UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TESLA, INC.

Petitioner

v.

PERCEPTIVE AUTOMATA LLC

Patent Owner

IPR2025-01576
U.S. Patent No.: 11,467,579 B2

PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-20 OF U.S. PATENT NO. 11,467,579

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

TABLE OF CONTENTS

Page

PETITIONER'S EXHIBIT LIST

Ex[No.]	Description of Documents
1001	U.S. Patent No. 11,467,579 ("'579 Patent")
1002	File History of U.S. Patent No. 11,467,579
1003	Declaration of Dr. Jason Janét
1004	U.S. Patent Application Publication No. 2019/0049970 ("Djuric")
1005	U.S. Patent Application Publication No. 2020/0160535 ("AA")
1006	U.S. Provisional Application 62/802,151
1007	U.S. Provisional Application 62/822,269
1008	International Publication No. WO2014/210334 ("Cox")
1009	U.S. Patent Application Publication No. 2018/0247200 ("Rolfe")
1010	Complaint in *Perceptive Automata LLC v. Tesla, Inc.*, Case No. 2:25-cv-00742 (E.D. Tex. filed July 23, 2025)
1011	Excerpts from *Probabilistic Robotics*, Sebastian Thrun, Wolfram Burgard, and Dieter Fox, The MIT Press, 2006
1012	CV of Dr. Jason Janét
1013	Excerpts from *Robot Motion Planning*, Jean-Claude Latombe, Kluwer Academic Publishers, 1991
1014	Excerpts from *Principles of Robot Motion*, 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, the '579 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"). EX1010.

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. 1st 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. 1st St., Suite 1300 Austin, TX 78704-1296 Phone: (512) 322-2507 ellyar.barazesh@bakerbotts.com 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 '579 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 for navigating autonomous vehicles, comprising:
1[a]	training a probabilistic neural network, the probabilistic neural network configured to perform steps comprising:
1[a-i]	[the probabilistic neural network configured to perform steps comprising:] receiving as input, an image of traffic, the image displaying a traffic entity belonging to the traffic,
1[a-ii]	[the probabilistic neural network configured to perform steps comprising:] generating a feature vector for a plurality of features, the feature vector comprising values describing statistical distribution for each feature, and
1[a-iii]	[the probabilistic neural network configured to perform steps comprising:] generating output representing hidden context for the traffic entity, the output comprising a plurality of values, each value representing a likelihood of receiving a particular user response from a user presented with the image;
1[b]	receiving a new image captured by a camera mounted on an autonomous vehicle navigating through traffic;
1[c]	executing the probabilistic neural network to generate output representing hidden context for at least a traffic entity displayed in the new image;
1[d]	determining a measure of uncertainty for each of the plurality of values; and
1[e]	navigating the autonomous vehicle to avoid the traffic entity displayed

	in the new image, the navigation based on at least the measure of uncertainty generated by the probabilistic neural network.

Claim 2

2	The method of claim 1, wherein the values describing statistical distribution for each feature comprise a mean value and a standard deviation for the feature.

Claim 3

3	The method of claim 1, wherein training the probabilistic neural network comprises determining evidence lower bound between the plurality of values predicted and a plurality of values determined from user responses obtained from users presented with images from a training dataset.

Claim 4

4	The method of claim 1, wherein determining the measure of uncertainty for each of the plurality of values comprises: generating a plurality of samples for the input image using the probabilistic neural network; and determining a confidence interval for each of the plurality of values using the plurality of samples.

Claim 5

5	The method of claim 1, wherein navigating the autonomous vehicle comprises ensuring that the autonomous vehicle stays at least a threshold distance away from the traffic entity displayed in the new image, the threshold distance determined based on the measure of uncertainty generated by the probabilistic neural network.

Claim 6

6	The method of claim 5, wherein the threshold distance is determined to be a value directly related to the measure of uncertainty generated by the probabilistic neural network.

Claim 7

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

Claim 8

8	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 9

9	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 10

10	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 11

11	The method of claim 1, wherein navigating the autonomous vehicle comprises: generating signals for controlling the autonomous vehicle based on motion parameters and the hidden context of the traffic entity; and sending the generated signals to controls of the autonomous vehicle.

Claim 12

12	The method of claim 1, wherein the probabilistic neural network is a probabilistic convolutional neural network.

Claim 13

13[Pre]	A non-transitory computer readable storage medium storing instructions, that when executed by a processor, cause the processor to perform steps comprising:
13[a]	training a probabilistic neural network, the probabilistic neural network configured to perform steps comprising:
13[a-i]	[the probabilistic neural network configured to perform steps comprising:] receiving as input, an image of traffic, the image displaying a traffic entity belonging to the traffic,
13[a-ii]	[the probabilistic neural network configured to perform steps comprising:] generating a feature vector for a plurality of features, the feature vector comprising values describing statistical distribution for each feature, and
13[a-iii]	[the probabilistic neural network configured to perform steps comprising:] generating output representing hidden context for the traffic entity, the output comprising a plurality of values, each value representing a likelihood of receiving a particular user response from a user presented with the image;
13[b]	receiving a new image captured by a camera mounted on an autonomous vehicle navigating through traffic;
13[c]	executing the probabilistic neural network to generate output representing hidden context for at least a traffic entity displayed in the new image;
13[d]	determining a measure of uncertainty for each of the plurality of values; and
13[e]	navigating the autonomous vehicle to avoid the traffic entity displayed in the new image, the navigation based on at least the measure of uncertainty generated by the probabilistic neural network.

Claim 14

14	The non-transitory computer readable storage medium of claim 13, wherein the values describing statistical distribution for each feature comprise a mean value and a standard deviation for the feature.

Claim 15

15	The non-transitory computer readable storage medium of claim 13, wherein training the probabilistic neural network comprises determining evidence lower bound between the plurality of values predicted and a

	plurality of values determined from user responses obtained from users presented with images from a training dataset.

Claim 16

16	The non-transitory computer readable storage medium of claim 13, wherein determining the measure of uncertainty for each of the plurality of values comprises: generating a plurality of samples for the input image using the probabilistic neural network; and determining a confidence interval for each of the plurality of values using the plurality of samples.

Claim 17

17	The non-transitory computer readable storage medium of claim 13, wherein navigating the autonomous vehicle comprises ensuring that the autonomous vehicle stays at least a threshold distance away from the traffic entity displayed in the new image, the threshold distance determined based on the measure of uncertainty generated by the probabilistic neural network.

Claim 18

18	The non-transitory computer readable storage medium of claim 17, wherein the threshold distance is determined to be a value directly related to the measure of uncertainty generated by the probabilistic neural network.

Claim 19

19	The non-transitory computer readable storage medium of claim 13, wherein navigating the autonomous vehicle comprises: generating signals for controlling the autonomous vehicle based on motion parameters and the hidden context of the traffic entity; and sending the generated signals to controls of the autonomous vehicle.

Claim 20

20[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:
20[a]	training a probabilistic neural network, the probabilistic neural network configured to perform steps comprising:
20[a-i]	[the probabilistic neural network configured to perform steps comprising:] receiving as input, an image of traffic, the image displaying a traffic entity belonging to the traffic,
20[a-ii]	[the probabilistic neural network configured to perform steps comprising:] generating a feature vector for a plurality of features, the feature vector comprising values describing statistical distribution for each feature, and
20[a-iii]	[the probabilistic neural network configured to perform steps comprising:] generating output representing hidden context for the traffic entity, the output comprising a plurality of values, each value representing a likelihood of receiving a particular user response from a user presented with the image;
20[b]	receiving a new image captured by a camera mounted on an autonomous vehicle navigating through traffic;
20[c]	executing the probabilistic neural network to generate output representing hidden context for at least a traffic entity displayed in the new image;
20[d]	determining a measure of uncertainty for each of the plurality of values; and
20[e]	navigating the autonomous vehicle to avoid the traffic entity displayed in the new image, the navigation based on at least the measure of uncertainty generated by the probabilistic neural network.

III. INTRODUCTION

Tesla, Inc. ("Petitioner" or "Tesla") petitions for *inter partes* review of claims 1-20 (the "Challenged Claims") of U.S. Patent No. 11,467,579 (EX1001) (the "'579 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 '579 Patent is available for IPR, and that Petitioner is not barred or otherwise 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 '579 Patent was filed on February 6, 2020, and claims priority to U.S. Provisional Application 62/802,151, filed on February 6, 2019, and U.S. Provisional Application 62/822,269, filed on March 22, 2019. EX1001, Cover Page; EX1006; EX1007. Without conceding that the '579 Patent is entitled to these priority dates, the Djuric, AA, Cox, Rolfe, and Ross references are all prior art to the '579 Patent before February 6, 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. 2020/0160535 (Ali Akbarian, abbreviated "AA") (EX1005); filed on November 15, 2018; prior art under 35 U.S.C. § 102(a)(2).
- International Publication No. WO2014/210334 ("Cox") (EX1008); 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 Application Publication No. 2018/0247200 ("Rolfe") (EX1009); published August 30, 2018, and filed August 18, 2016; prior art under 35 U.S.C. § 102(a)(1) and § 102(a)(2).
- 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, AA, Cox, Rolfe, or Ross were cited or otherwise discussed during prosecution.

[1] Cox lists Harvard as applicant and Samuel Anthony as an inventor. EX1008, Cover Page. This inventor appears to be the same Samuel Anthony listed as an inventor of the '579 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	1-20	§ 103	Djuric and AA
2	1-20	§ 103	Djuric, AA, and Cox
3	1-20	§ 103	Djuric, AA, and Rolfe
4	1-20	§ 103	Djuric, AA, Cox, and Rolfe
5	1-20	§ 103	Djuric, AA, and Ross
6	1-20	§ 103	Djuric, AA, Cox, and Ross

VII. SUMMARY OF THE '579 PATENT

A. Prosecution History

The Office issued a first action notice of allowance for the '579 Patent, stating that the "prior art of record fail[ed][] to teach or suggest" the following claim elements:

> "generating output representing hidden context for the traffic entity, the output comprising a plurality of values, each value representing a likelihood of receiving a particular user response from a user presented with the image,
>
> …

determining a measure of uncertainty for each of the plurality of values; and

navigating the autonomous vehicle to avoid the traffic entity displayed in the new image, the navigation based on at least the measure of uncertainty generated by the probabilistic neural network."

EX1002, 124-131. But as demonstrated by the grounds of this petition, these limitations—like the Challenged Claims as a whole—were well known. As such, the Challenged Claims should not have been allowed.

B. Technology Background

Probabilistic neural networks are computer programs that analyze input data and generate an output based on the data. EX1003, ¶58. Such models are "probabilistic" because, for example, they generate outputs with a particular statistical distribution. *Id*. It is well known that such distributions have uncertainties associated with them, both in the field of neural networks and more generally in the field of statistics. *Id.*

Neural networks, including probabilistic neural networks, may be trained using labeled training data that is known to represent a correct output. EX1003, ¶59. 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' decisions when labeling training data. *Id*.

For example, as was commonly known in the art, for training a traffic

prediction probabilistic neural network, human annotators label all instances of training data images that display objects moving in traffic. EX1003, ¶60. The labels used by human annotators correspond to the output predictions made by the neural network. *Id*. Therefore, after training the neural network with the labels, the neural network can be used to determine the probability of different movement patterns by vehicles represented in new, non-training data, such as camera images. *Id*. The neural network's predictions for vehicles reflect how the human annotators would have labeled the new images because the model was trained based on the human annotators' decisions when labeling training data. *Id*. For example, different humans may label images with some degree of variation, which would be reflected by the output from the probabilistic neural network. *Id*.

C. The Alleged Invention

The alleged invention of the '579 Patent relates to (1) training a probabilistic neural network (PNN) to generate output representing "hidden context," (2) determining a measure of uncertainty of the output, and (3) navigating an autonomous vehicle based on the determined uncertainty. EX1001, 2:6-18, 6:45-63, 7:37-55, 14:11-37, 15:11-26. But, as illustrated by the cited art, this invention was well known for many years in the fields of both probabilistic neural networks and autonomous vehicle control and reflects standard training and operations of conventional probabilistic networks. Section X; EX1003, ¶61; *see also generally* EX1011.

The substance of the '579 Patent is nothing new. The Patent explains how a PNN is trained to predict "hidden context" for traffic entities, such as vehicles and pedestrians, but this language is just jargon; such "hidden context" may simply "represent[] the behavior of the traffic entities," for example. EX1001, Abstract, 2:1-4, 2:64-3:15. The PNN 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, 2:10-18, 4:22-35, 4:63-5:3, 7:4-21, 8:39-44, Figs. 4, 8; EX1003, ¶62; EX1004, ¶0053; EX1005, ¶0035 (explaining how manual or supervised training techniques may be used to train models to predict future behavior). Indeed, the *inventor's own prior art* (the Cox reference, EX1008) shows that the '579 Patent should not have been issued. Cox discloses to ensure that machine learning model predictions are "more consistent with the decisions of the human annotators" and "better mimic human performance." EX1008, ¶¶0004, 0032.

Thereafter, the trained PNN is used to make a prediction based on "live" input data, such as images of traffic entities. EX1001, 10:15-28, 13:26-48. Using probabilistic algorithms to make predictions based on live input data, including images, was well known, for example, at least before 2005. *See generally* EX1011, pp. 250, 350-353. EX1003, ¶63; *see also* EX1004, ¶¶0030-0032, 0061-0068, 0100-0102, 0111-0112; EX1005, ¶¶0073, 0079-0081 (describing the input of current images into a model that predicts future movement of objects depicted in the images.)

The Patent also describes determining a measure of uncertainty for the PNN output, which was a well-known concept of all probabilistic algorithms, for example, before 2005. *See, e.g.,* EX1011, pp.25-26 ("The key idea in probabilistic robotics is to represent uncertainty explicitly"). EX1001, 5:62-65, 6:45-7:3, 15:27-65, 19:35-48; EX1003, ¶63; EX1004, ¶¶0034, 0085-0088, 0093-0094, 0112, 0136.

Next, the Patent describes controlling an autonomous vehicle based on the prediction and the measure of uncertainty. EX1001, 2:36-40, 16:48-50, 18:60-65. However, control of "cars that travel by themselves" using the well-known concept of "probabilistic algorithms" was also a well-known concept dating back to at least 2005. *See, e.g.,* EX1011 pp. 24-26, 351, 391, 494, 508. EX1003, ¶64. Indeed, controlling an autonomous vehicle based on the probabilistic predictions of a neural network was well-known, as demonstrated by at least Djuric and AA. EX1003, ¶64; EX1004, ¶¶0008, 0021-0024, 0035-0036, 0050, 0078, 0095, 0103-0104, 0114; EX1005, ¶¶0002, 0030, 0050, 0073, 0084.

The '579 Patent describes nothing more than well-known machine learning concepts. EX1003, ¶65. 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 '579 Patent as of February 6, 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, ¶66.

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 9 1303 (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, ¶67.

X. THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. Ground 1: Claims 1-20 are unpatentable under 35 U.S.C. § 103 over Djuric and AA

Djuric alone renders obvious most limitations of the Challenged Claims, as explained below. Further, AA explicitly discloses: (1) the description of the neural network as a "probabilistic neural network" (*see* element 1[a]), and (2) the discussion of a "feature vector" in element 1[a-ii], and it would have been obvious to a POSITA to combine the teachings of Djuric and AA. Sections X.A.3.b, X.A.3.d (elements 1[a] and 1[a-ii]); EX1003, ¶68.

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

1. Summary of Prior Art

a. 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 that captures camera images of its surroundings, which include objects such as other vehicles and pedestrians, and generates predictions of object motion using a predictive model, such as a neural network. EX1004, ¶¶0027, 0047, 0053, 0061-0067. The model is trained using supervised training techniques with manually labeled object trajectories. 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, ¶69.

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. The model also generates a trajectory confidence level for each of the plurality of predicted trajectories. EX1004, ¶¶0089-0093. The trajectory confidence levels indicate "a per-trajectory probability, quantifying how likely an object is to follow a certain trajectory." EX1004, ¶0089.

Additionally, the predicted trajectories each include a plurality of waypoints. EX1004, ¶0083. Djuric's neural network model determines "a way-point confidence level for each way-point," where the way-point confidence levels for a predicted trajectory "help quantify the confidence in how exactly the object is going to follow a certain trajectory." EX1004, ¶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 '579 Patent, from the same field of endeavor as the Patent (e.g., predictive neural network model techniques), 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"). EX1004, Abstract, ¶¶0032-0034, 0081; EX1001, 1:38-62, 4:8-22; EX1003, ¶73.

b. AA-EX1005

AA discloses using "probabilistic generative models in an artificial neural network to predict subject body poses and subject movement intent." EX1005, ¶0001. AA describes using the movement intent predictions "in autonomous driving applications" to predict pedestrian movement. EX1005, ¶¶0050, 0073.

Specifically, AA describes training a convolutional neural network that is a "trained probabilistic model." EX1005, ¶¶0005-0008, 0034-0035, 0052-0058, 0065-0073. AA's teachings align with the '579 Patent, which explains that its neural network can be a "probabilistic convolutional neural network." EX1001, 5:56-60; EX1003, ¶75.

AA also discloses "feature vectors" that "may include a number that corresponds to a possible feature of the image[.]" EX1005, ¶¶0034-0037. AA further describes that applying a softmax function to one of these feature vectors can convert the numbers in the feature vector "to a probability" for each feature represented by the values in the vector. EX1005, ¶0037, EX1003, ¶76. AA's application of a softmax function to the generated feature vector aligns with the '579 Patent's purported invention, which describes processing a feature vector by applying a softmax function to produce probability values. EX1005, 18:23-33, 19:7-10, 19:29-31, Fig. 9, Fig. 10, EX1003, ¶76.

AA is analogous art to the '579 Patent, from the same field of endeavor as the Patent (e.g., predictive neural network model techniques) 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"). EX1005, ¶¶0034-0044, 0050, Fig. 2D; EX1001, 1:38-62, 4:8-22; EX1003, ¶77.

2. Motivation to Combine Djuric and AA

A POSITA would have been motivated to combine Djuric and AA such that:

- The model in the Djuric-AA combination teaches a **probabilistic** convolutional neural network, as taught by Djuric and AA. EX1004, ¶¶0027, 0033, 0053, 0069, 0072, 0079-0094, 0100, 0110-0113, 0126, 0132-0139; EX1005, ¶¶0005-0008, 0034-0035, 0052-0058, 0065-0073; EX1003, ¶78.
- The model in the Djuric-AA combination generates a feature vector as described in AA. EX1005, ¶¶0034-0037; EX1003, ¶78.

A POSITA would have been motivated to adopt AA's disclosure of a "probabilistic model" in Djuric at least because Djuric teaches that its model may be a Bayesian network, and AA's disclosures regarding probabilistic model simply provide more implementation detail. EX1004, ¶¶0126, 0132; *see also id.*, 0052-0053, 0081–0094, 0113; EX1003, ¶79. A POSITA would have recognized that implementing the probabilistic approach from AA would help Djuric "account for the uncertainty of future movements" and "account for the uncertainty of plausible future poses in predicting subject motion… in real-time." EX1005, ¶¶0051, 0052, 0065, 0084–0086. EX1003, ¶79. Thus, a POSITA would have been motivated to combine AA's PNN with Djuric's model, which itself describes various examples

of PNN's. EX1003, ¶79.

Djuric even details probabilistic training targets (e.g., Gaussian waypoint uncertainty and per-trajectory classification with "one additional output… to estimate the per-trajectory probability (resulting in 5N+1 output for a trajectory)"). EX1004, ¶¶0086-0090 (parenthetical in original). Refining Djuric's Bayesian network or convolutional neural network as a probabilistic convolutional neural network, per AA is a natural, predictable refinement that aligns with Djuric's approach. EX1003, ¶80. This refinement would also improve Djuric's prediction results, as AA explains that its probabilistic generative approach "account[s] for the uncertainty of plausible future poses in predicting subject motion." EX1005, ¶¶0051-0052; EX1003, ¶80.

A POSITA would also implement AA's "feature vector" stage to make Djuric's model more effective. EX1003, ¶81. Using a compact feature-vector representation after Djuric's rasterized, multi-channel inputs would reduce spatial redundancy and thereby improve compute efficiency. EX1003, ¶81. This directly serves Djuric's stated goal that "the combined data set can allow the model's network to be smaller and, thus, use less computational resources (e.g., processing, storage, etc.)[.]" EX1004, ¶0037. Implementation would also be straightforward. Djuric already "input[s] the combined data set into the machine-learned model" to predict trajectories of objects around the autonomous vehicle (EX1004, ¶0037), and it would be straightforward to use this data to generate a feature vector as taught by

AA. EX1003, ¶81. Thus, this modification of Djuric would have been well within the skill of a POSITA, and a POSITA would have an expectation of success in the modification. EX1003, ¶81.

In view of the above, the combination of Djuric and AA would have been the predictable combination of well-known prior art elements (e.g., Djuric's autonomous vehicle and convolutional neural network; AA's probabilistic convolutional neural network and feature vector generation) according to known methods to yield predictable results (an autonomous vehicle with a predictive convolutional neural network; the Djuric-AA model using a feature vector). *KSR Int'l Co. v. Teleflex Inc.,* 550 U.S. 398, 415-421 (2007); EX1003, ¶82. The combination would have further used known techniques (known probabilistic convolutional neural networks as taught by AA) to improve similar devices (Djuric's autonomous vehicle and convolutional neural network) in the same way (using AA's aforementioned probabilistic neural network). *Id*.

Lastly, this combination would also have been well within the skill of a POSITA, and a POSITA would have had a reasonable expectation of success. EX1003, ¶83. In particular, because Djuric discloses that its "machine-learned model can be or can otherwise include one or more various model(s)" to "observe its surrounding environment" and AA discloses a "probabilistic model" capable of accounting for the uncertainty of plausible future movements of detected objects, a POSITA would have understood that it would be a simple incorporation of the

"probabilistic model" of AA in the model of Djuric to achieve the results touted by AA. EX1004, ¶¶0003, 0027, 0053, 0126, 0132; EX1005, ¶¶0051, 0052, 0065, 0084–0086; EX1003, ¶83. The incorporation of a feature vector as taught by AA would be similarly within the skill of a POSITA, especially if the probabilistic model from AA is being used. EX1003, ¶83. Similarly, since Djuric already discloses that similar models (e.g., a Bayesian network) may be used, a POSITA would have an expectation of success in making the combination with AA. *Id*. .

3. Claim 1

Djuric and AA render obvious claim 1. EX1003, ¶84.

a. 1[Pre]

To the extent limiting, Djuric and AA teach the preamble. EX1003, ¶85. Djuric describes systems and methods that "enable an autonomous vehicle to predict the respective motion of… objects and plan the vehicle's motion accordingly" (*[a] method for navigating autonomous vehicles*). EX1004, ¶¶0020-0024, 0035-0036, 0044-0046, 0050, 0078, 0095, 0103-0104, 0114, 0118, Fig. 9 (item 920); EX1003, ¶85.

b. 1[a]

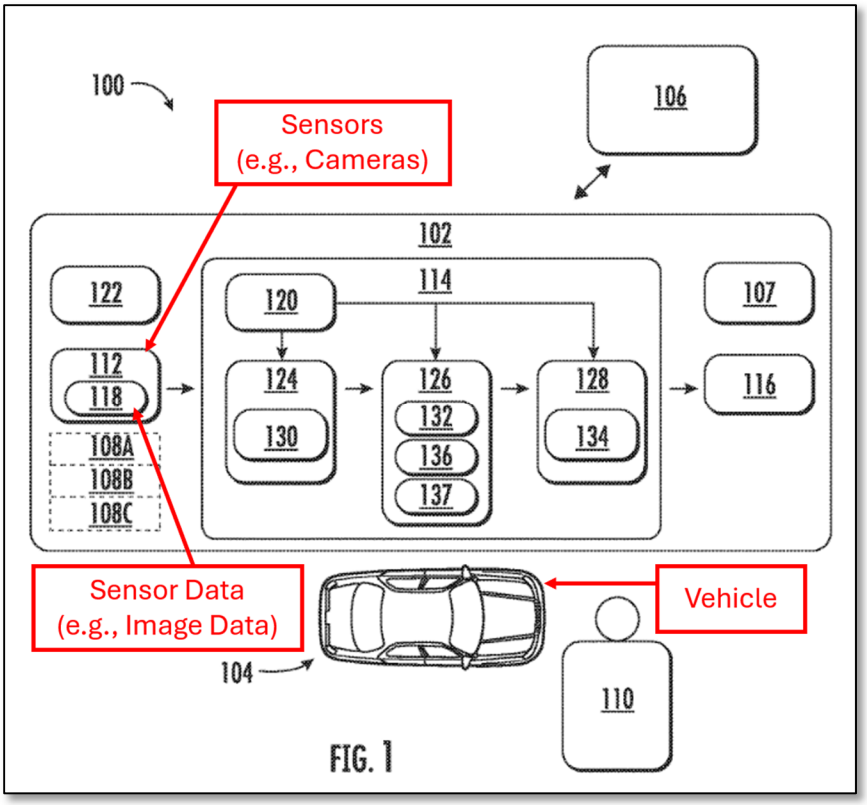
Djuric and AA teach element 1[a]. EX1003, ¶86. Djuric discloses that its system uses "a model," which is trained by "supervised training techniques." EX1004, ¶¶0027, 0053, 0132-0139. Djuric's trained model may be various types of "neural network[s]," including a "Bayesian network[]," which is a type of

probabilistic neural network (PNN) (*training a probabilistic neural network*). EX1004, ¶¶0027, 0053, 0126, 0132; EX1003, ¶86. The model is configured to perform a number of steps, including those discussed, *infra*, in Sections X.A.3.c (element 1[a-i]) through X.A.3.e (element 1[a-iii]) (*the probabilistic neural network configured to perform steps comprising*). EX1003, ¶86.

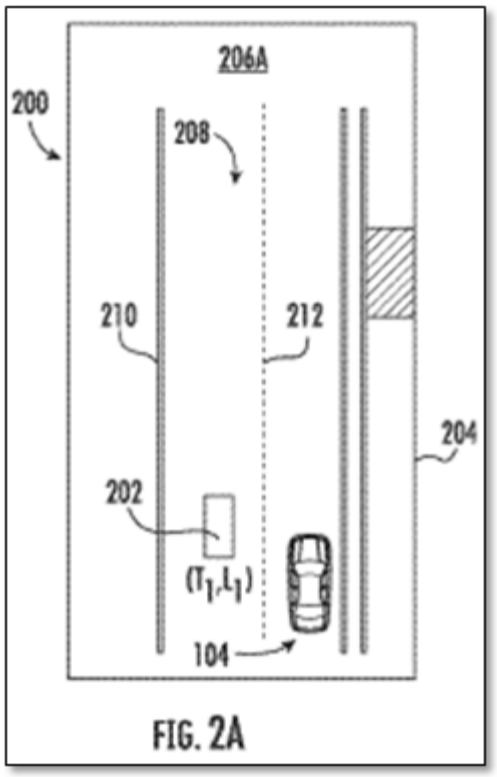
To the extent not explicitly disclosed by Djuric, AA describes training a convolutional neural network, as a "trained probabilistic model," that "generates a probability distribution" to "account for the uncertainty of plausible future poses in predicting subject motion." EX1005, ¶¶0005-0008, 0034-0035, 0052-0058, 0065-0073. Thus, AA's convolutional neural network that is trained as a probabilistic model is *a probabilistic neural network. Id.*; EX1003, ¶87. AA's teachings align with the '579 Patent, which explains that its neural network can be a "probabilistic convolutional neural network." EX1001, 5:56-60. In view of this teaching and as described above, it would have been obvious to a POSITA to implement Djuric using a PNN as taught by AA (*a probabilistic neural network*). EX1003, ¶87. As explained above, Djuric further teaches using "supervised training techniques" on the model (*training a probabilistic neural network*). EX1004, ¶¶0027, 0033, 0053, 0069, 0072, 0079-0094, ("training the model 136 to evaluate which trajectory of a set of trajectories is closest to a ground truth") 0100, 0110-0113, 0126, 0132-0139; EX1005, ¶¶0005-0008, 0034-0035, 0052-0058, 0065-0073; EX1003, ¶87.

c. 1[a-i]

Djuric and AA teach element 1[a-i]. EX1003, ¶88. Djuric discloses that autonomous vehicle 104 includes cameras that generate image data. EX1004, ¶¶0022, 0028-0035, 0046-0059, 0095-0103, 0104-0114, Figs. 1, 2A-2C, 6, 9. The image data includes "objects that are within the surrounding environment of the vehicle 104," where the objects in the images include "vehicles, pedestrians, [and] bicycles" that are part of traffic in a roadway and on a sidewalk; and the image data, as part of a "combined data set," is input into the model ([*the probabilistic neural network configured to perform steps comprising:*] *receiving as input, an image of traffic, the image displaying a traffic entity belonging to the traffic*). *Id.*; *see id.*, ¶¶0021-0023, 0028-0032 ("[t]he prediction system can input the combined data set (e.g., the image data) into the machine-learned model"), 0051, 0059-0063, 0079-0081; Section X.A.3.b (element 1[a]); EX1003, ¶88.



EX1004, Fig. 1 (annotated)



EX1004, Fig. 2A

Djuric describes that image data is input to (or received by) the model for training purposes as well as during operation. EX1004, ¶¶0021-0022, 0027-0037, 0046-0063, 0079-0094, 0095-0103, 0104-0114, 0134-0139, Figs. 1, 2A-2C, 6, 9; Section X.A.3.b (element 1[a]). For example, Djuric explains that "sensor data," which includes image data from cameras, is used during training of the model. EX1004, ¶¶0053, 0079, 0134-0135. Moreover, "supervised training techniques" using, for example, "a set of labeled training data," and the "backwards propagation of errors" are used to train the model. *Id.*, ¶¶0027, 0053, 0134. And further, data from previous events (e.g., "a previously **observed** trajectory of the object") are used as training data, which "allow[s] the training data to train a model based on

previously recorded real-world object motion and data associated therewith." *Id.*, ¶0053 (emphasis added); *see id.*, ¶¶0021, 0027, 0053, 0079, 0086-0103, 0134-0135; EX1003, ¶89.

d. 1[a-ii]

Djuric and AA teach element 1[a-ii]. EX1003, ¶90. AA discloses the model can generate a "first feature vector" and a "second feature vector," each of which "may include a number that corresponds to a possible feature of the image 226[.]" ([*the probabilistic neural network configured to perform steps comprising:*] *generating a feature vector for a plurality of features*). EX1005, ¶¶0034-0037; Section X.A.3.b (element 1[a]); EX1003, ¶90. AA further discloses that a softmax function may convert the numbers of a feature vector to a probability. EX1005, ¶0037. Because the numbers of the feature vector can be used to generate a "probability" for each feature, these values "*describ[e] statistical distribution for each feature*." EX1003, ¶90. Applying a softmax function to the generated feature vector aligns with the '579 Patent, which also describes processing a feature vector by applying a softmax function to produce probability values. EX1001, 18:23-33, 19:7-10, 19:24-33.

As described above in claim element 1[a-i], the image data in Djuric includes traffic entities "that are within the surrounding environment of the vehicle 104." Section X.A.3.c (element 1[a-i]); EX1004, ¶¶0022, 0028-0035, 0046-0059, 0095-0103, 0104-0114. Similar to AA, Djuric discloses that there are "features

associated" with each object (i.e., traffic entity). EX1004, ¶¶0022, 0028-0035, 0057; EX1005, ¶¶0034-0037; EX1003, ¶91. Accordingly, the Djuric–AA combination teaches that each object in the image data may have one or more feature vectors generated as taught by AA (([*the probabilistic neural network configured to perform steps comprising:*] *generating a feature vector for a plurality of features*) and a softmax function applied to those feature vectors will produce probabilities for each of the features (*the feature vector comprising values describing statistical distribution for each feature*). EX1003, ¶91.

e. 1[a-iii]

Djuric and AA teach element 1[a-iii]. EX1003, ¶92. Djuric teaches that the model generates an output including both predicted trajectories for an object ([*the probabilistic neural network configured to perform steps comprising:*] *generating output representing hidden context for the traffic entity*) and associated confidence levels for each trajectory (*the output comprising a plurality of values*). EX1004, ¶¶0005-0007, 0021, 0026-0027, 0032-0043, 0046-0056, 0071-0084, 0085-0095, 0096-0103, 0104-0114, 0126-0139, Claims 1, 2, Figs. 6, 9; Section X.A.3.b (element 1[a]); EX1003, ¶92.

Djuric teaches *generating output representing hidden context for the traffic entity* by teaching that the model generates predicted trajectories in consideration of "the nominal pathway(s) of an object." EX1004, ¶¶0081-0084. Using such nominal pathways "can allow the model 136 to consider a 'goal' that an object may aim to

accomplish." *Id*, ¶¶0032-0043, 0081-0085, 0107-0113; EX1003, ¶93. Djuric further describes "goal-based prediction" where the nominal pathways correspond to "common patterns of object travel." EX1004, ¶¶0032-0043, 0081-0085, 0107-0113. Djuric's disclosure of "goal-based prediction" aligns with the '579 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, 2:64-3:3.

The model output in Djuric also *compris[es] a plurality of values* because the output includes "a trajectory confidence level for **each** predicted trajectory." EX1004, ¶¶0082, 0089-94 (emphasis added); EX1003, ¶94. Because Djuric teaches generating a "plurality of predicted trajectories for an object" there would also be a plurality of confidence levels. EX1004, ¶0082; EX1003, ¶94.

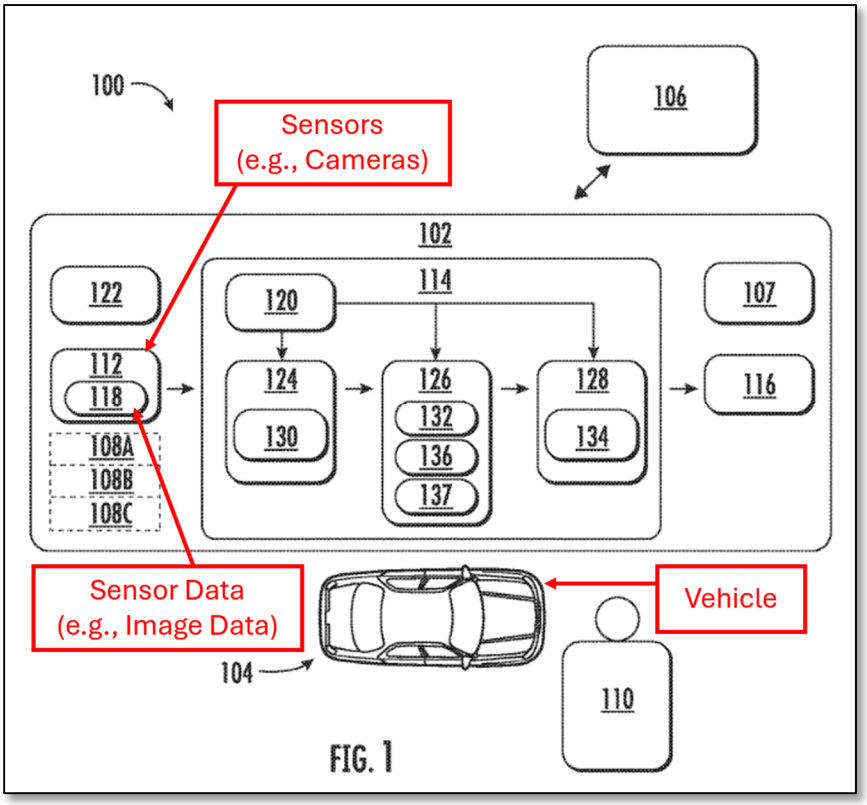
These trajectory confidence levels (*each value*) in Djuric also "*represent[] a likelihood of receiving a particular user response from a user presented with the image*." EX1003, ¶95. First, the trajectory confidence levels each "indicate a per-trajectory **probability** [(*a likelihood*)], quantifying how likely an object is to follow a certain trajectory." EX1004, ¶¶0088-0093 (emphasis added). Second, Djuric teaches using "supervised training techniques," including using "labeled" data. *Id*., ¶0053; *see also id*., ¶¶0027, 0053, 0090, 0134-0136. Such labeled data used in supervised training is data labeled by a human, e.g., a "*user response from a user*

presented with the image." EX1003, ¶95.

From the training techniques disclosed in Djuric, a POSITA would have further understood that the model in Djuric is trained such that the predicted probabilities for a given trajectory align with the probability that a human labeler would apply if given the same image. EX1004, ¶¶0021, 0027, 0053, 0079, 0085-0103, 0134-0136; EX1003, ¶96. In particular, Djuric teaches that during training, a loss function is "optimized" to "train the network to output 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 human-labeled training data, and are ultimately used to train the model to more accurately reflect human-labeled training data. EX1004, ¶¶0027, 0053, 0085-0094, 0134-0136; EX1003, ¶96. Djuric discloses that training the model "can include training the model 136 to evaluate which trajectory of a set of trajectories is closest to a ground truth." EX1004, ¶0094. A POSITA would have understood that "ground truth"—in the context of a supervised trained model, using labeled data—reflects how a human should respond when presented with an image. EX1003, ¶96. Thus, a POSITA would have understood that the purpose of this training in Djuric is to bring the model's predicted trajectories and confidence levels in line with how humans would likely respond to the same image (*each value representing a likelihood of receiving a particular user response from a user presented with the image*). EX1003, ¶96.

f. 1[b]

Djuric and AA teach element 1[b]. EX1003, ¶97. Djuric discloses that an autonomous vehicle 104 includes cameras that generate image data at various time points, including new image data, as the vehicle navigates through traffic (*receiving a new image captured by a camera mounted on an autonomous vehicle navigating through traffic*). EX1004, ¶¶0021-0023, 0028-0032 ("[t]he prediction system can input the combined data set (e.g., the image data) into the machine-learned model"), 0046-0059, 0096-0103, 0104-0114, Figs. 1, 2A-2C, 6, 9; EX1003, ¶97.



EX1004, Fig. 1 (annotated)

For example, Djuric's vehicle computing system 102 processes a plurality of images taken at times T_1, T_2, and T_3, for example. EX1004, ¶¶0056-0063. The

images may be generated with input "image data from a front-facing camera [or] other cameras," *Id*. The images can be of an object 202 "within a relevant portion of the surrounding environment of the vehicle 102," including a roadway and lane boundaries. *Id.*, ¶0061, Figs. 2A-2C, 3A-3C; EX1003, ¶98.

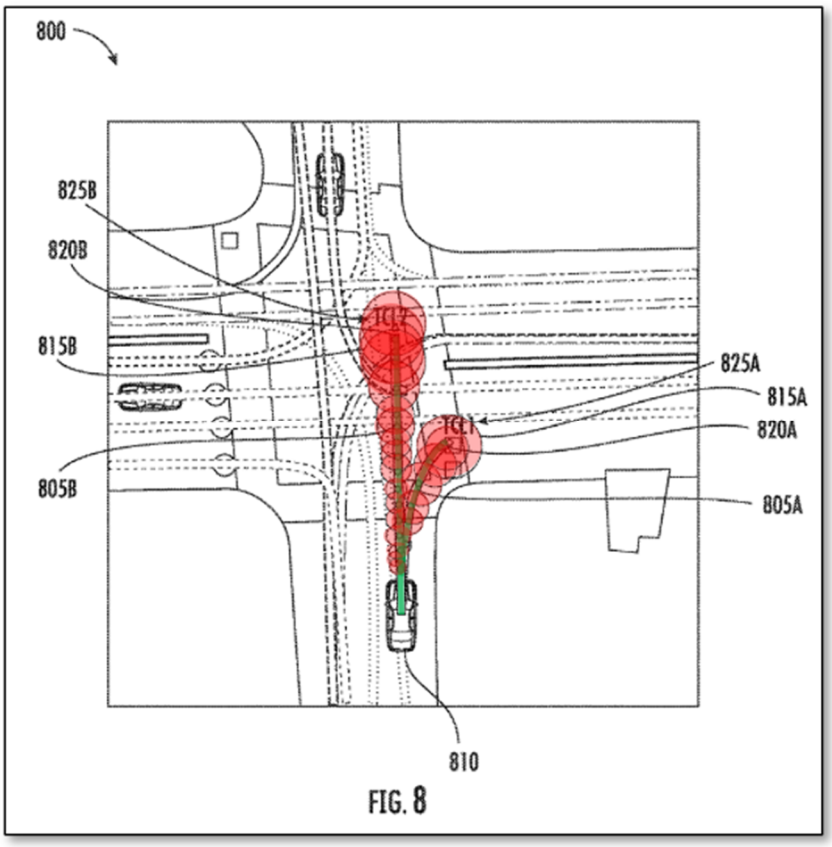
g. 1[c]

Djuric and AA teach element 1[c]. EX1003, ¶99. As explained above for element 1[a], Djuric and AA teach training a *probabilistic neural network*. Section X.A.3.b (element 1[a]); EX1003, ¶99. As explained above for element 1[a-iii], Djuric teaches that the model is configured to predict trajectories *representing hidden context* for objects in an image when executed. Section X.A.3.e (element 1[a-iii]); EX1004, ¶¶0005-0007, 0021, 0026-0027, 0032-0043, 0046-0056, 0071-0084, 0085-0095, 0096-0103, 0104-0114, 0126-0139, Claims 1, 2, Figs. 6, 9; EX1003, ¶99. Djuric teaches that the model may be executed over new images captured at times T_1, T_2, and T_3 "to determine a predicted trajectory of an object" in the images. EX1004, ¶¶0070-0071. Thus, for the reasons discussed in element 1[a-iii], when new images are provided to the model after the model is trained (as explained with respect to element 1[b]), the trained model will output predicted trajectories for the traffic entity detected in the new images (*executing the probabilistic neural network to generate output representing hidden context for at least a traffic entity displayed in the new image*). *Id*.; Sections X.A.3.e, X.A.3.f (elements 1[a-iii], 1[b]); EX1003, ¶99.

h. 1[d]

Djuric and AA teach element 1[d]. EX1003, ¶100. As discussed in Sections X.A.3.e (element 1[a-iii]) and X.A.3.g (element 1[c]), executing the model generates predicted trajectories and associated trajectory confidence levels for objects (e.g., vehicles, bicycles, or pedestrians) in a new image of the traffic surroundings of autonomous vehicle 104. *Id*.

Djuric describes determining a measure of uncertainty for each of the predicted trajectories and trajectory confidence levels by measuring a "way-point confidence level" each point along the trajectory (*determining a measure of uncertainty for each of the plurality of values*). EX1004, ¶¶0081-0088, 0112; EX1003, ¶101. Figure 8 (reproduced below) shows an example of each of the way-point confidence levels, where each way-point confidence level is reflected in red and the predicted trajectories for a vehicle are reflected in green. EX1004, ¶¶0081-0085, Fig. 8.



EX1004, Fig. 8 (annotated)

Djuric explains that "[e]ach of the predicted trajectories includes a plurality of predicted waypoints of the object." EX1004, ¶¶0007, 0083-0094, 0112, 0136, Claims 3, 4, 17, 19, and 20. According to Djuric, "each way-point can be associated with a respective way-point confidence level." *Id.*, ¶¶0085-0086. These way-point confidence levels "can be indicative of **the per-way-point uncertainty** for the way-point (e.g., predicted future location) within a predicted trajectory" and "can help quantify the confidence" in how exactly the object (the vehicle, pedestrian, bicycle) is going to follow a certain trajectory. *Id*. (parenthetical in original, emphasis added). Because each way-point confidence level quantifies the confidence in an object

following a predicted trajectory, the way-point confidence levels are thus a determined measure of uncertainty for the predicted trajectory and associated trajectory confidence level (*determining a measure of uncertainty for each of the plurality of values*). *Id*.; Section X.A.3.e (element 1[a-iii]); EX1003, ¶102.

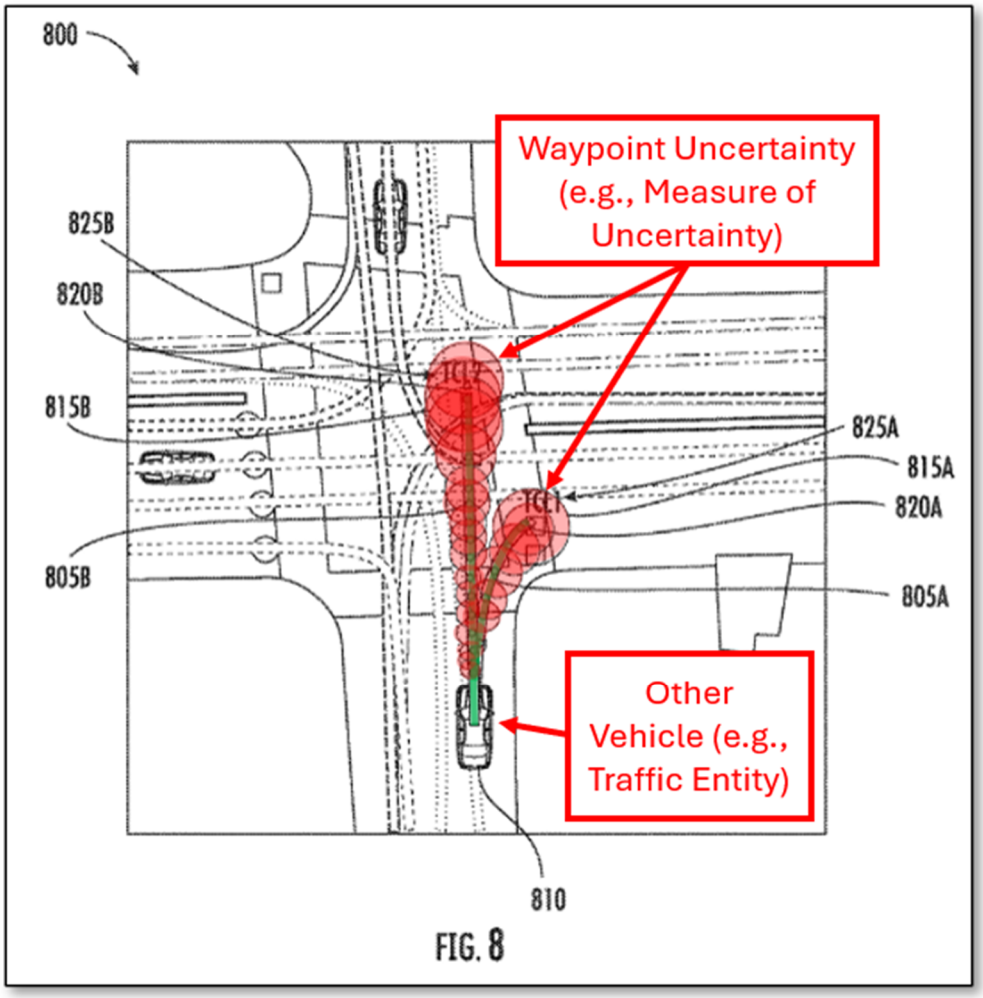
Djuric also teaches the *measure of uncertainty* for element 1[d] in an additional, independent way: Djuric calculates a "classification loss" for predicted trajectories, which "quantifies the accuracy of the per-trajectory probability" (i.e., the "trajectory confidence level"). EX1004, ¶¶0090-0091, 0136; EX1003 ¶103. Specifically, Djuric's model is "trained according to the loss function[]" that includes the classification loss of the trajectories, and the "loss function" is the difference between the training data and model outputs. EX1004, ¶¶0090-0091, 0136; EX1003, ¶103. Thus, when the model is executed, the classification loss in Djuric is a *measure of uncertainty* for each of the determined predicted trajectories and trajectory confidence levels. EX1004, ¶0090; EX1003, ¶103.

i. 1[e]

Djuric and AA teach element 1[e]. EX1003, ¶104. Djuric explains that "[g]iven knowledge of its surrounding environment, the autonomous vehicle can navigate through such surrounding environment." EX1004, ¶0003. Djuric's autonomous vehicle uses a "motion planning system, which can, for example, weigh each of the predicted paths and confidence levels when planning vehicle motion." EX1004, ¶¶0024, 0034-0035. After the combination's model produces a plurality of

predicted trajectories for an object (e.g., a vehicle, pedestrian, or bicyclist) in image data, including the new image data of Section X.A.3.f (element 1[b]), vehicle computing system 102 controls autonomous vehicle 104 "based at least in part" on the output predicted trajectories, where the vehicle uses the predicted future locations of the object within a predicted trajectory "to plan and control its motion (e.g., to avoid the object)" (*navigating the autonomous vehicle to avoid the traffic entity displayed in the new image*). EX1004, ¶¶0021, 0034, 0036, 0077, 0083-0094, 0102-0114, Figs. 6, 9 (parenthetical in original); Sections X.A.3.g (element 1[c]), X.A.3.h (element 1[d]); EX1003, ¶104.

As explained for element 1[d], each way-point confidence level forms *the measure of uncertainty*. Section X.A.3.h (element 1[d]); EX1003, ¶105. Djuric teaches that this way-point confidence level is used to predict object trajectories. EX1004, ¶0094 ("[A] predicted trajectory can be selected based at least in part on one or more of the trajectory confidence levels and/or **the way-point confidence levels**.") (emphasis added). These trajectories are in turn used in the navigation of the autonomous vehicle (e.g., avoiding crossing the traffic entity's trajectory at the same time, thereby causing a collision), which means "*the navigation [is] based on at least the measure of uncertainty generated by the probabilistic neural network*." *Id.*, ¶¶0083-0094, 0104-0114, Fig. 9; EX1003, ¶105.



EX1004, Fig. 8 (annotated)

Djuric also meets element 1[e] in an additional, independent way. As explained for element 1[d], Djuric's classification loss for trajectory confidence levels also provides a *measure of uncertainty*. *See* X.A.3.h (element 1[d]); EX1003, ¶106. Djuric's model is trained to minimize the determined classification loss. EX1004, ¶¶0090-0092. This means that the output predicted trajectories in Djuric are based on the determined classification loss (*measure of uncertainty*), and Djuric therefore controls its vehicle to avoid the object based on the determined

classification loss (*navigating the autonomous vehicle to avoid the traffic entity displayed in the new image, the navigation based on at least the measure of uncertainty generated by the probabilistic neural network*). EX1003, ¶106; EX1004, ¶¶0021, 0036, 0077, 0083-0094, 0104-0114, Fig. 9.

4. Claim 2

Djuric and AA render obvious claim 2. EX1003, ¶107. As discussed for element 1[a-ii], AA teaches a softmax function is applied to the numbers in the disclosed feature vector that "convert[s] the numbers" to a "probability;" the values of the feature vector therefore describe a probability distribution (i.e., probability for and conversely probability against) for each feature represented by each of the values (*the values describing statistical distribution for each feature*). EX1005, ¶0037; Section X.A.3.d (element 1[a-ii]); EX1003, ¶107. A POSITA would have found it obvious for such values to *comprise a mean value and a standard deviation for the feature*. EX1003, ¶107. This is at least because a mean and a standard deviation are well-known statistical measures that are routinely applied by POSITAs to data. *Id.* Moreover, Djuric teaches that its model can be trained using "an assumption […] that each way-point can be sampled from a 2D-Gaussian distribution, where the Gaussian distribution is parametrized by 2D **mean** (e.g., x-y positions) and 2x2 covariance matrix (e.g., modeling uncertainty)." EX1004, ¶0086 (emphasis added, parenthetical in original). In addition to directly disclosing the use of a *mean value*, a POSITA would have understood that the referenced "Gaussian distribution"

includes *a standard deviation*. EX1003, ¶107. In view of this teaching about the training of the model in Djuric, the Djuric-AA combination therefore teaches that the values from the feature vector taught by AA would also *comprise a mean value and a standard deviation for the feature*. EX1003, ¶107.

5. Claim 3

Djuric and AA render obvious claim 3. EX1003, ¶108. As discussed for element 1[a-iii], Djuric describes that the model training uses "supervised training techniques" using, for example, "a set of labeled training data," where the "backwards propagation of errors" is used to train the model. EX1004, ¶¶0027, 0053, 0134-0136; Section X.A.3.e (element 1[a-iii]). And furthermore, Djuric explains that model training is performed by using "a number of sets of data from previous events (e.g., previously obtained state data, previous obtained data associated with a geographic area)" and "a previously observed trajectory of the object," as well as "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, etc.)" are used as training data, which "allow[s] the training data to train a model based on previously recorded real-world object motion and data associated therewith." EX1004, ¶0053; *see id.*, ¶¶0021, 0027, 0053, 0079, 0086-0103, 0134-0136; EX1003, ¶108. A POSITA would thus have found it obvious that the training data in Djuric would include "*a plurality of values determined from user responses obtained from users presented with images from a training dataset*." EX1003, ¶108.

Djuric also discloses model training so that "the negative log-likelihood of the training data set can be minimized." EX1004, ¶0086; *see also id.*, ¶¶0085-0094, 0136. The use of an Evidence Lower Bound ("ELBO") was well-known for determining a lower bound of the log-likelihood of a data set to reduce the loss of a model. EX1003, ¶109. Thus, based on Djuric's stated goals, a POSITA would have understood to implement an ELBO to optimize Djuric's model (*training the probabilistic neural network comprises determining evidence lower bound between the plurality of values predicted and a plurality of values determined from user responses obtained from users presented with images from a training dataset*). EX1003, ¶109; *see also* EX1004, ¶¶0086, 0090-0092.

Alternatively, a POSITA would have found it obvious to implement an ELBO to optimize Djuric's model based on Djuric's stated goals (*Supra*) and because Djuric discloses that its model may be a Bayesian network, and the determination of an ELBO is an element of variational Bayesian methods. EX1004, ¶¶0126, 0132; EX1003, ¶110. Modifying Djuric's model to implement an ELBO would reduce the loss of the model when training, thereby improving the accuracy of the model. *Id.*

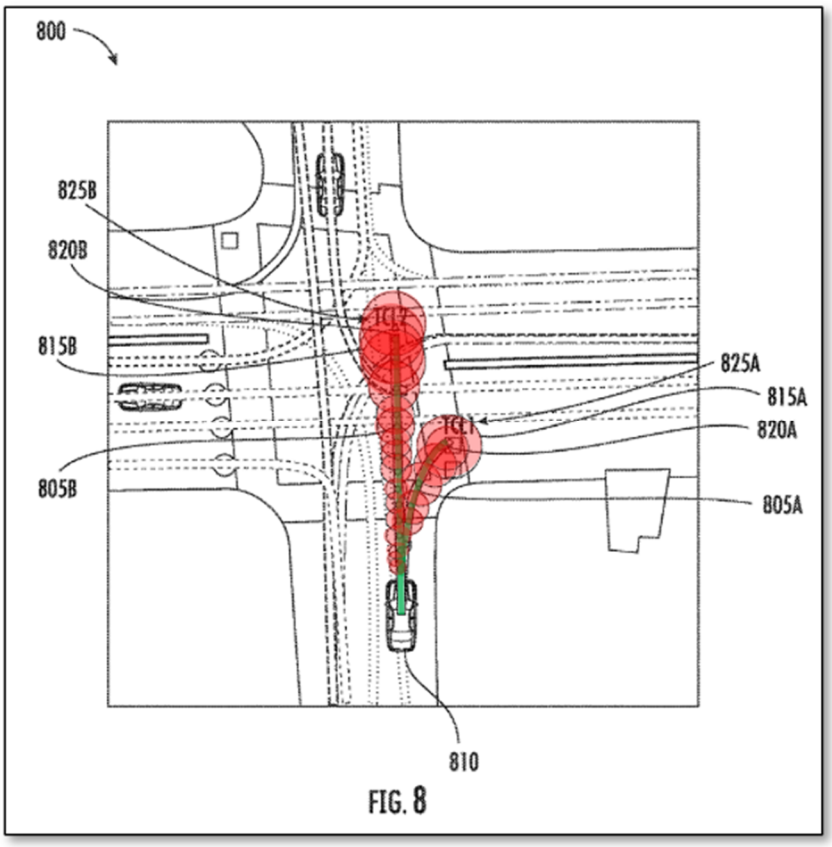
6. Claim 4

Djuric and AA render obvious claim 4. EX1003, ¶111. As explained above, Djuric describes determining way-point confidence levels as a measure of uncertainty for each predicted trajectory and the trajectory's associated trajectory confidence level (*determining the measure of uncertainty for each of the plurality of*

values). Section X.A.3.h (element 1[d]); EX1003, ¶111.

To generate the way-point confidence levels (*measure of uncertainty*), Djuric determines the waypoints of predicted trajectories as elements of predicted trajectories that are based on the input image (*generating a plurality of samples for the input image using the probabilistic neural network*). *See* Sections X.A.3.b-X.A.3.e (elements 1[a] through 1[a-iii]); EX1003, ¶112.

Djuric also teaches that generating the way-point confidence levels includes *determining a confidence interval for each of the plurality of values using the plurality of samples*. EX1003, ¶113. Specifically, Djuric teaches that way-point confidence levels can be determined by assuming that each waypoint can be sampled from a 2D-Gaussian distribution "parameterized by [a] 2D mean (e.g., x-y positions) and 2×2 covariance matrix (e.g., modeling uncertainty)." EX1004, ¶¶0085-0087 (parentheticals in original). The variance and covariance for each waypoint can also "be represented by a circle centered on [each] way-point 820A, where the radius of the circle is indicative of the confidence." EX1004, ¶0088, Fig. 8 (below). Because the variance and covariance for the way-point confidence level describe a range of confidence in the location of an object, the variance and covariance for each of the waypoints are "*a confidence interval for each of the plurality of values using the plurality of samples*." EX1003, ¶113.



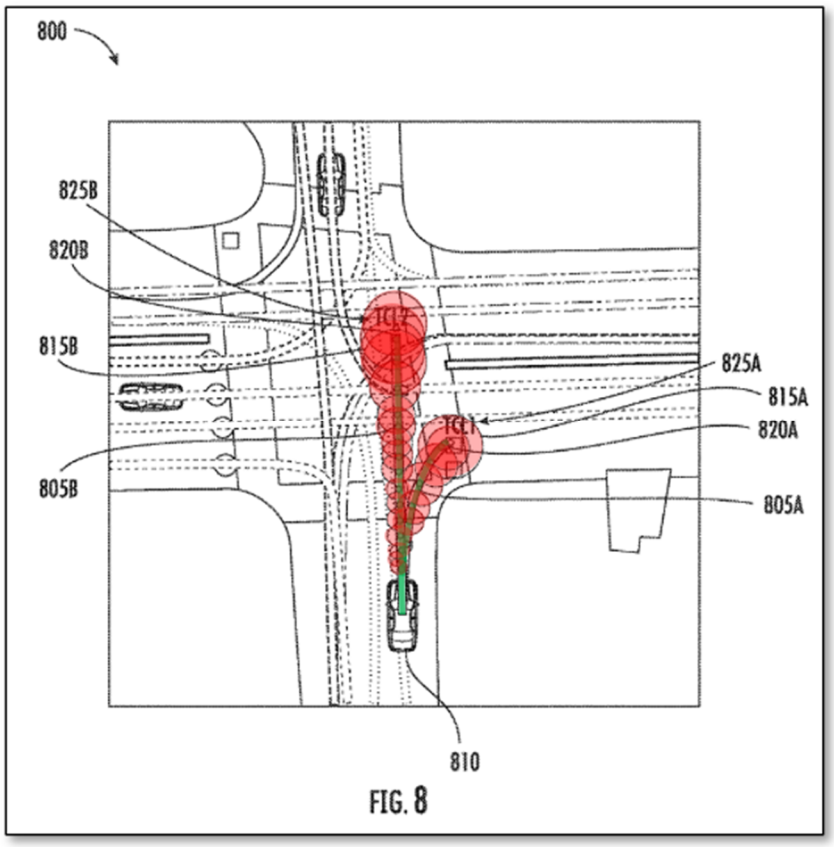
EX1004, Fig. 8 (annotated)

7. Claim 5

Djuric and AA render obvious claim 5. EX1003, ¶114. Djuric discloses that its autonomous vehicle uses an object's predicted object trajectories and "one or more predicted future locations of the object" to "plan and control its motion (e.g., to avoid the object)." EX1004, ¶0021 (parenthetical in original); *see also id*., ¶0036 (listing additional benefits "enhancing passenger/vehicle safety and improving vehicle efficiency by reducing collisions."). Djuric further teaches that the autonomous vehicle is navigated "relative to the objects 202, 302 at their **future** locations." *Id*., ¶0077 (emphasis added). From Djuric's stated goal of "reducing

collisions," a POSITA would have understood that the autonomous vehicle would navigate itself to stay at least a non-zero distance away from the predicted trajectories of objects in the new image data (Section X.A.3.f (element 1[b]), and their predicted trajectories (*wherein navigating the autonomous vehicle comprises ensuring that the autonomous vehicle stays at least a threshold distance away from the traffic entity displayed in the new image*). EX1004, ¶¶0021, 0034, 0036, 0077, 0083-0094, 0104-0114, Fig. 9 (parenthetical in original); Section X.A.3.i (element 1[e]); EX1003, ¶114. Failing to maintain a non-zero minimum distance from a predicted trajectory would result in a collision. *Id.*

Moreover, as discussed in Section X.A.3.h (element 1[d]), Djuric explains that "[e]ach of the predicted trajectories includes a plurality of predicted waypoints of the object." EX1004, ¶¶0007, 0083-0094, 0112, 0136, Claims 3, 4, 17, 19, and 20. A "way-point confidence level" (the *measure of uncertainty*) is determined "for each way-point within a predicted trajectory" such that "each way-point can be associated with a respective way-point confidence level." *Id.*, ¶¶0085-0086. These way-point confidence levels can be represented by a circle with a radius that is "indicative of the confidence/uncertainty associated with that particular predicted future location of the object." *Id.*, ¶0088, Fig. 8 (annotated below with way-point confidence in red).



EX1004, Fig. 8 (annotated)

Because the autonomous vehicle in Djuric is navigated based on the predicted trajectories along with the way-point confidence levels, (*see* Section X.A.3.i (element 1[e]); EX1004, ¶0094) it would have been obvious to a POSITA based on Djuric that the non-zero distance that Djuric's autonomous vehicle would have stayed away from objects in the new image data would have been based on the way-point confidence levels (*the threshold distance determined based on the measure of uncertainty generated by the probabilistic neural network*). EX1004, ¶¶0007, 0077, 0083-0094, 0112, 0136, Fig. 8, Claims 3, 4, 17, 19, and 20; EX1003, ¶116.

8. Claim 6

Djuric and AA render obvious claim 6. EX1003, ¶117. As discussed for claim 5, the non-zero distance Djuric's autonomous vehicle is kept away from the objects in image data, including the new image data (Section X.A.3.f (element 1[b])), and their predicted trajectories form *the threshold distance*, and the way-point confidence levels form *the measure of uncertainty generated by the* Djuric-AA *probabilistic neural network.* Section X.A.3.h (element 1[d]); Section X.A.7 (claim 5); EX1003, ¶117. It would further have been obvious to a POSITA that—consistent with basic driving techniques—the autonomous vehicle should give an object more space when there is more uncertainty about the object's predicted location. *Id*. In other words, the threshold distance maintained from an object should increase when the measure of uncertainty increases (*wherein the threshold distance is determined to be a value directly related to the measure of uncertainty*). Section X.A.3.h (element 1[d]); Section X.A.7 (claim 5); EX1003, ¶117.

9. Claim 7

Djuric and AA render obvious claim 7. EX1003, ¶118. As explained for claim 1, Djuric teaches that the predicted trajectories and associated trajectory confidence levels for an object—for example, for a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle "may aim to accomplish" that are associated with "nominal pathways," such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.3.g (element

1[c]); Section X.A.3.e (element 1[a-iii]); EX1003, ¶118. Because this *hidden context* represents the goals/intentions of a vehicle (and in turn, the vehicle's driver), the *hidden context* represents the state of mind as to where the vehicle's 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, ¶118. This is consistent with the '579 Patent's explanation that a state of mind may include "that a pedestrian is likely to cross the street." EX1001, 12:11-16; EX1003, ¶118.

10. Claim 8

Djuric and AA render obvious claim 8. EX1003, ¶119. As explained for claim 1, Djuric teaches that the predicted trajectories and associated trajectory confidence levels for an object—for example, for a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle "may aim to accomplish" that are associated with "nominal pathways," such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.3.g (element 1[c]); Section X.A.3.e (element 1[a-iii]); EX1003, ¶119. 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, ¶119. This is consistent with the '579 Patent's explanation that a task that a user is planning to accomplish may be "cross[ing] the street." EX1001, 3:3-15; 12:11-16; EX1003,

¶119.

11. Claim 9

Djuric and AA render obvious claim 9. EX1003, ¶120. As explained for claim 1, Djuric teaches that the predicted trajectories and associated trajectory confidence levels for an object—for example, for a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle "may aim to accomplish" that are associated with "nominal pathways," such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.3.g (element 1[c]); Section X.A.3.e (element 1[a-iii]); EX1003, ¶120. 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, ¶120. 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 the vehicle 104" that is relevant to vehicle 104 passing object 202. *See* EX1004, ¶0076; EX1003, ¶120. 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, ¶¶0033-0034, 0074-0077.

12. Claim 10

Djuric and AA render obvious claim 10. EX1003, ¶121. As explained for claim 1, Djuric teaches that the predicted trajectories for an object—for example, for a vehicle—represent a *hidden context* for the vehicle because they represent goals the vehicle "may aim to accomplish" that are associated with "nominal pathways," such as turning, merging, or continuing straight. EX1004, ¶¶0032-0043, 0081-0085, 0107-0113; Section X.A.3.g (element 1[c]); Section X.A.3.e (element 1[a-iii]); EX1003, ¶121. Because this *hidden context* represents the goals/intentions of a vehicle (and in turn, the vehicle's driver), 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*. EX1004, ¶¶0032-0043, 0071, 0081-0085, 0102, 0107-0113; EX1003, ¶121.

Moreover, Djuric discloses each trajectory "can include a plurality of predicted future locations of the object 302 over time (e.g., at future times)," which teaches the *goal* is expected to be achieved within a *threshold time interval* because the future times form an interval within which the object is predicted to reach the last waypoint of the predicted trajectory. EX1004, ¶¶0032-0043, 0071, 0081-0085, 0102, 0107-0113; EX1003 ¶122. 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 ¶122. Any prediction Djuric makes of a goal that a user will achieve occurs within the computational window or time horizon (*threshold time*) defined by Djuric's limited computational resources. *Id*. Thus, Djuric teaches *the user expects to achieve the goal within a threshold time interval*. *Id.*

13. Claim 11

Djuric and AA render obvious claim 11. EX1003, ¶123. As described in claim 1, Djuric describes that the autonomous vehicle is navigated (*navigating the autonomous vehicle*). Section X.A.3.i (element 1[e]); EX1003, ¶123. 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); EX1003, ¶123.

Djuric further teaches that the autonomous vehicle is controlled *based on motion parameters and the hidden context of the traffic entity*. EX1003, ¶124. First, Djuric teaches that "the vehicle computing system 102 can control the motion of the vehicle 104 based at least in part on the output 406 from the model 136." EX1004, ¶0078; *see also id*. ¶¶0024, 0034-0035. Second Djuric teaches that "[t]he output 406 can be indicative of a predicted trajectory of the object" (*hidden context*) and "the

output 406 can include a predicted velocity of the object" (*motion parameters*). EX1004, ¶¶0071, 0102; EX1003, ¶124.

Djuric discloses an additional *motion parameter* 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, ¶125. 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. Section X.A.3.g (element 1[c]); EX1004, ¶¶0104-0114. The navigation of the autonomous vehicle thus includes controlling the autonomous vehicle based on the predicted trajectories and also based on the state data describing object motion (*controlling the autonomous vehicle based on motion parameters and the hidden context of the traffic entity*). EX1004, ¶¶0021-0025, 0034-0036, 0051, 0071, 0077-0078, 0083-0094, 0102-0114, 0127, Fig. 9; X.A.3.i (element 1[e]); EX1003, ¶125.

14. Claim 12

Djuric and AA render obvious claim 12. EX1003, ¶126. As explained for element 1[a], AA teaches that the model may be a probabilistic neural network. Section X.A.3.b (element 1[a]); EX1005, ¶¶0005-0008, 0034-0035, 0052-0058,

0065-0073. Djuric further teaches that "[n]eural networks can include convolutional neural networks." EX1004, ¶0027; *see also id*., ¶¶0033, 0053, 0069, 0072, 0079-0094, 0100, 0110-0113, 0126, 0132-0139. As such, the Djuric-AA combination teaches that *the probabilistic neural network is a probabilistic convolutional neural network*. EX1003, ¶126.

15. Claim 13

Djuric and AA render obvious claim 13 for the same reasons discussed with respect to claim 1. *See* Section X.A.3 (claim 1); EX1003, ¶127.

a. 13[Pre]

See Section X.A.3.a (element 1[Pre]). 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. 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 steps comprising*). *Id*., ¶¶0117-0118; EX1003, ¶128.

b. 13[a]

See Section X.A.3.b (element 1[a]); EX1003, ¶129.

c. 13[a-i]

See Section X.A.3.c (element 1[a-i]); EX1003, ¶130.

d. 13[a-ii]

See Section X.A.3.d (element 1[a-ii]); EX1003, ¶131.

e. 13[a-iii]

See Section X.A.3.e (element 1[a-iii]); EX1003, ¶132.

f. 13[b]

See Section X.A.3.f (element 1[b]); EX1003, ¶133.

g. 13[c]

See Section X.A.3.g (element 1[c]); EX1003, ¶134.

h. 13[d]

See Section X.A.3.h (element 1[d]); EX1003, ¶135.

i. 13[e]

See Section X.A.3.i (element 1[e]); EX1003, ¶136.

16. Claim 14

Djuric and AA render obvious claim 14 for the same reasons discussed above in Section X.A.4 (claim 2). EX1003, ¶137.

17. Claim 15

Djuric and AA render obvious claim 15 for the same reasons discussed above in Section X.A.5 (claim 3). EX1003, ¶138.

18. Claim 16

Djuric and AA render obvious claim 16 for the same reasons discussed above

in Section X.A.6 (claim 4). EX1003, ¶139.

19. Claim 17

Djuric and AA render obvious claim 17 for the same reasons discussed above in Section X.A.7 (claim 5). EX1003, ¶140.

20. Claim 18

Djuric and AA render obvious claim 18 for the same reasons discussed above in Section X.A.8 (claim 6). EX1003, ¶141.

21. Claim 19

Djuric and AA render obvious claim 19 for the same reasons discussed above in Section X.A.13 (claim 11). EX1003, ¶142.

22. Claim 20

Djuric and AA render obvious claim 20 for the same reasons discussed with respect to claims 1 and 13. *See* Sections X.A.3, X.A.15 (claims 1, 13). EX1003, ¶143.

a. 20[Pre]

See Section X.A.15.a (element 13[Pre]); Djuric discloses the "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)" (*A computer system*). EX1004, ¶0116; *see also id.*, ¶¶0040, 0045, 0140; EX1003, ¶144.

b. 20[a]

See Section X.A.3.b (element 1[a]); EX1003, ¶145.

c. 20[a-i]

See Section X.A.3.c (element 1[a-i]); EX1003, ¶146.

d. 20[a-ii]

See Section X.A.3.d (element 1[a-ii]); EX1003, ¶147.

e. 20[a-iii]

See Section X.A.3.e (element 1[a-iii]); EX1003, ¶148.

f. 20[b]

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

g. 20[c]

See Section X.A.3.g (element 1[c]); EX1003, ¶150.

h. 20[d]

See Section X.A.3.h (element 1[d]); EX1003, ¶151.

i. 20[e]

See Section X.A.3.i (element 1[e]); EX1003, ¶152.

B. Ground 2: Claims 1-20 are unpatentable under 35 U.S.C. § 103 over Djuric, AA, and Cox.

To the extent Djuric alone does not render obvious the use of human labeled data to generate values *representing a likelihood of receiving a particular user response from a user presented with the image* as recited in element 1[a-iii], 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 into Djuric, along with the above described teachings from AA. EX1003, ¶153.

1. Summary of Prior Art: Cox-EX1008

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 '579 Patent. EX1001, Cover; EX1008, Cover. Cox is a material reference to the '579 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." EX1008, Abstract, ¶¶0001-0004. Cox teaches crowd-sourcing training data used to train machine learning algorithms. EX1008, Abstract, ¶¶0001-0004, 0015-0016, 0025-0034, 0044, 0060. Just like the '579 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* EX1008, ¶¶0029-0032, Figs. 1A-B, 3 *with* EX1001, 4:23-35, 8:39-44; Fig. 4. The goal of Cox's machine learning techniques is to "better mimic human performance." EX1008, ¶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. EX1008, ¶¶0056, 0060.

Cox is analogous art to the '579 Patent, from the same field of endeavor as the '579 Patent (e.g., predictive neural network model techniques), and is reasonably pertinent to the particular problem that the '579 Patent was trying to solve (e.g., how to "accurately predict motion of non-stationary traffic objects"), describing the gathering training data from a population of number of human labelers, and training a model using the training data, and executing it to output predictions. EX1008, ¶¶0004, 0029-0032, Figs. 1A-B, 3; EX1001, 1:38-62, 4:8-22; EX1003, ¶156.

2. Motivation to Combine Djuric, AA, and Cox

A POSITA would have been motivated to further combine Djuric and AA with Cox such that confidence levels (*each value*) from Djuric, quantifying how likely an object is to follow a certain trajectory (EX1004, ¶¶0088-0093), are generated in accordance with human-annotated/labeled training data. EX1003, ¶157. As taught by Cox, using human-annotated image data to train the Djuric-AA 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." EX1008, ¶¶0004, 0027-0032, 0051; EX1003, ¶157. A POSITA would have been motivated to combine Djuric and AA with Cox in this way for several reasons.

EX1003, ¶157.

First, 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.a; EX1003, ¶158. 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. EX1008, ¶¶0007-0010, 0015-0019, 0025-0038, 0051, 0060, Fig. 4B; EX1003, ¶158. 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." EX1008, ¶¶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, ¶158.

A POSITA seeking to implement the Djuric-AA system would also have looked to Cox to provide more instructions for training a machine learning model to predict human behavior. EX1003, ¶159. 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. EX1008, ¶¶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, ¶159.

A POSITA would have also been motivated to combine the teachings of the Djuric-AA system and Cox because the combination would have benefitted the Djuric-AA system. EX1003, ¶160. Enhancing the Djuric-AA system with Cox's human annotated training data that considers the state of mind of training subjects, as well as Cox's associated model training and optimization processes, would have further refined the Djuric-AA model so it better considers the state of mind of other road users and what they might be thinking when present in autonomous vehicle's environment. *Id.*; EX1008, ¶¶0007-0010, 0015-0019, 0025-0035, 0051, 0060. This would have improved the Djuric-AA system's goal-based predictions, and thus helped reduce collisions between the autonomous vehicle and human drivers or pedestrians. EX1003, ¶160.

All the reasons explained above establish that a POSITA would have been motivated to combine Djuric, AA, and Cox as described. EX1003, ¶161. This combination would have been the predictable combination of well-known prior art elements (e.g., the Djuric-AA model; Cox's model training and optimization

techniques) according to known methods to yield predictable results (the Djuric-AA model being a supervised learning model; the Djuric-AA-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). *KSR Int'l Co. v. Teleflex Inc.,* 550 U.S. 398, 415-421 (2007); EX1008, ¶0032; EX1003, ¶161. 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-AA model for goal-based predictions) in the same way (using Cox's aforementioned techniques). *Id*. Lastly, due to the overlap in the disclosures of Djuric, AA, and Cox and the fact that Djuric already generates and uses training data, using human-annotated training data from Cox with the teachings of Djuric and AA was well within the skill of a POSITA, and a POSITA would have had an expectation of success in this combination. *Id*.

3. Claim 1

Cox's teachings only impact how the model generates values *representing a likelihood of receiving a particular user response from a user presented with the image* in element 1[a-iii]. The combination of Djuric, AA, and Cox teaches element 1[a-iii] for the reasons explained below, as well as the reasons addressed in Ground 1. Sections X.A.3, X.A.3.e (element 1[a-iii]); EX1003, ¶162. Moreover, Djuric, AA, and Cox teach the remaining elements of claim 1 for the same reasons addressed in Ground 1. *Id.*

a. 1[a-iii]

Djuric, AA, and Cox teach 1[a-iii] for the same reasons addressed in Ground 1 (Section X.A.3.e), and in an additional way. EX1003, ¶163. Specifically, and as described further below, Cox teaches that human-labeled image data is used to train a model, so that the model's output *represent[][s] a likelihood of receiving a particular user response from a user presented with the image*. EX1008, ¶¶0019, 0029-0038, 0060; Fig. 4B; Section VI.A.3.e (Ground 1, element 1[a-iii]); EX1003, ¶163. So, implementing these teachings of Cox with Djuric, the trajectory confidence levels (*each value*) in Djuric would further "*represent[] a likelihood of receiving a particular user response from a user presented with the image*." EX1003, ¶163.

Cox uses "[a] crowd-sourc[ed][]" and "large population" of human annotators to annotate a "training object" such as an image, generate "statistical information" based on all of the responses received from the multiple different human annotators, and train a supervised machine learning model using the statistical information. EX1008, ¶¶0007-0010, 0015-0019, 0025-0035, 0051; EX1003, ¶164. "[T]he subject matter of each training object has certain features that may be classified by at least some of the human annotators 160." EX1008, ¶0017, Figs. 1A, 1B. For example, the human annotator responses may describe a state of mind of the training object, e.g., "an emotion that corresponds to the [sample training object] image, e.g., jealous, panicked, arrogant, or hateful." *Id*., ¶0016 (bracketing added); *id.*, ¶¶0007-0010,

0015-0019, 0025-0035, 0051; EX1003, ¶164. Cox discloses that the statistical information is "collected and applied to the training of computer-vision systems." EX1008, ¶0027. Cox's model can be utilized after training "to make predictions (based on various query objects) **that are more consistent with the decisions of the human annotators 160**." EX1008, ¶¶0027-0032 (parenthetical in original, emphasis added), 0051. The goal of Cox's machine learning techniques is to "better mimic human performance." *Id.*, ¶0004.

Cox additionally teaches that the model's output represents *a likelihood of receiving a particular user response from a user* because Cox's training data in the form of annotation responses and psychometric data from the human annotators 160 is utilized for training the model, "resulting in a human-weighted loss function 425 computationally derived by the penalization module 180." EX1008, ¶¶0031-0032, Fig. 4B; EX1003, ¶165. The loss function "includes penalties for misclassification of later presented query objects, as graphically illustrated in graph 430, and the magnitude of the penalties increases with increasing deviation from the classification data received from the human annotators 160." EX1008, ¶¶0032, 0034. That is, "[o]nce the human responses have been analyzed to quantify dominant patterns of error, the models may be translated into human-weighted loss functions for maximum-margin kernel machines and penalties for margins not consistent with human data." EX1008, ¶¶0032, 0034.

Similarly, Cox also describes that its model is trained to minimize the

difference between the model's predicted output and the human-labeled training data. EX1008, ¶¶0019, 0029-0038, 0060; Fig. 4B; EX1003, ¶166. Specifically, Cox discloses that its model training includes "classifying, by the computer processor, at least some of the training objects based at least in part on an initial loss function," which refers to executing the model on training data images. EX1003, ¶166; EX1008, ¶¶0008, 0019, 0029-0037; Fig. 4B (430). Cox, then teaches "comparing" the training data to the classification data output by the model to "identify, within the training data, features misclassified in comparison to the classification data," and "incorporating [] penalties for misclassification within the initial loss function to generate the human-weighted loss function." *Id*.; EX1003, ¶166. In particular, Cox discloses that the "penalties for misclassification may be assigned based at least in part on the psychometric data" and "the magnitude of the penalties increases with increasing deviation from the classification data received from the human annotators 160," which ensures that the trained model is "more consistent with the decisions of the human annotators." EX1008, ¶0008, 0032. Accordingly, Cox describes that its model is trained by predicting classifications of training data and optimizing the model such that the difference between the predicted classifications output by the model and the training data is minimized. EX1008, ¶¶0019, 0029-0038, 0060; Fig. 4B; EX1003, ¶166.

Thus, as applied to Djuric and AA, Cox teaches that the trajectory confidence levels quantifying how likely an object is to follow a certain trajectory (EX1004,

¶¶0088-0093) are generated in accordance with the human-annotated/labeled training data. EX1008, ¶¶0004, 0027-0032, 0051; EX1003, ¶167. The human-annotated/labeled training data includes annotated/labeled image data and is used to train the Djuric-AA model to provide predictions that are "more consistent with the decisions of the human annotators," and that "better mimic human performance." *Id*.; EX1003, ¶167. Thus, the model's outputs are representative of how humans would have annotated or labeled images input into the model if the images were part of the training data. EX1008, ¶¶0004, 0019, 0029-0038, 0051, 0060; Fig. 4B; EX1003, ¶167. As such, the Djuric-AA-Cox combination teaches that the model generates predicted trajectories for an object and associated trajectory confidence levels for the predicted trajectories ([*the probabilistic neural network configured to perform steps comprising:*] *generating output representing hidden context for the traffic entity, the output comprising a plurality of values*) (*see* Section X.A.3.e (element 1[a-iii]), and Cox, applied to Djuric and AA, further teaches the predicted confidence levels (*each value*) represent likelihoods of how humans would have annotated/labeled the input image data, thus teaching *each value representing a likelihood of receiving a particular user response from a user presented with the image*. EX1004, ¶¶0021, 0027, 0053, 0079, 0085-0103, 0134-0136; EX1008, ¶¶0004, 0019, 0029-0038, 0051, 0060; Fig. 4B; EX1003, ¶167; Section X.A.3.e (element 1[a-iii]).

4. Claims 13 and 20

Djuric, AA, and Cox render obvious claims 13 and 20 for the same reasons discussed above with respect to claim 1. The combination with Cox only impacts the analysis for elements 13[a-iii] and 20[a-iii]. Elements 13[a-iii] and 20[a-iii] are taught by Djuric, AA, and Cox for the reasons stated above for 1[a-iii]. *Supra*, Section X.B.3.a (element 1[a-iii]). Djuric, AA, and Cox further teach elements 13[a-iii] and 20[a-iii], as well as the remaining elements of claims 13 and 20, for the same reasons addressed for Djuric and AA in Ground 1. Sections X.A.15 (claim 13), X.A.22 (claim 20).

5. Claims 2-12 and 14-19

Djuric, AA, and Cox render obvious claims 2-12 and 14-19 for the same reasons discussed with respect to Djuric and AA in Section X.A above (Ground 1).

C. Grounds 3 and 4: Claims 1-20 are unpatentable under 35 U.S.C. § 103 over Djuric, AA, and Rolfe and/or Djuric, AA, Cox and Rolfe.

To the extent Djuric alone does not render obvious dependent claims 3 and 15, Rolfe confirms that the recited use of an evidence lower bound (ELBO) is a well-known statistical technique. The analysis for independent claims 1, 13, and 20 and dependent claims 2, 4-12, 14, and 16-19 is unchanged from Grounds 1 and 2. Claims 3 and 15 are thus rendered obvious over the further combination of Rolfe, with either Djuric and AA (Ground 3), or Djuric, AA, and Cox (Ground 4).

1. Summary of Prior Art: Rolfe-EX1009

Rolfe discloses quantum processors that may be used in machine learning. EX1009, Abstract. In this context, Rolfe discusses use of "the evidence lower bound (ELBO)" as an example approach "to optimize a lower bound on the log-likelihood function." *Id*., ¶0086.

Rolfe is analogous art to the '579 Patent, from the same field of endeavor as the Patent (e.g., predictive neural network model techniques), 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" using machine learning). EX1001, 1:38-62, 4:8-22; EX1003, ¶172.

2. Motivation to Combine Djuric, AA, Cox, and Rolfe

A POSITA would have been motivated to combine Rolfe with Djuric (along with AA and optionally Cox) to use the disclosed ELBO techniques to minimize the negative log-likelihood of the training data set in Djuric. EX1003, ¶173. Djuric teaches that the model is trained so that "the negative log-likelihood of the training data set can be minimized." EX1004, ¶0086; *see also id.*, ¶¶0085-0094, 0136. The disclosure of ELBO in Rolfe simply provides implementation details (already in the context of machine learning) that a POSITA could utilize to achieve the stated goal in Djuric. EX1009, ¶0086; EX1003, ¶173. As such, it would be well within the skill of a POSITA to use the implementation details from Rolfe for ELBO in the context of Djuric, and a POSITA would have an expectation of success in making this

combination. *Id*.

3. Claims 1-2, 4-14, and 16-20

Except for dependent claims 3 and 15 addressed below, Djuric, AA, and Rolfe (Ground 3) and Djuric, AA, Cox, and Rolfe (Ground 4) render obvious the challenged claims for the same reasons discussed above with respect to Djuric and AA (*Supra* Section X.A) and Djuric, AA, and Cox (*Supra* Section X.B). EX1003, ¶174.

4. Claim 3

Combining Rolfe with Djuric, AA (and potentially Cox) makes express what a POSITA would already have understood from Djuric alone: the use of an Evidence Lower Bound was well-known for determining a lower bound of the log-likelihood of a data set to reduce the loss of a model during training (*wherein training the probabilistic neural network comprises determining evidence lower bound between the plurality of values predicted and a plurality of values determined from user responses obtained from users presented with images from a training dataset*). EX1003, ¶175; EX1009, ¶0086. The further combination of Rolfe thus renders obvious Claim 3 for the same reasons discussed in Ground 1, with the added confirmation from Rolfe that a POSITA would already be familiar with the applications of an ELBO to the values recited in Claim 3. *Supra* Section X.A.5 (claim 3); EX1003, ¶175.

5. Claim 15

Djuric, AA, and Rolfe (Ground 3) and Djuric, AA, Cox and Rolfe (Ground 4) render obvious claim 15 for the same reasons discussed above with respect to claim 3. *Supra* Section X.C.4. EX1003, ¶176.

D. Grounds 5 and 6: Claims 1-20 are unpatentable under 35 U.S.C. § 103 over Djuric, AA, and Ross and/or Djuric, AA, Cox, and Ross

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. *Id*., 4:40-44.

Ross is analogous art to the '579 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"); Ross describes an model that provides "[r]easoning about the intents of objects, such as other road users." *Id*., 5:6-11, 15:1-16:52; EX1003, ¶178.

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

A POSITA would have been motivated to combine Djuric, AA, and/or Cox with Ross as discussed in Sections X.A.2 and X.B.2. EX1003, ¶179. A POSITA would have been further motivated to combine Ross with Djuric and AA (and optionally Cox) such that the Djuric-AA-Ross probabilistic neural network generates "goal-based prediction[s]" using Djuric's goals and Ross's intents or actions. EX1003, ¶179.

As discussed regarding elements 1[a-iii] and 1[c], Djuric describes generating "goal-based prediction[s]" of object trajectories in a vehicle's environment. EX1004, ¶¶0081-0085; Sections X.A.3.e, X.A.3.g (elements 1[a-iii], 1[c]). 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, ¶180. 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 scenarios Djuric's model is capable of predicting. EX1003, ¶180. 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. "Each possible intent of the set includes a corresponding next action to be initiated by the road user." *Id*. 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. EX1003, ¶180.

Thus, a POSITA would have been to combine Ross with Djuric and AA (and optionally Cox) as described. EX1003, ¶181. This combination would have been the predictable combination of well-known prior art elements (e.g., the Djuric-AA 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 the Djuric-AA 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, ¶181. The combination would have further used known techniques (Ross's intent determination for predictions) to improve similar devices (the Djuric-AA 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 an expectation of success in the modification. *Id*.

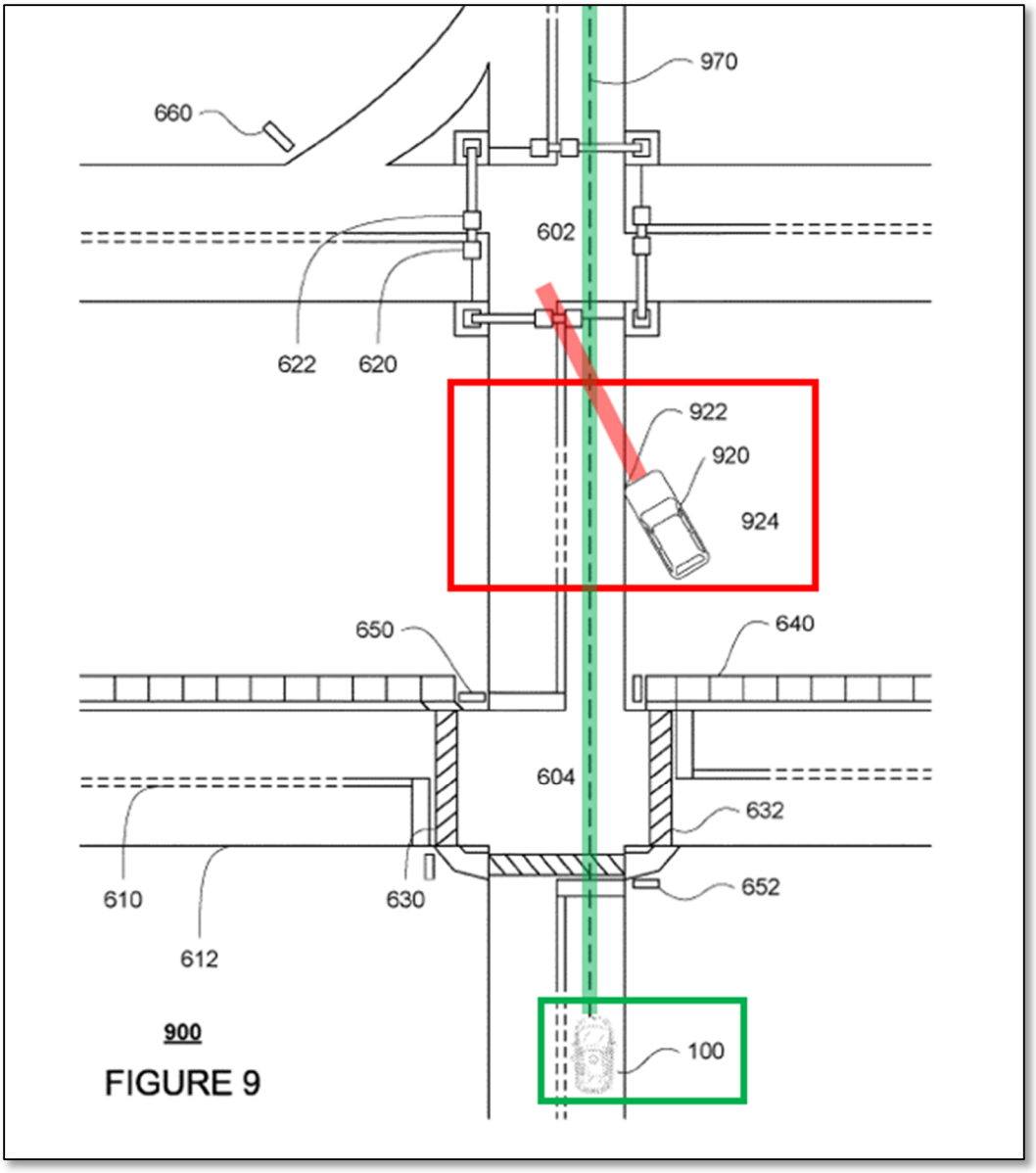
3. Claims 1-6, 8, and 10-20

Djuric, AA, and Ross (Ground 5) and Djuric, AA, Cox, and Ross (Ground 6) render obvious claims 1-6, 8, and 10-20 for the same reasons discussed above with respect to Djuric and AA (Ground 1, Section X.A) and Djuric, AA, and Cox (Ground 2, Section X.B).

4. Claim 7

Djuric, AA, and Ross render obvious claim 7 for the same reasons discussed in Section X.A.9 (Ground 1, claim 7). To the extent Djuric-AA does not explicitly disclose the hidden context represents a state of mind, the combination of Djuric-AA and Ross renders this obvious. EX1003, ¶183.

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 associated "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; EX1003, ¶184.



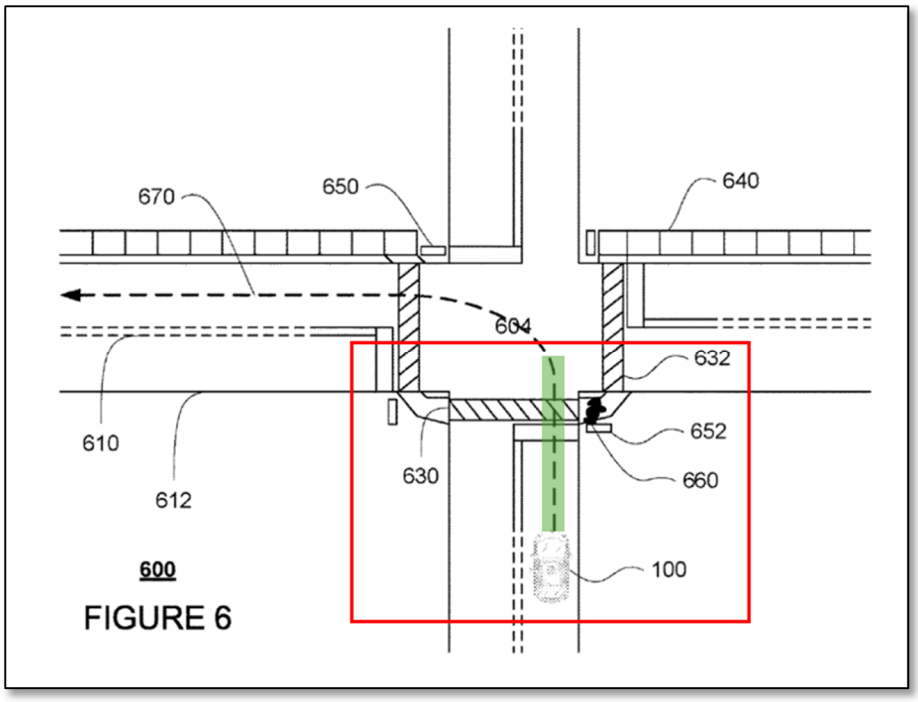
EX1015, Fig. 9 (in part, annotated)

Ross's teachings align with the '579 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, 12:11-16; see also EX1015, 16:5-41, Fig. 9; EX1003, ¶185.

5. Claim 9

Djuric and Ross render obvious claim 9 for the same reasons discussed in SectionX.A.11 (claim 9). To the extent Djuric-AA does not explicitly disclose the hidden context represents a degree of awareness of the autonomous vehicle by a user, the combination of Djuric-AA and Ross renders this obvious. EX1003, ¶186.

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, ¶188.

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-i], 13[a-i], 20[a-i] – "an image of traffic, the image displaying a traffic entity belonging to the traffic"
- 1[a-ii], 13[a-ii], 20[a-ii] – "the feature vector comprising values describing statistical distribution for each feature"
- 1[a-iii], 13[a-iii], 20[a-iii] – "output representing hidden context for the

traffic entity, the output comprising a plurality of values, each value representing a likelihood of receiving a particular user response from a user presented with the image"

- 1[c], 13[c], 20[c] – "output representing hidden context for at least a traffic entity displayed in the new image"
- 2, 14 – "values describing statistical distribution for each feature comprise a mean value and a standard deviation for the feature"
- 3, 15 – "user responses obtained from users presented with images from a training dataset"
- 7 – "the hidden context represents a state of mind of a user represented by the traffic entity"
- 8 – "the hidden context represents a task that a user represented by the traffic entity is planning on accomplishing"
- 9 – "the hidden context represents a degree of awareness of the autonomous vehicle by a user represented by the traffic entity"
- 10 – "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"

Each of the above limitations recites a data element (an image, a value, output, a response, etc.) and then defines that data element based on its content. EX1003,

¶191. But a patent cannot cover what an "image display[s]," or what an "output represent[s]" — 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 recitation in element 1[a-iii] about what each "value" represents should be ignored because the content of the claimed "values" is not patentable. Similarly, the recited limitations about what the "hidden context represents" in claims 7-10 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 in various data elements, 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, ¶192; *see also* 13[pre] and 20[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 9, 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 9, 2025, a complete and entire copy of the **PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-20 OF U.S. PATENT NO. 11,467,579** 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 '579 Patent and by email to counsel for Patent Owner in the EDTX Litigation, as included below:

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

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

Date: October 9, 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 13,662 words, within the 14,000 word limit allowed under §42.24(a)(1)(i).

Date: October 9, 2025

/*Roger Fulghum*/

Roger Fulghum
Reg. No. 39,678

Attorney for Petitioner, Tesla, Inc.