

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. IPR2023-01369
Patent No. 10,755,409

REPLY DECLARATION OF DR. HASSAN FOROOSH

TABLE OF CONTENTS

I. QUALIFICATIONS 1

II. LEGAL STANDARDS 2

III. LEVEL OF SKILL IN THE ART 2

IV. CLAIM CONSTRUCTION 4

 A. deep learning device.....4

 B. real time 6

V. THE CARRIER PROVISIONAL DISCLOSES GUIDING PATIENTS TO OBTAIN DENTAL IMAGES 7

VI. CLAIMS 1-15 WERE OBVIOUS AS ILLUSTRATED BY SALAH, CARRIER, AND MANINIS 9

 A. The Combination of Salah, Carrier, and Maninis is Not Inoperable.....10

 B. A POSA Would Have Had Reason to Use Carrier’s Alignment Technique in Combination with Salah.....12

 C. A POSA Would Have Been Motivated to use Deep Learning to Detect Edges and Classify Objects, as Taught by Maninis, to Guide a Mobile Device’s Image Acquisition.....18

 D. A POSA Would Not Have Been Deterred by Development Costs for the Proposed Combination.....28

 E. A POSA Would Have Been Motivated to Identify Individual Teeth.....33

VII. CONCLUDING STATEMENTS.....39

VIII. APPENDIX – CITED PETITIONER EXHIBITS 40

IX. APPENDIX – CITED PATENT OWNER EXHIBITS 44

I, Hassan Foroosh, declare as follows:

I. QUALIFICATIONS

1. I submitted an expert declaration in support of Petitioner Align's petition for *inter partes* review (IPR) of U.S. Patent No. 10,755,409 (the "'409 Patent," EX1001) in Case No. IPR2023-01369. I understand that an IPR was instituted and Patent Owner Dental Monitoring submitted a Patent Owner Response. I submit this expert declaration in support of Align's Reply to the Patent Owner Response.

2. I previously set forth my qualifications in my first declaration (EX1002) at paragraphs 1-13. I also submitted my curriculum vitae (EX1003) at that time, which further sets forth my qualifications.

3. In preparation for this declaration, I have reviewed the Declarations of Dr. John Mongan (EX2002 and EX2028) and the Declaration of Dr. Budi Kusnoto (EX2030). I have also reviewed the declaration of Dr. Maureen Valley, a dental expert whose opinions inform my analysis. EX1035; EX1036. Lists of exhibits I cited in my original declaration or in this declaration are provided in the Appendices.

II. LEGAL STANDARDS

4. I understand that my opinions in this declaration require applying the legal principles outlined in my first declaration. As was done previously, I have applied those same legal principles in forming my opinions here.

III. LEVEL OF SKILL IN THE ART

5. I previously set forth my understanding and opinions pertaining to a person of ordinary skill in the art in the relevant field (“POSA”) in my first declaration at paragraphs 69-70.

6. For example, in paragraph 70, I described that as of July 21, 2017, a POSA

would typically have an advanced degree, such as a Masters’ Degree or Ph.D. in computer science 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.

EX1002, ¶70.

7. I understand that Dr. Mongan has provided an alternative description of the POSA as follows:

In my opinion, a person of ordinary skill in the relevant field as of July 21, 2017 would have had a bachelor's degree or higher in computer science, bioinformatics, or a related engineering discipline, and several years of work experience relating to the development of machine learning or artificial intelligence models or algorithms. A POSA would have also worked as part of a multidisciplinary team, which would have included individuals familiar with medical imaging, dental imaging, dentistry, and/or orthodontics.

EX2028, ¶48.

8. I disagree with Dr. Mongan's description of the person of ordinary skill in the art. For example, Dr. Mongan omits a requirement that the POSA have any experience in computer vision, image analysis, medical imaging, or related areas. The claims of the '409 patent relate to computer vision and image analysis problems, and persons of ordinary skill working on such problems would typically have the type of experience I described in my original declaration.

9. However, I understand that Dr. Mongan has also expressed that in his view, the definition he provided is not materially different from the definition I provided in my original declaration. *See* EX2002, ¶49; *see also* EX2028, ¶50. Based on this clarification from Dr. Mongan, I interpret Dr. Mongan's characterization of the level of skill in the art as intended not to materially differ from mine. Applying this further clarification from Dr. Mongan that no material change to the level of skill I described is intended, the opinions I expressed in my

original declaration would remain the same under either my description of the level of skill in the art or under Dr. Mongan's.

IV. CLAIM CONSTRUCTION

10. I also set forth my understanding that, in the present proceeding, the claims should generally be given their ordinary and customary meaning. *See* EX1002, ¶71. I continue to apply that understanding in this declaration.

11. I also addressed certain claim terms in my first declaration. *See* EX1002, ¶¶73-76; *see also id.*, ¶72 (scope of claims also discussed throughout declaration). I continue to apply the constructions of those claim terms outlined in my first declaration.

12. I understand Dr. Mongan has provided constructions of certain claim terms. *See* EX2028, ¶¶150-58. These constructions would not change my conclusions with respect to the claims of the '409 patent. I address two of these proposed constructions below in more detail.

A. deep learning device

13. In my original declaration, I interpreted "deep learning device" to encompass a device comprising a deep learning neural network trained with a learning base, performing a function of analyzing an image (i.e., the analysis image). EX1002, ¶150.

14. I understand that Dr. Mongan opines that “the term ‘deep learning device’ refers to a device that, through training, is capable of analyzing an image and recognizing patterns therein.” EX2028, ¶151. Dr. Mongan further states that he “agree[s] that a neural network is an example of a deep learning device.” *Id.* The interpretation of deep learning device I applied in my original declaration encompassed devices comprising deep learning neural networks, and the prior art combination I discussed in my original declaration also used a neural network for multiple image analysis steps, including edge detection, object detection, and object classification. *See, e.g.*, EX1002, ¶¶77-88, 104-09, 118-121. Accordingly, this aspect of Dr. Mongan’s construction does not affect my analysis.

15. Dr. Mongan further opines that the word *device* in the term “deep learning device” does not refer to a physical device. EX2028, ¶¶151-53. Rather, Dr. Mongan asserts that the term device “refers to a computer program or algorithm(s), not to the hardware on which the algorithm runs.” EX2028, ¶152.

16. Dr. Mongan’s construction is not the plain and ordinary meaning that a POSA would have ascribed to the term “device,” which would typically refer to a physical apparatus; however, at best Dr. Mongan’s construction makes no difference. To function, a computer program or algorithm must be embodied in physical hardware. Dr. Mongan does not suggest that the ’409 patent limits which physical hardware components (or which combination of physical hardware

components) runs the deep learning neural network. Indeed, the prior art combination I described in my original declaration is consistent with performing the deep learning analysis on the image acquisition device, or on a separate device that is remotely connected to the image acquisition device. Accordingly, even if Dr. Mongan’s construction of “deep learning device” were adopted, it would not change the conclusions I set forth in my original declaration, and Dr. Mongan makes no suggestion to the contrary.

B. real time

17. I understand that Dr. Mongan has proposed a construction of “real time” as follows: “a POSA would have understood that the term ‘real time’ means that the analysis and checking occur ‘immediately,’ while the image acquisition process is still underway.” EX2028, ¶154. The term “immediate,” alone is subjective and encompasses a wide range of timescales, including delays of many seconds. However, Dr. Mongan’s clarification that his use of the word “immediate” means “while the image acquisition process is still underway” (EX2028, ¶154) makes it clear that his construction would not alter the conclusions of my original declaration. For each of the types of feedback I described in my original declaration, the feedback occurs while the image acquisition process is still underway (that is, while images are still being acquired). *See, e.g.*, EX1002, ¶¶77-79 (guidance during image acquisition process), 129 (prompts to “take

another photo” based on analysis performed while acquiring images), 130 (feedback while acquiring photos), 134 (combination would guide user to acquire a new image).

V. THE CARRIER PROVISIONAL DISCLOSES GUIDING PATIENTS TO OBTAIN DENTAL IMAGES

18. I understand Dr. Mongan has opined that “the Carrier Provisional only discloses a method for dental professionals, as opposed to the patient themselves, to use a mobile phone to capture ‘high quality dental images,’” and that as a result “the Carrier Provisional is directed to a fundamentally different invention from that claimed in claim 1 of Carrier.” EX2028, ¶¶104-05. In particular, Dr. Mongan states that “there is nothing in” the paragraphs I cited “or the Carrier Provisional as a whole, that suggest the reference to a “user” in the context of the mobile phone app refers to a *patient*.” EX2028, ¶106 (emphasis original). Dr. Kusnoto provides similar opinions. See EX2030, ¶¶72-75. I disagree.

19. The Carrier Provisional plainly contemplates that the user it refers to may be a patient. For example, in one of the paragraphs I cited in my original declaration, the Carrier Provisional discloses that its “overlay may be based on a generic image of teeth, or it may be customized to the user’s teeth, or to a patient-specific category (by patient age and/or gender, and/or diagnosis, etc.).” EX1006, ¶55. As Dr. Kusnoto acknowledges, this paragraph of the Carrier Provisional “uses

the term ‘user’ in a context where ‘user’ implies ‘patient’ but not a dental professional.” EX2030, ¶75.

20. Dr. Kusnoto suggests that the “user” in other contexts “refers to a dental professional to the exclusion of patients.” EX2030, ¶75. For example, Dr. Kusnoto identifies discussion of DSLR cameras as a context in which the user is a dental professional. *See id.* This paragraph discusses the prior art’s use of DSLR cameras by dental professional. The subsequent disclosure addresses providing guidance to capture images on a user’s cell phone, for which the Carrier Provisional teaches that the user may be a patient or dental professional.

21. The disclosure Dr. Kusnoto identifies does not undermine the fact that when discussing the providing of guidance via a user’s cell phone, the Carrier Provisional indicates that the patient can be the user. For example, the Carrier Provisional teaches adapting “a user’s own handheld electronics apparatus having a camera (e.g., smartphone)...so that the user’s device guides the user in taking high-quality images,” including by providing an overlay. EX1006, ¶54. In the very next paragraph, the Carrier Provisional indicates that the overlay may be customized to the user’s (i.e., patient’s) teeth. EX1006, ¶55; *see also id.*, ¶62 (twice referring to the user as the patient). That the Carrier Provisional indicates elsewhere that the user may be, in some instances, a dental professional does not negate the Carrier Provisional’s disclosure of the elements recited in claim 1 of the

Carrier Publication, including the elements reciting a “mobile application on a mobile device associated with a patient” (e.g., the smartphone of the patient), which is “guiding the patient through the user interface to take a first plurality of photos of teeth of the patient” as well as a “second plurality of photos of the teeth.” EX1005, claim 1; *see also* EX2028, ¶¶72, 74 (emphasizing these elements of claim 1 as allegedly lacking support).

22. I further understand from Dr. Valley that, contrary to Dr. Kusnoto’s opinions, a dental professional reading the Carrier Provisional would not interpret its disclosure as limited to users being dental professionals, but rather that the Carrier Provisional discloses and enables the guidance of smartphone users, including patients, to acquire images of their teeth. EX1035, ¶¶22-31. I further understand from Dr. Valley that patients would be able to take many of the photos taught by the Carrier Provisional, even without assistance. EX1035, ¶¶32-35.

VI. CLAIMS 1-15 WERE OBVIOUS AS ILLUSTRATED BY SALAH, CARRIER, AND MANINIS

23. As I explained in my original declaration, it is my opinion that each of claims 1-15 was obvious in view of Salah, Carrier, and Maninis. In this section, I address certain particular arguments raised by Drs. Mongan and Kusnoto regarding this combination.

A. The Combination of Salah, Carrier, and Maninis is Not Inoperable

24. Dr. Mongan opines that the combination of Salah, Carrier, and Maninis that I described in my original declaration would have been inoperable. As I discuss below, Dr. Mongan has misunderstood the teachings of the references and how I explained that a POSA would have combined them. Properly understood, the combination of the three references would have been operable.

25. To begin with, Dr. Mongan repeatedly refers to the combination as purportedly incorporating only the “low-level” parts of Maninis’s deep learning architecture, which Dr. Mongan refers to as “the Maninis COB neural network.” *E.g.*, EX2028, ¶¶130, 132-36, 168-72, 182-99. Dr. Mongan’s analysis is premised on a mistaken understanding that I relied solely on one part of the Maninis reference (what he calls the “COB neural network”) while not relying on Maninis’s remaining teachings of performing object classification in images.

26. In my declaration, I did not limit my reliance on the low-level portions of Maninis’s neural network architecture. On the contrary, when discussing object classification, I repeatedly cited to the discussion of high-level object classification in Maninis. For example, when discussing how “Maninis’s deep learning algorithm uses the contours it generates to define boundaries around regions and to identify the objects in each region of the image,” I cited to both Maninis’s abstract (which discusses both low- and high-level tasks) and to pages

11-12 of Maninis (which discusses object classification). EX1002, ¶117. In this latter section, Maninis discusses the use of neural networks that perform the task of “Semantic Boundaries” (which “requires not only detecting the boundaries, but also associating a semantic class to them”). EX1007, 11.

27. Maninis expressly teaches using the neural network and other components performing COB (as discussed earlier in Maninis) in combination with other neural network elements for object classification. EX1007, 1-2 (COB coupled to “well-known pipelines”). For example, Maninis teaches using “COB combined with the Dilated Convolution network on Semantic Boundaries,” using a combined system referred to as COB-dil to achieve “state-of-the-art results” in multiple metrics for semantic boundary detection. *Id.*; *see also id.*, 11-12 (discussing further implementations for performing object classification that employ COB as an element of the overall architecture, such as applications using COB and PSPNet). I cite to Maninis’s disclosure of these high-level applications repeatedly throughout my declaration. *See, e.g.*, EX1002, ¶¶68, 83-85, 116-17, 141-42, 181, 197. Thus, Dr. Mongan is mistaken in his assertion that I limited the proposed combination to what he calls the low-level “Maninis COB neural network” and to the exclusion of Maninis’s teaching of performing high-level tasks including object classification. EX2028, ¶172.

28. Because Dr. Mongan fails to address the combination I actually discussed in my declaration, including Maninis’s teaching to use deep learning when performing object classification, his conclusion of inoperability derives from his own misunderstanding, not a problem with the art or the combination I discussed.

B. A POSA Would Have Had Reason to Use Carrier’s Alignment Technique in Combination with Salah

29. I disagree with Dr. Mongan’s claim that a POSA would not have had a reason to “combine the alignment techniques disclosed in Carrier with the methods and acquisition apparatuses disclosed in Salah.” EX2028, ¶173; *see id.*, ¶¶174-81.

30. Dr. Mongan has misunderstood my statement that the aspects of Salah and Carrier used in the combination “would each perform substantially the same function as they provide alone.” EX1002, ¶82; *see* EX2028, ¶175. Carrier provides useful teachings that complement Salah’s techniques, including providing functionality not expressly disclosed in Salah. This includes not only allowing alignment to a dentition image without requiring registration marks, but also ensuring that all required teeth are visible. *See* EX1002, ¶¶81-82. These aspects would provide similarly-useful functions for guiding image acquisition in a combination of Carrier and Salah as they provide in Carrier alone, giving a POSA a reason to implement them in combination with the teachings of Salah.

31. I also disagree with Dr. Mongan's characterization of what he calls Salah's "distinct advantages" over Carrier. EX2028, ¶177. For example, Dr. Mongan suggests that using registration marks would provide "efficiency and accuracy" benefits, whereas he claims that "the Carrier technique does not provide the same functionality as the registration marks in determining the acquisition conditions because there is no standardized, physical marker present in the image of the patient's teeth in those techniques from which the acquisition conditions (e.g., angle of the camera in relation to the patient's mouth) could be determined." EX2028, ¶178; *see also id.*, ¶180-81 (claiming that Carrier's technique would be unnecessary, detrimental, and "offers no benefit over" Salah's approach).

32. Contrary to Dr. Mongan's characterization, Carrier offers multiple benefits relative to registration marks. To begin with, an important purpose of the imaging contemplated by both Salah and Carrier is to acquire images that provide visual information regarding the patients' teeth. Carrier aligns based on teeth actually detected in acquired images, whereas Salah's registration marks are attached to retractors placed in a patient's mouth. Salah's registration marks are thus only indirectly related to the position of the patient's teeth. When seeking to capture an image of an object, aligning to the object itself is a more straightforward method than aligning to other objects. As a matter of common sense, if you want to

capture an image of certain teeth, a natural way to ensure success is to detect the teeth themselves.

33. By contrast, capturing an image of a retractor with registration marks does not necessarily ensure that the image contains teeth at all—as a trivial example, the image reproduced by Dr. Kusnoto of a retractor captures all registration marks, yet it does not contain teeth at all (because the retractor is not worn in a patient’s mouth). *See* EX2030, ¶37. More generally, a POSA would expect variation between the location of registration marks and the position of a tooth, whereas the position of a tooth does not vary relative to itself. In other words, using a retractor introduces a source of error that is avoided by directly measuring teeth. In particular, I understand from Dr. Valley that typical variation in the way retractors sit in the mouths of patients would tend to move registration marks like those shown by Dr. Kusnoto substantially in comparison to the example threshold values (e.g., 1 mm or less) that Carrier teaches may be used for approximate alignment to an overlay. EX1035, ¶44.

34. In contrast to the opinions expressed by Drs. Mongan and Kusnoto (*see* EX2028, ¶179; EX2030, ¶92), a POSA would not consider the tendency of teeth to vary in size and position between patients or to move during orthodontic treatment to provide a reason not to align imaging to patients’ teeth using Carrier’s teachings.

35. Regarding intra-patient variation, as I discussed above, Carrier teaches that the overlay may be patient-specific, and such a patient-specific overlay would automatically account for the actual location of the edges of the patient's teeth to the extent they may vary from the expected (generic) arrangement and sizes. *See* EX1006, ¶55 (overlay "customized to the user's teeth"); EX1005, ¶100.

36. Regarding movement during treatment, Salah teaches taking images of patients' teeth after various short periods of treatment, with examples being after "1 week, 2 weeks, 1 month." EX1008, 53:6-7. I understand from Dr. Valley that a tooth typically moves at most about 0.5 to 1 mm per month. EX1035, ¶43. This amount of movement would not present a problem to a POSA looking to implement Carrier's technique. Indeed, Carrier teaches various thresholds for alignment that would allow for such variations from the overlay generated from a previous patient tooth arrangement; for example, Carrier teaches only requiring alignment within a threshold that may vary from 0.05 to 1 mm, and allowing for fewer than 100% of the edges of the patient's teeth (e.g., greater than 50% of teeth's edges) to be aligned. *See* EX1006, ¶63; EX1005, ¶108.

37. Further, as Dr. Mongan concedes, Carrier teaches that patients' tooth movement during treatment may be accounted for when generating patient-specific overlays. *See* EX2028, ¶116; EX1006, ¶64; EX1005, ¶109; EX1035, ¶¶41-42 (predicted tooth movement accounts for treatment effects). Thus, a POSA would

have reasonably expected techniques like Carrier's to be able to accommodate variation in patient tooth configurations, both from patient-to-patient and for a given patient at different times. In conjunction, given the ability to use patient-specific overlays, account for tooth movement for each patient, and allow for residual error via alignment thresholds, a POSA would have been confident that the techniques of Carrier, as applied in combination with those of Salah and Maninis as I described in my original declaration, would enhance the ability to align an imaging apparatus to capture dental images, including the type of images taught by Salah.

38. Dr. Mongan's proposed use of registration markers would also have drawbacks relative to Carrier's alignment techniques. For example, the registration marks visible on Dr. Kusnoto's retractor image are very small, so will appear in an image as a relatively small number of pixels, with relatively small areas and perimeters. *See* EX2030, ¶37. Such a marker provides limited reliability in determining position and orientation of the imaging device in comparison to larger reference objects such as the contours of teeth that provide many more pixels for image alignment and registration. *Compare id.* (registration marks small compared to size of retractor), *with* EX1006, Fig. 7B (showing teeth taking up large fraction of image relative to size of retractor); EX1005, Fig. 7B. Because it does not need to capture parts of a retractor containing registration marks, Carrier's technique

also allows closer images of teeth, thus allowing for more per-pixel detail. *See* EX1006, Fig. 7B (image aligned with teeth taking up large portion of frame, without capturing outer portions of retractor); EX1005, Fig. 7B; EX2030, ¶37 (registration marks on outer portions of retractor).

39. Dr. Mongan is also mistaken to suggest that “Salah expressly teaches away from measuring the patient’s teeth.” EX2028, ¶179; *see id.*, ¶77 (citing EX1008, 31:26-27). This is not correct. I understand that (as Dr. Mongan acknowledges) expression of a preference for one alternative is not a teaching away; rather, to teach away a reference must “criticize, discredit, or otherwise discourage investigation into the claimed invention.” EX2028, ¶42. Salah contemplates that measurements may be performed on teeth, merely expressing a preference not to use them. *See* EX1008, 31:26-27 (“Preferably, no measurements are taken on the teeth to bring the acquisition apparatus into the matching position.”). A POSA would not conclude from Salah’s expression of preference that performing measurements on teeth would be unsuccessful; rather, a POSA would have understood from Salah’s mention of using tooth measurements that Salah was acknowledging that measuring teeth when aligning is a known option. Salah does not teach that basing measurements on teeth would be unsuccessful, and regardless of what Salah may prefer, a POSA would have expected such

measurements, including those taught by Carrier, to be beneficial and successful.

Thus, Salah does not teach away from alignment techniques like Carrier's.

C. A POSA Would Have Been Motivated to use Deep Learning to Detect Edges and Classify Objects, as Taught by Maninis, to Guide a Mobile Device's Image Acquisition

40. Dr. Mongan suggests that the type of deep learning techniques taught by Maninis would be too slow to provide feedback in time to guide image acquisition. See EX2028, ¶¶191-96. For example, Dr. Mongan compares processor speeds of a Titan X GPU to those of a cell phone, estimating that if Maninis's deep learning architecture were to be run using only the processor of that cell phone, without any modifications, it "would take approximately 6 seconds rather than 0.8 seconds to analyze a single image." EX2028, ¶194; *see also id.*, ¶¶141-44 (describing assumptions), 195-96 (speculating that object classification would further double processing time).

41. Dr. Mongan is mistaken for multiple reasons. First, the combination I described in my original declaration was not limited to being performed on a cell phone alone. While I described that a cell phone may be used as the image acquisition device (*see, e.g.*, EX1002, ¶217), I described the image analysis step as being performed on a "deep learning device," which employs a neural network. *E.g.*, EX1002, ¶105. While such a neural network might optionally be run on a cell phone, a POSA would have readily appreciated numerous ways to implement the

combination I discussed, including configurations with a deep learning device separate from the cell phone as well. The claims do not limit the location in which the image analysis of the deep learning device is performed, so whether a POSA would have chosen to perform the image analysis on the cell phone or on a separate device (such as one or more computers in communication with the cell phone) makes no difference to the obviousness of the claims (that is, either choice would meet every limitation of the claims).

42. A POSA would find implementing the deep learning analysis I described in my original declaration remotely would have been straightforward. For example, POSAs would readily appreciate that images could be transmitted near-instantly to remote servers for analysis, with the appropriate response (e.g., indication of the guidance to provide the user for alignment or re-taking photos) sent just as quickly. Such remote computers would easily have sufficient processing power to generate feedback as rapidly as desired. And as for transmitting images, live video streaming (e.g., FaceTime) was commonplace on cell phones by 2017, so the image transmission aspects would also have been simple and predictable to the POSA. *See, e.g.*, EX1045¹

¹ EX1045 is a June 7, 2010, press release from Apple issued at <https://www.apple.com/newsroom/2010/06/07Apple-Presents-iPhone-4/>, as

(<https://www.apple.com/newsroom/2010/06/07Apple-Presents-iPhone-4/>) (2010 press release announcing inclusion of FaceTime in iPhone 4). Applying this type of ordinary implementation option would have been well within the skill of the POSA and would eliminate any concern Dr. Mongan raises regarding sufficiently fast feedback.

43. Moreover, even if a POSA were to choose to perform all operations on a single device, such as a cell phone, Dr. Mongan's concerns would not present a significant obstacle. By 2017, POSAs were well versed in methods for adapting deep learning applications that might otherwise be processor-intensive to devices with more limited resources, such as cell phones. This includes deep neural networks. For example, I was a co-author of a paper published in 2015 entitled "Sparse Convolutional Neural Networks" that describes how a neural network may be "sparsified" to dramatically reduce processing cost without sacrificing accuracy.

archived by the Internet Archive's Wayback Machine on May 25, 2017. This exhibit is a true and accurate copy of the archived webpage found at the following URL:

<http://web.archive.org/web/20170525092213/https://www.apple.com/newsroom/2010/06/07Apple-Presents-iPhone-4/>.

See EX1039,² 806 (“[E]xpressing the filtering steps in a convolutional neural network using sparse decomposition can dramatically cut down the cost of computation, while maintaining the accuracy of the system.”). For example, the paper describes a process by which over 90 percent of the parameters in the layers of a trained neural network used for “image classification and object detection” may be eliminated (set to zero), producing dramatically reduced processing cost with only a 1% loss of accuracy. *See id.*, Abstract, 813 (“5× speedup of fully connected layers with almost no drop in accuracy,” plus additional improvements “further providing over 2× speedup”).

44. A POSA would have been aware of various techniques like the one I described above to reduce processing cost and speed up image analysis. For example, a further technique known and used by POSAs was quantization, which allows neural networks to operate on mobile devices with sped-up computation and

² EX1039 is a true and accurate copy of “Sparse Convolutional Neural Networks” by Liu et al., of which I am a co-author. This paper was published in 2015, in the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

lower memory requirements. *See* EX1040,³ Abstract, 2 (4-6× speed-up with 15-20× compression with accuracy reduction of less than 1%). Google had also released the MobileNets models by early 2017, which employ global hyper-parameters to shrink, factorize, and/or compress trained networks to allow them to run efficiently and quickly on mobile devices. EX1041,⁴ 1-2, 6, Table 8 (reducing size by 32 times and computational burden by 27 times with comparable accuracy in image classification relative to non-optimized neural networks); *see also id.*, 7 (improvements object detection efficiency). To the extent there were any concern

³ EX1040 is “Quantized Convolutional Neural Networks for Mobile Devices,” a paper by Wu et al. published on arXiv.org in 2016. *See* EX1040 (showing submission date of May 16, 2016). This exhibit is a true and accurate copy of this paper, obtained from arXiv.org via the following download link: <https://arxiv.org/pdf/1512.06473v3>.

⁴ EX1041 is “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” a paper by Howard et al. published on arXiv.org in April 2017. *See* EX1041 (showing submission date of April 17, 2017). This exhibit is a true and accurate copy of this paper, obtained from arXiv.org via the following download link: <https://arxiv.org/pdf/1704.04861v1>.

regarding the ability of the deep learning device (whether it was a cell phone or other device) to compute sufficiently quickly, a POSA would have employed such techniques to predictably reduce any delay to an acceptable level for the task required. Dr. Mongan's analysis assumes that a POSA would have ignored such readily-available tools, thus assuming a problem into existence where a POSA would have seen none.

45. Indeed, the '409 patent fails to disclose any techniques to reduce processing costs for neural networks, whether separate from or implemented on a cell phone. The '409 patent merely cites generic, off-the-shelf neural network architectures, saying that even the listed options are not limiting. *See* EX1001, 16:15-43; *see also* EX2002, ¶84 (concluding the "only material difference between Maninis" and the '409 patent's neural networks is that "the Maninis neural network can generate a hierarchical organization of detected contours and regions"). Nor does the '409 patent specify whether its deep learning device is embodied in a cell phone (or other image acquisition device) or whether it is operated on a distinct computer. As I described above, and unlike the '409 patent, the prior art in 2017 included ample techniques to allow for various types of implementation, including adapting neural networks like Maninis's to run on cell phones and using remote computers communicatively coupled to the cell phone to perform any processor-intensive computational steps.

46. Dr. Mongan's analysis is also at odds with POSAs' understanding of software development, including development of applications employing deep learning. Dr. Mongan raises concerns that the proposed combination might be too resource-intensive relative to the computing capacity of the particular model of iPhone sold in 2017. However, POSAs developing such applications would not have been discouraged by the concerns Dr. Mongan raises. Rather, POSAs developing such software were well aware of the longstanding trend that computing devices, including smartphones, would increase exponentially in computing capacity. Thus, POSAs typically developed software with the expectation that it would be running on the more powerful devices being developed at the same time. In keeping with this trend, the document relied on by Dr. Mongan shows that the next model of iPhone, released just a year later, had about ten times the capacity for neural network processing as the version of iPhone discussed by Dr. Mongan. *See* EX2043, 9 (increase from 0.6 TFLOPS to 5 TFLOPS in one year, from iPhone 8 to iPhone XS); EX1046⁵ (iPhone XS on sale

⁵ EX1046 is a press release from Apple issued at <https://www.apple.com/newsroom/2018/09/iphone-xs-and-iphone-xs-max-bring-the-best-and-biggest-displays-to-iphone/>, as archived by the Internet Archive's Wayback Machine on September 12, 2018. This exhibit is a true and accurate copy

by September 2018). POSAs in 2017 would have been well aware of the trend of rapid increase in processing power, and based on this expectation, they would not have viewed concerns like those raised by Dr. Mongan as a reason to be discouraged from developing deep-learning approaches to image analysis for mobile devices, including a combination based on the teachings of Salah, Carrier, and Maninis.

47. Dr. Mongan also speculates that it “would take longer—likely, significantly longer—to apply a second neural network to determine the semantic contours on which Dr. Foroosh’s combination depends, and to analyze multiple dental images.” EX2028, ¶144. Dr. Mongan provides no quantification of what he means by “significantly longer,” but even simple methods of estimation show that any processing time associated with object classification would be minimal. Maninis explains that some portions of its neural network architecture identify boundaries and further identify instances of each object in an image. *See* EX1007, 7-9, Fig. 10 (showing boundaries for each identified object). Maninis discloses

of the archived webpage found at the following URL:

<http://web.archive.org/web/20180912213153/https://www.apple.com/newsroom/2018/09/iphone-xs-and-iphone-xs-max-bring-the-best-and-biggest-displays-to-iphone/>.

performing a final step of assigning a specific classification to each of these already-identified objects using further neural network operations; for example, using Fast R-CNN. EX1007, 11-12. Fast R-CNN is, as the name suggests, a very fast method for classifying objects—the paper cited by Maninis describes object classification performed on an image as quickly as 60 ms. *See* EX1042,⁶ Table 4 (reporting 0.06 s/image for Fast R-CNN using singular value decomposition (SVD), and .10 s/image even without SVD); *see also id.*, 11-12, 14 (identifying Faster R-CNN (reference [63]) as another available semantic segmentation pipeline); EX1001, 16:36-41 (identifying “R-CNN” and “Faster R-CNN” as appropriate neural networks for use in detecting objects). Given the minimal time required for object classification, a POSA would not have expected it to take significantly longer to assign classifications to the objects identified by Maninis.

48. Further, even if the timing concerns raised by Dr. Mongan were realistic, the combination I proposed is consistent with certain feedback being sent more slowly. For example, I explained how Salah’s teaching includes feedback

⁶ EX1042 is “Fast R-CNN,” a paper by Ross Girshick published on arXiv.org in 2015. *See* EX1042 (showing submission date of September 27, 2015). This exhibit is a true and accurate copy of this paper, obtained from arXiv.org via the following download link: <https://arxiv.org/pdf/1504.08083v2>.

that the device may “emit a signal to inform the user that...he/she must take another photo” (EX1008, 27:21-25), and consistent with this, Carrier’s teaching to use “image recognition algorithms to ensure that all required teeth are visible” (EX1006, ¶¶81; EX1005, ¶128) could be implemented using the object classification taught by Maninis. *See* EX1002, ¶¶129, 134; *see also id.*, ¶¶115-21 (describing use of object classification taught by Maninis to detect required teeth). This type of feedback would not need to be sent instantly to guide the user during the image acquisition process—even if the device took several seconds to analyze an image, as Dr. Mongan hypothesizes, a message after such a delay would still be sufficiently immediate to guide the user to take another photo. Indeed, such delays would also be consistent with Dr. Mongan’s definition of real-time, since they would come “immediately”—that is, “while the image acquisition process is still underway.” EX2028, ¶154.

49. Similarly, the parts of Carrier involving edge detection require only lower-level aspects of computer vision (*see* EX1006, ¶63 (“edge detection image processing”); EX1005, ¶108), which correspond to the outputs of earlier layers in Maninis’s neural network architecture. *See, e.g.*, EX1007, Fig. 2 (showing earlier layers producing fine feature maps and later layers producing coarse feature maps). Maninis expressly discloses that these different levels of feature maps from different layers are *all* outputs of its neural network. *See id.*, 3-4, Figs. 1-2. To the

extent faster processing time were needed for edge-detection tasks, using such side outputs would be quicker because the output comes prior to the remaining processing steps used by Maninis’s neural network architecture for coarse detection, object detection, and object classification.

D. A POSA Would Not Have Been Deterred by Development Costs for the Proposed Combination

50. Dr. Mongan opines that the “cost, time, and complexity” of assembling and labeling training images would have deterred a POSA from practicing the proposed combination, and Dr. Kusnoto similarly opines that a POSA would not have been motivated to assemble and label images for a learning base because he asserts that “process of obtaining a learning base for use in capturing dental photographs would have been expected to be costly and complicated.” EX2028, ¶205; EX2030, ¶50; *see* EX2028, ¶¶197-205; EX2030, ¶¶45-50. I disagree.

51. Dr. Mongan appears to estimate the costs of annotating images by assuming that POSAs would limit themselves to manual annotation of every pixel for every image in the training dataset. *See, e.g.*, EX2028, ¶201; *see also* EX2030, ¶¶48-49 (Dr. Kusnoto discussing costs for manual labeling of each of 10,000 or more images). However, there were numerous techniques known to POSAs in 2017 that were available to reduce the costs of generating labeled datasets. For example, in my original declaration, I described one technique that POSAs would

have been motivated to use: the mapping of an existing model representing the ground truth to a corresponding image. *See* EX1002, ¶¶183-211. As I explained in my declaration, Salah teaches techniques that allow this mapping to be performed automatically, and as a result provides a simple, cost-effective way of labeling images. *See id.*

52. There are also additional, complementary techniques to further reduce costs with which POSAs would have routinely employed. One such technique is data augmentation. In data augmentation, an original historical image is used to generate a set of numerous historical images by performing standard transformations on the original image. *E.g.*, EX1033, 1-2. For example, an image might be rotated to different orientations, with each rotation being used as a distinct historical image in a training dataset. The use of data augmentation was routine to persons of ordinary skill in the art and was well known in fields of application including medical imaging; indeed, I note that Dr. Mongan was a co-author of a paper that used such techniques to generate a set of over 150,000 labeled images from 1885 labeled patient images to use in training and validation of a deep convolutional neural network. EX1033, 2.

53. POSAs were also familiar with other techniques used to further reduce labeling costs. One example is pseudo-labeling, in which a subset of images are initially manually labeled, and then further images are labeled with the help of a

neural network trained on that subset, generating “pseudo-labels” for initially unlabeled images. *See* EX1043,⁷ 2-3. The result is to train a neural network to label images in a semi-supervised fashion, reducing the need to rely on manual labeling. *See id.*, 4-6. Further, non-deep-learning techniques were known to automatically segment images into semantically meaningful parts for easy application of labels, which can function with as few as just one labeled example. *See* EX1047,⁸ 373-74,

⁷ EX1043 is “Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks,” a paper by Dong-Hyun Lee, published on the ResearchGate website in 2015. *See* EX1043, 1 (showing submission date of August 1, 2015). This exhibit is a true and accurate copy of this paper, obtained from ResearchGate via the following download link:

https://www.researchgate.net/profile/Dong-Hyun-Lee/publication/280581078_Pseudo-Label_The_Simple_and_Efficient_Semi-Supervised_Learning_Method_for_Deep_Neural_Networks/links/55bc4ada08ae092e9660b776/Pseudo-Label-The-Simple-and-Efficient-Semi-Supervised-Learning-Method-for-Deep-Neural-Networks.pdf.

⁸ EX1047 is “Inducing Semantic Segmentation from an Example,” a paper by Schnitman et al., published as a conference paper in 2006, in the Computer Vision—ACCV 2006 Proceedings. EX1047 is a true and correct copy of this

381-83. Additionally, POSAs would have employed techniques to minimize the effects of any noise that arose when using automatic labeling, using techniques such as backward and forward correction of label noise. *See* EX1044,⁹ Abstract, 3-4.

54. Because Dr. Mongan and Dr. Kusnoto failed to consider routine cost-saving measures such as the ones I outlined above, the cost estimates they have generated substantially overestimate the actual costs of generating a labeled dataset of the kind I described in my original declaration.

55. Regardless, even accepting the final numbers Drs. Mongan and Kusnoto generated as realistic estimates, a POSA would not have been deterred by those costs. Dr. Mongan estimates the full cost of preparing the type of dataset I

paper, obtained from SpringerLink via a download link at https://link.springer.com/chapter/10.1007/11612704_38.

⁹ EX1044 is “Making Deep Neural Networks Robust to Label Noise: a Loss Correction Approach,” a paper by Patrini et al. published on arXiv.org in March 2017. *See* EX1044 (showing submission date of March 22, 2017). This exhibit is a true and accurate copy of this paper, obtained from arXiv.org via the following download link: <https://arxiv.org/pdf/1609.03683v2>.

described in my original declaration as being from “tens of thousands to hundreds of thousands of dollars depending on total number of training images required to achieve sufficient results.” EX2028, ¶149. The production of a learning base is essentially a one-time, upfront expense as part of the development process of a deep learning device. Costs in the tens to hundreds of thousands of dollars are quite small compared to typical development costs even of non-commercial products, and a POSA would not have been discouraged by the prospect of such development costs.

56. Dr. Kusnoto disagrees with Dr. Mongan’s cost estimate, suggesting that costs may rise as high as \$1.5 million, including as much as 2 years of dental student labor. EX2030, ¶49. This cost, too, would not have been seen by a POSA as a significant deterrent to development of such a learning base. Such costs are not particularly high even compared to typical budgets for academic research, to say nothing of the vastly larger budgets available for development of commercially-useful products. For example, my CV details research projects with funding totaling over \$14 million, including multiple multi-million-dollar projects, and I note that Dr. Mongan’s CV reports similar budgets for certain of his own funded development projects. See EX1003, 9-13; EX2029, 4, 28-32. Thus, a POSA would be motivated to combine Salah, Carrier, and Maninis as described in my original

declaration even if the associated costs corresponded to either the estimate of Dr. Mongan or the estimate of Dr. Kusnoto.

E. A POSA Would Have Been Motivated to Identify Individual Teeth

57. Dr. Mongan suggests that a POSA would not have been motivated to identify individual teeth when combining the teachings of Salah, Carrier, and Maninis. EX2028, ¶206. However, as I explain below, Dr. Mongan misunderstands my discussion of the combination as well as the teachings of the references themselves.

58. To begin with, when addressing the combination I discussed in my original declaration, Dr. Mongan conflates two aspects of Carrier: Carrier's teaching to align tooth edges to an overlay and Carrier's teaching that the apparatus may additionally use image recognition algorithms to ensure all required teeth are visible. See EX2028, ¶¶207-08. In particular, Dr. Mongan suggests that the combination of Salah, Carrier, and Maninis I proposed in my original declaration was limited to using semantic contours for alignment. See EX2028, ¶207.

59. As I explained in my original declaration, Maninis teaches generating contours at multiple scales and levels, from low-level multi-scale oriented contours (either generic or object-specific), to high-level contours labeled as containing a specific semantic class of object. See, e.g., EX1002, ¶¶106, 117; EX1007, 4-5 (fine to coarse contours), 7-9 (object detection), 11-12 (object classification and

semantic boundaries), Figs. 2 (showing fine-to-coarse contours), 10 (showing contours identifying object instances), 16 (showing use of Maninis’s techniques to identify class for each object contained in distinctly-detected contour boundaries); *see also* EX2028, ¶¶132 (Maninis discloses using neural networks in a combined architecture both “for determining boundaries in an image” and “to determine semantic contours”), 138.

60. Dr. Mongan suggests that not only could Maninis’s semantic contours be used in alignment, but also that alignment could be performed without semantic contours—such as with the lower-level generic contours of Maninis. *See* EX2028, ¶207. I agree that both options would have been understood as available by the POSA. Indeed, a POSA would have understood that, due to the structure of Maninis’s neural network architecture, lower-level contours would naturally be generated in the course of generating contours identifying object instances (also known as object proposal), as well as higher-level contours labeled with object classes. *See* EX1007, 4-5, 7-9, 11-12, Figs. 2, 10, 16. A POSA implementing a method based on the combined teachings of Salah, Carrier, and Maninis would thus weigh the benefits of different implementations, including using semantic contours for both alignment and identification of required teeth, or using lower-level contours for one function and higher-level contours for the other.

61. In my original declaration, I discussed the use of object classification and the associated semantic contours in relation to Carrier's teaching that, in addition to performing alignment with an overlay, images captured may also be assessed using image recognition algorithms to ensure that "all required teeth are visible." EX1006, ¶81; EX1005, ¶128; *see* EX1002, ¶¶80-82, 103, 117-21, 126, 134. Despite Carrier expressly teaching that it is desirable to use image recognition to ensure that all required teeth are visible, Dr. Mongan suggests that "determining a list of present teeth would not be valuable." EX2028, ¶208. I disagree.

62. Dr. Mongan's opinion is premised on the mistaken assumption that Carrier's teachings about ensuring required teeth are visible are directed to and limited to ensuring an appropriate alignment of the imaging apparatus. *See* EX2028, ¶¶208 (asserting Carrier's teachings for this aspect are "primarily concerned with the approximate alignment of the 'view'"), 209 (asserting Carrier's "overlay technique" performs the function of ensuring required teeth are visible). For example, Dr. Mongan expressly asserts that Carrier teaches that its "overlay technique" is an "image recognition algorithm to ensure that all required teeth are visible." EX2028, ¶209. However, nowhere does Carrier describe its overlay technique as an image recognition algorithm, nor does it teach that the overlay determines whether all required teeth are visible. Carrier teaches that its overlay technique may employ "edge detection image processing" (EX1006, ¶63; EX1005,

¶108), and when Carrier teaches using an “image recognition algorithm to ensure that all required teeth are visible,” it is discussing analysis of an image after being captured, to determine whether to tell the user to take another image. *See* EX1006, ¶¶80-81; *see also id.*, ¶81 (same paragraph teaching method can include determining photo is out of focus), Fig. 5 (showing instruction to retake photo because it was out of focus); EX1005, ¶¶127-28, Fig. 5. As Dr. Mongan admits, Carrier’s “overlay technique guides the approximate distance, angle, and view” for alignment (EX2028, ¶210), as opposed to determining whether all required teeth are visible.

63. Dr. Mongan suggests that Carrier’s statement that the “overlay with an outline of teeth in a predetermined view can further provide information such as required visible teeth in the predetermined view” implies that the overlay itself is the image recognition algorithm. EX1006, ¶60; EX1005, ¶105; EX2028, ¶209. However, this statement from Carrier is not describing an algorithm; rather, Carrier says that the overlay can convey to the user (visually) which teeth are required. Thus, Carrier describes the overlay as “provid[ing] information such as required visible teeth in the predetermined view,” by “show[ing] what teeth should be visible in the view” rather than saying that it ensures that all required teeth are visible (e.g., by determining whether each required tooth is visible). EX1006, ¶60; EX1005, ¶105. What should be is different from what is—even when Carrier

teaches that the overlay can optionally indicate what teeth should be visible, the step of ensuring that they are visible is a separate step that uses image recognition techniques, and for which a POSA would be motivated to employ the type of object classification techniques taught by Maninis to identify visible teeth, to ensure that all required teeth are actually visible, and to require a user to retake a photo if they were not. *See also* EX1035, ¶¶36-38 (“required teeth” refer to teeth intended to be captured in image).

64. Dr. Mongan’s statement that “[c]onfirming that a list of expected teeth are visible in the image is unhelpful if those teeth are visible from an inappropriate angle or view” also misses the point. EX2028, ¶209. Whether teeth are visible and whether the angle or view is appropriate are separate questions—teeth can be visible from an inappropriate angle or view or not visible despite the camera being at an appropriate angle or view. For example, I understand from Dr. Valley that even in an orthodontically proper view, patients’ teeth may be obstructed by objects such as a tongue, a lip, or a dental appliance (e.g., an inappropriately worn orthodontic appliance, a misposition retractor, etc.), or one or more teeth may be badly illuminated (e.g., due to poor lighting) so as to make them difficult to discern. EX1035, ¶39. This comports with my own understanding of common-sense issues that can arise in all types of imaging, and a POSA would similarly have appreciated that such concerns would mean that alignment to teeth and

visibility of teeth are distinct problems that Carrier addresses with its distinct teachings to align to an overlay and to also evaluate whether all required teeth are visible in the aligned image, and if not, to tell the user a new image is needed, as taught by Carrier and Salah.

65. Dr. Mongan also raises a concern that if a patient were missing teeth, then this might interfere with determining that all required teeth are visible. *See* EX2028, ¶209 (noting that Carrier’s overlay can “provide information such as **required visible teeth**” and observing that a patient’s arch may be missing a molar shown in an example overlay). This concern is misplaced. As Carrier teaches, the overlay can be customized to the user’s teeth. *See* EX1006, ¶¶55, 62-63; EX1005, ¶¶100, 106-07. I understand from Dr. Valley that standard dental records would reflect patient-specific dental information including indications of which teeth were missing, damaged, or otherwise modified, so the apparatus would reflect, at least as part of its patient-specific overlay-generating software, any teeth of the patient that would not be required to be detected (e.g., because they were missing). EX1035, ¶40. Dr. Mongan emphasizes Carrier’s teaching that the overlay may provide information about what teeth are required to be visible (EX2028, ¶209), and this same teaching, read in light of Carrier’s teaching that patient-specific overlays may be used, confirms that Dr. Mongan’s concerns are misplaced.

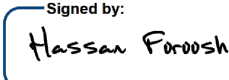
VII. CONCLUDING STATEMENTS

66. In signing this declaration, I understand that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I acknowledge that I may be subject to cross-examination in this case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

67. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: August 20, 2024

By:

Signed by:

0ECB4C4CE2B34A0...
Hassan Foroosh, Ph.D.

VIII. APPENDIX – CITED PETITIONER EXHIBITS

Exhibit No.	Description
1001	U.S. Patent No. 10,755,409 to Salah et al.
1002	Declaration of Hassan Foroosh, Ph.D.
1003	Curriculum Vitae of Hassan Foroosh, Ph.D.
1004	U.S. Patent Application No. 16/031,201 File History
1005	U.S. Patent Pub. No. 2021/0068923 to Carrier et al.
1006	Provisional Patent Application No. 62/417,985 to Carrier et al.
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	U.S. Patent Pub. No. 2013/0282351 to Martin Tank
1010	U.S. Patent Pub. No 2020/0229901 (Application No. 16/827,594)
1011	U.S Patent Pub. No. 2018/0125610 (Application No. 15/803,718)
1012	PCT Pub. No. WO 2016066651 to Salah et al.
1013	Xie et al., “Semantic Instance Annotation of Street Scenes by 3D to 2D Label Transfer,” <i>2016 IEEE Conference on Computer Vision and Pattern Recognition</i> (2016)
1014	Chen et al., “Beat the MTurkers: Automatic Image Labeling from Weak 3D Supervision,” <i>2014 IEEE Conference on Computer Vision and Pattern Recognition</i> (2014)
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	Wang et al., “Simultaneous Image Classification and Annotation,” <i>CVPR 1</i> , No. 2 (2009)
1019	Tao et al., “Similarity Voting based Viewpoint Selection for Volumes,” <i>Computer Graphics Forum</i> 35, No. 2 (2016)
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)
1033	Rajkomar et al., “High-Throughput Classification of Radiographs Using Deep Convolutional Neural Networks,” <i>J. Digital Imaging</i> , 30, 95-101 (Feb. 2017)

1035	Declaration of Dr. Maureen Valley
1036	Curriculum Vitae of Maureen Valley, D.M.D., M.P.H., M.S.
1039	Liu et al., “Sparse Convolutional Neural Networks,” 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
1040	Wu et al., “Quantized Convolutional Neural Networks for Mobile Devices,” <i>arXiv.org</i> (https://arxiv.org/pdf/1512.06473v3) (May 16, 2016)
1041	Howard et al., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” <i>arXiv.org</i> (https://arxiv.org/pdf/1704.04861v1) (Apr. 17, 2017)
1042	Ross Girshick, “Fast R-CNN,” <i>arXiv.org</i> (https://arxiv.org/pdf/1504.08083v2) (Sep. 27, 2015)
1043	Dong-Hyun Lee, “Pseudo-Label : The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks,” <i>ResearchGate</i> (https://www.researchgate.net/profile/Dong-Hyun-Lee/publication/280581078_Pseudo-Label_The_Simple_and_Efficient_Semi-Supervised_Learning_Method_for_Deep_Neural_Networks/links/55bc4ada08ae092e9660b776/Pseudo-Label-The-Simple-and-Efficient-Semi-Supervised-Learning-Method-for-Deep-Neural-Networks.pdf) (Aug. 1, 2015)
1044	Patrini et al., “Making Deep Neural Networks Robust to Label Noise: a Loss Correction Approach,” <i>arXiv.org</i> (https://arxiv.org/pdf/1609.03683v2) (Mar. 22, 2017)
1045	“Apple Presents iPhone 4,” https://www.apple.com/newsroom/2010/06/07Apple-Presents-iPhone-4/ (as archived May 25, 2017, at http://web.archive.org/web/20170525092213/https://www.apple.com/newsroom/2010/06/07Apple-Presents-iPhone-4/)
1046	“iPhone Xs and iPhone Xs Max Bring the Best and Biggest Displays to iPhone,” https://www.apple.com/newsroom/2018/09/iphone-xs-and-

	iphone-xs-max-bring-the-best-and-biggest-displays-to-iphone/ (as archived Sep. 12, 2018, at http://web.archive.org/web/20180912213153/https://www.apple.com/newsroom/2018/09/iphone-xs-and-iphone-xs-max-bring-the-best-and-biggest-displays-to-iphone/)
1047	Schnitman et al., “Inducing Semantic Segmentation from an Example,” <i>Computer Vision—ACCV 2006</i> (2006)

IX. APPENDIX – CITED PATENT OWNER EXHIBITS

Exhibit No.	Description
2002	Declaration of John Mongan
2028	Second Declaration of Dr. John Mongan
2029	Updated Curriculum Vitae of Dr. John Mongan
2030	Declaration of Dr. Budi Kusnoto
2043	M. Owen, “How iPhone Speeds Have Grown in the Last 5 Years,” <i>Apple Insider</i> https://appleinsider.com/articles/22/09/26/how-iphone-speedshave-grown-in-the-last-5-years