

Filed on behalf of: Align Technology, Inc.

By: Matthew A. Argenti (margenti@wsgr.com)
Michael T. Rosato (mrosato@wsgr.com)
Patrick M. Medley (pmedley@wsgr.com)

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ALIGN TECHNOLOGY, INC.,
Petitioner,

v.

DENTAL MONITORING SAS,
Patent Owner.

Case No. IPR2024-00052
Patent No. 11,049,248

**PETITION FOR INTER PARTES REVIEW OF
U.S. PATENT NO. 11,049,248**

TABLE OF CONTENTS

	<u>Page</u>
I. Introduction.....	1
A. Background	2
1. Image analysis and computer vision	3
2. Edge detection and contour detection	4
3. Deep learning	5
B. Cited Art	6
1. Kim.....	6
2. Invisalign.....	7
3. Salah.....	8
4. Maninis.....	10
C. The '248 Patent	10
D. Prosecution History	12
E. Level of Skill in the Art.....	14
II. Statement of the Precise Relief Requested for Each Claim Challenged	14
III. Claim Construction.....	15
A. tooth attribute	15
B. image attribute	16
C. image extracted from a film	18
D. Steps 1(b'') and 15(b'')	18
IV. Claims 1-2, 4, 7-8, and 13-16 are unpatentable as obvious under 35 U.S.C. §103 over Kim, Invisalign, Salah, and Maninis	20
A. Overview of Ground and Rationale to Combine with Reasonable Expectation of Success	20
B. Claim 1	27
1. Preamble.....	28
2. Step 1(a'')	28
3. Step 1(b'')	34

C.	Claim 2	41
D.	Claim 4	43
E.	Claim 7	44
1.	Step 7(1).....	45
2.	Step 7(2).....	47
3.	Step 7(3).....	47
4.	Step 7(4).....	49
F.	Claim 8	50
1.	Step 8(1')	51
2.	Step 8(2')	52
3.	Step 8(3')	53
G.	Claim 13	53
H.	Claim 14	54
I.	Claim 15	55
J.	Claim 16	57
V.	Discretionary Denial is Not Warranted	59
VI.	Grounds for Standing.....	61
VII.	Conclusion	61
VIII.	Mandatory Notices under 37 C.F.R. §42.8.....	63
IX.	Payment of Fees under 37 C.F.R. §§42.15(a) and 42.103	63
X.	Certificate of Compliance.....	64
XI.	Appendix – List of Exhibits.....	65

LISTING OF CHALLENGED CLAIMS

1. A method for assessing the shape of an orthodontic aligner, said method comprising the following steps:

a") more than 1 week after the start of the treatment with the aligner, acquisition of at least one image at least partially representing the aligner in a service position in which it is worn by a patient, called "analysis image", the analysis image being a photograph, or an image extracted from a film;

b") analysis of the analysis image by means of a deep learning device, trained by means of a learning base, so as to determine a value

for at least one tooth attribute of an "analysis tooth zone" representing, at least partially, a tooth on said analysis image, the tooth attribute relating to a separation between the tooth represented by the analysis tooth zone, and the aligner represented on the analysis image,

in the step a"), a cellphone is used to acquire the analysis image.

2. The method as claimed in claim 1, comprising, after the step b"), step c") as follows:

c") assessment of the suitability of the aligner as a function of the value of said tooth attribute.

4. The method as claimed in claim 2, comprising, after the step c"), a step d") as follows:

d") sending of an information message as a function of said assessment.

7. The method as claimed in claim 1, in which the step b") comprises the following steps:

1) creation of a learning base comprising more than 1000 images of dental arches, or "historical images", each historical image representing an aligner worn by a "historical" patient and comprising one or more zones each representing a tooth, or "historical tooth zones", to each of which, for at least one tooth attribute relating to a separation between the tooth represented by the historical tooth zone considered, and the aligner represented, a tooth attribute value is assigned;

2) training of at least one deep learning device, by means of the learning base;

3) submission of the analysis image to the deep learning device for it to determine at least one probability relating to: iii) for each of said teeth, determination of a distance score dependent on

the presence, in a location of said analysis image, of an analysis tooth zone, and

the attribute value of the tooth represented on said analysis tooth zone;

4) determination, as a function of said probability, of the presence of a separation of the aligner and of the tooth represented by said analysis tooth zone.

8. The method as claimed in claim 1, in which the step b'') comprises the following steps:

1') creation of a learning base comprising more than 1000 images of dental arches, or "historical images", each historical image comprising an attribute value for at least one image attribute, or "image attribute value", the image attribute relating to a separation between the tooth or teeth and the aligner represented on the historical image;

2') training of at least one deep learning device, by means of the learning base;

3') submission of the analysis image to the deep learning device for it to determine, for said analysis image, at least one probability relating to said image attribute value.

13. The method as claimed in claim 1, in which in step b''), the deep learning device is a neural network.

14. The method as claimed in claim 1, in which the step b'') comprises the following steps:

1) creation of a learning base comprising more than 1000 images of dental arches, or "historical images", each historical image representing an aligner worn by a "historical" patient and comprising one or more zones each representing a tooth, or "historical tooth zones", to each of which, for at least

one tooth attribute relating to a separation between the tooth represented by the historical tooth zone considered, and the aligner represented, a tooth attribute value is assigned;

2) training of at least one deep learning device, by means of the learning base;

3) submission of the analysis image to the deep learning device for it to determine at least one probability relating to:

the presence, in a location of said analysis image, of an analysis tooth zone; and

the attribute value of the tooth represented on said analysis tooth zone;

4) determination, as a function of said probability, of an amplitude of said separation.

15. A method for assessing the shape of an orthodontic aligner, said method comprising the following steps:

a") more than 1 week after the start of the treatment with the aligner, acquisition of at least one image at least partially representing the aligner in a service position in which it is worn by a patient, called "analysis image" the analysis image being a photograph, or an image extracted from a film;

b") analysis of the analysis image by means of a deep learning device, trained by means of a learning base, so as to determine a value for an image attribute of the analysis image, the image attribute relating to a separation between at least one tooth represented on the analysis image, and the aligner represented on said analysis image.

16. The method as claimed in claim 15, comprising, after the step b"), step c") as follows:

c") assessment of the suitability of the aligner as a function of the value of said image attribute.

Align Technology, Inc., (“Petitioner” or “Align”) requests review of United States Patent No. 11,049,248 to Philippe Salah et al. (“the ’248 patent,” EX1001), currently assigned to Dental Monitoring SAS (“Patent Owner”). This petition demonstrates by a preponderance of the evidence that claims 1-2, 4, 7-8, and 13-16 of the ’248 patent are unpatentable as obvious.

I. INTRODUCTION

The ’248 patent purports to have invented a method for assessing the shape of orthodontic aligners. However, the claims of the ’248 patent merely recite the use of generic image analysis with deep learning to perform long-known methods of assessment of orthodontic appliances.

The ’248 patent recites steps of acquiring images of the teeth of patients wearing orthodontic aligners and analyzing the images to generate a value that relates to a separation between the teeth and the aligner. However, assessing separation between the ends of teeth and the surface of an aligner worn over the teeth was long established in the published prior art as useful for assessing aligner fit. Further, the acquisition of cellphone images of patients’ teeth for orthodontic analysis, including images depicting aligners currently worn by the patient, was also taught as a method to allow remote assessment of orthodontic patients’ treatment status, such as the appropriateness of the worn aligner.

While the '248 patent also recites the use of deep learning at a high level of generality, the prior PCT publication of Salah et al. had already suggested using such deep learning when analyzing dental images, including detecting edges of objects such as teeth, and a prior art publication to Maninis et al. had shown how deep learning could be used for edge and object detection on a wide array of image subjects. Thus, as explained in more detail in the petition, the '248 patent fails to patentably distinguish over existing prior art techniques.

Claims 1-2, 4, 7-8, and 13-16 of the '248 patent should be found unpatentable and cancelled.

A. Background

This petition is supported by the Declaration of Dr. Hassan Foroosh, a Professor in the Department of Computer Science at the University of Central Florida. EX1002, ¶¶1-16; EX1003. As Dr. Foroosh explains and as addressed in further detail below, the involved claims would not have been considered new or non-obvious to a person of ordinary skill in the art (“POSA”) at the relevant time.

The '248 patent describes methods involving image analysis by a deep learning device, with specific focus on analysis of orthodontic images. The brief background described below provides a concise overview of some important elements of image processing and deep learning that are relevant to understanding

what a POSA would have known in these areas at the time of filing of the '248 patent. EX1002, ¶29.

1. Image analysis and computer vision

Image analysis is often described as the process of extracting information from images that may be subsequently used to perform certain tasks. EX1002, ¶30. Developing methods and computer algorithms to provide such capability has been the subject of study in computer vision in the past few decades. However, it was not until after the early years of last decade, with major breakthroughs in deep learning, that computers managed to also outperform humans in vision tasks. EX1015, 1-2, 8; EX1002, ¶31. Hardware developments made it possible for deep learning networks to be trained in an end-to-end manner and on much larger training sets, involving tens of thousands to millions of images. EX1002, ¶31.

Computer vision, in simple terms, is the science of making computers see. Vision tasks can be divided into three levels: low-level, mid-level, and high-level tasks. EX1016, 1538-39, 1541; EX1002 ¶32. Accordingly, computer vision algorithms were traditionally developed to handle low-level, mid-level, and high-level vision problems. Low-level computer vision methods process images at pixel-level (e.g., quantization, encoding, image restoration, image enhancement); mid-level computer vision methods produce feature-level representations of images (e.g., color features, edge features, 3D geometric features); and high-level

computer vision aims to understand image contents and contextual relationships (e.g., object classification, localization, and instance recognition). EX1002, ¶¶32-33. Boundary-based representations, such as edges and contours defining separation between regions in images, are common mid-level features. *Id.*

The development of deep learning led to a new class of computer vision methods that learn features from data, as well as taking advantage of high-level labels to learn better features, while still relying also on low-level image properties to extract features. These allow a device to receive raw image data and directly learn to accomplish a vision task. EX1002, ¶34.

2. Edge detection and contour detection

Edges, contours, boundaries, and discontinuities refer to similar concepts involving a representation of a discontinuity in an image. Edges and contours have been used for various computer vision, from high-level (e.g., classification, object recognition, annotation) to low-level (e.g., registration, alignment) tasks. EX1002, ¶35.

Due to the key role of edges in computer vision, researchers developed deep-learning based methods for edge and contour detection. EX1017, 1-2; EX1007, Abstract; EX1002, ¶36. Instead of binary labels (1 and 0) for edge locations, deep-learning methods lead to multiple labels in the contour image. EX1017, 5-9; EX1007, 8, 9-12; EX1002, ¶36. Each contour is labeled based on high-level labels

(e.g. class, instance, etc.). Thus, deep-learning methods allow categorization of different objects in an image into specific classes. EX1002, ¶37. These type of contours or boundaries are sometimes referred to as “category-specific” or “semantic.” *Id.*; EX1017, 9; EX1007, 11. Thus, a deep learning device trained to look for features based on a training set with high-level labels can return features (e.g., edges) specific to those labels. For instance, a machine trained to look for the contours of different types of teeth in intraoral images would detect edges as well as label edges and the regions they define based on which type of tooth they belong to. EX1002, ¶37.

3. Deep learning

Deep learning refers to methods that incorporate many feature-learning layers to model the often non-linear relationship between inputs (the training set) and outputs (probabilities of classes or instances). These models are referred to as “deep” because they have many “hidden” layers between input and output, with these layers potentially comprising millions of parameters to be adjusted through a training process using a training dataset, to produce accurate outputs for unseen test data. EX1002, ¶38.

By 2017, the best results in computer vision arose from learning models based on deep neural networks (DNNs), such as deep Convolutional Neural Networks (CNNs). EX1007, 2. For DNNs, hidden layers’ output may be viewed as

a feature map, with features learned through training. EX1002, ¶¶39-40. Earlier layers represent lower-level features, whereas deeper layers correspond to higher-level features. *Id.*; see EX1007, 3 (fine-scale features in early layers, large-scale features in later layers). In image-classification or object-recognition applications, the output feature map of the first layer corresponds almost to an edge map, while deeper feature maps correspond to higher-level features, such as object parts, shapes, and class. *Id.* By July 2017, DNNs had proven to be remarkably successful in computer vision and surpass human performance, due to their ability to extract multiple levels of features optimized by training with large training datasets. EX1002, ¶40; EX1007, 7-9 (“human” and “super-human” performance in various tasks).

B. Cited Art

1. Kim

U.S. Patent Publication No. 2017/0281313 (“Kim,” EX1005) is entitled “Method and System for Providing Remote Teeth Alignment Services and Computer-readable Medium on which the Method is Recorded.” Kim was filed on April 5, 2016, and published on October 5, 2017. Kim is prior art to the ’248 patent under 35 U.S.C. §102(a)(2).

Kim teaches methods and systems to provide “remote teeth alignment services.” EX1005, Abstract. Kim teaches methods by which an orthodontist can

be provided “accurate teeth alignment information using a device, e.g., a smartphone, of an orthodontic patient,” including information allowing the orthodontist to “diagnose the condition of the patient based on the teeth alignment information” transmitted as well as to “remotely provide information about an accurate step of a clear aligner to be currently worn by the orthodontic patient.” EX1005, ¶10. Kim refers to the patient’s smartphone as an “orthodontic patient terminal,” though Kim notes that a variety of other patient devices, such as other mobile devices, laptop computers, smart cameras or camcorders, and similar devices may also be used. EX1005, ¶¶41-42; EX1002, ¶¶44-45.

As part of Kim’s methods, teeth alignment information is obtained by taking a variety of photographs of the patient’s teeth from different angles. EX1005, ¶¶55-56. Additionally, Kim teaches that the patient’s smartphone can send “clear aligner information about a currently worn clear aligner” to the orthodontist. EX1005, ¶66. This information is transmitted “in such a manner that the orthodontist checks a photo or a video of the clear aligner currently worn by the orthodontic patient, and determines whether the clear aligner is normal and corresponds to an appropriate step.” *Id.*; EX1002, ¶46.

2. Invisalign

The Invisalign Clinical Monitoring Guide (“Invisalign,” EX1006) was published in 2003 through multiple channels including being available for

download from Align's Online Clinical Education Center, invisaligncec.com, as archived by the Internet Archive on September 26, 2003. EX1006, i; EX1011; EX1002, ¶47. Invisalign is prior art to the '248 patent under 35 U.S.C. §102(a)(1).

Invisalign teaches methods for monitoring progress of orthodontic treatment when using clear plastic aligners manufactured by Align Technology, Inc. EX1006, i. The guide deals with monitoring of treatment progress in patients after aligners have been received and are being worn by the patient. *Id.* Invisalign identifies key considerations for effective use of aligners, and describes the need to use high-quality records, including both dental impressions and photographs of the teeth of the patient wearing the aligners. EX1006, ii-iv; EX1002, ¶48.

Invisalign also provides specific examples of measurements to perform to track whether aligners are properly treating patients and discusses what those measurements indicate about whether the aligner is appropriate in terms of fit, seating, and other properties. *See, e.g.,* EX1006, 1-2. In one example, Invisalign illustrates how poor aligner seating may be assessed, illustrating how a digital photograph of an aligner on a patient's teeth can depict a poor fit by showing a gap of over 1 mm between teeth and aligner. *See* EX1006, 1; EX1002, ¶49.

3. Salah

PCT Publication No. WO2016066651 to Salah et al. ("Salah") is entitled "Method for Monitoring Dentition." Salah names five inventors, including all four

inventors listed on the '248 patent; however, Salah was published on May 6, 2016, more than 1 year before the earliest claimed priority date of the '248 patent. Salah is prior art under 35 U.S.C. §102(a)(1). As Salah was published in French (EX1012), a certified translation has been submitted as EX1008.

Salah discloses methods for monitoring patients' teeth, such as during orthodontic treatment. EX1008, Abstract, 1:12-22, 8:10-18. Salah teaches methods of obtaining 2D images of patient teeth, comparing those teeth to a 3D model, and using those comparisons to update the model and measure treatment progress. EX1008, 1:26-2:19, 46:15-19, Fig. 1. Salah's analysis can include using contour detection on an acquired image to aid mapping to a reference model, and Salah discloses that this contour detection can be performed using deep learning. EX1008, 34:11-14, 36:14.¹ These 2D images may be obtained by a patient or relative using a device such as a cell phone, which may also perform the evaluation process. EX1008, 2:21-27, 31:4-7; EX1002, ¶¶50-52.

¹ Salah uses the term "deepleraning," which a POSA would have recognized to be a typographical error meant to refer to "deep learning." *Id.*; EX1002, ¶96.

4. Maninis

Kevis-Kokitsi Maninis et al. published “Convolutional Oriented Boundaries: From Image Segmentation to High-Level Tasks” (“Maninis,” EX1007) no later than May 2, 2017, on arXiv.org. *See* EX1002, ¶¶53-58; EX1007, 1; EX1022 (arXiv.org publication policies); EX1023 (document unchanged); EX1026 (same); EX1024 (dissemination by authors); EX1025 (direct download link); EX1021 (earlier version with same material disclosure). Maninis is prior art under 35 U.S.C. §102(a)(1).

Maninis discloses methods for image classification and contour detection using a Convolutional Neural Network (CNN). EX1007, Abstract. Maninis teaches a “deep learning architecture” that produces significant improvements over state-of-the-art techniques. EX1007, 3, 8-9. Maninis uses a neural network trained on precisely annotated images to detect contours defining regions containing objects based on detected probabilities for contour locations. EX1007, 3-4, 5-6, 13, Fig. 2. Maninis also demonstrates the ability to label the specific object contained within each identified region, again improving over the state of the art. EX1007, 11-12, Tables 4-7, Fig. 16; EX1002, ¶59.

C. The '248 Patent

The '248 patent is entitled “Method for Analyzing an Image of a Dental Arch.” EX1001, Title. The '248 patent describes acquiring 2D images of aligners

worn on the teeth of patients. EX1001, Abstract, 28:13-15. The image can be acquired whenever aligner shape needs determination, such as over one week after the start of treatment. EX1001, 30:27-30.

The two independent claims, claims 1 and 15, recite acquiring at least one such image (an “analysis image”) showing the aligner worn in a service position on a patient’s teeth more than 1 week after the start of treatment. EX1001, 32:6-23, 34:43-57. A deep learning device is then used to analyze the image to determine a value for either a tooth attribute of an “analysis tooth zone” at least partially representing a tooth (for claim 1) or an image attribute of the analysis image (for claim 15). *Id.* The image or tooth attribute relates to a separation between a tooth and the aligner on the analysis image. *Id.* Claim 1 requires that a cellphone is used to acquire the analysis image, whereas claim 15 does not limit what device acquires the analysis image. *Id.*; EX1002, ¶¶41-42.

As to the remaining challenged claims, none adds anything beyond steps conventionally known in the art. EX1002, ¶43. Dependent claims 2 and 16 recite that the aligner’s suitability is assessed based on the tooth or image attribute. EX1001, 32:24-27, 34:58-61. Dependent claim 4 recites sending an information message as a function of the assessment in claim 2. EX1001, 32:33-36. Dependent claims 7 and 8 recite limitations regarding training of the deep learning device with images showing tooth or image attributes. EX1001, 32:45-33:15. Claim 14 recites

limitations similar to claim 7, and it further specifies that probabilities are determined relating to the presence of an analysis tooth zone and a tooth attribute for the corresponding tooth. EX1001, 34:21-42. Finally, claim 13 recites that the deep learning device is a neural network. EX1001, 34:19-20.

D. Prosecution History

The '248 patent was filed as U.S. Application No. 16/030,137 (“the '137 application”) on July 9, 2018, claiming priority to French Patent Application No. 1756956, filed July 21, 2017.

During prosecution, the Examiner issued three office actions rejecting pending claims over the prior art. Each office action asserted anticipation or obviousness of the rejected claims over U.S. Patent Publication No. 2013/0230818 to Matov et al., either alone or in combination with one or more of U.S. Patent Publication No. 2015/0132708 to Kuo, U.S. Patent Publication No. 2017/0300207 to Wen et al., and U.S. Patent Publication No. 2017/0049311 to Borovinskih et al. EX1004, 811-24, 858-71, 905-26. Applicant responded by arguing that Matov and Wen failed to disclose acquiring an image of the aligner in a service position more than one week after the start of treatment because Matov dealt with design of the aligner prior to its manufacture, and Wen only taught acquiring images of teeth to create tooth models prior to treatment, and rather than acquiring images of an aligner. EX1004, 947-51; *see also id.*, 842-45, 891-94 (similar arguments in prior

responses). Applicant further argued that “even if the method of Wen may be performed after an aligner has been worn, there is no reason to acquire an image comprising the aligner.” EX1004, 950.

Subsequently, Applicants initiated an Examiner Interview and agreed to an Examiner’s Amendment. EX1004, 959, 962. The Examiner amended one of the independent claims (claim 1) to further recite that “a cellphone is used to acquire the analysis image,” but canceled a claim depending from the other independent claim (pending claim 14, issued as claim 15) that had recited this limitation without incorporating it into that other independent claim. EX1004, 959-60. In the reasons for allowance, the Examiner stated that

neither Matov nor Wen teach or suggest, among other things, “more than 1 week after the start of the treatment with the aligner, acquisition of at least one “image at least partially representing the aligner” in a service position in which it is “worn” by a patient, “and tooth attribute relating to a separation between the tooth represented by the analysis tooth zone, and the aligner represented on the analysis image, a cellphone is used to acquire the analysis image.”

These key features in combination with the other features of the claimed invention are neither taught nor suggested by (Matov and Wen) prior art of record.

EX1004, 960-61.

E. Level of Skill in the Art

As Dr. Foroosh explains, a person of ordinary skill in the art (POSA) by July 21, 2017, would typically have an advanced degree, such as a Masters' Degree or Ph.D. in computer science, computer engineering, electrical engineering, or a related field, with at least two years of work in computer vision, image analysis, medical imaging, or related areas. A POSA could have less education but significant professional experience in one or more of these fields. A POSA could be a member of an interdisciplinary team including persons with backgrounds such as electrical engineering, computer science, optical imaging, and dentistry or orthodontics, and could call upon the knowledge of other team members as appropriate. *See* EX1002, ¶¶60-61.

II. STATEMENT OF THE PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED

Align requests review of claims 1-2, 4, 7-8, and 13-16 of the '248 patent under 35 U.S.C. §311 and AIA §6. The grounds for relief are as follows:

Ground	Claims	Description
1	1-2, 4, 7-8, and 13-16	Unpatentable under 35 U.S.C. §103 as obvious over Kim in view of Invisalign, Salah, and Maninis

III. CLAIM CONSTRUCTION

Claim terms should be given their ordinary and customary meaning as would be understood by a POSA in light of the specification. 37 C.F.R.

§42.100(b); *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312-13 (Fed. Cir. 2005) (en banc). Unless otherwise stated, this petition applies the ordinary and customary meaning of the claim terms. The following terms warrant discussion.

A. tooth attribute

Claim 1 recites “tooth attribute relating to a separation between the tooth represented by the analysis tooth zone, and the aligner represented on the analysis image.” The ’248 patent distinguishes image attributes and tooth attributes, with image attributes relating to an image as a whole and tooth attribute relating to a tooth. *See* EX1001, 1:22-28, 3:23-37. A tooth attribute may nevertheless be used to define an image attribute. EX1001, 3:59-62. For example, a tooth attribute can be a mean separation of a single tooth from an appliance, while an image attribute can be a mean separation of all teeth from the appliance. EX1001, 28:48-55. The ’248 patent describes a wide variety of example categories of tooth attribute. *See* EX1001, 7:64-8:42; *see also id.*, claim 6. Thus, on its own, a “tooth attribute” includes any attribute related to a tooth. EX1002, ¶¶62-64.

Regarding the recited separation between tooth and aligner, the ’248 patent states that the tooth attribute may be “chosen from the group formed by a maximal

separation along the free edge of the tooth [or] a mean separation along the free edge of the tooth,” but it indicates that these are preferred options, not an exhaustive list. EX1001, 28:48-55. Dependent claim 6 confirms that, rather than being limited to a measure of separation, the tooth attribute may be selected among many values only indirectly relating to the recited separation. Accordingly, the “tooth attribute” recited in claim 1 includes any attribute of a tooth that may relate to the tooth’s separation from an aligner, including attributes specific to the tooth or to the aligner, such as their respective positions or edges, which may bear a relation to the separation between the tooth and the aligner. EX1002, ¶65.

B. image attribute

Claims 8 and 15 recite limitations relating to an “image attribute.” The ’248 patent describes an image attribute as an attribute that has a value that “relates to the whole image and not to part of the image.” EX1001, 3:23-37. The ’248 patent lists various examples of what an image attribute may relate to. *See* EX1001, 3:38-49; *see also id.*, 26:3-14 (image attribute value may be list of values for multiple teeth). These various examples indicate that an image attribute includes an attribute having a value (which may be a set of values for multiple teeth) relating to an image as a whole. EX1002, ¶66.

However, claims 8 and 15 recite that the image attribute relates to a separation between one or more teeth and the aligner represented on either a

historical image (for claim 8) or an analysis image (for claim 15). In some examples, the '248 patent states that the image attribute may be “chosen from the group formed by a maximal separation along all of the teeth represented, a mean separation along the free edges of all of the teeth represented, [and] an overall acceptability of the separation of the teeth represented,” but it indicates that these are preferred options, not an exhaustive list. EX1001, 28:48-55. The '248 patent also teaches that an image attribute may be “determined as a function of the tooth attribute values.” EX1001, 26:3-14. An image attribute may be a single value or a list of values for multiple teeth, such as a “list of the tooth numbers represented” in an image. *Id.* An image attribute may also indicate “the ‘presence’ or ‘absence’ of a dental, preferably orthodontic, appliance.” EX1001, 26:11-14, *see also id.*, 29:30-53. In light of this disclosure, the “image attribute” recited in claims 8 and 15 includes any attribute of the image as a whole that relates to the separation between one or more teeth and the aligner. EX1002, ¶67. These attributes would include, for example, an indication of whether an aligner or gap was present in the image, a measure of individual or average separation an indication of whether the overall separation is acceptable, or a measure of another value that may be used to assess separation of teeth from the aligner. *Id.*

C. image extracted from a film

Claims 1 and 15 require an analysis image that is “a photograph, or an image extracted from a film,” and claim 1 also specifies that the analysis image is acquired with a cellphone. The specification states that a patient “preferably takes a photo or video of his or her teeth, preferably by means of a cellphone equipped with a camera,” and elsewhere it states that the “analysis image is preferably a photo or an image from a film, preferably with a cellphone.” EX1001, 19:41-44, 25:54-57. In light of this parallel disclosure and the claims’ requirement that the analysis image be acquired with a cellphone, the phrase “image extracted from a film” should be interpreted to refer to an image extracted from a video recording. EX1002, ¶68.

D. Steps 1(b'') and 15(b'')

The second step of claims 1 and 15 recite “analysis of the analysis image by means of a deep learning device, trained by means of a learning base, so as to determine a value” for either a tooth attribute (claim 1) or an image attribute (claim 15). Align intends to argue in district court that this is a means-plus-function limitation for which the specification does not provide sufficient structure to discern the claim’s full metes and bounds. Even if the Board agrees, that does not preclude an analysis that the claims fail to distinguish from the prior art.

Based on Patent Owner’s assertion of this claim against Align, Align understands Patent Owner does not view the term as indefinite. To date, Patent Owner has not identified whether it interprets this claim element as a means-plus-function element.² However, regardless of whether this step is construed as a means-plus-function limitation, neither the claim language nor the relevant disclosure in the specification distinguishes over the prior art, as explained below for claim 1. For example, the specification repeatedly provides a generic description of “the deep learning device [as] preferably a neural network,” which is met by the deep learning neural network of Maninis. EX1001, 1:24-25; *see* EX1007, 2-3.

For the purposes of this petition, Align identifies the recited function as analyzing an image and the structure to encompass a device comprising a deep learning neural network trained with a learning base. *See* EX1001, 1:24-25, 1:39-40, 1:51-2:3, 3:14-18, 4:5-6, 8:28-40, 13:53-58, 24:3-11, 24:34-35, 24:54-55,

² As discussed below in §V, claim construction briefing has not yet been scheduled in district court. Align will keep the Board apprised of any constructions for this term proposed by the parties or adopted by the district court.

28:17-28, 29:14-15, 29:45-46; *see also id.*, 16:4-43 (non-limiting examples of neural networks).

IV. CLAIMS 1-2, 4, 7-8, AND 13-16 ARE UNPATENTABLE AS OBVIOUS UNDER 35 U.S.C. §103 OVER KIM, INVISALIGN, SALAH, AND MANINIS

A. Overview of Ground and Rationale to Combine with Reasonable Expectation of Success

As discussed in detail below, Kim in view of Invisalign, Salah, and Maninis renders claims 1-2, 4, 7-8, and 13-16 obvious.

Kim discloses methods for providing remote tooth alignment services, including evaluation of a clear aligner to be worn by a patient and determination of, for example, “when a new or modified clear aligner is necessary.” EX1005, Abstract, ¶10. These services involve the use of an “orthodontic patient terminal,” such as a patient’s cellphone. EX1005, Abstract, ¶¶1, 10. 42. To facilitate these services, photographic or video images may be transmitted to an orthodontist to provide information about teeth alignment and the patient’s clear aligner. EX1005, ¶12. Kim also teaches that automatic recognition of contours of a patient’s face and teeth may be employed as part of the imaging process. EX1005, ¶14. In one example, an “orthodontist checks a photo or a video of the clear aligner currently worn by the orthodontic patient, and determines whether the clear aligner is normal and corresponds to an appropriate step.” EX1005, ¶66. Kim teaches that its techniques provide reduced diagnosis time and cost relative to traditional in-office

visits, as well as providing for enhanced accuracy. *See* EX1005, ¶¶8-10, 72; EX1002, ¶69-70.

Invisalign corroborates Kim’s teachings regarding imaging an aligner in a service position on teeth, showing how such an aligner may be assessed to determine if it fits appropriately. Moreover, to the extent Patent Owner may argue Kim does not sufficiently describe the aligner currently worn and imaged is in a service position on the patient’s teeth, taking such an image would have been readily apparent in view of Invisalign’s teachings. Invisalign teaches how to “successfully monitor Invisalign treatments and address issues that may arise between receipt of Aligners from Align to the completion of treatment.” EX1006, i. The first criterion identified by Invisalign addresses “General fit issues,” specifying that if an aligner is well seated, there will be a distance of less than 1 mm between the end of the teeth and the surface of the aligner. EX1006, 1. This is illustrated in a digital photograph, where a poorly-seated aligner worn by a patient in a service position is shown, with arrows indicating that the aligner surfaces corresponding to teeth have a separation of greater than 1 mm from the ends of the teeth. *See id.* The resulting gap is plainly visible on the photograph as a space between the contour defining the occlusal end of the teeth and the matching contour defining the occlusal surface of the clear aligner, as reproduced below:



EX1006, 1. Thus, Invisalign teaches that orthodontic assessment of an aligner may include assessing the separation between the edges of the aligner and the patient's teeth when the aligner is worn, as may be seen in a digital photograph showing the patient's teeth as well as the clear aligner being worn in a service position on the patient's teeth. EX1002, ¶71.

In view of Kim's disclosure of orthodontic checking of a photo or video of an aligner currently worn by the patient and Invisalign's teachings of how to evaluate aligner fit, illustrated by an image of an aligner worn on a patient's teeth, a POSA would have had reason to analyze images of the teeth of a patient wearing an aligner in a service position to evaluate whether the aligner fit appropriately. This combination would be applying known methods for orthodontic evaluation of

aligner fit based upon the appearance of the aligner worn on the patient's teeth, as taught by Invisalign, in order to analyze a photo or video of the aligner worn by the patient, as taught by Kim, thereby providing feedback regarding tooth-aligner separation to aid an orthodontist regarding whether the clear aligner is normal and corresponds to an appropriate step of treatment. As Kim indicates, employing such techniques would serve goals including enhancing accuracy, reducing the time and cost of treatment, and facilitating remote assessment. *See* EX1005, ¶¶8-10, 72. A POSA also would have understood that the type of image analysis suggested by Kim and Invisalign could be readily applied on images or videos showing a clear orthodontic aligner being worn on the teeth of a patient. EX1002, ¶72.

Not only are these types of images of aligners on teeth expressly depicted in Invisalign (*see* EX1006, 1), but the fact that POSAs were motivated to use such images for orthodontic evaluation was confirmed in publications by practicing orthodontists. For example, Dr. Scott Frey published a step-by-step guide on the Ortho Cosmos website explaining how to conduct a simple remote orthodontic monitoring system allowing orthodontists to analyze orthodontic patients' teeth and aligners over the internet. EX1009, 17-27. As explained in the guide, patients would be shown example photos or videos to show exactly what images they should capture, including one video "with teeth apart and aligners in." EX1009, 19. Patients would then record and transmit such videos with their smartphone,

providing orthodontists with video images sufficient to conduct typical orthodontic treatment evaluations. *See* EX1009, 20-27. Thus, published descriptions show that photographs or videos of the teeth of patients wearing aligners, including images taken with cellphones, were known to be useful to analyze for orthodontic assessment purposes. EX1002, ¶73.

A POSA would have known how to determine the width of such gaps using image processing software. Like Kim, Salah teaches application of computer image analysis to dental images for orthodontic treatment purposes. For example, Salah discloses methods useful for monitoring patient's dentitions, including by acquiring images of the patient's teeth at various times during orthodontic treatment. EX1008, Abstract, 8:19-28. The images are acquired by an acquisition apparatus such as a mobile phone. EX1008, 13:24-28, 31:4-9. Further paralleling Kim, Salah teaches that images may be analyzed to "measure the effectiveness of orthodontic treatment," including by "remotely monitor[ing] the progress of an orthodontic treatment." EX1008, 46:15-24. Salah discloses that analysis of acquired images may include determining contours of objects in an image, including by using deep learning for contour detection. EX1008, 34:5-36:14; EX1002, ¶74.

To the extent that Salah does not expressly teach using a deep learning device trained by means of a learning base for image analysis, this would have

been obvious to a POSA and is taught by Maninis. Although Maninis is not restricted to dental imaging, Maninis describes broadly applicable techniques for image classification and contour detection using a deep learning device trained using a learning base. EX1002, ¶¶75-76; *see* EX1007, Abstract (“Convolutional Oriented Boundaries (COB)” technique using “Convolutional Neural Networks (CNNs)” to detect boundaries), 1-2, 5-6, 11-12 (images divided into regions with objects classified within). Consistent with Maninis, Salah expressly teaches using a variety of known methods for contour detection, including deep learning. EX1008, 34:5-36:14; EX1002, ¶76. Maninis motivates employing deep learning techniques for object and contour detection. Image analysis trends favored adoption of deep learning techniques like Maninis’s. EX1002, ¶76. For example, Maninis teaches that the “latest wave of contour detectors takes advantage of deep learning to obtain state-of-the-art results.” EX1007, 2. Further, Maninis provides a flexible, “modular” technique that “can incorporate and benefit from future improvements in the base CNN.” *Id.* Maninis’ technique also achieves improvements relative to “recent approaches to contour detection using deep learning.” EX1007, 3, 9 (“Overall, comparing to previous state-of-the-art, we get a significant improvement at a fraction of the computation time (24.37 to 0.79 seconds).”), 11-12 (improved object identification); EX1002, ¶77.

Based on Maninis's flexibility and prospect to improve performance, a POSA would have been motivated to analyze images using a deep learning device as taught by Maninis to improve image acquisition and analysis apparatus as taught by Salah and Kim. For example, a POSA would have been motivated to adapt Salah and Kim's image acquisition and analysis apparatus to employ a deep learning architecture as taught by Maninis, trained on images relevant to the dental imaging applications of Salah, Kim, and Invisalign. Maninis's deep learning architecture would provide specific functionalities required by Salah, Kim, and Invisalign, including the ability to identify objects (such as the teeth and aligners depicted in each image) and to detect contours (such as the edges of teeth and aligners useful to assess aligner fit as taught by Invisalign). EX1002, ¶78.

Further, it would have been obvious to measure the separation between the ends of the patient's teeth and the corresponding edges of the aligner, as taught by Invisalign, using the deep learning techniques taught by Salah and Maninis to detect the contours of each object in an image depicting the patient wearing the aligner on the patient's teeth, and thereby identifying the size of the gap between aligners and teeth. A POSA would have had reason to use the techniques of Salah and Maninis for contour detection to detect the edges of teeth and aligners, as Invisalign teaches that the gap between the surfaces of teeth and an aligner should be below a threshold value (1 mm), and as Dr. Foroosh explains, a POSA would

have understood that the width of a gap between two objects in a photograph is equivalent to the separation between the two contours defining the objects' edges. In the case of an image of an aligner on teeth, the distances to measure would be the distances from the contour defining the end of the teeth to the contour defining the surface of the clear plastic aligner. EX1002, ¶¶71, 79.

The resulting combination would use each reference's teachings to perform substantially the same function as they provide alone, applying the contour detection taught by Salah and Maninis to detect edges of teeth and aligners to produce a measurement of distance between these surfaces, as taught by Invisalign, to assess whether the imaged aligner fits appropriately for the current step in patient treatment, as taught by Kim. POSAs would thus have had a reasonable expectation of success, as the combination would be one of complementary techniques, applied in understood ways, to produce a predictable outcome. EX1002, ¶¶80-81.

The discussion of claims 1-2, 4, 7-8, and 13-16 below provides further detail regarding POSAs' rationale to combine Kim, Invisalign, Salah, and Maninis with a reasonable expectation of success.

B. Claim 1

Kim, Invisalign, Salah, and Maninis teach each element of claim 1, as discussed below. EX1002, ¶¶82-109.

1. Preamble

1. A method for assessing the shape of an orthodontic aligner, said method comprising the following steps:

To the extent the preamble is limiting, Kim and Invisalign teach such a method. Kim teaches evaluating an aligner by reviewing an image of the aligner currently worn by the patient, to determine whether the aligner is normal and appropriate. EX1005, ¶66. Invisalign teaches that evaluating a currently-worn clear aligner includes assessing the fit of the aligner, and that this assessment involves determining a distance between the teeth and the aligner when worn by the patient, as depicted in photographs showing good and poor aligner seating. EX1006, 1. Kim and Invisalign also teach that the evaluation can determine that an aligner does not fit appropriately (e.g., because the aligner may have been formed using a distorted impression of teeth). *Id.*; EX1005, ¶¶66-67; EX1002, ¶¶83-85.

2. Step 1(a'')

a'') more than 1 week after the start of the treatment with the aligner, acquisition of at least one image at least partially representing the aligner in a service position in which it is worn by a patient, called “analysis image”, the analysis image being a photograph, or an image extracted from a film; ...

in the step a''), a cellphone is used to acquire the analysis image.

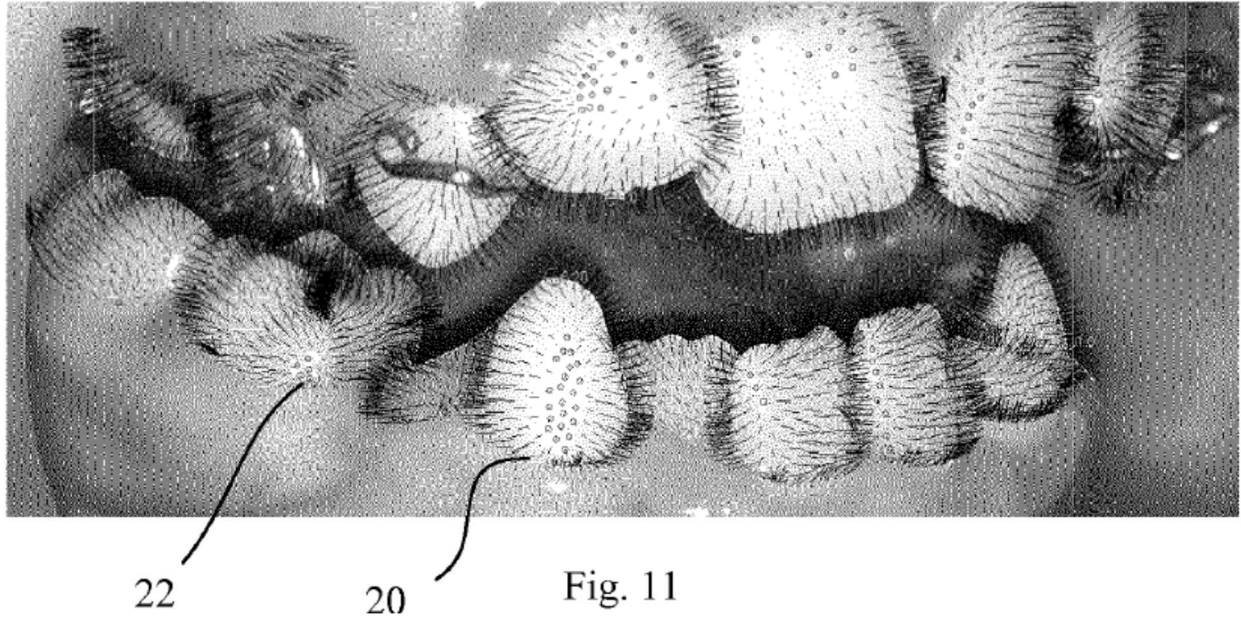
Step 1(a'') requires acquiring at least one image (the "analysis image") showing a patient's teeth and an aligner worn by the patient in a service position. The image must be a photograph or image extracted from film, and it must be acquired more than 1 week after starting treatment with the aligner. Claim 1 also specifies that in step (a'') a cellphone is used to acquire the analysis image. This step would have been obvious over Kim, Invisalign, and Salah. EX1002, ¶86.

Kim teaches that a patient's smartphone or similar device may be used to acquire orthodontic treatment information. *See* EX1005, Abstract, ¶¶1, 10, 42. This can include "using a camera" of the device to acquire photographs "showing a teeth alignment state of the orthodontic patient." EX1005, ¶¶55-56. This can further include aligner information, and in particular may include "receiving the clear aligner information about a currently worn clear aligner," enabling the orthodontist to "check[] a photo or a video of the clear aligner currently worn by the orthodontic patient," and in particular to "determine[] whether the clear aligner is normal and corresponds to an appropriate step" EX1005, ¶66. Thus, Kim teaches, or at minimum suggests, using a cellphone to acquire an analysis image of a patient's teeth wearing an aligner as recited in step 1(a''). EX1002, ¶87.

Invisalign confirms that analyzing an image of an aligner in a service position on the patient's teeth (e.g., a photo or video taken by a cellphone, consistent with the teachings of Kim) can produce measurements useful for

orthodontic assessment. Invisalign teaches that a large gap between a tooth and an aligner (e.g., more than 1 mm) indicates poor fit. EX1006, 1. Invisalign further illustrates that such a gap is visually distinguishable in a digital photograph of a patient's teeth that at least partially represents an aligner worn in a service position on the teeth. *See id.* Accordingly, Invisalign provides a reason to acquire an image of the patient's teeth with an aligner being worn in a service position: by assessing the gap visible between aligner and teeth, Invisalign teaches that it can be determined whether an aligner fits correctly (e.g., is normal and corresponds to an appropriate step), or whether it is seated or fit poorly. EX1002, ¶88.

Salah similarly teaches acquiring dental images with a cellphone, including of teeth fitted with orthodontic appliances. For example, Salah teaches that a user (such as a patient or relative) may acquire images of a patient's teeth using an image acquisition apparatus such as a cellphone with a camera. EX1008, 27:11-13, 31:4-7. Salah also teaches that the image of the teeth may include an orthodontic appliance; for example, Fig. 11 (below) depicts a reference map generated from an image of teeth of a patient wearing an orthodontic appliance. EX1008, 64:7-12, Fig. 11 (showing teeth with orthodontic wires visible). Thus, Salah also teaches acquiring images of a patient's teeth, including teeth of a patient wearing orthodontic appliances. EX1002, ¶89.

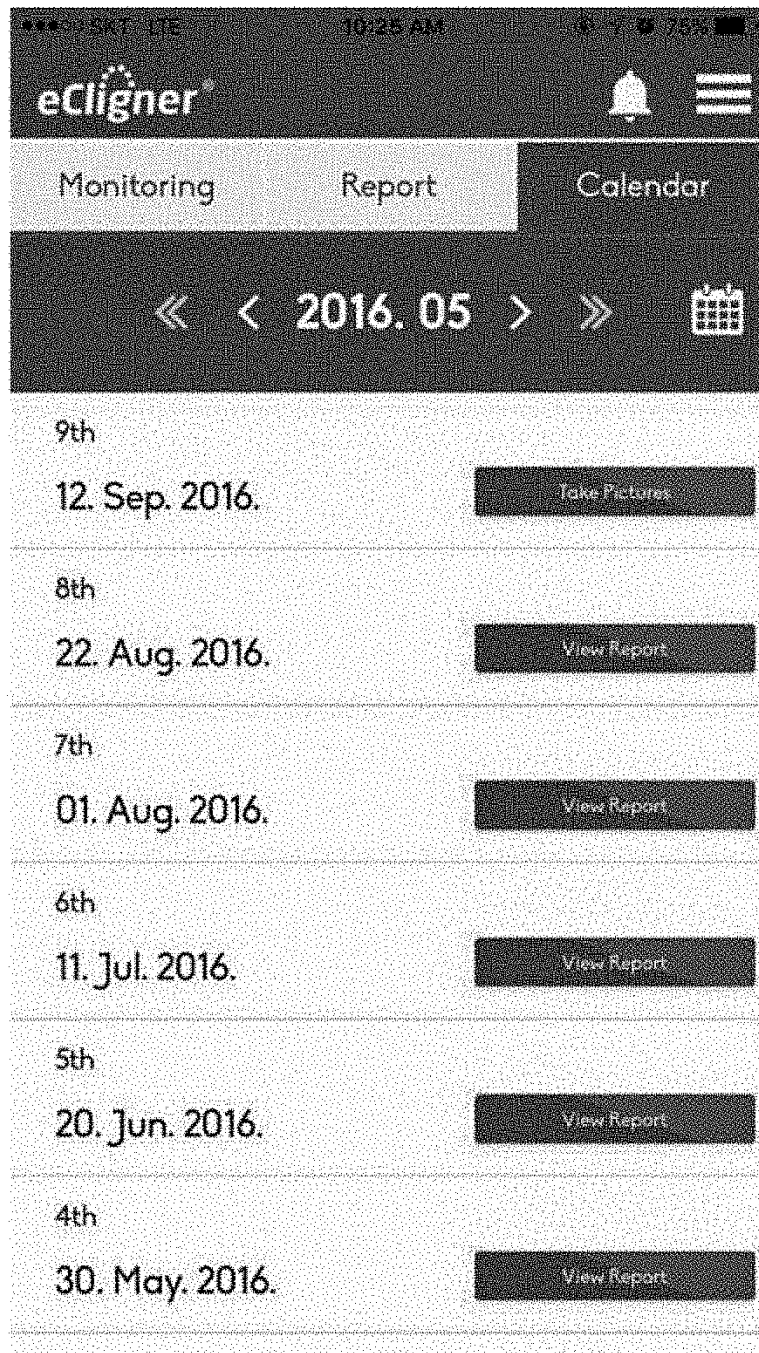


EX1008, Fig. 11.

Use of cellphones to acquire images depicting teeth wearing an aligner is also consistent with published methods of orthodontic practice. For example, a step-by-step guide for remote orthodontic monitoring published prior to the '248 patent's earliest claimed priority date described how to instruct patients to acquire images for use in assessment of orthodontic treatment status, including taking images of "teeth apart and aligners out," "biting down with aligners removed," and "with teeth apart and aligners in." EX1009, 19. This confirms that acquiring digital photographs of patient's teeth, including patient-acquired cellphone images of teeth with aligners in a service position, was known to be a useful practice for assessment of a patient's orthodontic treatment status. EX1002, ¶90.

Finally, Kim and Invisalign teach acquiring the image at least one week after the start of treatment with the aligner. Invisalign teaches evaluating “issues that may arise between receipt of Aligners from Align to the completion of treatment.” EX1006, i. These treatments naturally take course over many weeks, and Invisalign teaches that treatment is assessed over such periods. For example, when discussing assessing fit by comparing to a 1 mm threshold for separation between teeth and aligners, Invisalign teaches that the “[a]ligners should be fitting well at the end of the 2-week period.” EX1006, 1. Invisalign thus teaches assessing aligner fit (e.g., by acquiring images) more than one week after starting aligner treatment. EX1002, ¶91. Kim agrees, showing that orthodontic images may be taken recurrently based on calendared appointments. Such a schedule is shown in Fig. 15, showing a calendar with images recurrently scheduled to be taken over the course of months of treatment.

FIG. 15



EX1005, Fig. 15; *see also id.*, ¶83 (“For example, the orthodontic patient may view teeth alignment results per date as illustrated in FIG. 15....”). Thus, both Invisalign

and Kim teach assessing a patient's orthodontic treatment status, including by acquiring cellphone images of a patient's teeth wearing aligners, more than one week after starting treatment with the aligner. EX1002, ¶92.

Accordingly, Kim, Invisalign, and Salah teach step 1(a"), including claim 1's requirement that in step 1(a"), "a cellphone is used to acquire the analysis image."

3. Step 1(b")

b") analysis of the analysis image by means of a deep learning device, trained by means of a learning base, so as to determine a value

for at least one tooth attribute of an "analysis tooth zone" representing, at least partially, a tooth on said analysis image, the tooth attribute relating to a separation between the tooth represented by the analysis tooth zone, and the aligner represented on the analysis image

Step 1(b") requires using a trained deep learning device to determine a tooth attribute value for an "analysis tooth zone" representing a tooth, with the tooth attribute relating to a separation between the tooth and aligner. Kim, Invisalign, Salah, and Maninis teach step 1(b"). EX1002, ¶¶93-94.

Salah teaches that dental images may be analyzed to detect "discriminant information," which can include contours of objects depicted in the images, including teeth and other intraoral objects, such as gums. EX1008, 34:5-14, 55:21-

56:7. A contour map of detected tooth edges showing the resulting outlines of teeth is shown in Fig. 6b, reproduced below.

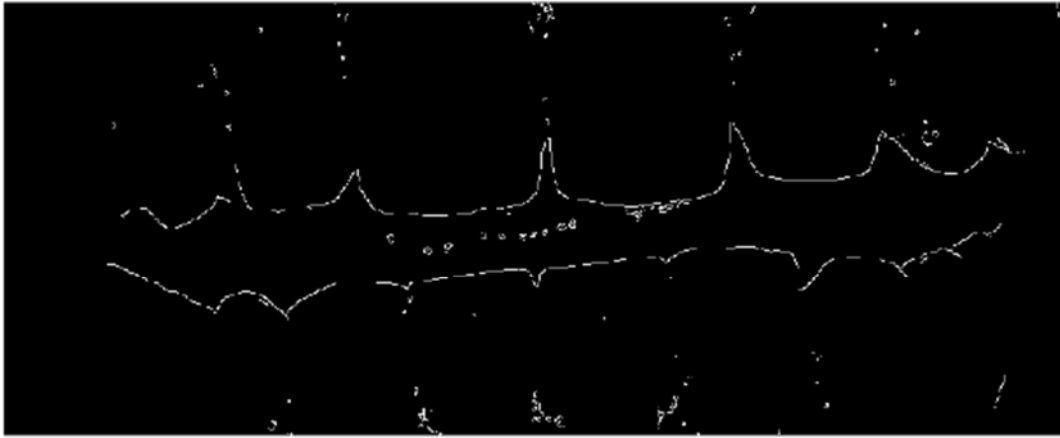


Fig. 6b

EX1008, 34:5-14, Fig. 6b. This discriminant information can be used to generate a “reference map representing the discriminant information,” and Salah provides an example of such a map generated for an image depicting teeth of a patient with an intraoral appliance. EX1008, 40:19-21, 64:7-8; *see id.*, Fig. 11; EX1002, ¶95.

Salah expressly discloses that the contours it describes may be detected using “image processing software,” which “can be used, for example, to detect high contrast regions and determine contours.” EX1008, 34:11-14. Salah further describes a list of such software, including both preferred options and others that can also be used. EX1008, 34:15-36:14. Among these options, Salah teaches that contours may be detected using deep learning. EX1008, 36:14 (“Deepleraning [*sic*]”); EX1002, ¶96.

Applicable deep learning techniques for contour and object detection are taught by Maninis. Maninis teaches methods of detecting edges (or “contours”) in images with a computing device using a deep learning neural network to perform image analysis, using a process referred to as “Convolutional Oriented Boundaries (COB).” EX1007, Abstract. Maninis’s “contour detection” produces “multiscale oriented contours and region hierarchies” using “Convolutional Neural Networks (CNNs).” *Id.* Maninis describes improvements on “recent approaches to contour detection using deep learning,” presenting a “CNN architecture to produce contour detection at different scales, ... all in a single CNN forward pass at the image level.” EX1007, 3-5 (“contour probability map” used to detect region borders based on probability strength), Figs. 3-5 (showing results); EX1002, ¶97.

Maninis’s “deep learning architecture,” illustrated below, shows generation of contours using multiple feature map layers in a CNN to detect image contents at different scales (fine and coarse detections) as well as different orientations.

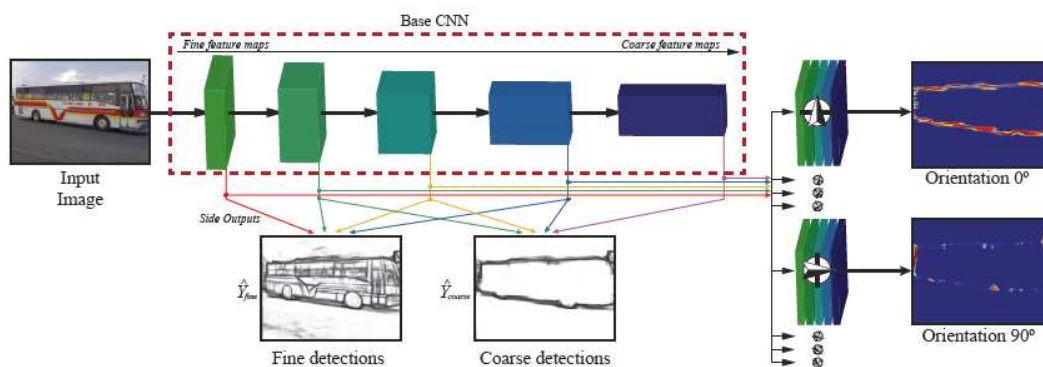


Fig. 2. Our deep learning architecture (best viewed in color). The connections show the different stages that are used to generate the multiscale contours. Orientations further require additional convolutional layers in multiple stages of the network.

EX1007, Fig. 2; EX1002, ¶98.

Maninis's use of a computing device comprising a trained deep-learning convolutional neural network to analyze images matches the '248 patent's preferred embodiment. *E.g.*, EX1001, 1:24-25; EX1007, 3-5, 9; *see supra* §III.C; EX1002, ¶99. *Compare* EX1001, 16:9-36 (listing "Microsoft ResNet" and "VGG Net" as non-limiting embodiments), *with* EX1007, 2 ("[W]e present results both using VGGNet and ResNet....").

Maninis's deep learning device is trained on a "training dataset" (i.e., a "learning base" as recited in the '248 patent). EX1007, 3; EX1002, ¶100. In one example, Maninis uses a training dataset 10,103 images with "localized pixel-wise semantic annotations," meaning pixel-by-pixel indications of each object's location in each image. EX1007, 5-6; EX1002, ¶100. For training and optimization purposes, Maninis divides these images into "train, validation, and test sets" to train and validate the deep learning device, substantially improving over prior classification techniques. *See* EX1007, 5-6, Table 3, Fig. 6; EX1002, ¶100.

Maninis also uses the contours it generates to define boundaries around regions and to identify the objects in each region of the image. *See* EX1007, Abstract, 11-12. Maninis describes such image recognition tasks as involving "image classification, object detection, and semantic segmentation" where detected contours are assigned to categories matching the "ground-truth" (i.e., identification of what actual objects are where in an image). EX1007, 1-2. Objects are identified

in regions by “label[ing] each of the detected pixels with the corresponding semantic class.” EX1007, 11-12; *see id.*, Tables 4-12 (identifying specific object types); EX1002, ¶101.

In view of Maninis’s teaching that its deep-learning-based contour detection techniques identify and classify objects in an image, a POSA would have been motivated to use those techniques to identify teeth and aligners depicted in orthodontic images by detecting their contours and identifying the objects found within each contour. Adjusting Maninis’s deep learning device to perform the specific task of identifying teeth and orthodontic aligners would have been straightforward and well within the skill of the POSA. EX1002, ¶102. These alterations would include providing an appropriately annotated training set and setting output layers to provide the required object identifications corresponding to the labels used to identify the locations of teeth, aligners, and the gaps between them. *Id.*

As Dr. Foroosh explains, by using deep-learning based image processing software to detect contours in an image representing a patient’s teeth with an aligner worn in a service position on the teeth, the edges of the objects depicted would be detected. EX1002, ¶¶102-03. For example, in an image of an aligner and teeth, such as the one reproduced below from Invisalign, boundaries corresponding to the edges of the teeth and the aligner (among other objects depicted in the

image) would be accurately detected, and the regions defined by these boundaries would be identified by what was found within each region, such as a particular tooth, aligner surface, or gap between them.



EX1006, 1; EX1002, ¶103; *see also* EX1007, Figs. 9-10 (showing actual object boundaries detected by Maninis's technique). As Dr. Foroosh explains, these boundaries match the contours a person notices naturally when viewing the image, just as the gap between the contours corresponds to the gap naturally perceived by a person looking at the image (or when viewing teeth and an aligner in real life). EX1002, ¶103.

As explained above in §IV.A, a POSA would have been motivated in view of Kim and Invisalign to apply the contour-detection techniques of Salah and

Maninis to detect edges of teeth and aligners depicted in images of patient's teeth so that the assessment of gap between aligner and teeth could be made, as taught by Invisalign. For example, because the detected edges of the aligner and teeth correspond to the edges of the physical objects depicted, a POSA would have understood that measuring the distance between the outlines of the teeth and the aligner as detected by Salah's image processing software would allow determination of the gap between each tooth and the corresponding part of the aligner. EX1002, ¶104.

Analyzing the image of a patient's teeth and aligner using the deep-learning contour detection taught by Maninis and Salah would constitute "analysis of the analysis image by means of a deep learning device," as Maninis teaches analysis with a trained deep-learning device. EX1002, ¶105.

Further, using the detected tooth and aligner edges to determine the separation between one or more teeth and the aligner would constitute using the deep-learning device to "determine a value for at least one tooth attribute of an 'analysis tooth zone' representing, at least partially, a tooth on said analysis image, the tooth attribute relating to a separation between the tooth represented by the analysis tooth zone, and the aligner represented on the analysis image." Maninis's contour detection detects hierarchies of regions containing objects. EX1007, Abstract, 3-5. Such a region containing a tooth (such as a region containing a tooth

alone or a region containing a tooth and an aligner surface separated by a gap) would constitute an analysis tooth zone at least partially representing a tooth on the analysis image. EX1002, ¶106. Additionally, this analysis would generate tooth attributes, including the locations of the surfaces of the tooth and aligner (corresponding to detected contours), as well as the measured separation between them (the distance between the contours), consistent with the teachings of Invisalign. EX1002, ¶107 These tooth attributes would each be a tooth attribute related to the separation between the represented tooth and aligner. *Id.*

Thus, Kim, Invisalign, Salah, and Maninis teach step 1(b"). EX1002, ¶108.

Accordingly, claim 1 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶¶82-109.

C. Claim 2

Claim 2 depends from claim 1 and further recites that the method comprises "after the step b"), step c") as follows: c") assessment of the suitability of the aligner as a function of the value of said tooth attribute." The prior art teaches this step. EX1002, ¶110.

As discussed for claim 1, Invisalign uses a threshold to assess aligner acceptability. For example, Invisalign teaches that when an aligner worn on a patient's teeth has a tooth-aligner gap of less than 1 mm, "this is natural deflection that should occur so the Aligner can move the tooth," whereas if the gap is greater

than 1 mm, this means that the “Aligners are not seating well.” EX1006, 1. In view of this teaching, a POSA would have been motivated to compare the values discussed above for claim 1 (that is, the values representing the surfaces of the patient’s teeth, the corresponding aligner surfaces, and the separation between them) to apply the acceptability threshold taught by Invisalign, thereby determining that where the threshold is exceeded, the aligner is not seated well on the teeth, or conversely that where the threshold is not exceeded, the aligner is acceptably seated. EX1002, ¶111. This assessment would be consistent with Kim’s teaching to “check[] a photo or a video of the clear aligner currently worn by the orthodontic patient, and determine[] whether the clear aligner is normal and corresponds to an appropriate step.” EX1005, ¶66. For example, if applying Invisalign’s threshold assessment indicated that one or more teeth had excessive gaps, that would be an indication that the aligner was not normal or did not correspond to an appropriate step. EX1002, ¶¶112-13; *see* EX1006, 1 (evaluation may indicate aligner has “excessive flash” or results from a “distorted impression” (meaning it was not normal)), 6 (lack of fit may indicate incorrect aligners or need for mid-course correction (meaning the aligner did not correspond to an appropriate step)); EX1008, 38:7-10 (registration marks with known size and shape, which a POSA would have known could be used to scale image).

Accordingly, claim 2 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶114.

D. Claim 4

Claim 4 depends from claim 2, and further recites that the method comprises “after the step c”), a step d”) as follows: d”) sending of an information message as a function of said assessment.” Claim 4 would have been obvious in view of Kim, Invisalign, Salah, and Maninis. EX1002, ¶115.

A POSA would have found it obvious to send an information message as a function of the assessment recited in claim 2 (i.e., an assessment of the aligner’s suitability). EX1002, ¶116. For example, it would have been expected to send a message indicating that the aligner was or was not suitable based on the assessment, because of the readily apparent benefit of making the patient (or treating orthodontist) aware of the assessment. *Id.*

As one example, the assessment could be used to provide guidance to a patient about wearing the aligner. Invisalign teaches that poor fit (indicated by a tooth-aligner separation greater than 1 mm) may be due to “non-compliance,” i.e., the “patient not wearing Aligners long enough for [a] tooth to move.” EX1006, 1. In this case, the guide teaches that the patient should be told to extend weartime of the aligner, and that orthodontists should emphasize to the patient “never to dispose of their Aligners and instruct them that the Aligners should be fitting well

at the end of the 2-week period.” *Id.* Sending this instruction as an information message would constitute the claimed step d”. EX1002, ¶117. For example, this message could be sent to the patient’s mobile device. This option is taught by Kim, which teaches that a “teeth alignment guidance information transmission program” can send information from the orthodontist to the patient’s mobile device that “advises, for example, an appropriate step of a clear aligner to be currently worn by the orthodontic patient.” EX1005, ¶63; *see also id.*, ¶64 (“[T]he teeth alignment guidance information may include at least one of a clear aligner information about a clear aligner to be worn in a current step among a clear aligner set, a wearing time information about a wearing time of the clear aligner, a replacement information about replacement of a part of or the whole clear aligner set, a clear aligner state information about the state of the clear aligner, a clear aligner usage guidance information....”); EX1002, ¶118.

Accordingly, claim 4 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶119.

E. Claim 7

Claim 7 depends from claim 1 and further recites that step (b”) comprises four additional steps. Kim, Invisalign, Salah, and Maninis teach each of these steps, as discussed below. EX1002, ¶120.

1. Step 7(1)

1) creation of a learning base comprising more than 1000 images of dental arches, or “historical images”, each historical image representing an aligner worn by a “historical” patient and comprising one or more zones each representing a tooth, or “historical tooth zones”, to each of which, for at least one tooth attribute relating to a separation between the tooth represented by the historical tooth zone considered, and the aligner represented, a tooth attribute value is assigned;

Regarding the number of images in a learning base, a POSA would have understood that a large number of images, often far more than 1000, would be typically used to train a deep learning device. EX1002, ¶121-22. For example, Maninis discloses training its neural network using over 10,000 images. *See* EX1007, 1, 5. Thus, a POSA would have found it natural to train a deep learning device such as a neural network with more than 1000 images. EX1002, ¶122.

A POSA would also have used historical “images of dental arches” when adapting Maninis’s contour and object detection to the dental imaging apparatus of Kim and Salah. Maninis discloses using a training set with 459 semantic categories, corresponding to the objects depicted in the images. EX1007, 5; *see id.*, 11-12 (specific classification of objects); EX1002, ¶123. Training such a neural network teaches it to detect the categories of objects it has seen during training. EX1002, ¶123. Thus, to train a deep-learning-based neural network to recognize

teeth, aligners, and their respective edges that delimit the gap between them, a POSA would have used images of patients' teeth wearing aligners in the training set. *Id.* To obtain optimal results, a POSA would have used images representative of the types of images likely to be seen during ordinary use of the imaging device, such as real historical photographs of the teeth of a patient wearing the aligner, as taken with the same device (i.e., a cellphone). *Id.*

Finally, a POSA would have included zones representing teeth to which tooth attributes were assigned. Maninis teaches a learning base should include pixel-wise annotations indicating the regions for each object and labeling the types of objects (i.e., objects' attributes) that should be identified in that region. EX1007, 5-6; EX1002, ¶124. Specifically, Maninis's "pixel-wise semantic annotations" indicate which pixels correspond to each object in a given image, with objects' edges indicated by changes from one annotation to another. EX1002, ¶¶124. In the context of Kim, Invisalign, and Salah, a POSA would have applied annotations to historical dental images to indicate each object's attributes, including assigning tooth attributes to each tooth in each image, as well as to other parts of the image including the aligner and the gap separating the aligner and the tooth. *Id.*

With labels to indicate the pixels corresponding to teeth, the pixels corresponding to aligners, and the pixels corresponding to the gap between the two, each historical image would include tooth attribute values assigned directly to

the image by the labels. EX1002, ¶¶125-26. For example, the boundaries between regions would directly indicate the edges of teeth, or of the aligner, both of which are tooth attributes relating to the separation between the tooth and the aligner. *Id.* Further, labeling the pixels between each tooth and the corresponding aligner surface (i.e., the gap), the image would include a direct label for the tooth of the separation between the tooth and the aligner—the labeled gap is literally a label of the separation to be detected in order to apply Invisalign’s teaching to assess the separation between teeth and the corresponding parts of the aligner.

Thus, Kim, Invisalign, Salah, and Maninis teach step 7(1).

2. Step 7(2)

2) training of at least one deep learning device, by means of the learning base;

Maninis’s deep-learning-based neural network is trained using a training dataset, which is a “learning base” in the context of the ’248 patent. *See, e.g.*, EX1007, 3-4, 7-8; EX1002, ¶¶127-28. In the context of Kim, Invisalign, and Salah, a learning base containing annotated historical images of dental arches of patients wearing an aligner in a service position would be used to train a neural network like Maninis’s, thereby training “at least one deep learning device.” EX1002, ¶128.

3. Step 7(3)

3) submission of the analysis image to the deep learning device for it to determine at least one probability relating to: the presence, in a location

of said analysis image, of an analysis tooth zone, and the attribute value of the tooth represented on said analysis tooth zone;

As discussed for claim 1, it would have been obvious to submit images of patients' dental arches, as captured by the cellphone of Kim and Salah, to be analyzed by a deep learning neural network as taught by Maninis to detect regions containing teeth, aligners, and the gap between, and to characterize their attributes, including the separation between teeth and the corresponding parts of the aligner. EX1002, ¶¶129-30. Maninis teaches the determination of probabilities related to these properties. *Id.*

As Dr. Foroosh explains, Maninis's contour detection procedure generates a "contour detection probability map" that is used to determine "a hierarchical boundary map" demarcating its regions. EX1007, 3-5; EX1002, ¶131. In essence, Maninis describes determining a probability that a contour is located along a given set of pixels. EX1002, ¶131; *see* EX1007, 3 (providing equations to generate probability). Maninis uses these probability maps to derive the contours most likely to be present in the image (EX1007, 4-5), with different probabilities treated as different "strengths" of boundaries. *See* EX1007, 4-5, Figs. 3-5 (strengths depicted by line thickness). These "strengths" are used to determine a "hierarchy of regions" dividing the different objects in an image. *Id.*; EX1002, ¶131.

In the context of Kim, Invisalign, and Salah, the probabilities relate to the presence of a zone at least partially representing a tooth because the edges corresponding to these probabilities define an analysis tooth zone that has been determined to contain a particular tooth. The attribute value is also “represented on said analysis tooth zone” because the attribute values discussed above (tooth and aligner edges and the size of the gap between) are represented as a property of the image and are assigned by applying a contour and object detection process as taught by Maninis. EX1002, ¶¶132-34.

4. Step 7(4)

4) determination, as a function of said probability, of the presence of a separation of the aligner and of the tooth represented by said analysis tooth zone.

Maninis detects regions based on probabilities, as discussed above. *See supra* §IV.E.3; EX1007, 3-5. Further, as discussed above for claim 1, the edges of the regions correspond to the edges of teeth and aligners, and the distance spanned by the gap between these edges corresponds to the separation between the tooth and the corresponding part of the aligner, which is what Invisalign teaches to measure to assess aligner fit. EX1006, 1; EX1002, ¶¶135-36. Determining the size of a separation would include determining the presence of a separation, as if a separation of finite size is detected, then the separation is present, whereas if a

separation is so small that it is not detectable (e.g., if no gap was present, such that the separation was zero), then the presence of a separation would not be detected. EX1002, ¶136. For example, in an image like that shown in Invisalign to demonstrate poor fit as exemplified by a gap greater than 1 mm, the presence of a separation would be detected. *Id.* Thus, applying the teachings of Kim, Invisalign, Salah, and Maninis would involve detecting the locations and edges of teeth and aligners using probabilities, and then determining based on those probabilities the presence of a separation of the aligner and the tooth. *Id.*

Accordingly, claim 7 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶137.

F. Claim 8

Claim 8 depends from claim 1 and further recites that step (b'') comprises three additional steps. These steps resemble the steps recited in claim 7, but, among other differences, the steps recite an “image attribute,” rather than the tooth attributes recited in claims 1 and 7. As discussed above in §III.B, the image attribute recited in claim 8 encompasses any attribute of the image as a whole that relates to the separation between one or more teeth and the aligner, and such an attribute may be a single value or a set of values, including a value or values derived from tooth attribute values. Kim, Invisalign, Salah, and Maninis teach each of the steps of claim 8, as discussed below. EX1002, ¶138.

1. Step 8(1')

1') creation of a learning base comprising more than 1000 images of dental arches, or “historical images”, each historical image comprising an attribute value for at least one image attribute, or “image attribute value”, the image attribute relating to a separation between the tooth or teeth and the aligner represented on the historical image;

Step 8(1') is the same as step 7(1), except the historical images must comprise an image attribute value rather than historical tooth zones and tooth attributes. A POSA would have been motivated to create a learning base of more than 1000 historical images of dental arches of patients wearing aligners to train a neural network. *Supra* §IV.E.1. These images would comprise attribute values for at least one image attribute. For example, Maninis teaches pixel-wise annotations of all objects in each image of its training database. *See* EX1007, 5-6. In the context of Kim, Invisalign, and Salah, these images would be images representative of the type acquired with Kim and Salah’s cellphones in view of the teachings of Invisalign—i.e., cellphone images of patient’s teeth with aligners in service positions. EX1002, ¶139. Thus, for example, each acquired historical image would have an aligner and teeth represented in the image, with annotations indicating the regions containing each tooth and corresponding aligner portion, as well as indicating the separation between each tooth and corresponding aligner part.

The values for these attributes, as reflected directly in the annotations themselves, would relate to the separation of the teeth and the aligner. For example, the '248 patent indicates that “presence” of something (such as an appliance) can be an image attribute value, and that an image attribute value can also relate to the presence of a separation. *See* EX1001, 26:11-14, 29:30-53. By annotating the images with an aligner (where present) or gap (where present), the image would be annotated with (and thus comprise) image attributes relating to the recited separation. EX1002, ¶140. Further, the set of locations of each tooth and each corresponding aligner part would also be image attributes relating to the overall tooth-aligner separation represented in the historical image, as that separation is a measure of the distance from the edge of the region containing the tooth to the edge of the region containing the aligner. EX1002, ¶¶140-41. Thus, the set of annotations would collectively represent an image attribute value as claimed, as they would indicate whether the aligner in the image was well or poorly fitted, under the standard taught by Invisalign. *Id.*

2. Step 8(2')

2') training of at least one deep learning device, by means of the learning base;

This limitation is identical to step 7(2), so it is taught by Kim, Invisalign, Salah, and Maninis for the reasons discussed in §IV.E.2. EX1002, ¶¶127-28, 142.

3. Step 8(3')

3') submission of the analysis image to the deep learning device for it to determine, for said analysis image, at least one probability relating to said image attribute value.

As discussed for step 7(3), Maninis determines probabilities that it uses to detect boundaries between regions and enable identification of objects visible in those regions. EX1007, 4-5, 11-12; *see supra* §IV.E.3. The probabilities therefore relate to the image attributes discussed above for step 8(1'), such as the presence of an aligner and a gap between teeth and aligner, as well as a determination of the distance between the teeth's edges and the corresponding edges of the aligner, and whether this indicated a good fit or a poor fit according to Invisalign's teachings. EX1002, ¶¶143-44; *see also infra* §IV.I (discussing additional image attributes that a POSA would also have had reason to detect).

Accordingly, claim 8 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶145.

G. Claim 13

Claim 13 depends from claim 1 and further recites that “the deep learning device is a neural network.” Maninis teaches that its deep learning image analysis for contour and object detection is performed using a convolutional neural network (CNN). EX1007, Abstract; *see also id.*, 1-2, 13, Figs. 1-2. Accordingly, claim 13

would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶¶146-48.

H. Claim 14

Claim 14 depends from claim 1 and further recites that step (b'') comprises four additional steps. Steps 14(1), 14(2), and 14(3) are identical to steps 7(1), 7(2), and 7(3), respectively, so Kim, Invisalign, Salah, and Maninis teach each of these steps for the reasons discussed above for claim 7. *See supra*, §§IV.E.1-IV.E.3; EX1002, ¶¶ 120-34, 149.

Step 14(4) differs from step 7(4), in that whereas claim 7(4) recites “determination, as a function of said probability, of the presence of a separation of the aligner and of the tooth represented by said analysis tooth zone,” step 14(4) recites “determination, as a function of said probability, of an amplitude of said separation.” Kim, Invisalign, Salah, and Maninis teach step 14(4) for reasons similar to those discussed for step 7(4). *See supra* §IV.E.4; EX1002, ¶150.

As discussed above in §IV.E, Maninis teaches detecting objects (which in the context of Kim, Invisalign, and Salah would include teeth, aligners, and spaces between teeth and aligners). EX1002, ¶151. Invisalign teaches assessing aligner fit by comparing the size of separation between teeth and the corresponding part of the aligner to a threshold, where, for example, a separation of greater than 1 mm would indicate a poor fit. EX1006,1; EX1002, ¶151. As discussed above for claim

1, a POSA would have been motivated by Invisalign to assess the aligner-tooth separation for the patient as reflected in images of the patient's teeth and aligner, so as to determine whether the patient's aligners fit well. *See supra* §IV.A; EX1002, ¶151. This determination would involve determining an amplitude of the separation between teeth and aligners, as a function of the probabilities recited in step 14(3), as Maninis teaches the generation of probabilities as part of a determination of the boundaries between regions corresponding to different objects in an image. *See supra* §IV.E.3 (discussing identical step 7(3)); EX1007, 3-5, Figs. 3-5. Thus, Kim, Invisalign, Salah, and Maninis teach step 14(4). EX1002, ¶151.

Accordingly, claim 14 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶152.

I. Claim 15

Claim 15 is an independent claim reciting similar limitations to claim 1. The preamble of claim 15 and step 15(a'') are identical to claim 1's preamble and step 1(a''), respectively, except that step 15(a'') is broader because the last limitation of claim 1 (requiring that step 1(a'')'s image be acquired with a cell phone) is omitted from claim 15. Thus, the prior art teaches the preamble of claim 15 and step 15(a'') for the reasons discussed above for the corresponding elements of claim 1. *See supra*, §§IV.A-IV.B.2; EX1002, ¶¶153-54.

Furthermore, step 15(b'') is the same as step 1(b''), except that instead of determining a value for a tooth attribute of a tooth in an analysis tooth zone, claim 15 requires that the analysis using a deep learning device must be used when determining a “value for an image attribute of the analysis image, the image attribute relating to a separation between at least one tooth represented on the analysis image, and the aligner represented on said analysis image.” EX1002, ¶155.

As discussed above for claims 1 and 8, it would have been obvious to train a neural network, as taught by Maninis and Salah, to identify and detect edges of teeth and aligners, so as to assess the gap between the teeth and aligner, as taught by Invisalign, in order to determine whether the clear aligner is normal and corresponds to an appropriate step, as taught by Kim. *See supra*, §§IV.A, IV.B.3, IV.F. This would include detecting image attributes related to the separation between the teeth represented in the image and the aligner. For example, the image attributes detected would include the presence of an aligner, corresponding teeth, and a gap between them; the tooth edges and aligner edges defining the gap, the width of the gap, and the distances of each tooth from its corresponding part of the aligner. EX1002, ¶156. A POSA would further have understood that the detection of gaps between teeth and aligner would allow detection of a more general image attribute value, such as a value characterizing the overall size of the gap (e.g., for

each tooth, as an average across teeth, or as a largest tooth gap), any of which would permit generating a value indicating whether the gap was larger or smaller than the 1 mm threshold taught by Invisalign. EX1002, ¶156. Detecting such image attributes, based on the deep learning analysis taught by Salah and Maninis, would have been well within the skill of the POSA and would represent no more than commonsense, routine ways of characterizing gaps between aligners and teeth. EX1002, ¶157. Each such attribute would thus be determined using a deep learning device trained by a learning base, as taught by Maninis, and the attribute would relate “to a separation between at least one tooth represented on the analysis image, and the aligner represented on said analysis image,” as recited in claim 15(b”).

Accordingly, claim 15 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶158.

J. Claim 16

Claim 16 depends from claim 15 and further recites that the method comprises “after the step b”), step c”) as follows: c”) assessment of the suitability of the aligner as a function of the value of said image attribute.” The prior art teaches this step. EX1002, ¶159.

Just as it would have been obvious to assess suitability of an aligner as a function of a tooth attribute value, as discussed for claim 2 (*supra* §IV.C), so too would it have been obvious to assess suitability of an aligner as a function of an

image attribute, such as those discussed above for claim 15. EX1002, ¶160. For example, Invisalign teaches that when an aligner worn on a patient's teeth has a tooth-aligner gap of less than 1 mm, "this is natural deflection that should occur so the Aligner can move the tooth," whereas if the gap is greater than 1 mm, this means that the "Aligners are not seating well." EX1006,1. In view of this teaching, a POSA would have been motivated to use the image attribute values discussed above for claim 15, each of which relates to separation between teeth and aligner, to evaluate the acceptability threshold taught by Invisalign, thereby determining that where the threshold is exceeded, the aligner is not seated well on the teeth, or conversely that where the threshold is not exceeded, the aligner is acceptably seated. EX1002, ¶161.

For example, the presence of the aligner and of a gap between the aligner and the teeth, the width of the gap, and the distances of each tooth edge from its corresponding part of the aligner each relate to separation, and would each usefully inform a measure of separation. Likewise, the other properties discussed above for claim 15 would similarly aid in making the assessment taught by Invisalign, as they indicate, for example, maximal or average gap size, or simply a list of all gaps' sizes. EX1002, ¶162. This assessment would be consistent with Kim's teaching to "check[] a photo or a video of the clear aligner currently worn by the orthodontic patient, and determine[] whether the clear aligner is normal and

corresponds to an appropriate step.” EX1005, ¶66. For example, where applying Invisalign’s threshold assessment indicated that one or more teeth had excessive gaps, either individually or collectively, that would be an indication that the aligner was not normal or did not correspond to an appropriate step. EX1002, ¶163; *see* EX1006, 1 (evaluation may indicate aligner has “excessive flash” or results from a “distorted impression” (meaning it was not normal)), 6 (lack of fit may indicate incorrect aligners or need for mid-course correction (meaning the aligner did not correspond to an appropriate step)).

Accordingly, claim 16 would have been obvious over Kim, Invisalign, Salah, and Maninis. EX1002, ¶¶164-65.

V. DISCRETIONARY DENIAL IS NOT WARRANTED

This petition does not implicate the Board’s discretion according to *Fintiv*. *Apple Inc., v. Fintiv, Inc.*, IPR2020-00019, Paper 11. *See generally* Memorandum on Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation (June 21, 2022) (*Fintiv* Memo). Patent Owner filed its district court complaint in the Northern District of California on November 18, 2022, asserting three patents including the ’248 patent. At a case management conference on July 14, 2023, the court ordered the parties to proceed with discovery relating to two claims from the three asserted patents, the claims to be selected by September 25, 2023, and the court scheduled a summary judgment

hearing on those claims for February 22, 2024. At a discovery hearing on October 3, 2023, the court moved the deadline for selecting claims to October 31, 2023, and ordered the parties to submit a new schedule with a later date for the summary judgment hearing. After the summary judgment hearing, the parties expect to proceed with discovery relating to remaining issues in the case.

The district court proceedings remain in early stages. For example, this petition was filed just three months after the initial case management conference, just over two months after Patent Owner's service of infringement contentions, and one month after the asserted claims were finalized. Expert reports have not been exchanged, no work on claim construction has been completed, and no Markman hearing has been scheduled. While the parties have agreed to be ready for trial by November 2024, no trial date has been scheduled. EX1027. Nonetheless, the median time from filing to trial in the Northern District of California is 37.1 months. EX1020, 66. The expected trial date is therefore late December 2025, whereas the expected date of a final written decision here would be mid-April 2025. *See Fintiv* Memo, 8-9 (institution favored where median time-to-trial indicates earlier FWD date). Moreover, the petition presents compelling merits, as discussed in detail in §IV.

This petition also does not implicate §325(d). The petition presents new art and argument, as none of Kim, Invisalign, and Maninis was disclosed or

considered during examination, and although Salah was cited in an IDS and in the specification of the '248 patent, the version submitted was in French, with only an English abstract, not an English translation. *See* EX1004, 80, 261-341, 979. Nor are the references relied on in the petition cumulative to the references applied during examination, as the discussion the reasons for allowance shows that the elements that the Examiner indicated were missing from the cited art were in fact found in prior art references reflecting common orthodontic and image analysis practices, such as Kim, Invisalign, Salah, and Maninis. Finally, the discussion in the ground above explains in detail why the claims of the '248 patent were obvious over the prior art, confirming that their allowance constituted material error. *See supra* §IV.

VI. GROUNDS FOR STANDING

Align certifies that, under 37 C.F.R. §42.104(a), the '248 patent is available for *inter partes* review, and Align is not barred or estopped from requesting *inter partes* review of the '248 patent on the grounds identified.

VII. CONCLUSION

For the reasons set forth above, claims 1-2 and 4, 7-8, and 13-16 of the '248 patent are unpatentable. Align therefore requests that an *inter partes* review of these claims be instituted.

Respectfully submitted,

Dated: October 16, 2023

/Matthew A. Argenti /
Matthew A. Argenti, Lead Counsel
Reg. No. 61,836

VIII. MANDATORY NOTICES UNDER 37 C.F.R. §42.8

Real Party-in-Interest (37 C.F.R. §42.8(b)(1)): Align Technology, Inc. is the real party-in-interest.

Related Matters (37 C.F.R. §42.8(b)(2)): Patent Owner has asserted the '248 patent against Petitioner in the United States District Court in the Northern District of California (3-22-cv-07335).

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

Lead Counsel: Matthew A. Argenti (Reg. No. 61,836)

Back-Up Counsel: Michael T. Rosato (Reg. No. 52,182); Patrick M. Medley (Reg. No. 74,735)

Service Information – 37 C.F.R. §42.8(b)(4). Petitioner hereby consents to electronic service.

Email: margenti@wsgr.com; mrosato@wsgr.com; pmedley@wsgr.com

Post: Wilson Sonsini Goodrich & Rosati, 650 Page Mill Rd., Palo Alto, CA 94304

Tel.: 650-354-4154

Fax: 650-493-6811

IX. PAYMENT OF FEES UNDER 37 C.F.R. §§42.15(A) AND 42.103

The required fees are submitted herewith. If any additional fees are due at any time during this proceeding, the Office is authorized to charge such fees to Deposit Account No. 23-2415.

X. CERTIFICATE OF COMPLIANCE

Pursuant to 37 C.F.R. §42.24(d), the undersigned certifies that this Petition complies with the type-volume limitation of 37 C.F.R. §42.24(a). The word count application of the word processing program used to prepare this Petition indicates that the Petition contains 13,121 words, excluding the parts of the brief exempted by 37 C.F.R. §42.24(a).

Dated: October 16, 2023

/Matthew A. Argenti /
Matthew A. Argenti, Lead Counsel
Reg. No. 61,836

XI. APPENDIX – LIST OF EXHIBITS

Exhibit No.	Description
1001	U.S. Patent No. 11,049,248 to Salah et al.
1002	Declaration of Dr. Hassan Foroosh
1003	Curriculum Vitae of Hassan Foroosh, Ph.D.
1004	U.S. Patent Application No. 16/030,137 File History
1005	U.S. Patent Pub. No. 2017/0281313 to Taewon Kim
1006	Align Technology, Invisalign Clinical Monitoring Guide (2003) (achived September 26, 2003, at http://web.archive.org/web/20030926024959/http://www.invisaligncec.com:80/consistent/pdfs/ICMG_Final.pdf)
1007	Maninis et al., “Convolutional Oriented Boundaries: From Image Segmentation to High-Level Tasks,” <i>arXiv.org</i> (Apr. 28, 2017)
1008	Translation of PCT Pub. No. WO 2016066651 to Salah et al.
1009	Scott Frey, Blog – The Ortho Cosmos, (archived May 4, 2017 at http://web.archive.org/web/20170504155213/http://theorthocosmos.com:80/blog)
1010	Declaration of Nathaniel E. Frank-Wright
1011	Declaration of Mitra Derakhshan
1012	PCT Pub. No. WO 2016066651 to Salah et al.
1013	Intentionally Left Blank
1014	Intentionally Left Blank
1015	Krizhevsky et al., “ImageNet Classification with Deep Convolutional Neural Networks,” <i>Advances in Neural Information Processing Systems</i> 25 (2012)

1016	Patrick Cavanaugh, “Visual Cognition” <i>Vision Research</i> 51 (2011)
1017	Prasad et al., “Learning Class-specific Edges for Object Detection and Segmentation,” <i>Computer Vision, Graphics and Image Processing: 5th Indian Conference, ICVGIP 2006</i> (2006)
1018	Intentionally Left Blank
1019	Intentionally Left Blank
1020	United States District Courts — National Judicial Caseload Profile, available at https://www.uscourts.gov/sites/default/files/data_tables/fcms_na_distprofile0331.2023.pdf (March 2023)
1021	Maninis et al., “Convolutional Oriented Boundaries: From Image Segmentation to High-Level Tasks,” <i>arXiv.org</i> (Jan. 17, 2017)
1022	<i>arXiv.org Submission Help Page</i> , https://arxiv.org/help/submit (as archived May 3, 2017, at https://web.archive.org/web/20170503071322/https://arxiv.org/help/submit)
1023	<i>arXiv.org Article Replacement Help Page</i> , https://arxiv.org/help/replace (as archived July 18, 2017, at https://web.archive.org/web/20170718173822/https://arxiv.org/help/replace)
1024	<i>COB - Convolutional Oriented Boundaries</i> , http://people.ee.ethz.ch/~cvlsegmentation/cob/ (as archived May 3, 2017, at http://web.archive.org/web/20170503233424/http://people.ee.ethz.ch/~cvlsegmentation/cob/)
1025	<i>arXiv.org Author Search Page for Luc van Gool</i> , https://arxiv.org/find/cs/1/au:+Gool_L/0/1/0/all/0/1 (as archived July 8, 2017, at https://web.archive.org/web/20170708002617/https://arxiv.org/find/cs/1/au:+Gool_L/0/1/0/all/0/1)
1026	<i>arXiv.org Article Replacement Help Page</i> , https://arxiv.org/help/replace (as archived September 8, 2016, at https://web.archive.org/web/20160908205209/https://arxiv.org/help/replace)

1027	Joint Stipulation and Proposed Order (Dkt. No. 73), <i>Dental Monitoring SAS v. Align Technology, Inc.</i> , No. 3-22-cv-07335-WHA (N.D. Cal.)
------	--

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §§42.6(e) and 42.105(a), this is to certify that I caused to be served a true and correct copy of the foregoing Petition for Inter Partes Review of U.S. Patent No. 11,049,248 (and accompanying Exhibits 1001-1012, 1015-1017 and 1020-1027) by overnight courier (Federal Express or UPS), on this 16th day of October, 2023, on the Patent Owner at the correspondence address of the Patent Owner as follows:

Cooper Legal Group LLC
1388 Ridge Road, Unit 1
Hinckley, OH 44233

Dated: October 16, 2023

/Matthew A. Argenti /
Matthew A. Argenti, Lead Counsel
Reg. No. 61,836