

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.

Petitioner

v.

PERCEPTIVE AUTOMATA LLC

Patent Owner

IPR2025-01577

U.S. Patent No.: 11,520,346 B2

**PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-23 OF U.S.
PATENT NO. 11,520,346**

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

TABLE OF CONTENTS

	Page
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	10
IV. PAYMENT OF FEES UNDER 37 C.F.R. § 42.103.....	10
V. CERTIFICATION OF GROUNDS FOR STANDING	10
VI. OVERVIEW OF CHALLENGE AND RELIEF REQUESTED	10
A. Prior Art Printed Publications	10
B. Identification of Challenge and Statement of Precise Relief Requested	12
VII. SUMMARY OF THE '346 PATENT	12
A. Prosecution History	12
B. Technology Background	14
C. The Alleged Invention.....	15
VIII. LEVEL OF ORDINARY SKILL IN THE ART	18
IX. CLAIM CONSTRUCTION.....	18
X. THE ASSERTED GROUNDS OF INVALIDITY	19
A. Ground 1: Claims 1-23 are unpatentable under 35 U.S.C. § 102 or 35 U.S.C. § 103 over Djuric.....	19

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

1.	Summary of Prior Art: Djuric-EX1004	19
2.	Claim 1	21
	a. 1[Pre].....	21
	b. 1[a]	21
	c. 1[b]	23
	d. 1[c]	24
	e. 1[d-i].....	24
	f. 1[d-ii]	25
	g. 1[d-iii]	29
	h. 1[d-iv].....	30
	i. 1[d-v]	33
	j. 1[e].....	34
3.	Claim 2.....	36
4.	Claim 3	38
5.	Claim 4.....	40
6.	Claim 5	41
7.	Claim 6.....	41
8.	Claim 7	43
9.	Claim 8.....	45
10.	Claim 9.....	46
11.	Claim 10.....	47
	a. 10[Pre].....	47
	b. 10[a]	47
	c. 10[b]	47

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

d. 10[c]	48
e. 10[d-i]	48
f. 10[d-ii]	48
g. 10[d-iii]	48
h. 10[d-iv]	48
i. 10[d-v]	48
j. 10[e]	48
12. Claim 11	48
13. Claim 12	48
14. Claim 13	48
15. Claim 14	49
16. Claim 15	49
17. Claim 16	49
18. Claim 17	49
19. Claim 18	49
20. Claim 19	49
a. 19[Pre]	49
b. 19[a]	50
c. 19[b]	50
d. 19[c]	50
e. 19[d-i]	50
f. 19 [d-ii]	50
g. 19[d-iii]	50
h. 19[d-iv]	50

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

i. 19[d-v]	50
j. 19[e]:	50
21. Claim 20	50
22. Claim 21	51
23. Claim 22	51
24. Claim 23	51
B. Ground 2: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Cox	51
1. Summary of Prior Art: Cox-EX1006	51
2. Motivation to Combine Djuric and Cox	53
3. Claim 1	56
a. 1[d-ii]	57
4. Claims 10 and 19	59
5. Claims 2-9, 11-18, and 20-23	59
C. Grounds 3 and 4: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Zhu or Djuric, Zhu, and Cox	60
1. Summary of Prior Art: Zhu-EX1005	60
2. Motivation to Combine Zhu with Djuric, or Djuric and Cox ...	61
3. Claim 1	63
a. 1[b]	64
b. 1[d-iv]	65
4. Claims 10 and 19	67
5. Claims 2-9, 11-18, 20-23	67
D. Grounds 5, 6, 7, and 8: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Ross (Ground 5), Djuric, Cox, and Ross	

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

(Ground 6), Djuric, Zhu, and Ross (Ground 7), or Djuric, Cox, Zhu, and Ross (Ground 8).....	68
1. Summary of Prior Art: Ross-EX1015.....	68
2. Motivation to Combine Djuric, Zhu, Cox, and Ross.....	69
3. Claims 1-3, 5, 7-12, 14, and 16-21	70
4. Claim 4.....	71
5. Claim 6.....	73
6. Claim 13.....	74
7. Claim 15.....	74
8. Claim 22.....	74
9. Claim 23.....	74
E. Printed Matter.....	75
XI. CONCLUSION	77

PETITIONER’S EXHIBIT LIST

Ex[No.]	Description of Documents
1001	U.S. Patent No. 11,520,346 (“’346 Patent”)
1002	File History of U.S. Patent No. 11,520,346
1003	Declaration of Dr. Jason Janét
1004	U.S. Patent Application Publication No. 2019/0049970 (“Djuric”)
1005	U.S. Patent Application Publication No. 2019/0164018 (“Zhu”)
1006	International Publication No. WO2014/210334 (“Cox”)
1007	U.S. Provisional Application 62/798,978
1008	U.S. Provisional Application 62/800,416
1009	U.S. Provisional Application 62/822,269
1010	Excerpts from <i>Probabilistic Robotics</i> , Sebastian Thrun, Wolfram Burgard, and Dieter Fox, The MIT Press, 2006
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
1013	Excerpts from <i>Robot Motion Planning</i> , Jean-Claude Latombe, Kluwer Academic Publishers, 1991
1014	Excerpts from <i>Principles of Robot Motion</i> , Howie Choset, et al., The MIT Press, 2005.
1015	U.S. Patent No. 10,496,091 (“Ross”)

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. 11,520,346 (the “’346 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. To the best knowledge of Petitioner, the ’346 Patent is not involved in any other post-grant proceedings.

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

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

	Matthew Williams, Reg. No. 78,734 Baker Botts L.L.P. 401 S. 1st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2550 matthew.williams@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 Paul Margulies, Reg. No. 59,580 Tesla, Inc. 800 Connecticut Ave. NW Washington, DC 20006 Phone: (202) 695-5388 pmargulies@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 '346 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 addresses:

Roger Fulghum (roger.fulghum@bakerbotts.com)
Mark Speegle (mark.speegle@bakerbotts.com)
Ellyar Y. Barazesh (ellyar.barazesh@bakerbotts.com)
Matthew Williams (matthew.williams@bakerbotts.com)
Lance J. Goodman (lance.goodman@bakerbotts.com)
Paul Margulies (pmargulies@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 method comprising:
1[a]	receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle, the sensor data comprising one or more images;
1[b]	generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data;
1[c]	identifying, one or more traffic entities based on the sensor data, the traffic entities representing non-stationary objects in traffic in which the autonomous vehicle is driving;
1[d-i]	for each of the one or more traffic entities: determining one or more motion parameters describing movement of the traffic entity;
1[d-ii]	[for each of the one or more traffic entities:] providing an image of the one or more images, the image showing the traffic entity as input to a machine learning model configured to receive an input image showing an input traffic entity and output summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image;
1[d-iii]	[for each of the one or more traffic entities:] determining the hidden context of the traffic entity based on the output of the machine learning based model;
1[d-iv]	[for each of the one or more traffic entities:] determining a region of the

	point cloud where the traffic entity is expected to reach within a threshold time interval; and
1[d-v]	[for each of the one or more traffic entities:] modifying the region based on the hidden context of the traffic entity; and
1[e]	navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities.

Claim 2

2	The method of claim 1, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in the direction having a component along the motion vector, extending the region along the direction of the motion vector.
----------	--

Claim 3

3	The method of claim 1, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in a direction having a component opposite to direction of the motion vector, decreasing a size of the region along the direction of the motion vector.
----------	--

Claim 4

4	The method of claim 1, wherein the hidden context represents a state of mind of a user represented by the traffic entity.
----------	---

Claim 5

5	The method of claim 1, wherein the hidden context represents a task that a user represented by the traffic entity is planning on accomplishing.
----------	---

Claim 6

6	The method of claim 1, wherein the hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity.
----------	---

Claim 7

7	The method of claim 1, wherein the hidden context represents a goal of a user represented by the traffic entity, wherein the user expects to achieve the goal within a threshold time interval.
----------	---

Claim 8

8	The method of claim 1, wherein navigating the autonomous vehicle comprises: generating signals for controlling the autonomous vehicle based on the one or more motion parameters and the hidden context of each of the one or more traffic entities; and sending the generated signals to controls of the autonomous vehicle.
----------	---

Claim 9

9	The method of claim 1, wherein the sensor data represents one or more images captured by a camera mounted on the autonomous vehicle or a lidar scan captured by a lidar mounted on the autonomous vehicle.
----------	--

Claim 10

10[Pre]	A non-transitory computer readable storage medium storing instructions, that when executed by a processor, cause the processor to perform steps comprising:
10[a]	receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle, the sensor data comprising one or more images;
10[b]	generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data;
10[c]	identifying, one or more traffic entities based on the sensor data, the traffic entities representing non-stationary objects in traffic in which the

	autonomous vehicle is driving;
10[d-i]	for each of the one or more traffic entities: determining one or more motion parameters describing movement of the traffic entity;
10[d-ii]	[for each of the one or more traffic entities:] providing an image of the one or more images, the image showing the traffic entity as input to a machine learning model configured to receive an input image showing an input traffic entity and output summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image;
10[d-iii]	[for each of the one or more traffic entities:] determining the hidden context of the traffic entity based on the output of the machine learning based model;
10[d-iv]	[for each of the one or more traffic entities:] determining a region of the point cloud where the traffic entity is expected to reach within a threshold time interval; and
10[d-v]	modifying the region based on the hidden context of the traffic entity; and
10[e]	navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities.

Claim 11

11	The non-transitory computer readable storage medium of claim 10, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in the direction having a component along the motion vector, extending the region along the direction of the motion vector.
-----------	--

Claim 12

12	The non-transitory computer readable storage medium of claim 10, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in a direction having a component
-----------	--

	opposite to direction of the motion vector, decreasing a size of the region along the direction of the motion vector.
--	---

Claim 13

13	The non-transitory computer readable storage medium of claim 10, wherein the hidden context represents a state of mind of a user represented by the traffic entity.
-----------	---

Claim 14

14	The non-transitory computer readable storage medium of claim 10, wherein the hidden context represents a task that a user represented by the traffic entity is planning on accomplishing.
-----------	---

Claim 15

15	The non-transitory computer readable storage medium of claim 10, wherein the hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity.
-----------	---

Claim 16

16	The non-transitory computer readable storage medium of claim 10, wherein the hidden context represents a goal of a user represented by the traffic entity, wherein the user expects to achieve the goal within a threshold time interval.
-----------	---

Claim 17

17	The non-transitory computer readable storage medium of claim 10, wherein navigating the autonomous vehicle comprises: generating signals for controlling the autonomous vehicle based on the one or more motion parameters and the hidden context of each of the one or more traffic entities; and sending the generated signals to controls of the autonomous vehicle.
-----------	---

Claim 18

18	The non-transitory computer readable storage medium of claim 10, wherein the sensor data represents one or more images captured by a camera mounted on the autonomous vehicle or a lidar scan captured by a lidar mounted on the autonomous vehicle.
-----------	--

Claim 19

19[Pre]	A computer system comprising: a processor; and a non-transitory computer readable storage medium storing instructions that when executed by the processor, cause the processor to perform steps comprising:
19[a]	receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle, the sensor data comprising one or more images;
19[b]	generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data;
19[c]	identifying, one or more traffic entities based on the sensor data, the traffic entities representing non-stationary objects in traffic in which the autonomous vehicle is driving;
19[d-i]	for each of the one or more traffic entities: determining one or more motion parameters describing movement of the traffic entity;
19[d-ii]	[for each of the one or more traffic entities:] providing an image of the one or more images, the image showing the traffic entity as input to a machine learning model configured to receive an input image showing an input traffic entity and output summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image;
19[d-iii]	[for each of the one or more traffic entities:] determining the hidden context of the traffic entity based on the output of the machine learning based model;
19[d-iv]	[for each of the one or more traffic entities:] determining a region of the point cloud where the traffic entity is expected to reach within a threshold time interval; and
19[d-v]	[for each of the one or more traffic entities:] modifying the region based on the hidden context of the traffic entity; and
19[e]	navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities.

Claim 20

20	The computer system of claim 19, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in the direction having a component along the motion vector, extending the region along the direction of the motion vector.
-----------	--

Claim 21

21	The computer system of claim 19, wherein the region is in a direction determined based on a motion vector of the traffic entity, wherein the traffic entity represents a user, wherein modifying the region based on the hidden context comprises: responsive to determining based on the hidden context that the user represented by the traffic entity is likely to move in a direction having a component opposite to direction of the motion vector, decreasing a size of the region along the direction of the motion vector.
-----------	--

Claim 22

22	The computer system of claim 19, wherein the hidden context represents a state of mind of a user represented by the traffic entity.
-----------	---

Claim 23

23	The computer system of claim 19, wherein the hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity.
-----------	---

III. INTRODUCTION

Tesla, Inc. (“Petitioner”) petitions for *inter partes* review of Claims 1-23 (the “Challenged Claims”) of U.S. Patent No. 11,520,346 (EX1001) (the “’346 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 ’346 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 ’346 Patent was filed on January 30, 2020, and claims priority to U.S. Provisional Application 62/798,978, filed on January 30, 2019 (the “’978 Provisional”) (EX1007). EX1001, Cover Page. The ’346 Patent also claims priority to U.S. Provisional Application 62/800,416, filed on February 1, 2019 (the “’416 Provisional”) (EX1008) and U.S. Provisional Application 62/822,269, filed on March 22, 2019 (the “’269 Provisional”) (EX1009). EX1001, Cover Page.

Petition for *Inter Partes* Review of U.S. Patent No. 11,520,346

Without conceding that the '346 Patent is entitled to these priority dates, the Djuric, Zhu, Cox, and Ross references are prior art to the '346 Patent before January 30, 2019.

- U.S. Patent Application Publication No. 2019/0049970 (“Djuric”) (EX1004); filed on September 5, 2018; prior art under 35 U.S.C. § 102(a)(2).
- U.S. Patent Application Publication No. 2019/0164018 (“Zhu”) (EX1005); filed on November 27, 2017; prior art under 35 U.S.C. § 102(a)(2).
- International Publication No. WO2014/210334 (“Cox”) (EX1006); 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. 10,496,091 (“Ross”) (EX1015); filed on August 17, 2016; prior art under 35 U.S.C. § 102(a)(2).

None of Djuric, Zhu, Cox, or Ross were cited or otherwise discussed during prosecution.

¹ Cox lists Harvard as the 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 '346 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	1-23	§ 103	Djuric
2	1-23	§ 103	Djuric and Cox
3	1-23	§ 103	Djuric and Zhu
4	1-23	§ 103	Djuric, Zhu, and Cox
5	1-23	§ 103	Djuric and Ross
6	1-23	§ 103	Djuric, Cox, and Ross
7	1-23	§ 103	Djuric, Zhu, and Ross
8	1-23	§ 103	Djuric, Cox, Zhu, and Ross

VII. SUMMARY OF THE '346 PATENT

A. Prosecution History

The Office issued a first non-final office action that provisionally rejected claims 1, 4-10, and 13-19 on the ground of non-statutory double patenting over co-pending U.S. Patent Application No. 16/777,386. EX1002, 126-129. The Office also rejected independent claims 1-3, 8-12, and 17-20 under 35 U.S.C. § 103 as being unpatentable over U.S. Patent Application Publication No. 2020/0193829 (“Cheng”) in view of U.S. Patent Application Publication No. 2017/0318360 (“Tran”). *Id.*, 130-

138. In response, the applicant amended claim 1 as follows, and additional independent claims 10 and 19 similarly:

“A method comprising:
receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle, the sensor data comprising one or more images;
generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data;
identifying, one or more traffic entities based on the sensor data, the traffic entities representing non-stationary objects in traffic in which the autonomous vehicle is driving;
for each of the one or more traffic entities:
determining one or more motion parameters describing movement of the traffic entity;
providing an image of the one or more images, the image showing the traffic entity as input to a machine learning model configured to receive an input image showing a traffic entity and output summary statistics of expected human responses describing a hidden context of the traffic entity shown in the input image;
determining [[a]] the hidden context of the traffic entity based on [[a]] the output of the machine learning based model, wherein the machine learning based model is trained based on feedback received from users responsive to presenting images or videos showing traffic scenarios to the users;
determining a region of the point cloud where the traffic entity is expected to reach within a threshold time interval;
modifying the region based on the hidden context of the traffic entity;
navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities.

EX1002, 186-193.

In response to the applicant’s amendments and related arguments, the Office issued a final office action that maintained the double patenting rejection, but withdrew the § 103 rejection. EX1002, 211-215. The applicant thereafter submitted

a terminal disclaimer and made other minor amendments. *Id.*, 237-247, 253-254. The Office then issued a notice of allowance without commenting on the claims. *Id.*, 260-264.

However, as demonstrated by the grounds of this petition, the added “providing...” claim limitation used to distinguish the claims was well known, as were the Challenged Claims as a whole.

B. Technology Background

The '346 Patent generally relates to machine learning models, which are computer programs that analyze input data and generate an output based on the data. EX1003, ¶59. Machine learning models may be trained using labeled training data that is known to represent a correct output. *Id.* Human annotators (also referred to as “labelers,” “reviewers,” or “annotators”) can be used to 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.*

For example, to develop a vehicle detection machine learning model, human annotators may label all instances of training data images that display a vehicle. EX1003, ¶60. The model is then trained using the labeled training data, and, after training, the model can detect vehicles in new, non-training data from a sensor system that may produce camera image data, LIDAR point cloud data, and RADAR data. *Id.* The model’s detection of vehicles reflects how the human annotators

would have labeled the new images because the model was trained based on the human annotators' decision-making when labeling training data. *Id.* The model output can be used by an autonomous vehicle to direct vehicle navigation. *Id.*

C. The Alleged Invention

The alleged invention of the '346 Patent relates to (1) gathering sensor data, including image data, with sensors of an autonomous vehicle, to generate a point cloud representation, (2) inputting the image data into a machine learning model trained to output summary statistics of expected human responses describing “hidden context,” (3) determining a region of the point cloud and modifying the region based on the hidden context, and (4) navigating the autonomous vehicle to avoid the modified region. EX1001, 2:1-57, 3:19-55, 15:57-16:36, 17:46-18:26. But this alleged invention was well-known for many years in the fields of both predictive models and autonomous vehicle control, as demonstrated by the grounds herein. Section X; EX1003, ¶61.

The substance of the '346 Patent is nothing new. EX1003, ¶62. The Patent describes the well-known process of using a sensor system to “perceive” the surroundings of an autonomous vehicle by gathering sensor data, including LIDAR scans and camera images, which was a well-documented concept at least as of 2005. EX1010, pp. 3, 151, 169-174 (describing “Robot Perception” using laser range scanners and cameras to control “cars that travel by themselves”). EX1001, 1:36-38, 3:53-55, 6:66-7:6, 7:49-50, 15:57-61; EX1004, ¶¶0022, 0047, 0049-0050, 0059;

EX1005, ¶¶0004, 0012, 0015, 0026, 0041. Then, a point cloud is generated based on the sensor data, which is another well-known step. EX1001, 3:19-24, 15:61-67; EX1004, ¶0022; EX1005, Abstract, ¶¶0012, 0015, 0026, 0032, 0041. The sensor data can be used to identify nearby moving objects and their associated motion parameters, yet another well-known step. EX1001, 2:44-47, 2:55-58, 7:7-15; EX1004, ¶0025, 0033, 0051.

The '346 Patent explains how a machine learning model uses image input to output summary statistics of expected human responses describing “hidden context” of vehicles and pedestrians, but this seemingly complex language is just jargon; “hidden context” may simply represent the behavior of other vehicles and pedestrians, for example. EX1001, Abstract, 2:3-5. The machine learning model may be trained with training data obtained and labeled by human annotators, who provide their understanding or prediction of how traffic entities will behave, but this concept was well-known. EX1001, EX1003, ¶63; EX1004, ¶0053; Indeed, the inventor’s own prior art shows that the '346 Patent should not have issued. The Cox reference that, like the '346 Patent, lists Samuel Anthony of Cambridge, Massachusetts, as an inventor (EX1006, Cover Page), discloses using human-annotated training data to ensure that machine learning model predictions are “more consistent with the decisions of the human annotators” and “better mimic human performance.” EX1006, ¶¶0004, 0032. Neither the patentee nor Mr. Anthony brought the Cox reference to the Patent Office’s attention during prosecution of

the '346 Patent. *See* EX1001; EX1002.

After training on the labeled data, the machine learning model in the '346 Patent generates outputs based on “live” input data. EX1001, 10:15-28, 13:26-48. This, too, was well-known at least as early as 2005, and the application of this step to autonomous vehicles does not save it. EX1010, pp. xvii, 3 (describing “perception and control” of “robotic systems include[ing] ... cars that travel by themselves”). EX1003, ¶64; EX1004, ¶¶0030-0032, 0061-0068, 0100-0102, 0111-0112; EX1005, ¶¶0073, 0079-0081 (describing inputting current images into a model that predicts future movement of objects depicted in the images).

The '346 Patent describes determining and modifying a region of the point cloud that a traffic entity will reach based on the hidden context. EX1001, 3:30-35, 17:46-18:12, 22:14-23:12. This is nothing more than the well-known step of predicting an object’s trajectory based on its expected behavior. EX1003 ¶65; EX1004, ¶¶0021, 0026, 0081-0094. For example, it was well-known to generate multiple predictions of different locations an object may reach and then modify the locations by selecting one most likely location. EX1004, ¶¶0034, 0093-0094, 0112-0114.

The '346 Patent further describes controlling an autonomous vehicle to stay a threshold distance away from the modified region. EX1001, 3:48-52, 18:13-26, 22:41-51. But it was well known to control an autonomous vehicle to avoid the predicted future locations of objects. EX1003, ¶66; EX1004, ¶¶0008, 0021-0024,

0035-0036, 0050, 0078, 0095, 0103-0104, 0114, Fig. 5; EX1010, pp. xvii, 3 (explaining a goal of robotic perception and control of autonomous cars as “making car accidents a notion of the past”).

Thus, the '346 Patent describes nothing more than the application of well-known machine learning concepts to the well-known field of autonomous vehicles. EX1003, ¶67. The grounds of this Petition demonstrate that the Challenged Claims should not have been issued and should now be canceled. *Id.*

VIII. LEVEL OF ORDINARY SKILL IN THE ART

A person having ordinary skill in the art (“POSITA”) relevant to the '346 Patent as of January 30, 2019 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, ¶68.

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.² EX1003, ¶69.

X. THE ASSERTED GROUNDS OF INVALIDITY

A. Ground 1: Claims 1-23 are unpatentable under 35 U.S.C. § 102 or 35 U.S.C. § 103 over Djuric.

Djuric discloses all the same concepts claimed in the '346 Patent, and as such Djuric renders obvious all of the Challenged Claims. EX1003, ¶70.

1. Summary of Prior Art: Djuric-EX1004

Djuric is generally directed to “predicting the motion of objects proximate to an autonomous vehicle and controlling the autonomous vehicle.” EX1004, ¶0002. Djuric describes an autonomous vehicle with a sensor system that includes cameras and LIDAR sensors. EX1004, ¶¶0022, 0046-0047, 0049, 0059. The sensor system captures sensor data, including camera image data and LIDAR point cloud data, associated with objects in the surroundings of the autonomous vehicle, such as other vehicles and pedestrians. *Id.* A perception system can process the sensor data to determine state data for the objects, such as “current and/or past location (also referred to as position), current and/or past speed/velocity, current and/or past acceleration, current and/or past heading, [and] current and/or past orientation.” EX1004, ¶¶0025 (parentheticals in original), 0028, 0051, 0057, 0072, 0080.

Djuric’s autonomous vehicle generates predictions of object motion using a

² Petitioner does not concede that the terms of the challenged claims are definite under 35 U.S.C. §112.

predictive model, such as a “machine-learned model.” EX1004, ¶¶0027, 0047, 0053, 0061-0067. Djuric’s training methodology additionally includes “training the model [] to evaluate which trajectory of a set of trajectories is closest to a ground truth,” so that the model produces an output closest to the human-labeled trajectories of the training data. EX1004, ¶0094; EX1003, ¶72.

Specifically, after capturing the images of objects in the autonomous vehicle’s environment, the autonomous vehicle inputs the images into the neural network model. EX1004, ¶0069. Using the images, the neural network model makes a “goal-based prediction” of future trajectories of the object, generating a plurality of predicted trajectories for the object. EX1004, ¶¶0080-0088. For example, for an object vehicle, the predicted future trajectories may be based on driving actions such as “continu[ing] according to its current lane; merg[ing] from its current lane into an adjacent lane; [or] turn[ing] from the current lane into an adjacent road portion.” EX1004, ¶0081. Additionally, the predicted trajectories each include a plurality of way-points and associated way-point confidence levels. EX1004, ¶¶0083, 0085.

Using the predicted trajectories, trajectory confidence levels, waypoints, and waypoint confidence levels, the autonomous vehicle generates a motion plan and the vehicle “travel[s] in accordance with this motion plan.” EX1004, ¶0114. To execute this, a “vehicle controller can translate the motion plan into instructions for the appropriate vehicle controls (e.g., acceleration control, brake control, steering control, etc.).” EX1004, ¶¶0035, 0078.

Djuric is analogous art to the '346 Patent, from the same field of endeavor as the '346 Patent (e.g., autonomous vehicle perception systems), and is reasonably pertinent to the particular problem that the Patent was trying to solve (e.g., how to “accurately predict motion of non-stationary traffic objects” in an autonomous vehicle’s surroundings). EX1004, Abstract, ¶¶0032-0034, 0081; EX1001, 1:34-67, 2:26-43, 3:19-35, 15:4-22; EX1003, ¶75.

2. Claim 1

Djuric renders obvious claim 1. EX1003, ¶76.

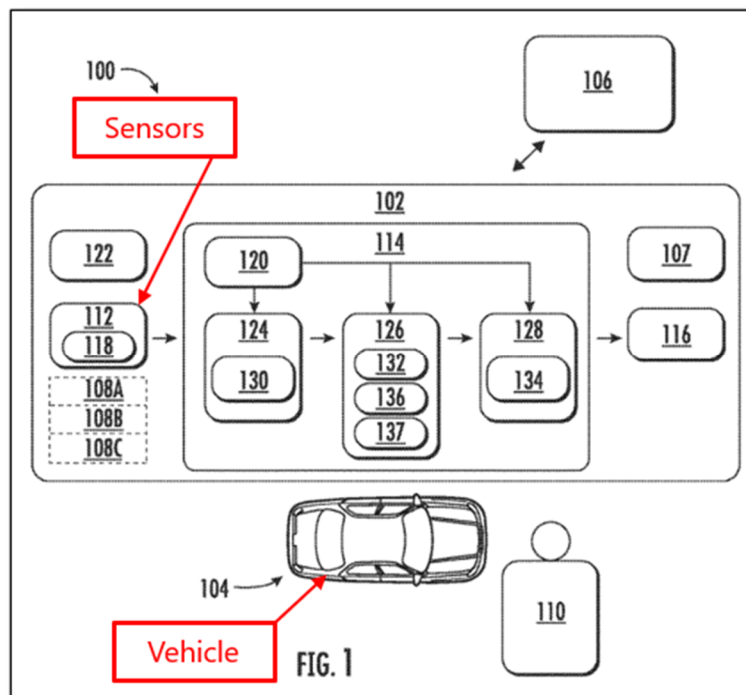
a. 1[Pre]

To the extent the preamble is limiting, Djuric teaches it. EX1003, ¶77. Djuric’s systems and methods “enable an autonomous vehicle to predict the respective motion of...objects and plan the vehicle's motion accordingly” (*[a] method*). EX1004, ¶¶0020-0024, 0035-0036, 0044-0046, 0050, 0078, 0095, 0103-0104, 0114, 0118, Fig. 9 (item 120); EX1003, ¶77.

b. 1[a]

Djuric teaches element 1[a]. EX1003, ¶78. Djuric describes a vehicle 104 that includes sensors 112, such as cameras, “and/or other types of imaging capture devices” (*sensors mounted on the autonomous vehicle*). EX1004, ¶¶0046-0047, 0059. The sensors 112 further include a “LIDAR system.” EX1004, ¶¶0022, 0047, 0049, 0059. The cameras and LIDAR system generate “sensor data 118” including “image data” from the cameras that show “objects that are within the surrounding

environment of the vehicle 104,” where the objects include “vehicles, pedestrians, [and] bicycles,” as well as LIDAR data that is a “LIDAR point cloud.” EX1004, ¶¶0022, 0032, 0046-0049, 0050-0051, 0059, Figs. 2A-2C. A POSITA would have understood that Djuric’s image data is formed by image frames; this is because image data is formed by image frames that have been captured by a camera. *Id.*; *see id.*, ¶¶0007, 0064, 0021, 0029-0033, 0037, 0058-0074, 0079-0080, 0099-0101; EX1003, ¶78. The sensor data 118 is received by autonomy computing system 114 of vehicle 104 (*receiving, by an autonomous vehicle, sensor data from sensors mounted on the autonomous vehicle, the sensor data comprising one or more images*). EX1004, ¶0050; EX1003, ¶78.



EX1004, Fig. 1 (annotated)

c. 1[b]

Djuric teaches element 1[b] as described below. EX1003, ¶79. Djuric explains “sensor data 118” is captured by one or more sensors and can include “image data, RADAR data, LIDAR data” including “point cloud(s).” EX1004, ¶¶0022, 0032, 0046-0051, 0059, Figs. 2A-2C; Section X.A.2.b (element 1[a]). A POSITA would have understood that the LIDAR data captured by vehicle 104’s LIDAR system would have been used to generate the point clouds of the vehicle’s surroundings, thus teaching *generating a point cloud representation...based on the sensor data*. *Id.*; EX1003, ¶79. This is because point clouds, which are images that display the depth/3D shape of an environment, were known as being generated using the LIDAR data formed by reflected light received by a LIDAR sensor system. *Id.*

Djuric further teaches that the sensor data is “associated with the surrounding environment of the vehicle.” EX1004, ¶0022; EX1003, ¶80. The “LIDAR point cloud(s) and/or other data” are “associated with one or more object(s) that are proximate to the autonomous vehicle” such as “within a field of view of the sensor(s)” as well as “one or more geographic characteristics of the geographic area” such as “curbs, lane markings, [and] sidewalks.” EX1004, ¶¶0022, 0031, 0036. Djuric thus teaches that its point cloud representation generated from LIDAR data is of the surroundings of vehicle 104 (*generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data*). *Id.*; EX1003, ¶80.

Djuric’s disclosure aligns with the ’346 Patent, which discloses a vehicle

“builds a point cloud representation of the surroundings of the autonomous vehicle” based on “lidar scans.” EX1001, 15:57-66; EX1003, ¶81.

d. 1[c]

Djuric teaches element 1[c]. EX1003, ¶82. Djuric’s autonomy computing system 114 includes a perception system 124 and “can identify one or more objects that are proximate to the vehicle 104 based at least in part on the sensor data 118” (*identifying, one or more traffic entities based on the sensor data*), where the objects include moving “vehicles, pedestrians, [and] bicycles” that are in traffic in a roadway where the autonomous vehicle 104 is driving (*the traffic entities representing non-stationary objects in traffic in which the autonomous vehicle is driving*). EX1004, ¶¶0022, 0025, 0047, 0050-0051, 0084, 0096, 0105, Figs. 2A-2C; EX1003, ¶82. The objects are non-stationary objects at least because they are associated with a non-zero “current...speed/velocity” and “current...acceleration.” EX1004, ¶¶0025, 0051, 0084, 0096, 0105.

e. 1[d-i]

Djuric teaches element 1[d-i]. Djuric’s perception system 124 processes “sensor data 118 to detect one or more objects that are within the surrounding environment of the vehicle 104 as well as state data 130 associated therewith.” EX1004, ¶¶0021, 0051. The state data indicates “at least a current or past state of an object” in the surrounding environment and “for each object” describes an estimate of the object’s current “speed/velocity” and “acceleration,” which are *motion*

parameters (for each of the one or more traffic entities: determining one or more motion parameters describing movement of the traffic entity). Id.; EX1003, ¶83.

f. 1[d-ii]

Djuric teaches element 1[d-ii]. EX1003 ¶84.

Djuric's vehicle 104 includes cameras that capture image data including "objects that are within the surrounding environment of the vehicle 104," where the objects include moving "vehicles, pedestrians, [and] bicycles" that are in traffic in a roadway where the autonomous vehicle 104 is driving. EX1004, ¶¶0022, 0025, 0030-0032, 0046-0051, 0080-0084, 0096, 0105, Figs. 2A-2C; EX1003, ¶85. The image data "can include a plurality of images 306A-C (e.g., rasterized images) that include the object." EX1004, ¶¶0007, 0064, 0021, 0029-0033, 0037, 0058-0074, 0079-0080, 0099-0101. A POSITA would have understood that Djuric's image data is formed by image frames of the surrounding environment; this is because image data is formed by image frames that have been captured by a camera. *Id.*; EX1003, ¶85. The image data including the object is provided as an input to Djuric's model (*[for each of the one or more traffic entities:] providing an image of the one or more images, the image showing the traffic entity as input to a machine learning model*), and the model receives the image data so is thus configured to receive such data (*machine learning model configured to receive an input image showing an input traffic entity*). *Id.*

Djuric teaches using *hidden context* because Djuric's model performs "a goal-

based prediction.” EX1004, ¶¶0080-0084. “Such an approach can allow the model 136 to consider a ‘goal’ that an object may aim to accomplish.” *Id.*, ¶0081. Djuric’s disclosure of goal-based predicted trajectories aligns with the ’346 Patent’s description of *hidden context*, which states, for example, that “[t]he hidden context may represent a goal of a user represented by the traffic entity” and “[t]he hidden context may represent a task that a user represented by the traffic entity is planning on accomplishing.” EX1001, 2:11-15, 2:26-30; EX1003, ¶86. The model in Djuric also predicts “how an object is likely to move relative to the characteristics of the geographic area” by generating “one or more predicted future locations of the object (e.g., a predicted object trajectory), which the autonomous vehicle can use to plan and control its motion (e.g., to avoid the object).” EX1004, ¶¶0021 (parentheticals in original), 0027, 0053, 0080-0084.

Djuric also teaches outputting associated confidence levels for each trajectory that are each a “per-trajectory probability, quantifying how likely an object is to follow a certain trajectory.” EX1004, ¶¶0032-0034, 0080-0084, 0089-0094. Because the confidence levels are a collection of numerical measurements summarizing how likely the trajectories are, a POSITA would have recognized that these values form *summary statistics*. EX1003, ¶87. Djuric’s model is thus configured to output predicted trajectories and associated confidence levels that summarize the likelihoods of the trajectories (*machine learning model configured to...output summary statistics*). *Id.*

Djuric further teaches that the output predicted trajectories and associated confidence levels (*output summary statistics*) are *of expected human responses describing a hidden context of the input traffic entity shown in the input image*. EX1003, ¶88. Djuric's model is trained using "supervised training techniques" including the use of "labeled" input training data. EX1004, ¶¶0027, 0053, 0134-0136. It would have been obvious for a POSITA to use, in supervised model training, training data having object trajectories that are labeled by a human; this is because it is well-known to use human labeled training data for model training, and Djuric even suggests using "manually" labeled (i.e., human labeled) training data. *Id.*; EX1003, ¶88.

Moreover, during model training, a loss function is "optimized" to "train the network to output [the] plurality of trajectories." EX1004, ¶0092. Loss functions were well known in the art as being used to determine the error between a model output and the labeled training data, and ultimately used to train the model to more accurately reflect labeled training data. EX1004, ¶¶0027, 0053, 0085-0094, 0134-0136; EX1003, ¶89. And additionally, Djuric discloses that training its model includes evaluating "which trajectory of a set of trajectories is closest to a ground truth." EX1004, ¶0094. A POSITA would have understood that such a "ground truth," in the context of a supervised trained model using labeled data, reflects a correct object trajectory that a human should label in an image when presented with the image. EX1003, ¶89. This is because ground truth is the true correct label for a

particular training image. *Id.*

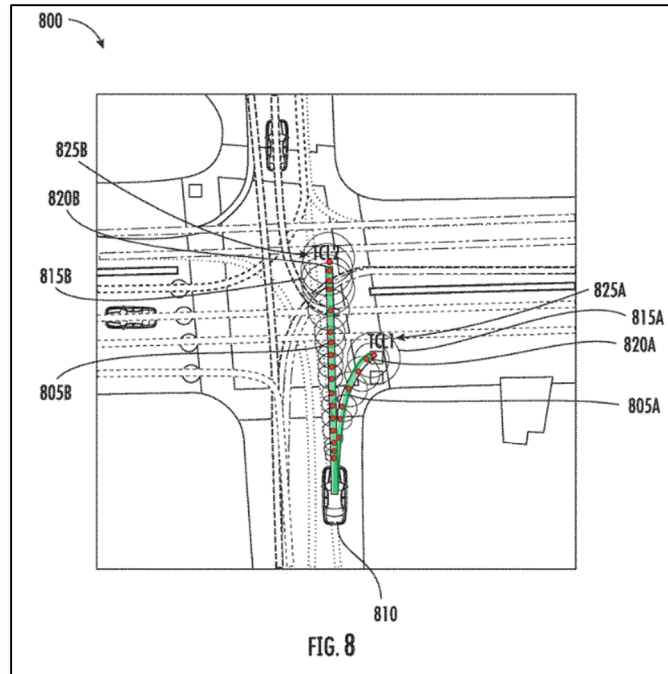
Thus, from the aforementioned training techniques disclosed in Djuric, a POSITA would have understood that Djuric's model is trained such that the predicted trajectories and associated confidence levels align with the expected responses from human labelers given the same input image. EX1004, ¶¶0021, 0027, 0053, 0079, 0085-0103, 0134-0136; EX1003, ¶90. Due to this training, the confidence levels (*output summary statistics*) for the predicted trajectories of objects in image data, as well as the predicted trajectories themselves, reflect expected responses from human labelers to the same input image data (*of expected human responses*). *Id.*

And as explained above, the responses (*expected human responses*) would have been *describing a hidden context of the input traffic entity shown in the input image* because they would reflect the potential goal of the object. EX1004, ¶¶0021, 0027, 0053, 0079, 0080-0103, 0134-0136; EX1003, ¶91. This is because Djuric's model performs "a goal-based prediction" that considers the goals that an object (vehicle, bicycle, or pedestrian) in image data may "aim to accomplish." *Id.* As such, Djuric's predicted trajectories are each representative of these goals, e.g., how an object may wish to travel in the environment around Djuric's autonomous vehicle. *Id.* A POSITA would thus have understood that the training data used to train Djuric's model would have included labels relating to such goals. *Id.*

g. 1[d-iii]

Djuric teaches element 1[d-iii]. Djuric's model outputs a plurality of predicted trajectories for an object and associated trajectory confidence levels that correspond to *hidden context* of the object. Section X.A.2.f (element 1[d-ii]); EX1003, ¶92. Because Djuric teaches using "goal-based prediction[s]," the plurality of predicted trajectories reflects one or more goals that the object may intend to pursue. *Id.*; *see also* EX1004, ¶0081. Djuric's system further evaluates the potential goals of a user when assigning confidence values to the predicted trajectories (*[for each of the one or more traffic entities:] determining the hidden context of the traffic entity based on the output of the machine learning based model*). EX1004, ¶¶0006, 0027, 0034, 0053, 0080-0094, 0112-0113, Figs. 8-9; EX1003, ¶92.

Djuric provides an example of this in relation to Figure 8, where a first predicted trajectory (805A) for an object vehicle predicts that the vehicle will perform a right-hand turn at an intersection and a second predicted trajectory (805B) for the object vehicle predicts that the vehicle will drive straight through the intersection. EX1004, ¶0093, Fig. 8.



Djuric, Fig. 8 (annotated)

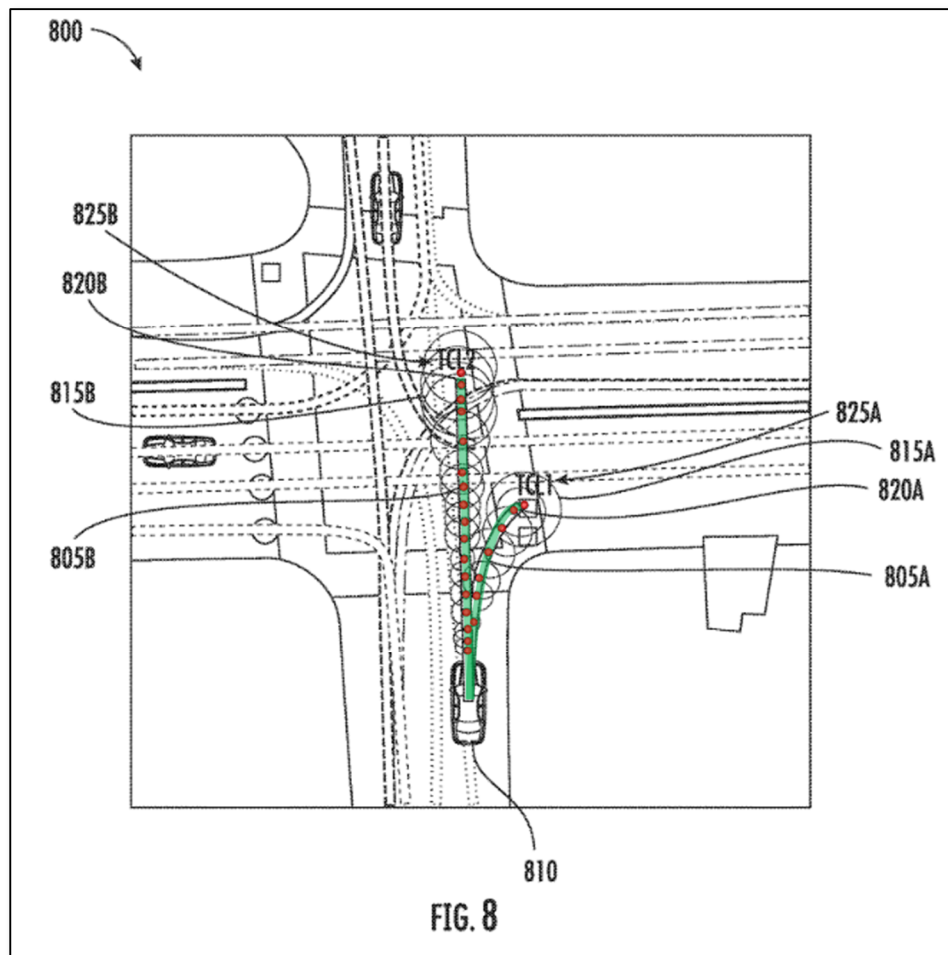
When the object vehicle is traveling in the right-hand lane but is not decelerating as it approaches the intersection, Djuric's model can be trained to determine that the confidence level for trajectory 805B (traveling straight through the intersection) is higher than the level for trajectory 805A (turning right). EX1004, ¶0093, Fig. 8; EX1003, ¶94. This example demonstrates that Djuric's system determines the likely goal(s) (*hidden context*) when assigning confidence levels to a predicted trajectory *[for each of the one or more traffic entities:] determining the hidden context of the traffic entity based on the output of the machine learning based model. Id.*

h. 1[d-iv]

Djuric teaches element 1[d-iv]. EX1003, ¶95. Djuric's predicted trajectories and associated confidence values identify *a region of the point cloud where the*

traffic entity is expected to reach within a threshold time interval as claimed.

EX1003, ¶95; Section X.A.2.f (element 1[d-ii]). The predicted trajectories are formed of waypoints that are each “indicative of a predicted future location of the object.” EX1004, ¶¶0007, 0082-0094. Figure 8 shows example future waypoint locations (annotated in red) where an object is expected to be based on the predicted trajectories (annotated in green). *Id.*, ¶¶0082-0094, Fig. 8. The image shown by Figure 8 is an “output 800 of the model 136.” EX1004, ¶0083.



Djuric, Fig. 8 (annotated)

Djuric’s predicted trajectories and waypoints teach *determining a region [...]*

where the traffic entity is expected to reach within a threshold time interval. EX1003, ¶96; EX1004, ¶¶0007, 0022, 0026, 0071, 0082-0094, Fig. 8. Djuric explains that each “predicted trajectory can include a plurality of predicted future locations of an object at one or a plurality of future times.” EX1004, ¶0071; *see also id.*, ¶¶0026, 0052, 0072-0073, 0102.

These “predicted future locations” of each “predicted trajectory” collectively define a determined *region [...] where the traffic entity is expected to reach* because they are the areas in the image of the environment that Djuric’s model predicts the object will potentially travel to. *Id.*; EX1003, ¶97. Moreover, the “plurality of future times” teaches the *threshold time interval* because they form an interval of time within which the object is predicted to reach the last waypoint of the predicted trajectory. EX1004, ¶¶0026, 0052, 0071-0073, 0102; EX1003, ¶97. Additionally, Djuric’s “vehicle computing system has limited computational resources onboard the vehicle and is running in real-time, as the vehicle is travelling.” EX1004, ¶0037. It would have been obvious to a POSITA that the limited computational resources define the time horizon of Djuric’s predictions. EX1003, ¶97. Any prediction Djuric makes of a region an object may reach occurs within the computational window or time horizon (*threshold time*) defined by Djuric’s limited computational resources. *Id.* Thus, the trajectories and waypoints further teach *a region [...] where the traffic entity is expected to reach within a threshold time interval* because they define the areas in the environment that Djuric’s model predicts the object will potentially

travel to within the computational window or time horizon defined by Djuric's limited computational resources. *Id.*

A POSITA would have further found it obvious that the trajectories and waypoints of objects in the environment image around Djuric's autonomous vehicle are "*a region of the point cloud.*" EX1003, ¶98. The sensor data used in Djuric includes "LIDAR point cloud(s) [...] associated with one or more object(s) [...] [and] one or more geographic characteristics of the geographic area." EX1004, ¶0022. Thus, a POSITA would have understood that the surroundings mapped in Djuric (e.g., as shown in Figure 8) and displayed in the image of the environment is LIDAR point cloud data captured by sensors. *Id.*; EX1003, ¶98. Thus, the specific trajectories and waypoints output from the model reflect future locations of the object in the LIDAR point cloud image (*[for each of the one or more traffic entities:] determining a region of the point cloud where the traffic entity is expected to reach within a threshold time interval*). EX1004, ¶¶0007, 0022, 0026, 0052, 0059, 0071, 0082-0094, Fig. 8; Section X.A.2.c (element 1[b]); EX1003, ¶98.

i. 1[d-v]

Djuric teaches element 1[d-v]. EX1003, ¶99. As explained for element 1[d-iv], the collection of predicted trajectories in Djuric defines a *region of the point cloud where the traffic entity is expected to reach within a threshold time interval*. See Section X.A.2.h. Djuric teaches that this *region* is modified because Djuric's system "determine[s] a selected predicted trajectory" from "among the plurality of

predicted trajectories of the object” by selecting the determined trajectory that has the highest confidence level. EX1004, ¶¶0006, 0034, 0080-0094, 0113, Fig. 9. Selecting a one “winner” trajectory out of the initial plurality involves eliminating some locations that the object was predicted to reach on other trajectories, thereby *modifying the region* that the object is predicted to reach. *Id.*; EX1003, ¶99. Moreover, the plurality of predicted trajectories for the object and associated trajectory confidence levels correspond to the *hidden context* of the object. Sections X.A.2.f (element 1[d-ii]), X.A.2.g (element 1[d-iii]); EX1003, ¶99. Accordingly, by selecting the determined predicted trajectory for each object (and by extension, discarding other trajectories for the object) based on the output confidence levels, Djuric teaches *[for each of the one or more traffic entities:] modifying the region based on the hidden context of the traffic entity*. EX1004, ¶¶0085-0094, 0113, Fig. 8; EX1003, ¶99.

j. 1[e]

Djuric teaches element 1[e]. After the model “output[s] one or more predicted future locations of the object (e.g., a predicted object trajectory),” the autonomous vehicle uses the trajectory to “plan and control its motion (e.g., to avoid the object).” EX1004, ¶¶0021 (parentheticals in original), 0034-0035, 0052, 0071, 0077-0078, 0103, 0112-0114; Sections X.A.2.h, X.A.2.i (elements 1[d-iv], 1[d-v]). Specifically, “the vehicle computing system 102 can control a motion of the vehicle 104 based at least in part on the selected predicted trajectory.” EX1004, ¶¶0113-0114. Djuric’s

system thus “best navigates the vehicle relative to the objects at their future locations.” *Id.*, ¶¶0035, 0077. Djuric also lists “enhancing passenger/vehicle/object safety” and “improving vehicle efficiency by reducing collisions” as benefits of its system. *Id.*, ¶¶0021, 0036. Thus, a POSITA would have understood that the vehicle is controlled such that it stays a distance greater than zero away from surrounding traffic (*navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities*). EX1003, ¶100; EX1004, ¶¶0021, 0077, 0103, 0112-0114, Figs. 6, 9.

Further, a POSITA would have found it obvious to control a vehicle such that it stays a distance greater than zero away from surrounding traffic and the selected predicted trajectory for an object (*navigating the autonomous vehicle so that the autonomous vehicle stays at least a threshold distance away from the modified region of each of the one or more traffic entities*) based on Djuric’s goals of “enhancing passenger/vehicle safety” and “reducing collisions.” EX1003, ¶101; EX1004, ¶¶0021, 0036, 0077, 0103, 0112-0114, Figs. 6, 9. This would have been a predictable combination of well-known prior art elements (e.g., Djuric’s autonomous vehicle control based on object trajectory prediction; the well-known method of not causing collisions by staying away from surrounding traffic) according to known methods to yield predictable results (Djuric’s model avoiding predicted trajectories to reduce collisions). *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007); EX1003, ¶101. A POSITA would have also had a reasonable

expectation of success because Djuric already explains that its system is designed to ensure that its autonomous vehicle does not collide with objects in its environment.

Id.

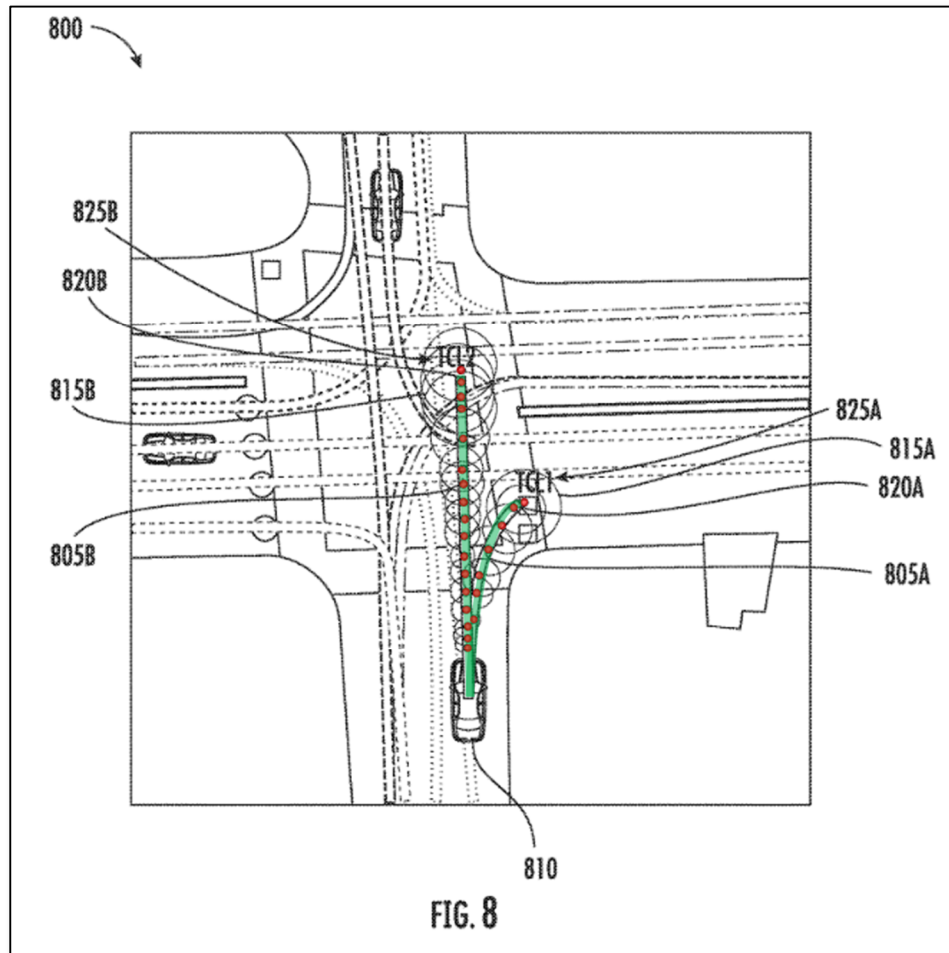
3. Claim 2

Djuric renders obvious claim 2. EX1003 ¶102. Djuric teaches that the “objects” being analyzed “can include, for example, other vehicles, pedestrians, bicycles, etc.” (*the traffic entity represents a user*). EX1004, ¶0022.

Djuric’s model outputs predicted trajectories and waypoints (*the region*) based on an input “combined data set” that includes “state data” of the object in image data (*the region is in a direction determined based on a motion vector of the traffic*). EX1004, ¶¶0025, 0079-0080, 0084, 0107-0109; *see* Section X.A.2.h (element 1[d-iv]). The state data for an object may include “an estimate of the object’s current and/or past location (also referred to as position), current and/or past speed/velocity, current and/or past acceleration, current and/or past heading, current and/or past orientation.” EX1004, ¶¶0025, 0051. At least the “speed/velocity” and “heading” combine to form a vector, and thus the output waypoints are *based on a motion vector of the traffic entity*. EX1003, ¶103.

As discussed above, Djuric’s model selects a predicted trajectory (which includes waypoints) based on a determined hidden context of the object (*modifying the region based on the hidden context comprises: responsive to determining based on the hidden context...*). EX1004, ¶¶0081, 0085-0088, 0094, 0113; *see* Section

X.A.2.i (element 1[d-v]). Djuric describes an example scenario where the model determines that a vehicle, a *traffic entity represented by a user*, traveling in a right-hand travel lane toward an intersection “may be likely to take a right-hand turn or travel straight through the intersection.” EX1004, ¶0084.



Djuric Fig. 8 (In part, annotated)

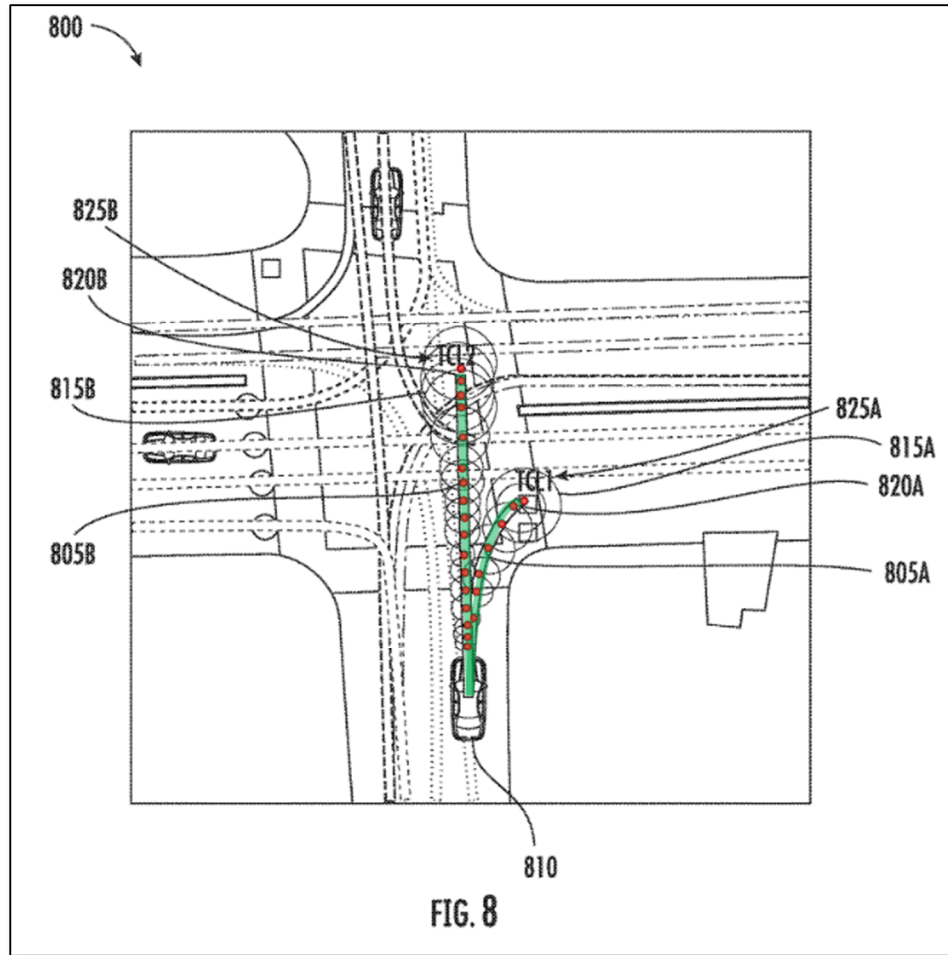
As Figure 8 shows, if the selected predicted trajectory corresponds to “travel[ing] straight through the intersection” (left trajectory, above), then the predicted trajectories and waypoints extend straight. EX1004, ¶0084, Fig. 8; EX1003, ¶105. In other words, *responsive to determining[...]*that the user**

represented by the traffic entity is likely to move “straight through the intersection” (in the direction having a component along the motion vector), Djuric teaches extending the predicted trajectories and waypoints (the region) in that same direction (along the direction of the motion vector). EX1004, Fig. 8; EX1003 ¶105.

4. Claim 3

Djuric renders obvious Claim 3. EX1003 ¶106. Claim 3 is similar to Claim 2 except that, instead of the user being likely to move in the same direction as its current motion vector, in Claim 3 the user is likely to move in an opposite direction. *Id*; see *supra* Section X.A.3. So, for the same reasons as in Claim 2, Djuric discloses “wherein the region is in a direction determined based on a motion vector of the traffic entity,” “wherein the traffic entity represents a user,” and “modifying the region based on the hidden context.” *Supra* Section X.A.3; EX1004, ¶¶0025, 0079-0080, 0084, 0107-0109; see also Sections X.A.2.h (element 1[d-iv]), X.A.2.i (element 1[d-v]); EX1003, ¶106.

Similar to Claim 2, Djuric also teaches *decreasing a size of the region along the direction of the motion vector* in response to *determining based on the hidden context that the user represented by the traffic entity is likely to move in a direction having a component opposite to direction of the motion vector*. EX1003, ¶107. Djuric’s example scenario where the model determines that a vehicle “may be likely to take a right-hand turn or travel straight through the intersection” is again informative. EX1004, ¶0084.



Djuric Fig. 8 (In part, annotated)

As shown in the above in the right-side trajectory, when the model determines that the vehicle is likely to turn right (instead of going straight) the predicted trajectories and waypoints do not extend as far in the direction that the vehicle is currently traveling (*decreasing a size of the region along the direction of the motion vector*). EX1004, ¶0084, Fig. 8; EX1003, ¶108. A POSITA would have understood that the right turn trajectory involves acceleration opposite in a direction relative to how the vehicle was traveling (e.g., braking), which is *a component* [of the predicted motion] *opposite to direction of the motion vector. Id.*

Additionally, it would have been obvious that a right-hand turn could be on a road that forms an acute angle with the current path, such that the motion vector after the turn has a component opposite to the original path of the car. For example, it would have been obvious to a POSITA that the vehicle could make a U-turn or a second right hand turn to fully travel in the opposite direction. EX1003, ¶109. Consistent with Figure 8, it would have been obvious to a POSITA that such a turn would further reduce the extent to which the predicted trajectories and waypoints (the *region*) extend in the direction of the vehicle’s current movement (*decreasing a size of the region along the direction of the motion vector*). EX1003, ¶109.

Additionally, a POSITA would have found it obvious that by selecting a right-hand turn trajectory with a motion vector that has a component opposite to the original path of the car, and by discarding the left-side trajectory of FIG. 8 that goes “straight through the intersection,” *a size of the region along the direction of the motion vector is decreas[ed]*. *Id.*, ¶¶0085-0094, Fig. 8; EX1003, ¶110.

5. Claim 4

Djuric renders obvious claim 4. EX1003, ¶111. As explained for claim 1, Djuric teaches that predicted trajectories for an object—for example, a vehicle—represent a *hidden context* for the vehicle because they each represent a “goal” the vehicle “may aim to accomplish,” such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.2.f (element 1[d-ii]). Because this *hidden context* represents the goals of a vehicle (and in turn, its driver),

the *hidden context* represents the state of mind regarding where the driver wants to go (*represents a state of mind of a user represented by the traffic entity*). EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; EX1003, ¶111. This aligns with the '346 Patent, which explains that a state of mind may include “that a pedestrian is likely to cross the street or that an oncoming vehicle is unlikely to be willing to yield.” EX1001, 12:7-12.

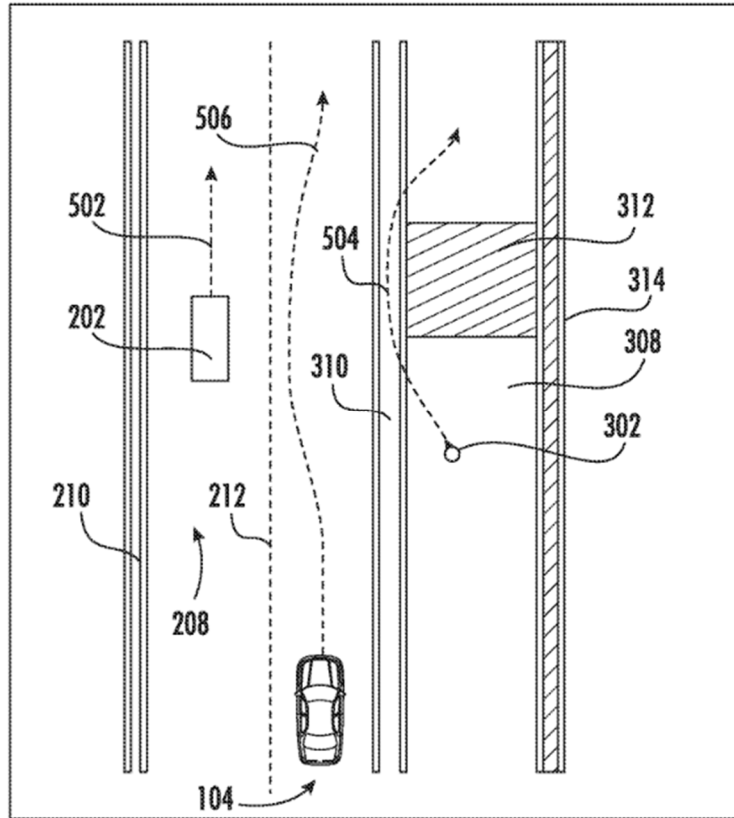
6. Claim 5

Djuric renders obvious claim 5. EX1003 ¶112. As explained for claim 1, Djuric teaches that the predicted trajectories for an object—for example, a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle “may aim to accomplish,” including tasks like turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.2.f (element 1[d-ii]). Because this *hidden context* represents the goals/intentions of a vehicle (and in turn, the vehicle’s driver), the *hidden context* represents a task that the driver is planning on accomplishing (*represents a task that a user represented by the traffic entity is planning on accomplishing*). EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; EX1003, ¶112.

7. Claim 6

Djuric renders obvious claim 6. EX1003 ¶113. As explained for claim 1, Djuric teaches that the predicted trajectories for an object—for example, a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle

“may aim to accomplish” such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.2.f (element 1[d-ii]). Because this *hidden context* represents the goals/intentions of a vehicle (and in turn, the vehicle’s driver), it would have been obvious to a POSITA that the *hidden context* can represent that driver’s awareness of the autonomous vehicle (*represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity*). EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; EX1003, ¶113. For example, a POSITA would have understood that another vehicle’s (e.g., object 202 in Figure 5) awareness of the autonomous vehicle 104 would impact “the potential risk to vehicle 104” that is relevant to vehicle 104 passing object 202. *See* EX1004, ¶0076; EX1003, ¶113. This is also consistent with the disclosure in Djuric that the model considers the awareness of other surroundings, like a pedestrian’s awareness of an obstruction in a sidewalk. *Id.*; EX1004, ¶¶0074-0077.



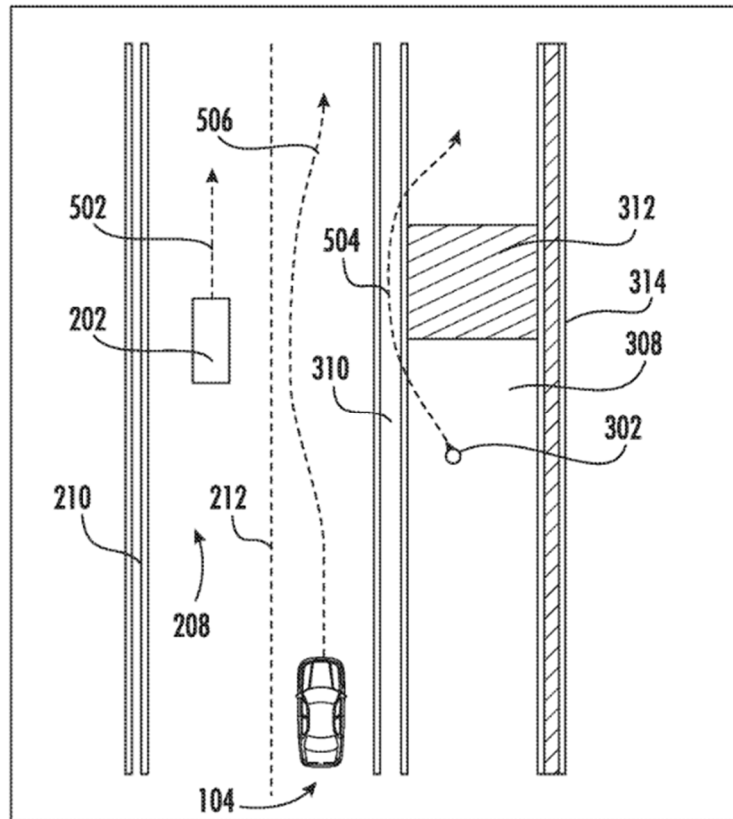
EX1004, Fig. 5

8. Claim 7

Djuric renders obvious claim 7. As explained for claim 1, Djuric teaches that the predicted trajectories for an object—for example, a pedestrian—represent a *hidden context* for the vehicle because they represent goals the pedestrian expects to achieve. EX1004, ¶¶0032-0043, 0074-0077, 0081-0085, 0107-0113; Section X.A.2.f(element 1[d-ii]). Because this *hidden context* represents the goals/intentions of a pedestrian, the *hidden context* represents a goal of a user represented by the traffic entity. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; EX1003, ¶114.

For example, Djuric discloses predicting a trajectory 504 based on a pedestrian’s goal to “bypass an obstruction 312 in a sidewalk.” EX1004, ¶¶0074-

0077, Fig. 5.



EX1004, Fig. 5

Djuric’s pedestrian (user) *expects to achieve the goal within a threshold time interval* because, as discussed for element 1[d-iv], the trajectory is predicted to occur within a threshold time interval. EX1003, ¶116; EX1004, ¶¶0007, 0022, 0026, 0071, 0082-0094, Fig. 8. For example, Djuric explains that each “predicted trajectory can include a plurality of predicted future locations of an object at one or a plurality of future times.” EX1004, ¶0071. These “predicted future locations” of each “predicted trajectory” define a region that the object is expected to reach, and the “plurality of future times” define the corresponding *threshold time interval*. *Id.*;

EX1003, ¶116. Djuric also teaches *a threshold time interval* in an additional way.

Djuric discloses that “the vehicle computing system has limited computational resources onboard the vehicle and is running in real-time, as the vehicle is travelling.”

EX1004, ¶0037. A POSITA would have found it obvious that the limited computational resources thus limit the time horizon of Djuric’s predictions. EX1003 ¶116. Therefore, Djuric’s prediction that the pedestrian plans to bypass an obstruction on the sidewalk occurs within the computational window or time horizon defined by Djuric’s limited computational resources. *Id.*

9. Claim 8

Djuric renders obvious claim 8. EX1003, ¶117. As described for claim 1, Djuric describes a method “of predicting object motion and controlling autonomous vehicles.” EX1004, ¶¶0095, 0104; *See* Section X.A.2.e (element 1[d-i]). Djuric further discloses that navigating the autonomous vehicle includes “send[ing] one or more **control signals** to the responsible vehicle control component (e.g., braking control system, steering control system, acceleration control system) to execute the instructions and implement the motion plan” (*generating signals for controlling the autonomous vehicle[...] and sending the generated signals to controls of the autonomous vehicle*). EX1004, ¶0078 (parenthetical in original, emphasis added); *see also id.*, ¶¶0035, 0079, 0105, 0107; EX1003, ¶117.

Djuric further teaches that the autonomous vehicle is controlled *based on one or more motion parameters and the hidden context of each of the one or more traffic*

entities. EX1003, ¶118. As discussed above, Djuric discloses *motion parameters* used in the model: “state data” for an object (vehicle, pedestrian, bicyclist) that “describe[s] an estimate of the object's current and/or past location (also referred to as position), current and/or past speed/velocity, current and/or past acceleration, current and/or past heading, current and/or past orientation” which are all motion (e.g., velocity, acceleration) parameters of the object (*motion parameters...of the traffic entity*). EX1004, ¶¶0025 (parenthetical in original), 0051, 0071, 0102, 0127; EX1003, ¶118; Section X.A.2.e (element 1[d-i]). The state data is used as part of a “combined data set” that is input to the model to generate the predicted trajectories for the objects. EX1004, ¶¶0104-0114. The act of *controlling the autonomous vehicle* in Djuric is thus based on the predicted trajectory (*hidden context*), and also based on the state data (*motion parameters*) of the detected objects (*each of the one or more traffic entities*). EX1004, ¶¶0021-0025, 0034-0036, 0051, 0071, 0077-0078, 0083-0094, 0102-0114, 0127, Fig. 9; Section XI.A.2.j (element 1[e]); EX1003, ¶118.

10. Claim 9

Djuric renders obvious claim 9. EX1003, ¶119. Djuric discloses that the autonomous vehicle includes one or more sensors, such as a LIDAR system and one or more cameras, which acquire sensor data associated with objects proximate to the autonomous vehicle. EX1004, ¶¶0022, 0047, 0051, 0064. The sensor data can include “image data” and “LIDAR data,” such as point cloud data (*the sensor data represents one or more images captured by a camera mounted on the autonomous*

vehicle or a lidar scan captured by a lidar mounted on the autonomous vehicle). *Id.*; EX1003, ¶119.

11. Claim 10

Djuric renders obvious claim 10 for the same reasons discussed with respect to claim 1, as indicated below. EX1003, ¶120; *See* Section X.A.2.

a. 10[Pre]

Djuric teaches element 10[Pre]. EX1003, ¶121. For example, Djuric discloses that “vehicle computing system 102 can include one or more computing device(s) 701. The computing device(s) 701 of the vehicle computing system 102 can include processor(s) 702 and a memory 704 (e.g., onboard the vehicle 104).” EX1004, ¶0116 (parenthetical in original). Additionally, “memory 704 can include one or more non-transitory computer-readable storage media,” (*[a] non-transitory computer readable storage medium*) and “the memory 704 can store instructions 706 that when executed by the one or more processors 702 cause the one or more processors 702 (the computing system 102) to perform operations.” (*storing instructions, that when executed by a processor, cause the processor to perform*). EX1004, ¶¶0117-0118; *See also* Section X.A.2.a (element 1[Pre]); EX1003, ¶121.

b. 10[a]

See Section X.A.2.b (element 1[a]); EX1003, ¶122.

c. 10[b]

See Section X.A.2.c (element 1[b]); EX1003, ¶123.

d. 10[c]

See Section X.A.2.d (element 1[c]); EX1003, ¶124.

e. 10[d-i]

See Section X.A.2.e (element 1[d-i]); EX1003, ¶125.

f. 10[d-ii]

See Section X.A.2.f (element 1[d-ii]); EX1003, ¶126.

g. 10[d-iii]

See Section X.A.2.g (element 1[d-iii]); EX1003, ¶127.

h. 10[d-iv]

See Section X.A.2.h (element 1[d-iv]); EX1003, ¶128.

i. 10[d-v]

See Section X.A.2.i (element 1[d-v]); EX1003, ¶129.

j. 10[e]

See Section X.A.2.j (element 1[e]); EX1003, ¶130.

12. Claim 11

Djuric renders obvious claim 11 for the same reasons discussed above in Section X.A.3 (Claim 2). EX1003, ¶131.

13. Claim 12

Djuric renders obvious claim 12 for the same reasons discussed above in Section X.A.4 (Claim 3). EX1003, ¶132.

14. Claim 13

Djuric renders obvious claim 13 for the same reasons discussed above in Section X.A.5 (Claim 4). EX1003, ¶133.

15. Claim 14

Djuric renders obvious claim 14 for the same reasons discussed above in Section X.A.6 (Claim 5). EX1003, ¶134.

16. Claim 15

Djuric renders obvious claim 15 for the same reasons discussed above in Section X.A.7 (Claim 6). EX1003, ¶135.

17. Claim 16

Djuric renders obvious claim 16 for the same reasons discussed above in Section X.A.8 (Claim 7). EX1003, ¶136.

18. Claim 17

Djuric renders obvious claim 17 for the same reasons discussed above in Section X.A.9 (Claim 8). EX1003, ¶137.

19. Claim 18

Djuric renders obvious claim 18 for the same reasons discussed above in Section X.A.10 (Claim 9). EX1003, ¶138.

20. Claim 19

Djuric renders obvious claim 19 for the same reasons discussed with respect to claims 1 and 10, as indicated below. *See* Sections X.A.2, X.A.11 (Claims 1, 10), EX1003, ¶139.

a. 19[Pre]

See Sections X.A.11.a (element 10[Pre]), X.A.2.a (element 1[Pre]); EX1003, ¶140.

b. 19[a]

See Section X.A.2.b (element 1[a]); EX1003, ¶141.

c. 19[b]

See Section X.A.2.c (element 1[b]); EX1003, ¶142.

d. 19[c]

See Section X.A.2.d (element 1[c]); EX1003, ¶143.

e. 19[d-i]

See Section X.A.2.e (element 1[d-i]); EX1003, ¶144.

f. 19 [d-ii]

See Section X.A.2.f (element 1[d-ii]); EX1003, ¶145.

g. 19[d-iii]

See Section X.A.2.g (element 1[d-iii]); EX1003, ¶146.

h. 19[d-iv]

See Section X.A.2.h (element 1[d-iv]); EX1003, ¶147.

i. 19[d-v]

See Section X.A.2.i (element 1[d-v]); EX1003, ¶148.

j. 19[e]:

See Section X.A.2.j (element 1[e]); EX1003, ¶149.

21. Claim 20

Djuric renders obvious claim 20 for the same reasons discussed above in Section X.A.3 (Claim 2). EX1003, ¶150.

22. Claim 21

Djuric renders obvious claim 21 for the same reasons discussed above in Section X.A.4 (Claim 3). EX1003, ¶151.

23. Claim 22

Djuric renders obvious claim 22 for the same reasons discussed above in Section X.A.5 (Claim 4). EX1003, ¶152.

24. Claim 23

Djuric renders obvious claim 23 for the same reasons discussed above in Section X.A.7 (Claim 6). EX1003, ¶153.

B. Ground 2: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Cox

To the extent Djuric alone does not teach *output[ing] summary statistics of expected human responses describing a hidden context* as recited in element 1[d-ii], the claimed operations are also disclosed by Cox. For the reasons described below, it would have been obvious to a POSITA to combine these features from Cox with Djuric. EX1003, ¶154.

1. Summary of Prior Art: Cox-EX1006

Cox is a PCT patent application by Harvard College, which lists Samuel Anthony as an inventor—the same Samuel Anthony listed as an inventor of the '346 Patent. EX1001, Cover; EX1006, Cover. Cox is a material reference to the '346 Patent that teaches several concepts relevant to the Challenged Claims but was never brought to the attention of the Patent Office during prosecution or otherwise

considered. EX1001, Cover; *see generally* EX1002.

Cox is directed toward “machine learning augmented by measurements acquired from human populations.” EX1006, Abstract, ¶¶0001-0004. Cox teaches crowd-sourcing training data used to train machine learning algorithms. EX1006, Abstract, ¶¶0001-0004, 0015-0016, 0025-0034, 0044, 0060. Just like the ’346 Patent, Cox discloses compiling an n-amount of images (training objects), transmitting the images to a plurality of human annotators (human labelers), compiling statistical information regarding the distribution of human responses (including the actual responses and psychometric data associated with the responses), and training a supervised learning model using gathered information “consistent with the decisions of the human annotators 160.” *Cf* EX1006, ¶¶0029-0032, Figs. 1A-B, 3 *with* EX1001, 2:31-43, 8:8-10:4; Fig. 4. The goal of Cox’s machine learning techniques is to “better mimic human performance.” EX1006, ¶0004. Moreover, 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 / semiautonomous automobile[]” applications. EX1006, ¶¶0056, 0060.

Cox is analogous art to the ’346 Patent, from the same field of endeavor as the Patent (e.g., autonomous vehicle perception systems), and is reasonably pertinent to the particular problem that the Patent was trying to solve (e.g., how to “accurately predict motion of non-stationary traffic objects” in an autonomous vehicle’s

surroundings), describing the gathering training data from a population of human labelers, training a model using the training data, and executing it to output predictions. EX1006, ¶¶0029-0032, Figs. 1A-B, 3; EX1001, 1:34-67, 2:26-43, 3:19-35, 15:4-22; Fig. 4; EX1003, ¶157.

2. Motivation to Combine Djuric and Cox

A POSITA would have been motivated to combine Djuric such that Djuric’s “goal-based prediction” (*hidden context*; EX1004, ¶¶0080-0084) and trajectory confidence levels (*summary statistics*) are generated in accordance with human-annotated/ manually-labeled training data, as disclosed by Cox. EX1004, ¶0053; EX1006, ¶¶0004, 0027-0032, 0051; EX1003, ¶158. A POSITA would have been motivated to combine Djuric and Cox in this way for several reasons. EX1003, ¶158. Djuric describes that the model can be trained with “supervised training techniques” and “using labeled state data and geographic data with known future location(s) and/or trajectories of objects (e.g., as captured in a data log, manually).” EX1004, ¶¶0027, 0053, 0134. As taught by Cox, using human-annotated image data to train Djuric’s model would enable the model to provide predictions that are “more consistent with the decisions of the human annotators,” and that “better mimic human performance.” EX1006, ¶¶0004, 0027-0032, 0051; EX1003, ¶158.

For example, Cox simply expands on Djuric’s disclosure of a model for an autonomous vehicle that is trained using supervised training and manual labeling. EX1004, ¶¶0027, 0053, 0134, 0137; Section X.A.1; EX1003, ¶159. Djuric discloses

generating training data, using the training data to train a model, and executing the model to predict trajectories for objects around the autonomous vehicle. EX1004, ¶¶0053, 0084-0086, 0090-0094, 0113, 0134-0139. 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. EX1006, ¶¶0007-0010, 0015-0019, 0025-0038, 0051, 0060, Fig. 4B; EX1003, ¶159. 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.” EX1006, ¶¶0004, 0060; see also *id.*, ¶¶0008-0010, 0015-0019, 0025-0035, 0051. Cox thus provides a natural expansion and improvement on Djuric’s base disclosure. EX1003, ¶159.

A POSITA seeking to implement Djuric’s model would also have looked to Cox to provide more instructions for training a machine learning model to predict human behavior. EX1003, ¶160. Both 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. EX1006, ¶¶0016, 0056. Cox therefore provides an explicit teaching,

suggestion, or motivation to make a combination with a system for controlling driverless (or autonomous) vehicles, just like Djuric. EX1003, ¶160.

A POSITA would have also been motivated to combine the teachings of Djuric and Cox because the combination would have benefited Djuric’s model EX1003, ¶161. Enhancing Djuric’s model 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 Djuric’s model so that it better considers the state of mind of other road users and what they might be thinking when present in the autonomous vehicle’s environment. *Id.*; EX1006, ¶¶0007-0010, 0015-0019, 0025-0035, 0051, 0060. This would have improved the Djuric model’s goal-based predictions and thus helped reduce collisions between the autonomous vehicle and human drivers or pedestrians. EX1003, ¶161.

All the reasons explained above establish that a POSITA would have been motivated to combine Djuric and Cox as described. EX1003, ¶162. This combination would have been the predictable combination of well-known prior art elements (e.g., the Djuric model being a supervised learning model; Cox’s model training and optimization techniques) according to known methods to yield predictable results (the Djuric-Cox model generating predictions for objects shown in image data that are “more consistent with the decisions of the human annotators” who annotated the training data). EX1006, ¶¶0032, 0037; *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398,

415-421 (2007). The combination would have further used known techniques (known machine learning model training and optimization techniques as taught by Cox) to improve similar devices (the Djuric model for goal-based predictions) in the same way (using Cox’s aforementioned techniques). *Id.* Using human-annotated image data in Djuric (as taught by Cox) would have been well within the skill of a POSITA, and a POSITA would have had an expectation of success in the modification because of the similarity to the training approach disclosed in Djuric. *Id.*; EX1003, ¶162.

3. Claim 1

Cox’s teachings only impact how a model generates *output summary statistics of expected human responses describing a hidden context* in element 1[d-ii]. The combination of Djuric and Cox teaches element 1[d-ii] for the reasons explained below, as well as the reasons addressed in Ground 1. Sections X.A.2, X.A.2.f (element 1[d-ii]); EX1003, ¶163. Moreover, the combination of Djuric and Cox teaches the remaining elements of claim 1 for the same reasons as addressed in Ground 1. *Id.*³

³ Where later elements (e.g., 1[d-iii]) refer back to the analysis in 1[d-ii], the additional analysis described in Ground 2 for 1[d-ii] is equally applicable to those later elements as well.

a. 1[d-ii]

The combination of Djuric and Cox teaches 1[d-ii] for the same reasons addressed in Ground 1 (Section X.A.2.f), and also in an additional way. EX1003, ¶164. Specifically, and as described further below, the Djuric-Cox model generates predicted trajectories (*hidden context*) and trajectory confidence levels (*summary statistics*) reflective of human-labeled image data that is used to train the model (*expected human responses*). Section X.A.2.f (element 1[d-ii]); EX1006, ¶¶0019, 0029-0038, 0060; Fig. 4B; EX1003, ¶164.

Cox teaches that human-labeled image data is used to train a prediction model. EX1006, ¶¶0019, 0029-0038, 0060; Fig. 4B; EX1003, ¶165. Training data is generated by crowd sourcing, where a “large population” of human annotators annotate a “training object” such as an image to generate “statistical information” based on all the responses received from the multiple different human annotators. EX1006, ¶¶0007-0010, 0015-0019, 0025-0035, 0051, Figs. 1A, 1B; EX1003, ¶165. For example, the human annotator responses describe a state of mind of a subject person in the image, e.g., “an emotion” such as “jealous, panicked, arrogant, or hateful.” EX1006, ¶0016; *id.*, ¶¶0007-0010, 0015-0019, 0025-0035, 0051; EX1003 ¶165. Cox’s supervised machine learning model is trained using the statistical information. EX1006, ¶¶0007-0010, 0015-0019, 0025-0035, 0051; EX1003, ¶165. Training may utilize a “loss function,” where the model predicts classifications of training data and is optimized such that the difference between the predicted

classifications output by the model and the training data is minimized. EX1006, ¶¶0008, 0019, 0029-0038, 0060, Fig. 4B; EX1003 ¶167. The model is then utilized after training “to make predictions” that are “**more consistent with the decisions of the human annotators 160.**” EX1006, ¶¶0027-0032 (emphasis added), 0051. The goal of Cox’s machine learning techniques is to “better mimic human performance.” *Id.*, ¶0004.

Thus, as applied to Djuric, Cox teaches that the predicted trajectories and trajectory confidence levels are generated in accordance with the human-annotated/manually-labeled training data (*summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image*). EX1004, ¶¶0080-0084; EX1006, ¶¶0004, 0027-0032, 0051; Section X.B.2. The human-annotated/labeled training data includes annotated/labeled image data and is used to train the Djuric model to provide predictions that are “more consistent with the decisions of the human annotators,” and that “better mimic human performance.” EX1006, ¶¶0004, 0032; EX1003, ¶168. This means the model’s outputs are representative of how humans would have annotated or labeled images input to the model if the images were part of the training data. EX1006, ¶¶0004, 0019, 0029-0038, 0051, 0060; Fig. 4B; EX1003, ¶168. As such, Djuric teaches that the model receives an image input and outputs predicted trajectories and trajectory confidence levels (*...output summary statistics*) (see Section X.A.2.f, element 1[d-ii]), and Cox, as applied to Djuric, teaches that the predicted confidence levels (*summary statistics*)

represent a likelihood of how humans would have annotated/labeled the input image data, thus teaching *summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image*. EX1003, ¶168; See Section X.A.2.f (element 1[d-ii]).

4. Claims 10 and 19

Djuric and Cox render obvious claims 10 and 19 for the same reasons discussed above with respect to claim 1. The combination with Cox only impacts the analysis for elements 10[d-ii] and 19[d-ii]. Elements 10[d-ii] and 19[d-ii] are taught by the Djuric-Cox combination for the reasons stated above for 1[d-ii]. *Supra*, Section X.B.3.a (Ground 2, element 1[d-ii]). Djuric and Cox further teach elements 10[d-ii] and 19[d-ii] as well as the remaining elements of claims 10 and 19 for the same reasons addressed for Djuric alone in Ground 1 above.⁴ Sections X.A.11 (claim 10), X.A.20 (claim 19); See EX1003, ¶169.

5. Claims 2-9, 11-18, and 20-23

Djuric and Cox render obvious claims 2-9, 11-18, and 20-23 for the same reasons discussed with respect to Djuric in Section X.A above.

⁴ Like for claim 1, where later elements (e.g., 10[d-iii] and 19[d-iii]) refer back to the analysis in 10[d-ii] or 19[d-ii], the additional analysis described in Ground 2 for 10[d-ii] and 19[d-ii] is equally applicable to those later elements as well.

C. Grounds 3 and 4: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Zhu or Djuric, Zhu, and Cox

To the extent Djuric alone (or the combination of Djuric and Cox) does not disclose or render obvious (1) using sensor data to generate a point cloud representation, and/or (2) determining and modifying a region of the point cloud, such operations are also disclosed by Zhu. Claims 1-23 would thus be obvious over the further combination of Zhu, with either Djuric (Ground 3) or with Djuric and Cox (Ground 4). EX1003, ¶171.

1. Summary of Prior Art: Zhu-EX1005

Zhu is generally directed toward generating a drivable road surface representation “based on data from cameras or other image generation devices and data from laser light detection and ranging (LIDAR) devices.” EX1005, ¶0004. The cameras and LIDAR devices are used by an autonomous vehicle “to accurately detect and map the drivable road surfaces proximate to or ahead of the location of the autonomous vehicle.” *Id.*, ¶0015. The autonomous vehicle can “produce accurate drivable road surface representations in a drivable road surface map,” and “use the drivable road surface map to safely and efficiently control the vehicle... while avoiding obstacles and safely controlling the vehicle.” *Id.*, ¶¶0013, 0015.

Zhu describes using “cameras or image capturing devices [that] may include one or more devices configured to capture a plurality of images of the environment of the vehicle.” EX1005, ¶0026. Zhu explains that LIDAR sensors generate “point

cloud data” that represents “measured distances at various points in three dimensional (3D) space in front of or adjacent to a vehicle on which the...LIDAR unit is mounted.” *Id.*, ¶¶0004, 0012, 0015, 0026. Zhu teaches to “combine” lidar point cloud data and image data by “project[ing]” the “3D point cloud data onto the 2D image data to produce mapped image/point cloud data” which thereby forms a “drivable road surface map that can be used to distinguish drivable areas of a roadway from non-drivable areas.” *Id.*, ¶¶0040, 0043-0049.

Zhu is analogous art to the ’346 Patent, from the same field of endeavor as the Patent (e.g., autonomous vehicle perception systems), and is reasonably pertinent to the particular problem that the Patent was trying to solve (e.g., how to “accurately predict motion of non-stationary traffic objects” in an autonomous vehicle’s surroundings); Zhu describes a sensor system that combines lidar point cloud data and image data to generate a vehicle environment representation. EX1005, ¶¶0004, 0012, 0013, 0037; EX1001, 1:34-67, 2:26-43, 3:19-35, 15:4-22; EX1003, ¶174.

2. Motivation to Combine Zhu with Djuric, or Djuric and Cox

A POSITA would have been motivated to combine Djuric and Zhu (and alternatively along with Cox)⁵ such that Djuric’s camera image data and LIDAR

⁵ The further inclusion of Cox does not impact the motivation to combine Djuric with Zhu, or a POSITA’s expectation of success in making the combination. EX1003, ¶¶175, 175.

point cloud data are combined to generate a drivable road surface map, where the point cloud frames captured by LIDAR devices are projected onto the image data captured by cameras to generate the drivable road surface map as taught by Zhu. EX1003, ¶175.

Both Djuric and Zhu are directed to autonomous vehicle vision systems and the control of such vehicles based on captured sensor data. Djuric already describes using LIDAR sensors and cameras to capture images of an autonomous vehicle's environment and detect objects within the environment. EX1004, ¶¶0022, 0032, 0046-0049, 0050-0051, 0059, Figs. 2A-2C. Zhu explains that combining the point cloud data output by LIDAR sensors and image data output by cameras allows for "accurate drivable road surface representations in a drivable road surface map." EX1005, ¶0013. This allows for an autonomous vehicle to "accurately detect and map the drivable road surfaces proximate to or ahead of the location of the autonomous vehicle." *Id.*, ¶0015.

A POSITA would have been motivated to use Zhu's teachings to further improve Djuric's object detection using LIDAR sensors and cameras. EX1003, ¶177. As taught by Zhu, combining the data allows objects to be detected and for drivable road surfaces (i.e., where objects are not located in a road) to be detected. *Id.* This improves the navigation of Djuric's autonomous vehicle in relation to predicted trajectories. *Id.*

This combination would have been the predictable combination of well-

known prior art elements (e.g., Djuric's autonomous vehicle with sensors that capture image data and LIDAR point cloud data; Zhu's techniques for LIDAR and image data combination) according to known methods to yield predictable results (generating a drivable road surface map by combining Djuric's camera image data and LIDAR point cloud data). *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007); EX1003, ¶178. The combination would have further used known techniques (the combination of LIDAR point cloud data and image data) to improve similar devices (Djuric's autonomous vehicle with sensors that capture image data and LIDAR point cloud data) in the same way (using Zhu's aforementioned techniques). *Id.* Lastly, because Djuric and Zhu disclose similar approaches to collecting sensor data (including LIDAR point cloud data) and making predictions, using Zhu's technique for combining LIDAR and image data in Djuric would have been well within the skill of a POSITA, and a POSITA would have had an expectation of success in the modification. *Id.*

3. Claim 1

Zhu's teachings only impact the analysis relating to point clouds in elements 1[b] and 1[d-iv]. The combination of Djuric and Zhu (for Ground 3) and the combination of Djuric, Cox, and Zhu (for Ground 4) teaches elements 1[b] and 1[d-iv] for the reasons explained below, as well as the reasons addressed in Ground 1. Sections X.A.2, X.A.2.c (element 1[b]), X.A.2.h (element 1[d-iv]); EX1003, ¶179. Moreover, the combination of Djuric and Zhu (Ground 3) and the combination of

Djuric, Cox, and Zhu (Ground 4) teach the remaining elements of claim 1 for the same reasons as addressed in Ground 1. *Id.*⁶ Additionally, for Ground 4, the combination of Djuric, Cox, and Zhu teaches element 1[d-ii] for the same reasons discussed in Ground 2. Section X.B.3.a.

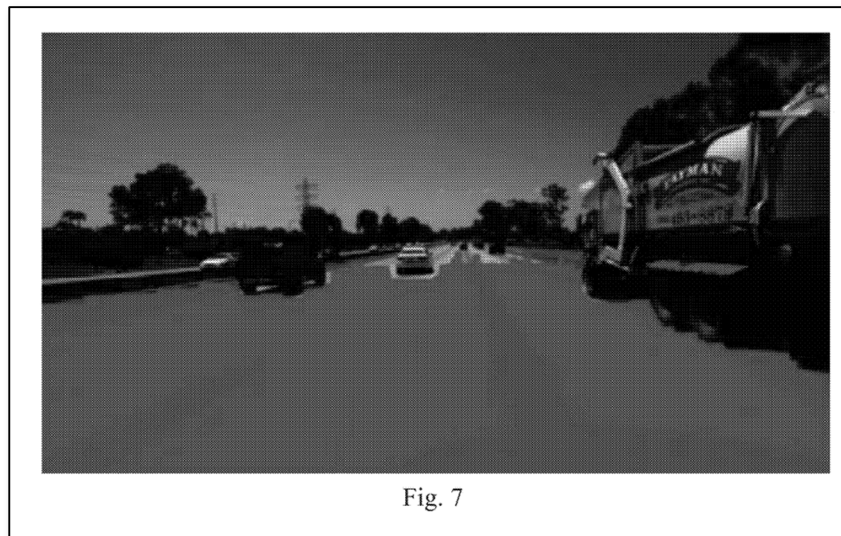
a. 1[b]

Zhu describes an autonomous vehicle having a vision system for the vehicle's surroundings that uses cameras and LIDAR. EX1005, Abstract, ¶¶0004-0015, 0021-0027, 0037-0047, Figs. 1-3. Zhu explains that LIDAR sensors generate “point cloud data” that represents “measured distances at various points in three dimensional (3D) space in front of or adjacent to a vehicle on which the...LIDAR unit is mounted.” *Id.*, ¶¶0004, 0012, 0015, 0026. Zhu teaches a process to “combine” the lidar point cloud data and the image data by “project[ing]” the “3D point cloud data onto the 2D image data to produce mapped image/point cloud data,” which thereby forms a “drivable road surface map that can be used to distinguish drivable areas of a roadway from non-drivable areas.” *Id.*, ¶¶0040, 0043-0049.

Thus, applying Zhu's teachings to Djuric, the Djuric-Zhu combination would have specified that that Djuric's camera image data and LIDAR point cloud data

⁶ Where later elements refer back to the analysis in 1[b] or 1[d-iv], the additional analysis described in Grounds 3/4 for 1[b] or 1[d-iv] is equally applicable to those later elements.

would have been combined to generate a drivable road surface map, where the point cloud frames captured by lidar devices are projected onto the image data captured by cameras to generate the drivable road surface map (*generating a point cloud representation of the surroundings of the autonomous vehicle based on the sensor data*). EX1005, Abstract, ¶¶0004-0015, 0021-0027, 0037-0047, Figs. 1-3; *See* Section X.A.2.b (element 1[a]); EX1003, ¶181. Figure 7 below shows an example drivable road surface map. EX1005, ¶0046, Fig. 7.



EX1005, Fig. 7

b. 1[d-iv]

As explained in Ground 1, Djuric teaches *determining a region [...] where the traffic entity is expected to reach within a threshold time interval* by generating predicted trajectories and waypoints where a traffic entity is expected to reach in a threshold time interval. *See* Section X.A.2.h. As explained below, in combination with Zhu it would have been obvious to a POSITA that this *region* would be on the

drivable surface generated using LIDAR point cloud data and camera image data (*a region of the point cloud*). EX1004, ¶¶0007, 0022, 0026, 0052, 0059, 0071, 0082-0094, Fig. 8; See Section X.C.3.a (element 1[b]); EX1003, ¶182.

Djuric explains that the image shown by Figure 8 is an “output 800 of the model 136.” EX1004, ¶0083. And Djuric explains that such an output may be formed using “image data from a front-facing camera, other cameras, RADAR data, LIDAR data, and/or data from other sensors.” *Id.*, ¶¶0022, 0059. As taught by Zhu, such LIDAR point cloud data and camera image data can be combined to form a drivable road surface map (the point cloud representation). See Section X.C.3.a (element 1[b]). Thus, by using Djuric’s model 136 to analyze Zhu’s drivable surface map (which is formed by LIDAR point cloud data and camera image data, as contemplated by Djuric), a POSITA would have understood that Djuric’s vehicle computing system would have determined the locations of the predicted trajectories and waypoints on the drivable surface map. EX1004, ¶¶0007, 0022, 0026, 0052, 0059, 0071, 0082-0094, Fig. 8; See Section X.C.3.a (element 1[b]); EX1003, ¶182.

Djuric’s predicted trajectories and waypoints would define *where the traffic entity is expected to reach within a threshold time interval* for the same reasons addressed in Ground 1. See Section X.A.2.h (element 1[d-iv]). So, when plotted onto the drivable surface map as taught by Zhu, the predicted trajectories and waypoints are “*a region of the point cloud where the traffic entity is expected to reach within a threshold time interval.*” EX1003, ¶184

4. Claims 10 and 19

Djuric and Zhu (Ground 3) and Djuric, Cox, and Zhu (Ground 4) render obvious claims 10 and 19 for the same reasons discussed above with respect to claim 1. The combinations with Zhu only impact the analysis for elements 10[b], 10[d-iv], 19[b], and 19[d-iv]. Elements 10[b], 10[d-iv], 19[b], and 19[d-iv] are taught by the Djuric-Zhu combination and the Djuric-Cox-Zhu combination for the reasons stated above for elements 1[b] and 1[d-iv]. Sections X.C.3.a and X.C.3.b. Djuric and Zhu (Ground 3) and Djuric, Cox, and Zhu (Ground 4) further teach elements 10[b], 10[d-iv], 19[b], and 19[d-iv], as well as the remaining elements of claims 10 and 19, for the same reasons addressed for Djuric alone in Ground 1.⁷ Sections X.A.11 (claim 10), X.A.20 (claim 19). Additionally, for Ground 4, Djuric, Cox, and Zhu teach elements 10[d-ii] and 19[d-ii] for the same reasons discussed in Ground 2. *Supra*, Sections X.B.3, X.B.4.

5. Claims 2-9, 11-18, 20-23

Djuric and Zhu (Ground 3) and Djuric, Cox and Zhu (Ground 4) render obvious claims 2-9, 11-18, and 20-23 for the same reasons discussed for Djuric in Section X.A above. *See* Section X.B.5.

⁷ Like for claim 1, where later elements refer back to the analysis in 10[b], 10[d-iv], 19[b] or 19[d-iv], the additional analysis described in Grounds 3/4 for 10[b], 10[d-iv], 19[b] or 19[d-iv] is equally applicable to those later elements as well.

D. Grounds 5, 6, 7, and 8: Claims 1-23 are unpatentable under 35 U.S.C. § 103 over Djuric and Ross (Ground 5), Djuric, Cox, and Ross (Ground 6), Djuric, Zhu, and Ross (Ground 7), or Djuric, Cox, Zhu, and Ross (Ground 8)

To the extent Djuric alone (or any of the Djuric-Cox, Djuric-Zhu, or Djuric-Cox-Zhu combinations) does not render obvious dependent claims 4, 6, 13, 15, 22, and 23, further combination with Ross renders these dependent claims obvious. The analysis for the remainder of the claims is unchanged from Grounds 1 through 4. Claims 1-23 would thus be obvious over the further combination of Ross, with either Djuric (Ground 5), Djuric and Cox (Ground 6), Djuric and Cox (Ground 7), or Djuric, Cox, and Zhu (Ground 8). EX1003, ¶187.

1. Summary of Prior Art: Ross-EX1015

Ross is generally directed to techniques for “maneuvering” an autonomous vehicle. EX1015, 1:30-31, 3:39-42, 3:64-5:5. Ross describes an autonomous vehicle control system that provides “[r]easoning about the intents of objects, such as other road users” (e.g., other vehicles, pedestrians, bicyclists) so the autonomous vehicle can “function in a more ‘human-like’ or ‘polite’ way.” *Id.*, 5:6-11, 15:1-16:52. For example, Ross explains that an action of a “pedestrian waiting to cross the road” can correspond to an intent of the pedestrian “to cross the road when clear,” i.e., when it is safe to cross.

Ross is analogous art to the ’346 Patent, from the same field of endeavor as the Patent (e.g., autonomous vehicle perception systems), and is reasonably pertinent

to the particular problem that the Patent was trying to solve (e.g., how to “accurately predict motion of non-stationary traffic objects” in an autonomous vehicle’s surroundings); Ross describes an model that provides “[r]easoning about the intents of objects, such as other road users.” EX1015, 5:6-11, 15:1-16:52; EX1003, ¶189.

2. Motivation to Combine Djuric, Zhu, Cox, and Ross

A POSITA would have been motivated to combine Djuric and Zhu and/or Cox as discussed in Sections X.B.2 and X.C.2. A POSITA would have been further motivated to combine Ross with Djuric (and optionally Zhu and/or Cox) such that the Djuric-Ross model generates “goal based prediction[s]” using Djuric’s goals and Ross’s intents or actions. EX1003, ¶190.

As discussed regarding elements 1[d-ii]-1[d-iii], Djuric describes generating “goal-based prediction[s]” of object trajectories in a vehicle’s environment. EX1004, ¶¶0081-0085; *See* Sections X.A.2.f, X.A.2.g (elements 1[d-ii], 1[d-iii]). A POSITA would have understood that Djuric’s model is not limited to predicting trajectories based solely on the example goals disclosed by Djuric. EX1003, ¶191. Instead, a POSITA would have found it obvious to supplement Djuric’s use of goals with the prediction of intent and/or actions disclosed by Ross, which would expand the goal scenarios Djuric’s model is capable of predicting. EX1003, ¶191. Ross describes an intent model that 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.”

EX1015, 1:55-2:13, 4:33-46, 9:13-40, 15:1-21. “[E]ach possible intent of the set includes a corresponding next action to be initiated by the road user.” *Id.*, 2:43-44 A POSITA would have found it obvious to supplement Djuric’s listed goals with the intents and/or actions disclosed by Ross to expand Djuric’s predictive capabilities, leading to more accurate predictions across a broader range of circumstances. EX1003, ¶191.

Thus, a POSITA would have been motivated to combine Djuric and Ross (and optionally Zhu and/or Cox) as described. This combination would have been the predictable combination of well-known prior art elements (e.g., Djuric’s goal-based object movement prediction model; Ross’s techniques for intent-based predictions of road users) according to known methods to yield predictable results (using Djuric’s model to predict object motion based on intents as determined based on Ross’s teachings). *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007); EX1003, ¶192. The combination would have further used known techniques (Ross’s intent determination for predictions) to improve similar devices (Djuric’s goal-based prediction model) in the same way (using Ross’s aforementioned techniques). *Id.* Lastly, because Djuric is already using goal-based predictions, further using Ross’s technique for determining intent would have been well within the skill of a POSITA, and a POSITA would have had an expectation of success in the modification. *Id.*

3. Claims 1-3, 5, 7-12, 14, and 16-21

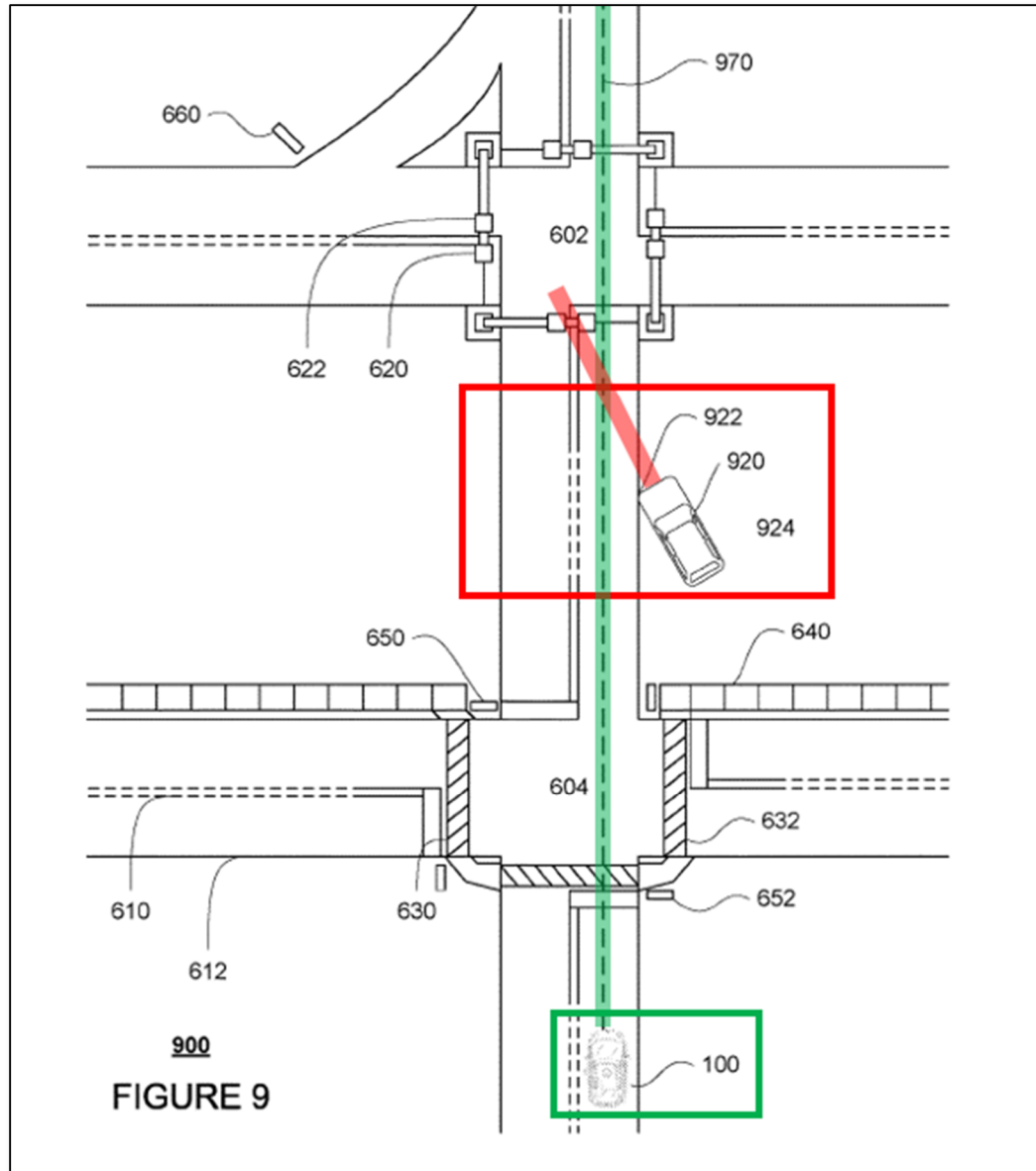
Djuric (Ground 1), Djuric and Cox (Ground 2), Djuric and Zhu (Ground 3),

and Djuric, Cox, and Zhu render obvious claims 1-3, 5, 7-12, 14, and 16-21 for the same reasons discussed above in Grounds 1 through 4. *See* respective sections of X.A (Ground 1), X.B (Ground 2), and X.C (Grounds 3 and 4); EX1003, ¶193.

4. Claim 4

Djuric and Ross render obvious claim 4 for the same reasons discussed in Section X.A.5. To the extent Djuric does not explicitly disclose the hidden context represents a state of mind, the combination of Djuric and Ross renders this obvious. EX1003, ¶194.

Ross describes a model that 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.” EX1015, 1:55-2:13, 4:33-46, 9:13-10:10, 15:1-16:52. Such actions reflect the “possible intent[s]” (*hidden context*) associated with the object. 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.” (*the hidden context represents a state of mind of a user represented by the traffic entity*). EX1015, 16:5-41, Fig. 9.



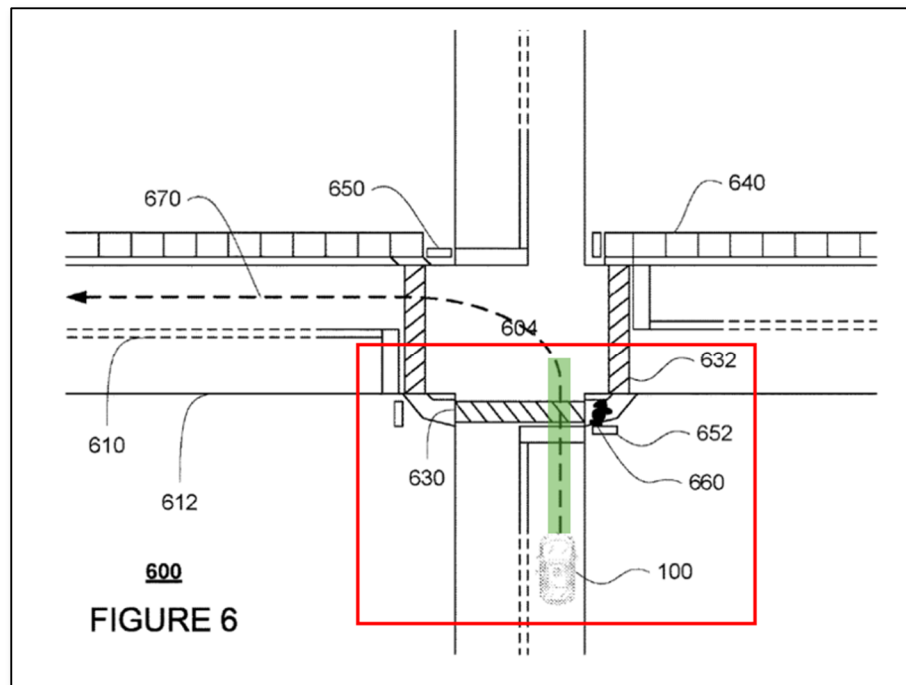
EX1015, Fig. 9 (in part, annotated)

Ross's teachings align with the '346 Patent, which explains that a state of mind may include "that a pedestrian is likely to cross the street or that an oncoming vehicle is unlikely to be willing to yield." EX1001, 12:7-12; *see also* EX1015, 16:5-41, Fig. 9; EX1003, ¶196.

5. Claim 6

Djuric and Ross render obvious claim 6 for the same reasons discussed in Section X.A.7. To the extent Djuric does not explicitly disclose the hidden context represents a degree of awareness of the autonomous vehicle by a user, the combination of Djuric and Ross renders this obvious. EX1003, ¶197.

Another predicted intent described by Ross is that a pedestrian “would like to cross the roadway” that a vehicle 100 is traveling on. EX1015, 16:23-41, Fig. 6.



EX1015, Fig. 6 (in part, annotated)

As described by Ross, a pedestrian may “cross[] a roadway” on which vehicle 100 is traveling or “cross[] a roadway in a crosswalk” on the road vehicle 100 is traveling, and when the crossing traverses the path of the vehicle (shown in green above), the vehicle must stop for the pedestrian to cross. EX1015, 4:47-65, 9:41-

10:10, 16:23-41, Fig. 6. Moreover, Ross explains that an action of a “pedestrian waiting to cross the road” can correspond to an intent “to cross the road when clear[.]” *Id.*, 15:50-16:22. Ross further explains a pedestrian “may want to cross but... won't step out in front of a vehicle unless that vehicle slows down and allows for them to cross.” *Id.*, 3:53-56. In such a case, the predicted intent (*hidden context*) thus demonstrates an awareness of the pedestrian of the vehicle driving toward the pedestrian (*represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity*). EX1003, ¶199.

6. Claim 13

Djuric and Ross render obvious claim 13 for the same reasons discussed above in Section X.D.4 (Claim 4). EX1003, ¶200.

7. Claim 15

Djuric and Ross render obvious claim 15 for the same reasons discussed above in Section X.D.5 (Claim 6). EX1003, ¶201.

8. Claim 22

Djuric and Ross render obvious claim 22 for the same reasons discussed above in Section X.D.4 (Claim 4). EX1003, ¶202.

9. Claim 23

Djuric and Ross render obvious claim 23 for the same reasons discussed above in Section X.D.5 (Claim 6). EX1003, ¶203.

E. Printed Matter

Many limitations of the Challenged Claims are not entitled to patentable weight because they merely recite the content of information, i.e., printed matter. *Praxair Distribution, Inc. v. Mallinckrodt Hosp. Prods. IP Ltd.*, 890 F.3d 1024, 1032 (Fed. Cir. 2018) (“Claim limitations directed to the content of information and lacking a requisite functional relationship are not entitled to patentable weight because such information is not patent eligible subject matter under 35 U.S.C. § 101.”). In particular, the following limitations are printed matter:

- 1[a], 10[a], 19[c] – “the sensor data comprising one or more images”
- 1[b], 10[b], 19[b] – “a point cloud representation of the surroundings of the autonomous vehicle”
- 1[d-i], 10[d-i], 19[d-i] – “motion parameters describing movement of the traffic entity”
- 1[d-ii], 10[d-ii], 19[d-ii] – “[the]/[an input] image showing [the]/[a] traffic entity,” “summary statistics of expected human responses describing a hidden context of the input traffic entity shown in the input image”
- 1[d-iv], 10[d-iv], 19[d-iv] – “a region of the point cloud where the traffic entity is expected to reach within a threshold time interval”
- Claims 4, 13 – “hidden context represents a state of mind of a user represented by the traffic entity”
- Claims 5, 14 – “hidden context represents a task that a user represented by

the traffic entity is planning on accomplishing”

- Claims 6, 15 – “hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity”
- Claims 7, 16 – “hidden context represents a goal of a user represented by the traffic entity, wherein the user expects to achieve the goal within a threshold time interval”
- Claims 9 and 18 – “the sensor data represents one or more images captured by a camera mounted on the autonomous vehicle or a lidar scan captured by a lidar mounted on the autonomous vehicle”

Each of the above limitations recites a data element (an image, a parameter, data, etc.) and then defines that data element based on its content. But a patent cannot cover what an “image show[s],” or what “parameters” and “statics ... 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 requirement in limitation 1[a] that “the sensor data compris[e] one or more images” should be ignored because the content of the sensor data is not patentable. Similarly, the recited limitations about what the predicted “hidden context represents” in claims 4-7 and 13-16 should be ignored—the patentee cannot claim the content of this “hidden context.” *See Praxair*, 890 F.3d at 1031-32. Because the above limitations are defined by the information contained, 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, ¶207; *see also* 10[pre] and 19[pre]. Rather than impacting how the computer memory operates, the printed matter limitations in the Challenged Claims “merely informs 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).

Respectfully Submitted,

Date: October 10, 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 October 10, 2025, a complete and entire copy of the **PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-23 OF U.S. PATENT NO. 11,520,346** including exhibits and testimony relied upon and Power of Attorney were served on Patent Owner via FedEx overnight at the correspondence address of record for the '346 Patent and 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: October 10, 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 13,941 words, within the 14,000 word limit allowed under §42.24(a)(1)(i).

Date: October 10, 2025

/Roger Fulghum/
Roger Fulghum
Reg. No. 39,678

Attorney for Petitioner, Tesla, Inc.