with the trajectory of autonomous vehicle 105. EX1006, ¶¶0051, 0057-0061, Figs. 4, 8. Such behavior also indicates an awareness

of the driver as to the autonomous vehicle 105, with a predicted trajectory that the driver will turn their vehicle into vehicle 105's path indicating that the driver has a low level of awareness of vehicle 105 or has a higher level of awareness of vehicle 105 and believes it can safely make the turn. *Id.*; EX1003, ¶185.

7. Claim 8

Zhang renders obvious claim 8. EX1003, ¶186. *See* Section X.C.6 (Claim 7).

8. Claim 9

Zhang renders obvious claim 9. EX1003, ¶187. *See* Section X.C.6 (Claim 7).

9. Claim 10

Zhang renders obvious claim 10 for the same reasons discussed with respect to claim 1. EX1003, ¶188.

a. 10[Pre]

See Section X.C.2.a (element 1[Pre]). Moreover, Zhang describes that its system is implemented using at least one processor that executes instructions stored in a non-transitory computer readable medium such as a storage device (*[a] non-transitory computer readable storage medium storing instructions that when executed by one or more processors, cause the one or more processors to perform the steps*). EX1006, ¶¶0003, 0006, 0018-0032, 0036, 0052-0061, 0073, Figs. 1, 5, 6, 9, 11; EX1003, ¶189.

b. 10[a]

See Section X.C.2.b (element 1[a]).

c. 10[b]

See Section X.C.2.c (element 1[b]).

d. 10[c]

See Section X.C.2.d (element 1[c]).

e. 10[d]

See Section X.C.2.e (element 1[d]).

10. Claim 11

Zhang renders obvious claim 11 for the same reasons discussed above in Section X.C.3 (claim 2). EX1003, ¶194.

11. Claim 12

Zhang renders obvious claim 12 for the same reasons discussed above in Section X.C.4 (claim 3). EX1003, ¶195.

12. Claim 13

Zhang renders obvious claim 13 for the same reasons discussed above in Section X.C.5 (claim 4). EX1003, ¶196.

13. Claim 16

Zhang renders obvious claim 16 for the same reasons discussed above in Section X.C.6 (claim 7). EX1003, ¶197.

14. Claim 17

Zhang renders obvious claim 17 for the same reasons discussed above in Section X.C.7 (claim 8). EX1003, ¶198.

15. Claim 18

Zhang renders obvious claim 18 for the same reasons discussed above in Section X.C.8 (claim 9). EX1003, ¶199.

16. Claim 19

Zhang renders obvious claim 19 for the same reasons discussed with respect to claim 1. EX1003, ¶200.

a. 19[Pre]

See Sections X.C.2.a (element 1[Pre]), X.C.9.a (element 10[Pre]).

b. 19[a]

See Section X.C.2.b (element 1[a]).

c. 19[b]

See Section X.C.2.c (element 1[b]).

d. 19[c]

See Section X.C.2.d (element 1[c]).

e. 19[d]

See Section X.C.2.e (element 1[d]).

17. Claim 20

Zhang renders obvious claim 20 for the same reasons discussed above in Section X.C.6 (claim 7). EX1003, ¶206.

18. Claim 21

Zhang renders obvious claim 21 for the same reasons discussed above in Section X.C.7 (claim 8). EX1003, ¶207.

19. Claim 22

Zhang renders obvious claim 22 for the same reasons discussed above in Section X.C.8 (claim 9). EX1003, ¶208.

D. Printed Matter

Many limitations of the Challenged Claims are “directed to the content of information and lacking a requisite functional relationship,” i.e., printed matter, and therefore “are not entitled to patentable weight.” *Praxair Distribution, Inc. v. Mallinckrodt Hosp. Prods. IP Ltd.*, 890 F.3d 1024, 1032. In particular, the following limitations are printed matter:

- 1[a], 10[a], 19[a] – “the sensor data displaying an object on the road”
- 1[b], 10[b], 19[b] – “input sensor data displaying a particular object on a road”
- 1[b], 10[b], 19[b] – “statistical summary characterizing a distribution of user responses expected to be received responsive to presenting the sensor data to a plurality of users”
- 1[c], 10[c], 19[c] – “statistical summary data characterizing a distribution of user responses expected to be received responsive to presenting the sensor data captured by the sensor installed on the autonomous vehicle to users”
- 4, 13 – “the user responses describe an action associated with the object, the action including one of the object staying in place, changing lanes,

or crossing a street”

- 5, 14 – “the parameter including at least one of a content of a response, a time associated with entering a response, and a position of an eye of a human observer associated with the response, the position being measured with respect to a display associated with the user interface”
- 6, 15 – “the parameter is further associated with at least one of a central tendency, a variance, a skew, a kurtosis, a scale, and a histogram”
- 7, 16, 20 – “user responses describing a state of mind of a road user associated with the object”
- 8, 17, 21 – “user responses describing an intention of the road user to move into the path of the autonomous vehicle”
- 9, 18, 22 – “user responses describing awareness of the road user of the autonomous vehicle”

The above limitations each recites a data element (sensor data, input sensor data, statistical summary, statistical summary data, parameter, user responses, etc.) and then defines that data element based on its content. But a patent cannot cover what “sensor data display[s],” what a “statistical summary characterize[s],” or what “user responses describ[e]”—such limitations merely cover “information claimed for its communicative content.” *C R Bard Inc. v. AngioDynamics, Inc.*, 979 F.3d 1372, 1381 (Fed. Cir. 2020). For example, the recitations in limitation 1[b] describing the content that is input to and output from the model should be ignored.

Even if the “model” may be patentable, the content of the input and output is not. Similarly, the recited limitations about what the predicted “user responses describ[e]” in claims 7-9, 16-18, and 20-22 should be ignored—the patentee cannot claim the content of user responses. *See Praxair*, 890 F.3d at 1031-32. Because the above limitations are defined by the information content, they are patent ineligible printed matter.

Printed matter is only entitled to patentable weight if it is functionally related to the substrate on which the information is present. *See In re Marco Guldenaar Holding B.V.*, 911 F.3d 1157, 1161 (Fed. Cir. 2018) (markings on dice are “not functionally related to the substrate of the dice”). The above printed matter limitations do not have any such functional relationship with the “substrate,” which in this case is computer memory. EX1003, ¶211; *see also* 10[Pre] and 19[Pre]. Rather than impacting how the computer memory operates, the printed matter limitations in the Challenged Claims “merely inform[] people of the claimed information[.]” *C R Bard Inc.*, 979 F.3d at 1381. As such, none of these limitations can be relied on to distinguish the Challenged Claims from the cited prior art.

XI. CONCLUSION

Petitioner respectfully requests institution of IPR and that the Challenged Claims be canceled as unpatentable pursuant to 35 U.S.C. §318(b).

Petition for *Inter Partes* Review of U.S. Patent No. 10,614,344

Respectfully Submitted,

Date: September 30, 2025

/Roger Fulghum/

Roger Fulghum
Reg. No. 39,678

Attorney for Petitioner, Tesla, Inc.

CERTIFICATE OF SERVICE

In accordance with 37 C.F.R. §§42.6(e) and 42.105, the undersigned certifies that on September 30, 2025, a complete and entire copy of the **PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-22 OF U.S. PATENT NO. 10,614,344** including exhibits and testimony relied upon and Petitioner's power of attorney were served on Patent Owner via FedEx overnight at the correspondence address of record for the '344 Patent and by email to counsel for Patent Owner in the EDTX Litigation, as included below:

758 - FENWICK & WEST LLP
c/o Rajendra B. Panwar (rpanwar@fenwick.com)
Silicon Valley Center
801 California Street
Mountain View, CA 94041

Patrick J. Conroy (pat@nelbum.com)
Andrea L. Fair (andrea@millerfairhenry.com)

Date: September 30, 2025

/Roger Fulghum/

Roger Fulghum
Reg. No. 39,678

Attorney for Petitioner, Tesla, Inc.

CERTIFICATION UNDER 37 C.F.R. §42.24(d)

Pursuant to 37 C.F.R. §42.24(d), the undersigned hereby certifies that the word count under §42.24(a)(1) for the foregoing Petition for *Inter Partes* Review totals 13,993 words, within the 14,000 word limit allowed under §42.24(a)(1)(i).

Date: September 30, 2025

/Roger Fulghum/

Roger Fulghum
Reg. No. 39,678

Attorney for Petitioner, Tesla, Inc