

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

MICROSOFT CORPORATION,
Petitioner,

v.

REALD SPARK, LLC,
Patent Owner.

IPR2023-01046
Patent 10,740,985 B2

Before JEAN R. HOMERE, CHARLES J. BOUDREAU, and
CHRISTOPHER L. OGDEN, *Administrative Patent Judges*.

HOMERE, *Administrative Patent Judge*.

JUDGMENT
Final Written Decision
Determining All Challenged Claims Unpatentable
35 U.S.C. § 318(a)

I. INTRODUCTION

A. Background

Microsoft Corporation (“Petitioner”) filed a Petition for *inter partes* review of claims 1–14 and 16–19 (“the challenged claims”) of U.S. Patent No. 10,740,985 B2 (Ex. 1001, “the ’985 patent”). Paper 1 (“Pet.”), 1. RealD Spark, LLC (“Patent Owner”) filed a Preliminary Response. Paper 6. Upon consideration of the parties’ contentions and supporting evidence, we instituted an *inter partes* review pursuant to 35 U.S.C. § 314, as to the challenged claims of the ’985 patent with respect to all grounds set forth in the Petition. Paper 7 (“Inst. Dec.”).

After institution of trial, Patent Owner filed a Response (Paper 12, “PO Resp.”); Petitioner filed a Reply to Patent Owner’s Response (Paper 18, “Pet. Reply”); and Patent Owner filed a Sur-Reply to Petitioner’s Reply (Paper 21). In addition, Petitioner filed a first set of objections to Patent Owner’s evidence (Paper 9) and a second set of objections to Patent Owner’s evidence (Paper 13). At the request of both parties, we held an oral hearing on September 20, 2024. A transcript of the hearing has been entered into the record. Paper 30 (“Tr.”).

We have jurisdiction under 35 U.S.C. § 6. This Final Written Decision is entered pursuant to 35 U.S.C. § 318(a). For the reasons that follow, we determine that Petitioner has demonstrated by a preponderance of the evidence that challenged claims 1–14 and 16–19 of the ’985 patent are unpatentable.

B. Real Parties in Interest

Petitioner identifies itself as the sole real party in interest. Pet. 1.

Patent Owner identifies itself as the real party in interest and also identifies “the following entities that are members of RealD Spark, LLC and therefore may also be a real party of interest: NEW RD INTERMEDIATE, LLC, HPS Investment Partners, LLC, and Falcon Strategic Partners V, LP.” Paper 3, 2.

C. Related Matters

The parties indicate that the ’985 patent is the subject of the following district court proceeding:

RealD Spark, LLC v. Microsoft Corporation, Case No. 2:22-cv-00942 (W.D. Wash., filed July 7, 2022).

Pet. 1; Paper 3, 2.

D. The Challenged ’985 Patent

The ’985 patent relates to “[m]ethods and devices for generating reference data for adjusting a digital representation of a head region, and methods and devices for adjusting the digital representation of a head region.” Ex. 1001, code (57). The ’985 patent purports “to correct the perceived gaze and/or to improve other aspects of a computer-generated display of the head region . . . with minimal demands on computer hardware and/or power such that they provide results at or near input data frame rate or user feedback requirements.” *Id.* at 1:46–55.

“FIG. 1 and FIG. 2 illustrate how incorrect gaze is perceived when a stereoscopic pair of images of a head is captured by” a source device and is displayed on a destination device. Ex. 1001, 4:19–24. Figure 1 is reproduced below.

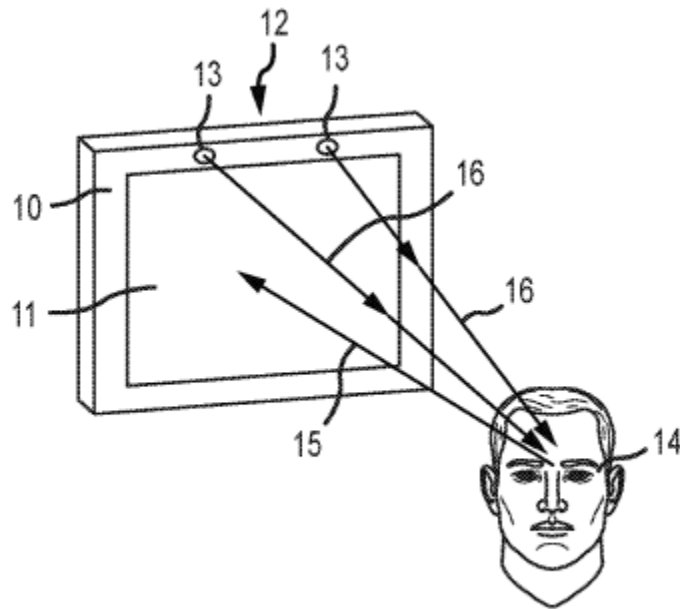


FIG.1

Figure 1 is a perspective view of capture device 10 using cameras 13 in camera system 12 to capture a stereoscopic pair of images of the head of source observer 14. *Id.* at 3:51–52, 4:33–37. Figure 2 is reproduced below.

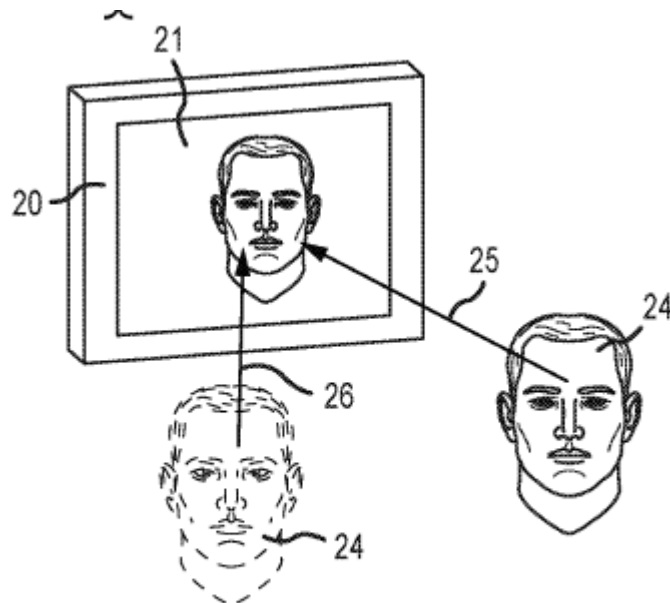


FIG.2

Figure 2 depicts destination device 20 displaying in display 21 the stereoscopic pair of images captured by device 10. *Id.* at 4:46–58. In this example, because cameras 13 look down on source observer 14 along line 16, destination observer 24 perceives the gaze of source observer 14 as being downwards, rather than looking at destination observer 24. *Id.* Additional perceived gaze errors can occur “[i]f the destination observer 24 is located in an offset viewing position, as shown by the dotted outline of the destination observer 24 so that the destination observer 24 views the display 21 along line 26,” or “if the destination observer 24 is located in the normal viewing position along line 25, but the displayed image (or stereoscopic pair of images in this example) is displayed on the display 25 in a position offset from the center of the display 25.” *Id.* at 4:64–5:8.

The ’985 patent describes correcting perceived gaze errors by adjusting the digital representations of the source observer using an image processor executing a computer program or some combination of software and hardware. Ex. 1001, 5:35–39. Figure 4 reproduced below depicts the processing of the stereoscopic pair of captured images:

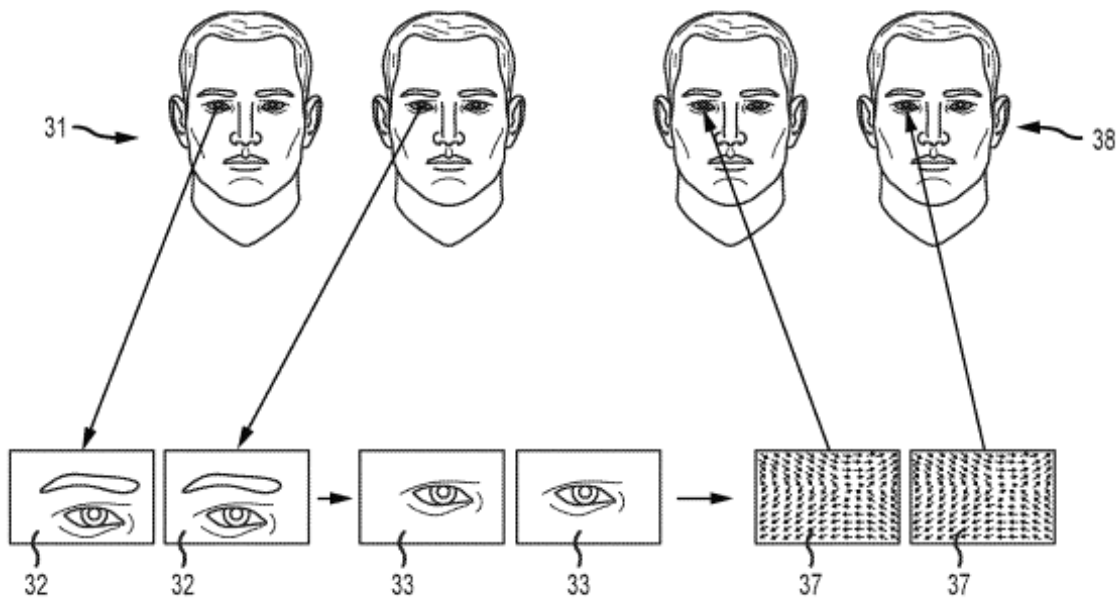


FIG.4

Figure 4 shows an example of stereoscopic pair of images 31 corresponding to image patches 32 located around the eyes of source observer 14.

Ex. 1001, 6:38–39, 7:1–5. Image patches 32 are transformed to align the points of the eye to produce image patches 33 in a normalized coordinate system in accordance with displacement vector fields 37 from which stereoscopic pair of adjusted images 38 with corrected gaze is produced. *Id.* at 7:14–25, 11:38–42, 11:65–12:3.

First, the image processor analyzes stereoscopic pair of images 31 of source observer 14 to detect the location of the source observer's head and eyes. Ex. 1001, 6:40–46. Second, the processor uses landmarks in the images to identify representative portions of a digital representation of the head region as “image patches” (e.g., the right eye, the left eye, etc.). *Id.* at 6:66–7:13. Third, the processor transforms the image patches into a

normalized coordinate system by aligning landmarks in the image patches with predetermined locations in the coordinate system. *Id.* at 7:14–25.

Next, the processor derives feature vectors representing parts of the image patches from local image descriptors of the image patches in the coordinate system. Ex. 1001, 7:36–48. The processor then uses the feature vectors to look up reference data for adjusting the image patches, which may be of any type that is expected to allow for discrimination between different individuals. *Id.*; *see also id.* at 9:45–56. Subsequently, the processor analyzes the reference data using any suitable machine learning technique as described in Figure 8. *Id.* at 7:55–56, 10:49–52.

Figure 8 is reproduced below.

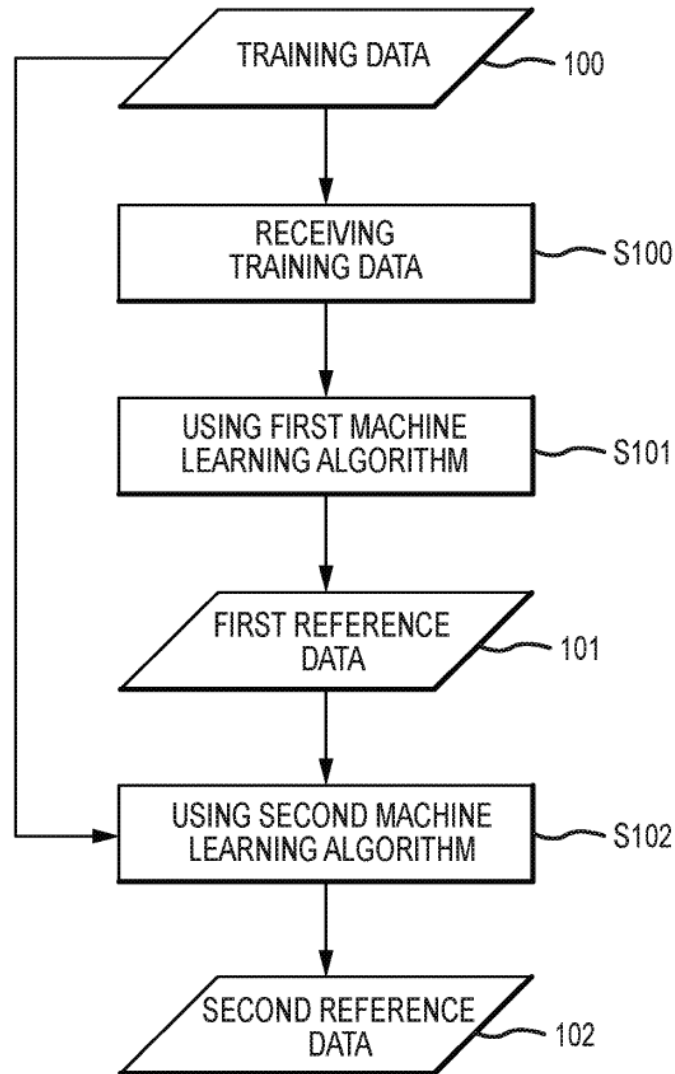


FIG.8

Figure 8 is a “flow chart of a method of generating reference data” for adjusting a digital representation of a head region. Ex. 1001, 3:63–64; *see also id.* at 12:62–65.

According to the method in Figure 8, first, a communications interface provides training data 100 including a set of input image patches or output

image patches to a data processing unit that performs the method (step S100). Ex. 1001, 13:32–41. The input image patches represent target features (e.g., an eye region, a nose region, etc.) of the digital representation prior to adjustment of the head region (*id.* at 13:41–57), and the output image patches represent adjusted or corrected target features of the input image patches (*id.* at 14:29–44). Second, a first machine learning algorithm generates first reference data 101 comprising editing instructions for adjusting the digital representation for a range of possible digital representations of the head region (e.g., different positions and/or orientations) using the training data 100 (step S101). *Id.* at 14:45–54. Third, a second machine learning algorithm generates second reference data 102 comprising editing instructions for adjusting the digital representation for a range of possible digital representations of the head region using training data 100 and first reference data 101 (step S102). *Id.* at 14:55–67. The second reference data is applied to the image patches of the target features to adjust the digital representation of the head region and the corrected digital representation is displayed. *Id.* at 2:54–67, 3:26–44, 11:38–45.

The '985 patent reports that “[t]he described use of two machine learning algorithms allows an improved balance to be achieved between accuracy of the adjustment process and speed of execution.” Ex. 1001, 2:17–19. “The first machine learning algorithm can be configured to provide highly detailed first reference data, which provides high accuracy,” but use of the first reference data directly in a processing technique “would be relatively expensive in terms of computational resources because of the high level of detail.” *Id.* at 2:19–27. According to the '985 patent, by providing a second machine learning algorithm with the original training

data and first reference data, “it is possible to benefit to an extent from the high accuracy of the first machine learning algorithm whilst also providing reference data that is less detailed and thus easier to process efficiently when performing the adjustment of the digital representation of the head region,” and “[t]he quality of the reference data provided by the second machine learning algorithm is found to be significantly improved.” *Id.* at 2:27–43. Efficiency can be further improved by providing the editing instructions in the second reference data in a compressed representation because “[t]he use of a compressed representation reduces data storage and bandwidth requirements.” *Id.* at 2:44–50.

E. The Challenged Claims

Petitioner challenges claims 1–14 and 16–19 of the ’985 patent. Pet. 1. Of the challenged claims, 1 and 16–19 are independent. Claim 1 is illustrative of the challenged claims and is reproduced below with key disputed limitations emphasized.

1. [1P] A method of generating reference data for adjusting a digital representation of a head region, the method comprising:
 - [1A] receiving training data comprising:
 - [1B] a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch; and
 - [1C] a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region;

[1D] using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region; and

[1E] *using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and the first reference data generated by the first machine learning algorithm*, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region,

[1F] *wherein the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch; and*

[1G] *wherein the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch.*

Ex. 1001, 26:31–27:3 (bracketed designations added by Petitioner (see Pet. 86)).

F. Prior Art and Asserted Grounds of Unpatentability

The Petition relies on the following prior art references:

Name	Reference	Exhibit
Chalom	US 10,664,949 B2, issued May 26, 2020	1005
Ganin	Ganin et al., <i>DeepWarp: photorealistic image resynthesis for gaze manipulation</i> , ECCV 2016, Part II, LNCS 9906, pp. 311–326 (2016)	1006
Aslan	US 2017/0132528 A1, published May 11, 2017	1007

Name	Reference	Exhibit
Grau	US 2018/0158246 A1, published June 7, 2018	1009
Pace	US 8,942,283 B2, issued January 27, 2015	1010
Liu	Liu et al., <i>Optical flow and principal component analysis-based motion detection in outdoor videos</i> , EURASIP Journal on Advances in Signal Processing (2010)	1016

Petitioner asserts claims 1–14 and 16–19 are unpatentable based on the following grounds (Pet. 3):

Claim(s) Challenged	35 U.S.C. ¹ §	Reference(s)
1–3, 5, 10, 12–14, 18	102(a)	Ganin
16, 19	103	Ganin, Liu
1–3, 5, 6, 8–14, 16–19	103	Chalom, Aslan
4	103	Chalom, Aslan, Grau
7	103	Chalom, Aslan, Pace

Pet. 3. Petitioner submits a declaration of Lina Karam, Ph.D. (Ex. 1003, “Karam Declaration”) in support of its contentions.

¹ Because the claims at issue have an effective filing date (August 7, 2017) after March 16, 2013, the effective date of the applicable provisions of the Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284 (2011) (“AIA”), we apply the post-AIA version of 35 U.S.C. § 103 in this Decision.

II. PATENTABILITY ANALYSIS

A. Principles of Law

1. Inter Partes Review

“In an [*inter partes* review], the petitioner has the burden from the onset to show with particularity why the patent it challenges is unpatentable.” *Harmonic Inc. v. Avid Tech., Inc.*, 815 F.3d 1356, 1363 (Fed. Cir. 2016) (citing 35 U.S.C. § 312(a)(3) (requiring *inter partes* review petitions to identify “with particularity . . . the evidence that supports the grounds for the challenge to each claim”)). Petitioner bears the burden of persuasion to prove unpatentability of each challenged claim by a preponderance of the evidence. 35 U.S.C. § 316(e) (2018). This burden never shifts to Patent Owner. *Dynamic Drinkware, LLC v. Nat’l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015).

2. Anticipation

“Under 35 U.S.C. § 102 a claim is anticipated ‘if each and every limitation is found either expressly or inherently in a single prior art reference.’” *King Pharms., Inc. v. Eon Laby’s, Inc.*, 616 F.3d 1267, 1274 (Fed. Cir. 2010) (quoting *Celeritas Techs. Ltd. v. Rockwell Int’l Corp.*, 150 F.3d 1354, 1360 (Fed. Cir. 1998)). “Anticipation requires the presence in a single prior art disclosure of all elements of a claimed invention arranged as in the claim.” *Therasense, Inc. v. Becton, Dickinson & Co.*, 593 F.3d 1325, 1332 (Fed. Cir. 2010) (quoting *Connell v. Sears, Roebuck & Co.*, 722 F.2d 1542, 1548 (Fed. Cir. 1983)).

A reference inherently discloses an element of a claim “if that missing characteristic is necessarily present, or inherent, in the single anticipating

reference.” *Schering Corp. v. Geneva Pharms.*, 339 F.3d 1373, 1377 (Fed. Cir. 2003) (citing *Cont’l Can Co. USA, Inc. v. Monsanto Co.*, 948 F.2d 1264, 1268 (Fed. Cir. 1991)). “Inherency, however, may not be established by probabilities or possibilities. The mere fact that a certain thing *may* result from a given set of circumstances is not sufficient.” *Therasense*, 593 F.3d at 1332 (quoting *Cont’l Can*, 948 F.2d at 1269).

3. Obviousness

A patent claim is unpatentable under 35 U.S.C. § 103 “if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains.” 35 U.S.C. § 103. The question of obviousness is resolved on the basis of underlying factual determinations including (1) the scope and content of the prior art, (2) any differences between the claimed subject matter and the prior art, (3) the level of skill in the art, and (4) when in evidence, any objective evidence of obviousness or nonobviousness.² *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

B. Level of Ordinary Skill in the Art

The level of ordinary skill in the art is “a prism or lens” through which we view the prior art and the claimed invention. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001). The person of ordinary skill in the art is a hypothetical person presumed to have known the relevant art at the time of the invention. *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995). In

² The parties have not directed us to any such objective evidence.

determining the level of ordinary skill in the art, we may consider certain factors, including: “(1) the educational level of the inventor; (2) type of problems encountered in the art; (3) prior art solutions to those problems; (4) rapidity with which innovations are made; (5) sophistication of the technology; and (6) educational level of active workers in the field.” *Best Med. Int’l, Inc. v. Elekta Inc.*, 46 F.4th 1346, 1353 (Fed. Cir. 2022) (citations omitted). “The patent’s purpose can also be informative.” *Id.*

Petitioner contends that a person having ordinary skill in the art at the time of the invention “would have had at least an undergraduate degree in computer science, electrical engineering, computer engineering, or equivalent education, and at least two years of work experience using machine learning for image processing.” Pet. 4. Petitioner asserts that “[a]dditional education might substitute for less work experience, and *vice versa*.” *Id.* (citing Ex. 1003 ¶ 22).

Patent Owner does not contest Petitioner’s proffered definition of the level of ordinary skill in the art, but offers a definition of its own. PO Resp. 8. Patent Owner contends that a person having ordinary skill in the art at the time of the invention “would have had at least a bachelor’s or master’s degree in electrical engineering, computer engineering, electrical and computer engineering, computer science, or similar educational background; or equivalent on-the-job training, including approximately two to three years of experience in machine learning and/or pattern recognition.” *Id.* (citing Ex. 2008 ¶ 22). Patent Owner asserts that the “level of skill is approximate and more experience would compensate for less formal education and *vice versa*.” *Id.*

We note that the differences between the parties' descriptions are immaterial and find Petitioner's description to be consistent with the problems and solutions disclosed in the '985 patent and prior art of record, and adopt it as our own for purposes of this Decision, albeit without the two instances of "at least," which introduce ambiguity. *See, e.g., In re GPAC Inc.*, 57 F.3d at 1579 (approving the determination of the level of ordinary skill in the art by appeal to the references of record).

C. Claim Construction

In an *inter partes* review, claims are construed using the same claim construction standard that would be used to construe the claims in a civil action under 35 U.S.C. § 282(b), including construing the claims in accordance with the ordinary and customary meaning of such claims as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent. 37 C.F.R. § 42.100(b) (2023). "[T]he ordinary and customary meaning of a claim term is the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention" and "after reading the entire patent." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1313, 1321 (Fed. Cir. 2005) (en banc). In addition to the specification and prosecution history, we also consider use of the terms in other claims and extrinsic evidence including expert and inventor testimony, dictionaries, and learned treatises, although extrinsic evidence is less significant than the intrinsic record. *Id.* at 1312–17. Usually, the specification is dispositive, and it is the single best guide to the meaning of a disputed term. *Id.* at 1315.

“The Board is required to construe ‘only those terms . . . that are in controversy, and only to the extent necessary to resolve the controversy.’” *Realtime Data, LLC v. Iancu*, 912 F.3d 1368, 1375 (Fed. Cir. 2019) (alteration in original) (quoting *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)).

Petitioner contends that it “applied plain and ordinary meanings of all terms and no express constructions are required here.” Pet. 5.

Patent Owner agrees with Petitioner that there are no claim terms that require construction for purposes of the Board’s final decision. PO. Resp. 8–9. We therefore determine that we do not need to expressly construe any term for purposes of this Decision. *See Realtime Data*, 912 F.3d at 1375.

D. Overview of the Asserted Prior Art

1. Ganin (Ex. 1006)

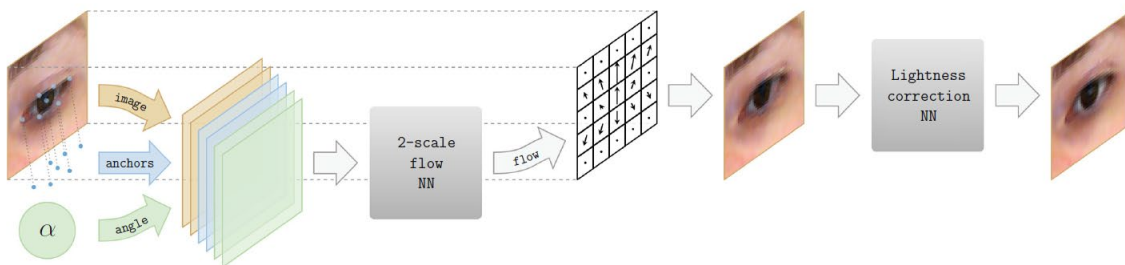
Ganin relates to “the task of generating highly-realistic images of a given face with a redirected gaze” using a neural network. Ex. 1006, 1.³ Ganin explains that “there is a need to digitally alter the appearance of eyes in a way that changes the apparent direction of the gaze,” such as during video conferencing. *Id.* at 2. Ganin “consider[s] the task of learning deep architectures that can transform input images into new images in a certain way (deep image resynthesis)” and recognizes that “using deep architectures for image generation has become a very active topic of research.” *Id.* at 1. Ganin purports to improve upon the prior art by providing “a new deep feed-

³ Page number citations for Ex. 1006 correspond to the page numbers added by Petitioner in the bottom right corner of each page of “Exhibit 1006 Corrected.”

forward architecture that combines several principles of operation (coarse-to-fine processing, image warping, intensity correction)” and “is trained end-to-end in a supervised way using a specially collected dataset that depicts the change of the appearance under gaze redirection in real life.” *Id.* at 2.

Ganin purports that its “deep architecture can synthesize very high-quality eye images . . . at several frames per second” and “[c]ompared to several recent methods for deep image synthesis, the output of [its] method contains [a] larger amount of fine details (comparable to the amount in the input image).” *Id.* at 3.

Ganin’s method uses a deep model for resynthesis in which “[t]he model is trained on pairs of images corresponding to eye appearance before and after the redirection. The redirection angle serves as an additional input parameter that is provided both during training and at test time.” Ex. 1006, 4. Ganin’s method is illustrated in Figure 2, reproduced below.



Ganin Figure 2

Figure 2 is a flow chart depicting Ganin’s proposed system, which takes an input eye region, feature points (anchors) as well as a correction angle α and sends them to the multi-scale neural network . . . predicting a flow field. The flow field is then applied to the input image to produce an image of a redirected eye. Finally, the output is enhanced by processing with the lightness correction neural network

Id. at 4 (emphasis and citations omitted). Ganin states that its model uses predictions based on a multi-stage method organized in a coarse-to-fine manner in which “the decisions at the fine scale are being informed by the result of the coarse stage,” and “the photorealism of the result is improved by performing pixel-wise correction of the brightness where the amount of correction is again predicted by the network.” *Id.* The model is trained using “pairs of images containing eyes of the same person looking in two different directions separated by a known angle α ” in which “[t]he head pose, the lighting, and all other nuisance parameters are (approximately) the same between the two images in the pair.” *Id.* at 5. In particular, Ganin discloses two warping modules, wherein “[e]ach of the two warping modules takes as an input the image, the position of the feature points, and the redirection angle.” *Id.* (emphasis added). More particularly, Ganin discloses:

Coarse Warping. The last convolutional layer of the first (half-scale) warping module produces a pixel-flow field (a two-channel map), which is then upsampled $\mathbf{D}_{\text{coarse}}(I, \alpha)$ and applied to warp the input image by means of a bilinear sampler $\mathbf{S}$ that finds the *coarse estimate*:

$$O_{\text{coarse}} = \mathbf{S}(I, \mathbf{D}_{\text{coarse}}(I, \alpha)) . \quad (1)$$

Here, the sampling procedure $\mathbf{S}$ samples the pixels of O_{coarse} at pixels determined by the flow field:

$$\begin{aligned} O_{\text{coarse}}(x, y, c) = I\{x + \mathbf{D}_{\text{coarse}}(I, \alpha)(x, y, 1), y + \\ \mathbf{D}_{\text{coarse}}(I, \alpha)(x, y, 2), c\}, \end{aligned} \quad (2)$$

Fine Warping. In the fine warping module, the rough image estimate O_{coarse} and the upsampled low-resolution flow $\mathbf{D}_{\text{coarse}}(I, \alpha)$ are concatenated with the input data (the image, the angle encoding, and the feature point encoding) at the original scale and sent to the $1\times$ -scale network which predicts another

two-channel flow $\mathbf{D}_{\text{res}}$ that amends the half-scale pixel-flow (additively):

$$\mathbf{D}(I, \alpha) = \mathbf{D}_{\text{coarse}}(I, \alpha) + \mathbf{D}_{\text{res}}(I, \alpha, O_{\text{coarse}}, \mathbf{D}_{\text{coarse}}(I, \alpha)) , \quad (3)$$

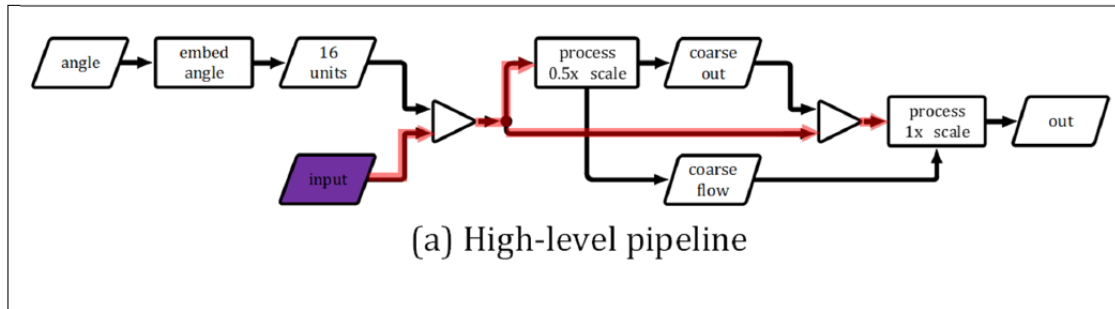
The amended flow is used to obtain the final output (again, via bilinear sampler):

$$O = \mathbf{S}(I, \mathbf{D}(I, \alpha)) . \quad (4)$$

The purpose of coarse-to-fine processing is two-fold. The half-scale (coarse) module effectively increases the receptive field of the model resulting in a flow that moves larger structures in a more coherent way. Secondly, the coarse module gives a rough estimate of how a redirected eye would look like. This is useful for locating problematic regions which can only be fixed by a neural network operating at a finer scale.

Id. at 5–6 (endnotes omitted).

Ganin's method is further illustrated in Figure 10(a), reproduced below.



Ganin Fig. 10 (a)

Figure 10(a) above depicts the basic warping architecture including a high-level pipeline that takes an input eye region augmented with eye feature points information (input) as well as a correction angle to produce an image of the redirected eye. Ex. 1006, 18. The model contains three main blocks: angle embedding module (embed angle) calculating a vector representation of the correction angle and two warping modules (process 0.5×-scale 10(b)

and process $1\times$ -scale 10(c)) predicting and applying pixel-flow to the input image. *Id.* at 18. In particular, the embedding vector is fed as an additional input to the warping network by replicating the 16-dimensional $F_{\text{angle}}(\alpha)$ for every spatial location creating a tensor of size $16 \times H \times W$ which is then concatenated to the rest of the input. *Id.* at 17. Additionally, Ganin discloses using a lightness correction module including a convolutional neural network to increase the lightness of selected regions. *Id.* at 17–18.

2. *Chalom (Ex. 1005)*

Chalom relates to eye contact correction in real time using machine learning and “[t]echniques related to eye contact correction to provide a virtual user gaze aligned with a camera while the user views a display.” Ex. 1005, codes (54), (57). Chalom purports to improve upon the prior art by “provid[ing] real time eye contact correction (or gaze correction) using machine learning in a cameral image signal processing unit for applications such as videotelephony, video conferencing, or the like.” *Id.* at 3:16–20.

Chalom explains that its methods “may include a training component (e.g., performed offline) and a real time component (e.g., performed during implementation or run time) that provides eye contact correction.” Ex. 1005, 3:21–24. In Chalom’s offline training component, “a database of faces with a known gaze angle may be used to train a machine learning system that learns a model.” *Id.* at 3:24–32. Chalom states that these techniques “may include determining features (e.g., a feature set) corresponding to uncorrected eye regions and determining motion vectors corresponding to the features using a pretrained classifier such as an optimal random forest tree search.” *Id.* at 3:32–36. Chalom also explains that

training may be performed based on a training set of images having known gaze angle difference therebetween. . . . During training, the training set images may be cropped to eye regions and the eye regions may be aligned to provide pairs of training eye region images. Based on the pairs of training eye region images, a likelihood map may be generated for each pixel of each of the first images of the pairs of training eye region images such that the likelihood map includes a sum of absolute differences (SAD) for each of multiple candidate motion vectors corresponding to the pixel.

Id. at 4:21–33. Motion vectors may be used to adjust the input image to a desired output image while running in real time. *Id.* at 3:24–32. For example, “[t]he motion vector field may be used to warp the eye region to generate the desired output image having corrected eye contact.” *Id.* at 3:36–38.

Chalom’s Figure 3 is reproduced below.

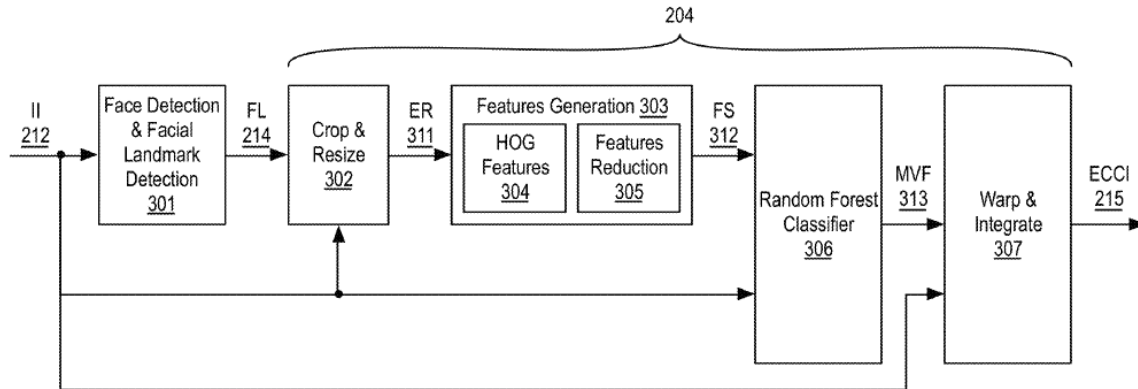


FIG. 3

Figure 3 depicts a flow chart of “an example eye contact correction module for providing eye contact correction.” Ex. 1005, 1:42–43; *see also id.* at 7:12–15. According to Figure 3, first, face detection and facial landmark detection module 301 receives input image (II) 212 from an image source and generates facial landmarks (FL) 214. *Id.* at 7:17–25; *see also id.* at 6:41–47. Second, eye contact correction module 204 receives input image

212 and facial landmarks 214, starting with crop and resize module 302. *Id.* at 7:21–30, 8:1–4. Third, crop and resize module 302 generates eye regions (ER) 311 using any suitable technique or techniques. *Id.* at 8:1–4.

Fourth, features generation module 303 comprising histogram of oriented gradient (HOG) features module 304 and features reduction module 305 receives the eye regions 311. Ex. 1005, 8:26–34, 8:49–51. Features generation module 303 generates feature sets (FS) 312 such that each corresponds to an eye region 311 and “include[s] any suitable data or data structure representative of compressed or condensed features based on the [HOG features].” *Id.* at 8:26–30, 9:1–4. Then, HOG features module 304 generates or determines HOG features for each eye region 311 using any suitable technique, such as counts of occurrences of gradient orientations in localized portions of each eye region 311 that are binned into histograms which may include a feature size of the width of the eye region by the height of the eye region. *Id.* at 8:30–45. Then, features reduction module 305 reduces the histograms, such as the dimensionality or number of HOG features, using any suitable technique including principal component analysis (PCA). *Id.* at 8:49–59. For example, PCA techniques may be applied such that the PCA “includes pretrained scaling and weighting factors determined during a training phase, as is discussed with respect to FIG. 8,” which is described further below. *Id.* at 8:63–9:1.

Fifth, random forest classifier 306 receives input image 212 and feature sets 312. Ex. 1005, 9:8–9. “[R]andom forest classifier 306 may be a pretrained classifier that applies the pretrained classifier to a feature set o[r] feature sets [(FS)] 312 to generate a corresponding (and optimal) motion vector field [(MVF)] 313,” but “any suitable pretrained classifier may be

applied such as, for example, a decision tree learning model, a kernel random forest model, or the like.” *Id.* at 9:9–17. Sixth, a warp and integrate module 307 receives the input image 212 and the motion vector fields 313. *Id.* at 9:26–30. The warp and integrate module 307 warps the eye regions 311 based on motion vector fields 313 and integrates the warped eye regions into the input image 212. *Id.*

Chalom’s Figure 8 is reproduced below.

800

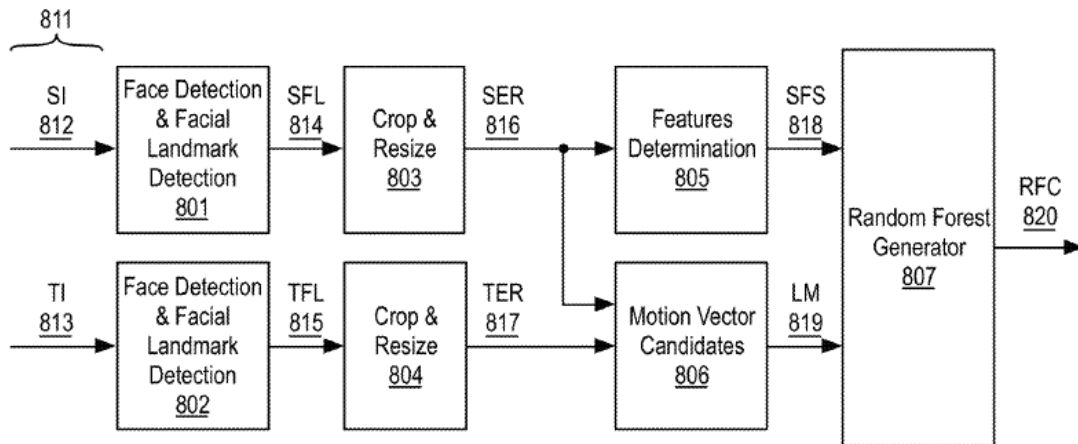


FIG. 8

Figure 8 depicts a flow chart of “an example system for pretraining an eye contact correction classifier.” Ex. 1005, 1:49–50; *see also id.* at 11:1–3. System 800 includes “a face detection and facial landmark detection module 801, a face detection and facial landmark detection module 802, a crop and resize module 803, a crop and resize module 804, a features determination module 805, a motion vector candidates module 806, and a random forest generator module 807.” *Id.* at 11:4–10. According to Figure 8, first, face detection and facial landmark detection module 801 and face detection and facial landmark detection module 802 each receive training images 811

comprising source images (SI) 812 and target images (TI) 813, respectively. *Id.* at 11:24–29. Source images 812 are paired with target images 813 and the pairs have a known gaze angle difference therebetween. *Id.* Then, face detection and facial landmark detection module 801 (which received source images 812) provides source facial landmarks (SFL) 814, and face detection and facial landmark detection module 802 (which received the target images 813) provides target facial landmarks (TFL) 815 in the same manner discussed above in the context of the face detection and facial landmark detection module 301 in Figure 3. *Id.* at 11:45–55.

Second, crop and resize module 803 receives source facial landmarks 814 and provides source eye regions (SER) 816, and crop and resize module 804 receives target facial landmarks (TFL) 815 and provides target eye regions (TER) 817 in the same manner discussed above in the context of the crop and resize module 302 in Figure 3. Ex. 1005, 12:1–9. Third, features determination module 805 receives SER 816 and determines source feature sets (SFS) 818 based on SER 816 in the same manner discussed above in the context of the features generation module 303 in Figure 3. *Id.* at 13:8–17.

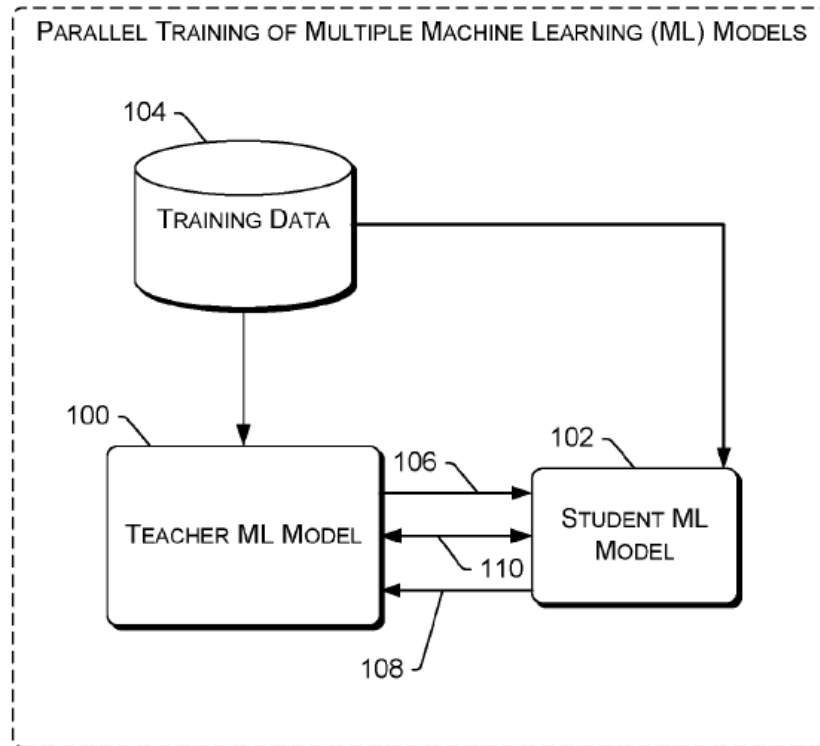
Fourth, motion vector candidates module 806 receives SER 816 and TER 817 and determines likelihood maps (LM) 819 that provide candidate motion vectors or inverse probabilities based on SER 816 and TER 817. *Id.* at 12:23–27, 12:39–41. Fifth, random forest generator module 807 receives SFS 818 and LM 819 and trains and determines random forest classifier (RFC) 820 based on SFS 818 and LM 819. *Id.* at 13:45–49. Then, random forest classifier 807 contains the motion vector field that is used to warp a source image to generate an eye contact corrected image. *Id.* at 17:17–33.

3. *Aslan (Ex. 1007)*

Aslan relates to “[a]n example process for jointly training multiple machine learning models [that] includes providing a set of machine learning models that are to learn a respective task, the set of machine learning models including a first machine learning model and a second machine learning model.” Ex. 1007, code (57). In particular, Aslan relates to “transform[ing]’ a machine learning model from a first type to a second type that mimics the first type of machine learning model,” which “can allow for model compression, where the second type of machine learning model that mimics the first type can, at the completion of the joint training, have a reduced size (in terms of storage footprint)” and can “allow[] for more flexible use of the second type of machine learning model in implementations where storage space and/or computational power is at a premium without significant loss in accuracy of the second model’s output.” *Id.* ¶ 4.

According to Aslan, joint training describes “techniques for training two or more machine learning models in parallel, wherein at least one of the machine learning models influences the training of the other machine learning model,” in contrast to sequential training in which “a first machine learning model is fully trained prior to initiating the training of a second machine learning model” such that “the second machine learning cannot influence the training of the first machine learning model.” Ex. 1007 ¶ 5. In a joint training process, “the second machine learning model can access information about the outputs of the first machine learning model based on the first model’s processing of the training data as input prior to the first model completing its training.” *Id.* ¶ 6.

Aslan's Figure 1 is reproduced below.



Aslan Fig. 1

Figure 1 shows “a schematic diagram of an example technique for joint training of multiple machine learning models.” Ex. 1007 ¶¶ 12, 21.

Figure 1 includes a first machine learning model, or “teacher ML model” 100, and a second machine learning model, or “student ML model” 102. *Id.* ¶ 21. Teacher ML model 100 and student ML model 102 make up a set of machine learning models that can be implemented as any type of machine learning model and are to be trained in parallel such that either model can be capable of learning from the other. *Id.* ¶¶ 21–22. Figure 1 also includes training data 104 which comprises any suitable data, such as image data, speech data, text data, or video data. *Id.* ¶¶ 23–24. In Figure 1, “both machine learning models 100 and 102 can receive at least some of the

training data 104,” but the training data can also be used to train just one of the machine learning models. *Id.*

According to Aslan, machine learning models 100 and 102 pass information between each other, as indicated in Figure 1 as paths 106, 108, and 110. Ex. 1007 ¶ 28. Passing information “comprises formulating an objective function for the multiple machine learning models in a set of models so that each model can have access to unlabeled data, and/or the training data 104, and/or outputs generated by at least one other model through one or more terms of the objective function.” *Id.* In one example, “the first model 100 can learn from the training data 104” and the second model 102’s training “can be influenced by what the first model 100 is learning from the training data 104 while the first model 100 is being trained, and/or before the first model 100 completes its training,” as shown in path 106. *Id.* ¶ 26.

In another example, “the second model 102 is trained in parallel with the training of the first model 100 by providing some or all of the training data 104 to the second model 102,” as shown in path 108, such that “the first (teacher) model 100 can ‘see’ what the second (student) model 102 is learning while the second model 102 trains, and/or before the second model 102 completes its training.” Ex. 1007 ¶ 29. “This can allow the first (teacher) model 100 to adapt what it learns to better match what the second (student) model 102 is learning or is capable of learning.” *Id.*

In a third example, “knowledge can be bi-directionally transferred between the first model 100 and the second model 102 during joint training,” as shown in path 110. Ex. 1007 ¶ 31. In this example, “data can be processed by each model 100 and 102, and the objective function used for

joint training of the models 100 and 102 can determine the degree to which the models 100 and 102 agree with each other, and can ‘push’ the models toward agreement.” *Id.*

Aslan explains that its joint training techniques are helpful in “model compression, which allows for compact representations of deep (i.e., many layers) machine learning models that generally are allocated a large amount of memory to maintain, are complex in architecture, and use a high amount of processing power to operate at runtime” and are also useful when the models “differ in: (i) the learning methods they employ during training, (ii) their respective speed of operation at runtime, (iii) their ability to be distributed across many different machines for use in parallel processing environments, or (iv) their ‘understandability,’” i.e., “one model is in a language more comprehensible to humans than the other.” Ex. 1007 ¶¶ 44, 46.

E. Asserted Obviousness Based on Chalom, Aslan, Grau, and Pace

Petitioner argues that claims 1–3, 5, 6, 8–14, and 16–19 would have been obvious over the combined teachings of Chalom and Aslan; claim 4 would have been obvious over the combined teachings of Chalom, Aslan, and Grau; and claim 7 would have been obvious over the combined teachings of Chalom, Aslan, and Pace. Pet. 41–85. In support of its showing, Petitioner relies upon the Karam Declaration. *Id.* (citing Ex. 1003). Patent Owner opposes. PO Resp. 21–38. Upon consideration of Petitioner’s explanations and supporting evidence on the record before us, we are persuaded that Petitioner has demonstrated by a preponderance of the evidence that the challenged claims would have been obvious over the

combination of Chalom and Aslan; Chalom, Aslan, and Grau; or Chalom, Aslan, and Pace as set forth above. Our analysis below focuses on the disputed claim elements.

Claim 1: 1[P]: A method of generating reference data for adjusting a digital representation of a head region

Independent claim 1 recites, in the preamble (1[P]): “A method of generating reference data for adjusting a digital representation of a head region, the method comprising: . . .” Ex. 1001, 26:30–33.

Petitioner argues that Chalom teaches the preamble of claim 1. Pet. 47–50 (citing Ex. 1003 ¶¶ 230–242; Ex. 1005, 3:21–24, 4:58–60, 9:17–22, 9:26–30, 13:49–52) (asserting that “Chalom discloses ‘[a] method of generating reference data [e.g., motion vector fields generated during training, including the recited “first reference data” in 1[D] and the recited “second reference data” of 1[E]] for adjusting a digital representation of a head region [e.g., gaze correction].”). *Id.* at 50.

Patent Owner does not specifically respond to these arguments. *See generally* PO Resp. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Without determining whether the preamble is limiting, we find Petitioner has shown by a preponderance of the evidence that Chalom (and, thus, the combination of Chalom with Aslan) teaches the preamble.

Claim 1: 1[A] Receiving Training Data comprising:

Independent claim 1 further recites 1[A]: “receiving training data.” Ex. 1001, 26:34.

Petitioner argues that Chalom teaches this limitation. Pet. 50–51 (citing Ex. 1003 ¶¶ 243–245; Ex. 1005, 4:19–22, 4:25–28) (asserting “Chalom discloses ‘receiving training data [e.g., pairs of training eye region images, including source eye regions 816 and target eye regions 817].’”).

Patent Owner does not specifically respond to these arguments. *See generally* PO Resp. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

We find Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches limitation 1[A].

Claim 1: 1[B] A Set of Input Patches

Independent claim 1 further recites that the training data comprises 1[B]: “a set of input patches, each input patch comprising a target feature of a digital representation of a head region prior to adjustment of the digital representation of the head region, wherein the target feature is the same for each input patch.” Ex. 1001, 26:35–39.

Petitioner argues that Chalom teaches these limitations. Pet. 51–53 (citing Ex. 1003 ¶¶ 246–249; Ex. 1005, 11:13–14, 11:24–26, 12:1–22) (asserting “Chalom discloses ‘a set of input patches [e.g., source eye regions 816], each input patch comprising a target feature [e.g., eye region] of a digital representation [e.g., image] of a head region prior to adjustment [e.g., before gaze correction] of the digital representation of the head region, wherein the target feature is the same for each input patch.’”).

Patent Owner does not specifically respond to these arguments. *See generally* PO Resp.

We find Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches limitation 1[B].

Claim 1: 1[C] A Set of Output Patches

Independent claim 1 further recites that the training data also comprises 1[C]: “a set of output patches in one-to-one correspondence with the input patches, each output patch comprising the target feature of the digital representation of the head region after adjustment of the digital representation of the head region.” Ex. 1001, 26:40–44.

Petitioner argues that Chalom teaches these limitations. Pet. 53–54 (citing Ex. 1003 ¶¶ 250–255; Ex. 1005, 11:36–39, 12:10–22, Figs. 8, 11) (asserting that “Chalom discloses ‘a set of output patches [e.g., target eye regions 817] in one-to-one correspondence with the input patches, each output patch comprising the target feature [e.g., eye region] of the digital representation of the head region after adjustment [e.g., after gaze correction, e.g., “with correct gaze”] of the digital representation of the head region.’”).

Patent Owner does not specifically respond to these arguments. *See generally* PO Resp.

We find Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches limitation 1[C].

Claim 1: 1[D]–1[E] Generating First and Second Reference Data

Independent claim 1 further recites 1[D]: “using a first machine learning algorithm to generate first reference data using the training data, the first reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital

representations of the head region.” Ex. 1001, 26:45–49. Additionally, independent claim 1 recites 1[E]:

using a second machine learning algorithm to generate second reference data using the same training data as the first machine learning algorithm and *the first reference data generated by the first machine learning algorithm*, the second reference data comprising editing instructions for adjusting the digital representation of the head region for a range of possible digital representations of the head region.

Ex. 1001, 26:50–57 (emphasis added).

Petitioner argues that Chalom teaches these limitations. Pet. 54–59. In particular, Petitioner argues that Chalom’s disclosure of system 800 using source eye regions SER 816 and target eye regions TER 817, candidate motion vectors, and likelihood maps LM 819 teaches using training data to train a learning machine algorithm. *Id.* at 54–55 (citing Ex. 1003 ¶¶ 256–258; Ex. 1005, 11:29–30, 12:23–27, 13:8–11, 13:45–49, Figs. 8, 11). According to Petitioner, because Chalom’s trained RFC 820 outputs a motion field, which may be used to warp the eye region thereby generating the desired output image having corrected eye contact, “Chalom’s training results in each leaf of the trees in Chalom’s RFC 820 providing a motion likelihood map for every pixel in [the] source eye region and a corresponding motion vector field.” *Id.* at 55–56 (citing Ex. 1003 ¶¶ 259–260; Ex. 1005, 3:36–38, 4:62–64, 9:9–22, 13:49–55).

Further, Petitioner argues that Chalom’s disclosure of an eye contact correction module being capable of implementing eye contact correction for multiple orientations or relative positions between camera 104 and display 102 teaches “adjusting the digital representation of the head region

for **a range of possible digital representations of the head region.**” *Id.* at 56 (citing Ex. 1003 ¶¶ 261–263; Ex. 1005, 5:40–50, 10:10–14).

Additionally, Petitioner proposes to modify Chalom’s system to use Aslan’s teacher-student architecture MLA such that Chalom’s training processes involve the two MLAs. *Id.* at 56–57 (citing Ex. 1003 ¶ 264; Ex. 1007 ¶ 23). According to Petitioner, “[i]n Chalom-Aslan, the training data would have been the source eye regions and target eye regions from Chalom, and this training data would have covered a range of possible digital representations of the head region based on the different camera angles utilized (as explained [in the Petition]).” *Id.* at 57 (citing Ex. 1003 ¶¶ 365–368). “In the Chalom-Aslan system, the ‘task’ is generating motion vector fields for gaze correction, and the teacher’s output, e.g., first editing instructions including motion vector fields, is used to train the student that also generates motion vector fields, i.e., second editing instructions.” *Id.* at 58 (citing Ex. 1003 ¶¶ 269–272). Therefore, Petitioner asserts the following:

[T]he Chalom-Aslan combination discloses “using a first machine learning algorithm [e.g., Chalom-Aslan’s teacher model] to generate first reference data using the training data, the first reference data comprising editing instructions [e.g., motion vector fields generated by the teacher model] for adjusting the digital representation of the head region [e.g., gaze correction] for a range of possible digital representations of the head region [e.g., different head orientations],” as well as “using a second machine learning algorithm [e.g., Chalom-Aslan’s student model] to generate second reference data using the same training data as the first machine learning algorithm [e.g., training data for the teacher model] and the first reference data generated by the first machine learning algorithm [e.g., motion vector fields generated by the teacher model], the second reference data comprising editing instructions [e.g., motion vector fields generated by the student model] for

adjusting the digital representation of the head region for a range of possible digital representations of the head region.”

Id. at 58–59 (citing Ex. 1003 ¶¶ 256–274).

Motivation to Combine

Petitioner contends that Chalom and Aslan are in the same or closely related fields, in that they disclose similar techniques of using machine learning for image processing (e.g., using image data to train MLAs to correct the gaze of a video-conferencing user). Pet. 41–42 (citing Ex. 1005, 3:21–32; Ex. 1007 ¶ 24). Petitioner also contends both “Chalom and Aslan are at least reasonably pertinent to a problem identified in the ’985 Patent of high accuracy MLAs often consuming more computational resources and being slower to run, as compared to efficient MLAs tending to provide less accurate predictions.” *Id.* at 42 (citing Ex. 1001, 16:24–30; Ex. 1007 ¶ 45). Thus, “Chalom and Aslan are at least reasonably pertinent to this efficiency-versus-accuracy problem.” *Id.* (citing Ex. 1005, 1:16–25, 9:48–55, 10:40–60, Fig. 3; Ex. 1007 ¶ 45; Ex. 1001, 16:24–44, 17:14–17; Ex. 1003 ¶ 212). Petitioner thus proposes to incorporate Aslan’s teacher-student architecture into Chalom’s training of its RFC to achieve efficiency gains. *Id.*

Petitioner asserts that the references provide express motivation for the proposed combination. Pet. 43. According to Petitioner, Chalom’s express instruction of training its random forest identifier (RFC) using any suitable technique(s) would have led a person having ordinary skill in the art (“POSITA” or “PHOSITA”) to use Aslan’s teacher-student architecture for compressing large models (e.g., Chalom’s RFC) into smaller models so they can be more readily employed on computing devices with limited resources.

Id. at 43–45 (citing Ex. 1003 ¶¶ 216–224; Ex. 1005, 15:4–7; Ex. 1007 ¶ 44; Ex. 2009, 191:4–23). Therefore, Petitioner submits the following:

One way is using Chalom’s RFC as a teacher model to train an even more compact MLA, e.g., a smaller random forest. A PHOSITA would have understood that Chalom’s RFC is a machine-learning ensemble with a collection of decision trees. Aslan explains that ensembles “typically require a relatively **large storage footprint and powerful processing resources to execute at runtime**, they are not well suited for implementations where storage space and/or computational power is at a premium (such as with smart phones, wearables, hearing aids, etc.).” EX-1007, ¶3. Aslan therefore adopts an improved teacher-student architecture to solve this problem. Thus, a PHOSITA would have been motivated to incorporate Aslan’s teacher-student training architecture by **using Chalom’s already trained RFC as a teacher** to train a more compact RFC with fewer and/or less complex decision trees, thereby maintaining accuracy of the original RFC while achieving efficiency gains. EX-1003, ¶¶219-222.

Id. at 44.

Another way a PHOSITA would have been motivated to incorporate Aslan’s teacher-student architecture is by using Chalom’s RFC as a student model, which learns from a different teacher model. EX-1003, ¶223. It was well known at the time of the ’985 Patent that using a random forest resulted in less optimal prediction accuracy as compared to deep convolutional neural networks for image processing/classification tasks. A PHOSITA looking to improve the accuracy for Chalom’s RFC would have been motivated to incorporate Aslan’s teacher-student training architecture by **training a new accurate teacher model first, which is then used to train Chalom’s RFC as a student model**. As taught by Aslan, such new teacher model could be, e.g., a deep neural network or another type of machine learning model. EX-1007, ¶4 5; EX-1003, ¶224.

Id. at 45.

Thus, a PHOSITA would have been motivated by Chalom's explicit teachings to consult literature such as Aslan for techniques to compress Chalom's RFC, e.g., using Aslan's teacher-student architecture for compressing the student model, as [Patent Owner's] expert concedes. Ex-1061, 86:18-87:3; Ex-1062, 239:16-22. Upon locating Aslan, a PHOSITA also would have known that Chalom's RFC is an ensemble that typically "requires a lot of time and resources" for training as indicated directly in Aslan and confirmed by [Patent Owner's] expert (Ex-1062, 227:3-5, 227:20-228:3), and therefore would have sought to use Aslan to address the known issues with ensembles. This all establishes an express motivation to combine, which alone suffices to establish that a PHOSITA would have combined Chalom with Aslan.

Pet. Reply 12.

Further, Petitioner contends that the proposed combination of Chalom with Aslan involves the combination of known elements (e.g., classifiers) that perform their ordinary functions (e.g. classifying by providing a motion vector field) to predictably improve the accuracy and efficiency of the ensuing Chalom-Aslan system. Pet. 45; Pet. Reply 13.

Patent Owner argues that Petitioner has not provided sufficient motivation to combine Chalom with Aslan. PO Resp. 21–34. First, Patent Owner argues that Petitioner's motivation rests on the incorrect assumption that Chalom is not suitable in a limited resource, low power environment. *Id.* at 21. In particular, Patent Owner asserts that Petitioner's expert incorrectly opined that a POSITA "would look to Aslan because Chalom uses 'different techniques' that cannot function in a low power, mobile device scenario." *Id.* at 21–22 (citing Ex. 2008 ¶¶ 116–117). According to Patent Owner, because Chalom provides a high-quality, real-time MLA that performs perfectly well in a limited, low power environment (e.g., mobile

device), there would not be any such need to combine it with Aslan. *Id.* at 22–23.

Second, Patent Owner argues that a POSITA would not be motivated to combine Chalom and Aslan in the manner proposed by Petitioner because nothing in Chalom indicates that it is inaccurate or inefficient. *Id.* at 23–24. (citing Ex. 2008 ¶¶ 105–115). According to Patent Owner, “adding another machine learning algorithm does not improve accuracy . . . a POSITA would improve an algorithm accuracy by increasing the size/diversity of the training data or by changing certain features of the algorithm, such as parameter or resolution.” *Id.* at 24. More particularly, Patent Owner asserts:

If a POSITA were truly concerned with improving Chalom’s accuracy, they would exercise one of the many other options on which both parties’ experts agree, such as ensuring adequate training data and adding sufficient structure, weights, and parameters to the first (and only) machine learning algorithm. [Ex. 2008] ¶ 111. The alleged inaccuracies of Chalom therefore do not motivate a POSITA to look to Aslan or a second machine learning algorithm for a solution.

Id.

Further, Patent Owner argues that the proposed combination of adding Aslan’s teacher-student into Chalom’s system would instead increase the complexity with more or larger models, and thereby reduce efficiency by teaching incorrect gaze correction parameters, which reduce accuracy. *Id.* at 24–27. More particularly, Patent Owner asserts:

If a POSITA desired a more efficient way to deploy Chalom, they would reduce the size of the model by having fewer trees and branches, or compress the model by selecting the most probable trees and branches. EX-2008 at ¶¶ 113-115. This is precisely what Chalom suggests. *Id.* at ¶ 115 (citing EX-1005 at 15:4-7). There is no reason to look for a second

algorithm, or a teacher-student system, to improve the efficiency of Chalom.

Id. at 26.

Third, Patent Owner argues that Chalom's motion vectors do not lend themselves to a combination with Aslan because these systems are directed to different techniques for different outcomes (i.e., Chalom performs a gaze correction task, whereas Aslan performs an image classification task). *Id.* at 27–28 (citing Ex. 2008 ¶¶ 119–120).

We disagree with Patent Owner's arguments. As set forth above, in its contentions for this limitation, Petitioner refers back to its analysis of limitations 1[D] and 1[E]. *See* Pet. 54–59. In its analysis of limitation 1[D], Petitioner, relying on Figure 8, argues that Chalom's disclosure of providing SFS 818, candidate motion vectors, and LM 819 to RFG 807 to train RFC 820, which outputs a motion field for warping the eye region of a desired output image, teaches using training data to train a learning machine algorithm to adjust an input gaze of a user. *Id.* at 54–55. Petitioner then proposes modifying Chalom's system with Aslan's teacher-student MLAs to train the two MLAs with the same input data to produce a motion field output, as well as to provide the motion vector correction from the teacher MLA to the student MLA for further refinement. *Id.* at 56–57 (contentions as to limitation 1[E]).

Because Chalom explicitly indicates that its RFC can be trained using any suitable techniques(s) including compressing the classifier, we agree with Petitioner that a POSITA would have looked to Aslan's teacher-student architecture that proposes compressing large models such as Chalom's RFC into smaller models so they can be more readily deployed on computing

devices with limited resources. Pet. Reply 11–12. As correctly noted by Petitioner, the experts of the parties agree that both Aslan and Chalom involve classifiers. More particularly, as acknowledged by Patent Owner’s representative during the hearing, compression is desirable for Chalom’s RFC. *See* Tr. 56:21–57:4.

Accordingly, we agree with Petitioner that a POSITA would have looked to a similar compression operation of classifiers in Aslan for training a more compact MLA at runtime or offline. Pet. Reply 11–12. For these same reasons, we find unavailing Patent Owner’s argument that Chalom and Aslan are directed to different objectives or techniques.

Further, we do not agree with Patent Owner’s argument that replacing Chalom’s large RFC with Aslan’s teacher-student MLA pair would have rendered the ensuing system more complex, and thereby inefficient. As correctly noted by Petitioner, Chalom indicates that the RFC can be embodied in a mobile computing device. *Id.* at 15 (citing Ex. 1005, 22:33–37, 19:65–20:6). And Dr. Karam persuasively explains that using the teacher model to train the student model is more efficient because the teacher’s knowledge allows the student model to converge faster. *Id.* at 17 (citing Ex. 2009, 178:14–24).

Accordingly, we find that a POSITA would have found sufficient motivation to replace Chalom’s RFC with Aslan’s teacher-student models to enhance the efficiency of the MLA pair in a mobile computing environment. On this record, we agree with Petitioner that the proposed combination of Chalom and Aslan would predictably result in the teacher MLA generating motion vector fields for user gaze correction (*i.e.*, first reference or editing instructions) that are passed to the student MLA, which subsequently

generates motion vector fields (*i.e.*, second reference or editing instructions) for further refining the received first reference data. *Id.* at 58.

Consequently, we reiterate that because the proposed combination involves known elements performing their ordinary functions to produce a predictable result, a POSITA would have appreciated that the proposed combination is proper.⁴

After considering Patent Owner’s arguments, we find Petitioner has shown by a preponderance of the evidence that the proposed combination of Chalom and Aslan is sufficiently supported by the requisite motivation to teach limitations 1[D] and 1[E].

Claim 1: 1[F]–1[G] First Editing/Selection Instructions and Second Editing/Selection Instructions

Independent claim 1 further recites 1[F]:

[W]herein the first reference data comprise first editing instructions for a range of possible configurations of the target feature and first selection instructions for selecting editing instructions for a particular input patch from the first editing instructions based on the configuration of the target feature of the input patch.

Ex. 1001, 26:58–63.

⁴ It has been held that where the claimed subject matter involves “more than the simple substitution of one known element for another or the mere application of a known technique to a piece of prior art ready for the improvement,” a holding of obviousness *must* be based on “an apparent reason to combine the known elements in the fashion claimed.” *KSR*, 550 U.S. at 417–18. Such reasons can be based on interrelated teachings of multiple patents, the effects of demands known to the design community or present in the marketplace, and the background knowledge possessed by a person having ordinary skill in the art. *Id.* at 418.

Additionally, independent claim 1 recites 1[G]:

[W]herein the second reference data comprise second editing instructions for a range of possible configurations of the target feature and second selection instructions for selecting editing instructions for a particular input patch from the second image editing instructions based on the configuration of the target feature of the input patch.

Ex. 1001, 26:64–27:3.

Petitioner argues that Chalom teaches these limitations. Pet. 59–63 (citing Ex. 1003 ¶¶ 275–286; Ex. 1005, 8:59–61, 10:35–39, 12:10–22, Figs. 8, 11). In particular, Petitioner asserts the following:

Chalom-Aslan discloses that “the first reference data comprise first editing instructions [e.g., teacher’s motion vector fields generated during training] for a range of possible configurations of the target feature [e.g., different gaze directions] and first selection instructions [e.g., trained teacher model] for; selecting editing instructions [e.g., optimal motion vector field] for a particular input patch from the first editing instructions [e.g., teacher’s motion vector fields generated during training] based on the configuration of the target feature of the input patch [e.g., Chalom’s feature set 312],” and that “the second reference data comprise second editing instructions [e.g., student’s motion vector fields generated during training] for a range of possible configurations of the target feature [e.g., different gaze directions] and second selection instructions [e.g., trained student model] for selecting editing instructions [e.g., optimal motion vector field] for a particular input patch from the second image editing instructions [e.g., student’s motion vector fields generated during training] based on the configuration of the target feature of the input patch [e.g., Chalom’s feature set 312].”

Id. at 62–63 (citing Ex. 1003 ¶¶ 275–286).

Patent Owner argues that the Petition does not establish that the Chalom-Aslan combination discloses a second MLA that receives the first

selection instructions (i.e., providing to the second MLA the first reference data including first editing instructions and *first selection instructions*), as required by claim 1. *Id.* at 28–29. According to Patent Owner, the Petition only establishes that Chalom’s trained teacher model [MLA1] teaches the “first selection instructions” without indicating whether the “trained teacher model” is provided to the second MLA, let alone indicating how the second MLA uses that the trained teacher model. *Id.* at 29.

This argument is unavailing because it is not commensurate with the scope of the claim. As correctly noted by Petitioner, neither the corpus of the ’985 patent nor the disputed claim limitation supports the argument that the first selection instruction must be passed from the MLA1 to the MLA2. Pet. Reply 20. We agree with Petitioner that although the ’985 patent indicates that MLA1 outputs first reference data including first selection instructions LA1 and first editing instructions ED1, it simply indicates that MLA2 receives first editing instructions ED1. *Id.* at 21 (citing Ex. 1001, 17:5–6). It is nonetheless silent regarding whether selection instructions LA1 are also passed to or received by MLA2. *Id.* Accordingly, we agree with Petitioner that it suffices that the Chalom-Aslan combination teaches the teacher model/MLA1 passing the editing instructions (e.g., motion vector fields generated by teacher model) to the student model/MLA2. *Id.* at 22–23.

Further, Patent Owner argues that the Petition does not establish that the Chalom-Aslan combination discloses a second MLA that uses the reference data including the editing instructions from the first MLA. *Id.* at 30. Patent Owner asserts that Petitioner fails to sufficiently explain what information passes between the teacher and student models, how the output

of the teacher model is used by the student model, or how the “teacher’s output” results in a smaller, more compact student model. *Id.* at 30–31 (citing Ex. 2008 ¶¶ 122–129). More particularly, Patent Owner asserts:

Without further detail, a POSITA is left to do as Petitioner commands: “follow the instructions provided in Aslan.” Petition at 45, 47. When a POSITA does so, they do not arrive at a combination that practices the ’985 claims. Indeed, they arrive at a teacher-student system that exchanges an “objective function,” which cannot be the claimed “first reference data.” EX-2008 at ¶ 130-133; *see also* EX-1007 at ¶ 28 (explaining that “‘passing information’ between machine learning models [] comprises formulating an objective function.”); *id.* at ¶ 39 (explaining that “passing information between the models 100 and 102 [occurs] via the objective function”).

Id. at 32.

According to Patent Owner, the ’985 patent’s editing instructions “adjust the digital representation of the head region,” whereas Aslan’s objective function is a “performance measure” of an algorithm’s error used during training of the MLA. *Id.* That is, because the information sent from the teacher to the student MLA is made during training of the algorithm, the disclosed editing instructions are not for adjusting the digital representation of the head region as in the ’985 patent. *Id.* at 33 (citing Ex. 2008 ¶ 133). This argument is not persuasive. As noted above, Petitioner has established that the Chalom-Aslan combination discloses the teacher model/MLA1 passing the editing instructions (e.g., motion vector fields generated by the teacher model) to the student model/MLA2. Pet. Reply at 22–23. Petitioner has further established that, similarly to the displacement vector field used in the ’985 patent, the teacher model uses the motion vector field to train the student model to perform gaze correction. *Id.* at 24–26. Accordingly, we

agree with Petitioner that a POSITA would likewise have been motivated to use Aslan's teacher-student model to correct the gaze of a user. *Id.*

Additionally, Patent Owner argues that Petitioner has not established that the combination of Chalom and Aslan discloses that the first MLA generates first reference data and that the second MLA generates second reference data, wherein the reference data outputs are different from each other. PO Resp. 34–36 (citing Ex. 2008 ¶ 136). According to Patent Owner, because the Chalom-Aslan combination uses the same inputs provided to the same modules, the teacher model and the student model would return the same outputs. *Id.* at 36 (citing Ex. 2008 ¶¶ 135–137).

Patent Owner's argument is not persuasive. Although the teacher model and the student model receive the same training data in the Chalom-Aslan combination, they do not have the same inputs or outputs. As discussed above, the teacher model receives the training data as an input to produce a vector field, which is forwarded along with the training data to the student model to produce the corrected gaze. We thus agree with Petitioner that the output of the teacher model (reference 1) is different from the output of the student model (reference 2). Pet. Reply 27.

After considering Patent Owner's arguments, we find Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches limitations 1[F]–1[G].

Conclusion Regarding Independent Claim 1

After considering Patent Owner's arguments, we find Petitioner has shown by a preponderance of the evidence that independent claim 1 is unpatentable under § 103 over the combination of Chalom and Aslan.

Independent Claims 16–19

For essentially the same reasons as for independent claim 1, Petitioner contends that independent claims 16–19, which differ only in certain uncontested limitations that Petitioner also addresses would have been obvious over the combined teachings of Chalom and Aslan. Pet. 75–80. Patent Owner does not separately argue independent claims 16–19 but relies on its arguments for limitations 1[E] and 1[F] discussed above for corresponding limitations of those claims. See PO Resp. 30–37.

For the same reasons as set forth above for independent claim 1, we determine that Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches the contested limitations in claims 16–19 and has provided persuasive rationale for combining the references. We also determine that Petitioner has shown that the additional, uncontested limitations of claims 16–19 are taught by Chalom and Aslan. See Pet. 75–80. We determine, therefore, that Petitioner has demonstrated by a preponderance of the evidence that claims 16–19 are unpatentable under § 103 over the combination of Chalom and Aslan.

Dependent Claims 2–14

Petitioner also contends that claims 2, 3, 5, 6, and 8–14, which depend directly or indirectly from independent claim 1, would have been obvious over Chalom and Aslan. Pet. 4, 63–75. Further, Petitioner contends that the combination of Chalom, Aslan, and Grau teaches the limitations in dependent claim 4; and the combination of Chalom, Aslan, and Pace teaches the limitations in dependent claim 7. *Id.* at 80–85. Patent Owner does not separately argue the dependent claims. See PO Resp.

We find Petitioner has shown by a preponderance of the evidence that the combination of Chalom and Aslan teaches the limitations in claims 2, 3, 5, 6, and 8–14; the combination of Chalom, Aslan, and Grau teaches the limitations in claim 4; and the combination of Chalom, Aslan, and Pace teaches the limitations in claim 7. Accordingly, we determine that the information presented in the Petition demonstrates by a preponderance of the evidence that dependent claims 2, 3, 5, 6, and 8–14 are unpatentable over the combination of Chalom and Aslan; claim 4 is unpatentable over the combination of Chalom, Aslan, and Grau; and claim 7 is unpatentable over the combination of Chalom, Aslan, and Pace.

F. Asserted Anticipation Based on Ganin

Petitioner argues that claims 1–3, 5, 10, 12–14, and 18 are anticipated by Ganin, and that claims 16 and 19 are rendered obvious based on the combined teachings of Ganin and Liu. Pet. 6–41. In support of its showing, Petitioner relies upon the Karam Declaration. *Id.* (citing Ex. 1003)

1. Claim 1

Petitioner argues that Ganin’s disclosure of a training process wherein a warping field (i.e., editing instructions) involving a deep convolutional network is applied to an input eye region to produce an image of a redirected eye teaches limitations 1[P] through 1[G]. Pet. 8–26. According to Petitioner, because each of Ganin’s coarse warping module and fine warping module includes a neural network having convolutional layers, the disclosed modules teach the first MLA and the second MLA, respectively, such that the coarse warping field (first instruction) obtained from the coarse warping module serves as a reference input to the fine warping module, which

subsequently outputs a fine warping field for refining the correction of the input image. *Id.* at 13–18. More particularly, Petitioner contends the following regarding limitations 1[D] and 1[E]:

A PHOSITA would have understood that the synthesized output O_{output} could refer to both coarse estimate O_{coarse} (output of the coarse warping module) and final output O (output of the fine warping module). Accordingly, the ground-truth O_{gt} refers to the “output images” having redirected gazes, which are provided to both the coarse warping module (to calculate the ℓ_2 -distance between coarse estimate O_{coarse} and O_{gt}) and the fine warping module (to calculate the ℓ_2 -distance between final output O and O_{gt}) during training. Thus, Ganin’s coarse warping module is trained using the claimed “training data” that includes both “input images” and “output images.” EX-1003, ¶¶115-117.

Id. at 16–17.

Therefore, Ganin discloses “using a first machine learning algorithm [e.g., coarse warping module including the $0.5\times$ -scale neural network] to generate first reference data [e.g., coarse warping fields $\mathbf{D}_{\text{coarse}}(I, \alpha)$] using the training data [e.g., “input maps” including input images, and output images as ground truth], the first reference data comprising editing instructions [e.g., pixel-flow fields and coarse warping fields $\mathbf{D}_{\text{coarse}}(I, \alpha)$ up-sampled therefrom] for adjusting the digital representation of the head region [e.g., gaze redirection] for a range of possible digital representations of the head region [e.g., different head poses].” EX-1003, ¶¶111-123.

Id. at 18.

Ganin’s fine warping module takes two inputs during runtime. First is the input maps... The other input to Ganin’s $1\times$ -scale NN is the coarse warping fields $\mathbf{D}_{\text{coarse}}(I, \alpha)$. These two inputs are provided to Ganin’s fine warping module during both runtime and training. EX-1003, ¶125.

In addition, Ganin teaches using “a regular ℓ_2 -distance between the synthesized output O_{output} and the ground-truth O_{gt}

as the objective function.” EX-1006, p.8. The ground-truth O_{gt} refers to the “output images” having redirected gazes, which are provided to the fine warping module during training. Thus, *Ganin’s fine warping module is trained using the claimed “training data” that includes both “input images” and “output images.”* EX-1003, ¶126.

Id. at 19–20 (emphasis added).

Therefore, Ganin discloses “using a second machine learning algorithm [e.g., fine warping module including the $1\times$ -scale NN] to generate second reference data [e.g., residual flow $\mathbf{D}_{res}$ and final warping field $\mathbf{D}(I, \alpha)$] using the same training data as the first machine learning algorithm [e.g., “input maps” including input images, and output images as ground truth] and the first reference data generated by the first machine learning algorithm [e.g., coarse warping fields $\mathbf{D}_{coarse}(I, \alpha)$], the second reference data comprising editing instructions [e.g., residual flow $\mathbf{D}_{res}$ and final warping field $\mathbf{D}(I, \alpha)$] for adjusting the digital representation of the head region [e.g., gaze redirection] for a range of possible digital representations of the head region [e.g., different head poses].” EX-1003, ¶¶124-130.

Id. at 21–22.

Patent Owner argues that Ganin does not disclose two separate MLAs, each receiving the same training data. PO Resp. 9. According to Patent Owner, Ganin’s coarse/fine warping modules relate instead to a two-step process executed as part of a single MLA containing two warping modules that are trained together with a single training data set as part of the whole system training. *Id.* at 10–17. First, Patent Owner argues that because Ganin discloses using one training dataset for the two modules (instead of one training data set per module, each performing a different task), it teaches one MLA with one set of training data performing one task. *Id.* at 10–11 (citing Ex. 2008 ¶¶ 85–87). This argument is not persuasive because it is not

commensurate with the scope of the claim. As persuasively argued by Petitioner, this argument belies the disputed claim language requiring two MLAs receiving *the same training data* to perform two different tasks. Pet. Reply 3. Thus, although the disputed claim limitation recites two MLAs, it does not require two different training datasets for the MLAs. Instead, it plainly requires using a single and the same training dataset for the two MLAs.

Second, Patent Owner argues that because Ganin discloses one objective function that assesses one output (i.e., redirected gaze image), it teaches one MLA. PO Resp. 13 (citing Ex. 2008 ¶¶ 94, 95). That is, although Ganin discloses a coarse warping module and a fine warping module, it does not assess the output at each of the modules, but instead redirects the gaze image at the output of the two modules together. *Id.* (alleging that Petitioner's expert conceding this point); Ex. 1003 ¶ 116. This argument is not persuasive. As correctly noted by Petitioner, the '985 patent does not mention an objective function at all, and Patent Owner's expert, Eli Saber, Ph.D., acknowledged that Aslan describes using an objective function for joint training of multiple MLAs. Pet. Reply 7. Accordingly, Patent Owner's asserted requirement for using each MLA having its own training data is not supported by the record before us. Further, as discussed in conjunction with Figure 10(a) above, Ganin's coarse warping module processes training data along with input maps (eye region + anchor points extracted from eye region + desired angle) to produce a first coarse output of the redirected image gaze, which is then concatenated with the training data and submitted as input to the fine warping module to produce the second fine output. Ex. 1003 ¶ 112; Ex. 1006, 18. Because Ganin's coarse warping

module produces a first output of a coarse redirected image gaze, which is provided to the fine warping module to subsequently produce a second output of the fine redirected image gaze, each of Ganin's modules performs a distinct and objective function.

Third, Patent Owner argues that Ganin's text repeatedly states that it includes one MLA. PO Resp. 14–15. According to Patent Owner, Ganin discloses a single algorithm using a two-step process, as opposed to two MLAs. *Id.* This argument is not persuasive. As a preliminary matter, we note that both Patent Owner's expert and Petitioner's expert appear to agree that the plain and ordinary meaning of machine learning algorithm (MLA) is a computer model or algorithm that trains itself or learns to perform a specific task. Ex. 1003 ¶ 23; Ex. 2008 ¶ 42. Further, as noted above, Ganin indicates that a module (e.g., lightness correction module) includes a neural network capable of being trained with input training data to output a corrected or redirected gaze image. Ex. 1006, 18. Because Ganin's coarse warping module and fine warping module are able to be trained with input training data to perform a coarse image redirection task and a fine image redirection task, the ordinarily-skilled artisan would appreciate that these modules qualify as two separate MLAs.

Fourth, Patent Owner argues that Ganin does not disclose a first MLA and a second MLA that use the same training data. PO Resp. 18. According to Patent Owner, Ganin's coarse warping module and fine warping module are trained on different sets of data, as opposed to the same training data, as claimed. *Id.* at 18–19 (citing Ex. 2008 ¶¶ 85–87, 94–103). This argument also is not persuasive. As discussed above, Figure 10(a) of Ganin depicts that both the coarse warping module and the fine warping module receive

the same training data. In particular, the coarse warping module receives the training data along with the input maps and the fine warping module receives the training data along with the coarse output and other reference data. Although the training data is concatenated with other types of data before being forwarded to the coarse or fine warping modules, the training data forwarded to these modules remains the same.⁵ Concatenating the training data with other types of data does not change the training data itself. Because the disputed claim limitation does not preclude concatenating other types of data with the training data before passing the training data to the modules, we agree with Petitioner that Ganin’s disclosure does describe passing the same training data to the two MLAs. Furthermore, as correctly noted by Petitioner, because Patent Owner’s expert understands Ganin to disclose the coarse and fine warping modules receiving the same input maps (training data), the record before us supports that Ganin discloses using the same training data for both MLAs. Pet. Reply 10–11.

After reviewing Petitioner’s assertions and supporting evidence in light of Patent Owner’s arguments, we are persuaded that Petitioner has demonstrated by a preponderance of the evidence that claim 1 is anticipated by Ganin.

⁵ Microsoft Computer Dictionary, 5th Ed. (2002) at 122 (defining “concatenate” as “to join sequentially”); the Merriam-Webster Online Dictionary, <https://www.merriam-webster.com/dictionary/concatenate> (defining “concatenate” as “to link together in a series or chain”), or the Cambridge Dictionary, <https://dictionary.cambridge.org/us/dictionary/english/concatenate> (defining “concatenate” as “to put things together as a connected series”).

2. Claims 2, 3, 5, 10, 12–14, and 18

Petitioner contends that Ganin teaches the limitations of claim 2, 3, 5, 10, 12–14, and 18. Pet. 26–38. Patent Owner does not present specific arguments regarding these claims. PO Resp. 20. We have reviewed Petitioner’s assertions and supporting evidence, and in light of the foregoing, we determine that Petitioner has demonstrated by a preponderance of the evidence that claims 2, 3, 5, 10, 12–14, and 18 are anticipated by Ganin.

G. Obviousness Based on Ganin and Liu

Petitioner contends that Ganin in combination with Liu teaches the limitations of claims 16 and 19. Pet. 38–41. Petitioner also articulates reasons to combine the prior art’s teachings. *Id.* Patent Owner does not present specific arguments regarding these claims. PO Resp. 20–21. We have reviewed Petitioner’s assertions and supporting evidence, and in light of the foregoing, we determine that Petitioner has demonstrated by a preponderance of the evidence that claims 16 and 19 are unpatentable under § 103 as obvious over the combination of Ganin and Liu.

III. CONCLUSION

For the reasons discussed above, Petitioner has shown by a preponderance of the evidence that claims 1–14 and 16–19 of the ’985 patent are unpatentable.⁶

⁶ Should Patent Owner wish to pursue amendment of the challenged claims in a reissue or reexamination proceeding subsequent to the issuance of this decision, we draw Patent Owner’s attention to the April 2019 *Notice Regarding Options for Amendments by Patent Owner Through Reissue or Reexamination During a Pending AIA Trial Proceeding*. See 84 Fed. Reg.

In summary:

Claim(s)	35 U.S.C. §	Reference(s)/Basis	Claim(s) Shown Unpatentable	Claim(s) Not Shown Unpatentable
1–3, 5, 10, 12– 14, 18	102	Ganin	1–3, 5, 10, 12–14, 18	
16, 19	103	Ganin, Liu	16, 19	
1–3, 5, 6, 8–14, 16–19	103	Chalom, Aslan	1–3, 5, 6, 8–14, 16–19	
4	103	Chalom, Aslan, Grau	4	
7	103	Chalom, Aslan, Pace	7	
Overall Outcome			1–14, 16–19	

IV. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that Petitioner has shown by a preponderance of the evidence that claims 1–14 and 16–19 of the '985 Patent are unpatentable;
and

FURTHER ORDERED that, because this is a Final Written Decision, parties to this proceeding seeking judicial review of this Decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

16,654 (Apr. 22, 2019). If Patent Owner chooses to file a reissue application or a request for reexamination of the challenged patent, we remind Patent Owner of its continuing obligation to notify the Board of any such related matters in updated mandatory notices. *See* 37 C.F.R. § 42.8(a)(3), (b)(2).

IPR2023-01046
Patent 10,740,985 B2

For PETITIONER:

Alexander Stein
Andrew Devkar
MORGAN, LEWIS & BOCKIUS LLP
alexander.stein@morganlewis.com
andrew.devkar@morganlewis.com

For PATENT OWNER:

Jennifer A. Mauri
Timothy J. Gorry
MICHELMAN & ROBINSON, LLP
jmauri@mrllp.com
tgorry@mrllp.com