CO/

IN THE HIGH COURT OF JUSTICE
QUEEN'S BENCH DIVISION
ADMINISTRATIVE COURT

BETWEEN

THE QUEEN
(on the application of EDWARD BRIDGES)

Claimant

-and-

THE CHIEF CONSTABLE OF SOUTH WALES POLICE

Defendant

- and -

THE SECRETARY OF STATE FOR THE HOME DEPARTMENT

Interested Party

EXPERT REPORT OF DR ANIL JAIN

I, Dr Anil Jain of [redacted], will say as follows:

1. I am a University Distinguished Professor in the Department of Computer Science & Engineering at Michigan State University ("**MSU**"), in the United States.

2. I make this Report in order to assist the Court by providing factual information with which I am familiar due to the experience I describe immediately below. I have been asked by Edward Bridges to produce this Report. I am not authorised to, and I do not by saying anything below, waive any privilege on behalf of Edward Bridges.

3. The contents of this Report are within my knowledge, save where I indicate otherwise. Where the contents of this Report are within my knowledge, I confirm that they are true. Where they are not, I have identified the source of the relevant information, and they are true to the best of my knowledge and belief.

4. I confirm that I understand that I owe a duty to assist the court in matters within my expertise, and that this duty overrides any obligation to those by whom I was instructed. I confirm that I have complied with this duty, and will continue to comply with this duty.

5. I have been made aware of the requirements of Part 35 of the Civil Procedure Rules, Practice Direction 35 of the Civil Procedure Rules, and the Guidance for the Instruction of Experts in Civil Claims 2014, and I confirm that I have complied with these requirements, and will continue to comply with relevant requirements under these rules and guidance.

6. I exhibit to this Report a consecutively paginated bundle of documents labelled "**AJ1/x**", where 'x' is the exhibit number.

Background and experience

7. Throughout my education and career, I have gained detailed knowledge of biometric recognition technology, and I am a leader in this field. I have been working on automated facial recognition ("**AFR**") technology in particular since the mid-1990s.

8. I have been a University Distinguished Professor at MSU since 1992, after having joined MSU in 1974, first as an Assistant Professor, then Associate Professor, Professor and Chair of the Department. Prior to this, I was working towards my Ph.D at Ohio State University in 1973, in the field of Electrical Engineering. I have also taught as a Visiting Professor at 13 other universities worldwide. At MSU, my particular research focus is on Pattern Recognition, Image Processing and Biometrics.

9. I have received a significant number of prestigious awards and recognitions. Most notably, in the context of biometric recognition technology: ISI Highly Cited Researcher (2017), IAPR Senior Biometric Investigator Award (2014) and the King-Sun Fu Prize by the International Association of Pattern Recognition (2008).

10. Between 2007 and 2008, I was a Member of the Biometrics Defense Support Team, and between 2008 and 2010, I was a Member of the Defense Science Board, providing independent advice to the U.S. Department of Defense on scientific and technical matters. Other relevant boards and groups that I have been a member of include: AAAS Latent Fingerprint Study Group (2015-2017); NIST Forensic Science Standards Board (2014-2016); and the National Academies Panel on Biometrics (2005-2009). I am currently a member of the Idiap Research Institute Scientific Advisory Committee, since joining in 2013. I was elected to the United States National Academy of Engineering in 2016 based on my contributions to pattern recognition and biometrics.

11. I have co-authored 14 books and 263 journal articles on the subject of pattern recognition, image processing and biometrics, and have been sole author of three journal

articles on the same. I also served as Editor-in-Chief of the IEEE Transactions on Pattern Analysis and Machine Intelligence (1991-1994), and have since served as Associate Editor for a number of journals that publish in my area of scientific expertise.

12. I delivered keynote speeches on biometric recognition technology at the 11th International Conference on Biometrics (2018); 13th International Conference on Signal Image Technology & Internet Based Systems (2017); The Royal Society, "The Paradigm Shift for UK Forensic Science" (2015); International Joint Conference on Biometrics (2014); National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (2013); International Conference of the Biometrics Special Interest Group (2013); International Conference on Biometrics (2013) and International Workshop on Biometrics and Forensics (2013), among many others.

13. As regards AFR technology in particular, I have co-authored 128 publications since 1998. I developed a system called FaceSketchID for comparing composites (sketches drawn by forensic artists based on descriptions provided by witnesses and/or victims) with police custody images, which has since been licenced to Safran Morpho. I also developed a system called Face-Search, which allows for large-scale searches within a database with hundreds of millions of faces, which has since been licensed to NEC Corporation ("NEC"). I am not aware of how these companies are using this software, or how they integrated FaceSketch ID and Face-Search into their existing AFR software.

14. The expertise of MSU's Department of Computer Science & Engineering's laboratory is often sought by those working in the biometrics industry, from start-ups to large industry-leading companies. I have previously worked with NEC – the manufacturer of South Wales Police's technology, "AFR Locate" – in the capacity of consultant, and I have also received research gifts in the form of funding from NEC. I currently serve as a consultant to NEC Laboratory, Princeton, NJ on the topic of tattoo image recognition.

15. My full Curriculum Vitae is exhibited to this statement at AJ 1/1.

Purpose of this Report

16. This Report covers the following matters:

    a) Firstly, I explain the technical functioning of AFR technology as a general framework, and how the technology works in an operational setting.

    b) Secondly, I explain the variables which may affect the accuracy of an AFR system.

    c) Thirdly, I explain how AFR systems are trained. I also explain how the characteristics of the dataset used for training the AFR system can affect its accuracy.

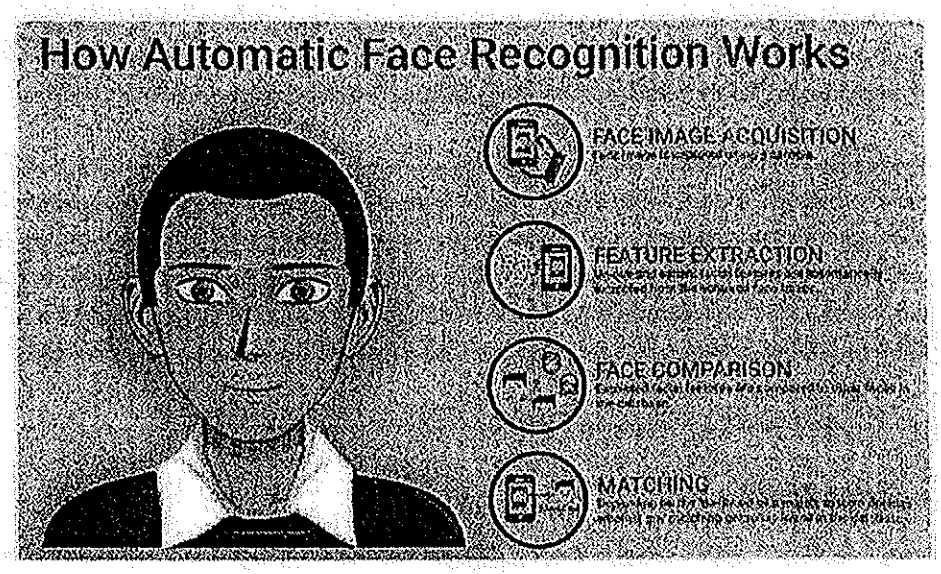
d) Finally, I explain the extent to which AFR technology is fully developed, and how it is being developed further.

17. This statement is not intended to comment directly on AFR Locate as I do not know the exact design of the AFR system,[1] or the face datasets used for training this software. However, all AFR systems follow the same general framework (as I explain below). Further, I understand that NEC has developed multiple AFR software solutions around a 'core' face recognition engine which remains the same for all of their solutions. For these reasons, I am confident that my comments on general AFR frameworks will also apply to AFR Locate.

How AFR works

General framework

18. All AFR systems follows the same general framework, using a four-stage approach to facial recognition, outlined in the figure below:



19. The operation of AFR technology slightly differs from a closed circuit television ("**CCTV**") camera. A CCTV camera simply captures video recordings, whereas AFR technology that uses a live feed captures such recordings and then, from those recordings, isolates individual faces, extracts features from those faces, compares those features with other images in a database, and indicates matches between faces captured through an AFR camera recording and those contained in images already held by the entity that deployed AFR.

[1] For example, state-of-the-art AFR systems like AFR Locate generally employ deep learning networks, which are tweaked depending on the system, the details of which are trade secrets.

20. AFR technology can be used either in a static context (such as passport checks at borders where the subject stands stationary looking at a camera) or in live context, using a live feed from a surveillance camera to identify faces from a moving image. The process is the same for each, but where appropriate I have focussed this report on the live context which I understand is in issue in these proceeding. I comment later in my report on the difference in accuracy between the static and live contexts and the reasons for that.

21. I will now comment on each of the stages outlined in the figure above.

Face image acquisition

22. When used in a live context, an AFR system works in real time to capture faces of individuals in public settings. First, the AFR camera captures a video of people in a public setting. A face detector then detects whether a face(s) exists in each still frame. Therefore, the same face might be detected multiple times, if it appeared in multiple still frames. If a face(s) is detected, the face image(s) is isolated from the video feed, cropped and aligned. The process of isolating a face from a video feed is called Face Detection.

23. During Face Detection, not every single still frame is considered due to the inevitability that some frames will be redundant. Instead, a 'frame rate' is chosen, for example, Face Detection is to be carried out for 15 frames per second. This frame rate is usually set at a default value by the manufacturer, but can be changed by the end user. So if the frame rate is set to 15 frames per second, for example, and it takes a person 60 seconds to walk past an AFR camera, this person may therefore have their face detected, and Features (defined below) extracted and compared with those of database images, up to 900 times.

24. AFR systems can have the functionality where the end user can know how many faces have been scanned and how many Features have been extracted on any given deployment.

25. AFR systems do not require the storage of face images captured by an AFR camera. This is because the matching process relies only on the Features extracted from the images. I am, however, aware that certain AFR systems do store face images for the purpose of providing evidence in legal proceedings; the length of time that face images are stored depends on the entity using AFR. However, this is not a requirement for an AFR system to work.

Feature extraction

26. Once Face Detection has occurred, a set of numerical values which make up an individual's face, known as "*Features*", are automatically extracted by the technology from the face image. It is these Features that are used for the comparison of two faces. This extraction process happens to each face image identified by the Face Detection

process, meaning that the same face may have its Features extracted multiple times. The details of how exactly feature extraction happens are trade secrets and not made public by manufacturers.

27. Like with face images, Features do not need to be stored for AFR systems to work and can be deleted immediately after the matching process is complete. However, just as face images can be stored (as described as §25 above), it is also possible for Features to be stored for a certain period of time or even permanently. Whether this happens depends on what is requested by the entity deploying AFR.

Face comparison

28. Once the Features have been extracted from a still image, they are then compared to the Features from face images in a database.

29. When two faces are compared via their respective Features, a score is generated: this estimates the likelihood that the faces in both images belong to the same person, and is known as the "*Similarity Score*". The Similarity Score is typically in a range 0 to 1, where 0 indicates a perfect non-match, and 1 indicates a perfect match. Therefore, a higher Similarity Score indicates a higher likelihood of a match.

Matching

30. In order to distinguish a potential positive match from a non-match, a threshold is established, which is known as the "*Threshold Value*". The Threshold Value is a chosen Similarity Score which distinguishes between a potential match and a non-match.

31. Where the Similarity Score is above the Threshold Value, a potential match is deemed to be made. Where the Similarity Score is below the Threshold Value, no match is deemed to be made.

32. Potential matches, i.e. matches generating a Similarity Score above the Threshold Value, are retrieved from the system and may then be shown to the operator to be verified. Whether potential matches are or are not verified by an operator depends on the particular use of the AFR system. When an AFR system is being used for surveillance, it is likely that operators are shown potential matches in order to manually verify the match before taking action like arresting.

33. If potential matches are shown to the operator, it is the operator who makes the decision as to whether they do indeed match, by reviewing the two face images being compared side by side on a computer screen.

34. The fixing of the Threshold Value can create two types of errors. If the Threshold Value is low, images of different people can wrongly be considered to be matches more often;

the rate at which this happens is called the "*False Alarm Rate*". By contrast, if the Threshold Value is high, images of the same person may wrongly be considered not to be matches more often; the rate at which this happens is called the "*False Reject Rate*". The Threshold Value is often chosen in order to fix one of these two potential errors.

35. The Threshold Value is generally suggested by the manufacturer, and depends on the intended use of the AFR system. It is common to suggest setting the Threshold Value so that the False Alarm Rate is 0.1%, 0.01% or 0.0001%. Most AFR systems, however, allow the end user to change the Threshold Value to whatever they choose.

36. I have been told by the Claimant's lawyers that South Wales Police ("SWP") is able to configure its Threshold Value either for an entire watch list, or for an individual in the watch list ("user-specific threshold"). User-specific thresholds are not commonly used in operational settings. This is because setting a Threshold Value for an individual requires a large amount of data (images and videos of that person's face) to determine the appropriate value. Also, user-specific thresholds can lead to an AFR system having a discriminatory impact on individuals or certain groups of individuals.

37. The end user will typically change the Threshold Value to account for the security requirement of their application. For example, applications demanding high security, such as border control or surveillance cameras, require a Threshold Value such that the probability of two different people matching is low. Therefore, the Threshold Value must be higher. On the other hand, applications on an iPhone, such as FaceID to unlock one's iPhone, require a Threshold Value such that the probability of the iPhone user being deemed a non-match is low. Therefore, the Threshold Value must be lower.

Accuracy of AFR

38. The performance of AFR technology is affected by a number of variables:

    a) *Training datasets:* I provide information on this in the section below, 'Training of AFR'.

    b) *Pose/illumination/expression ("PIE") variations:* The accuracy of AFR technology is affected by the pose, illumination and expression of face images, including the images on which the technology is trained, the quality of database images against which images captured by an AFR camera are matched, and the quality of the images acquired from the live video feed. If an AFR system has been trained on a limited number of images with different poses or expressions, or in different levels of illumination, then it will be less able to match faces with different poses or expressions, or in different levels of illumination. For this reason, most databases used by police forces consist of mugshots where the subjects are asked to provide frontal views with neutral expression in an indoor setting where the illumination is generally controlled.

c) *Face obstructions (also referred to as "occlusion"):* The accuracy of AFR technology is affected if the faces in the live video feed, the database for comparison or the training database are covered by scarves, hats, sunglasses, scars, marks, tattoos, etc.

d) *Spoofing:* The accuracy of AFR technology can be intentionally affected by the subjects of AFR cameras disguising their faces to look like somebody else. This can be achieved simply through heavy make-up.

e) *Aging:* The accuracy of AFR technology decreases with an increase in the time gap between the capture of the database image and the image being acquired from the live AFR video feed.[2]

f) *Image quality.* The accuracy of AFR technology is affected by the use of poor quality images, either in the database for comparison, the training database or acquired from the live video feed. Poor quality images can be caused by low image resolution, motion blur, etc. Using an AFR camera on top of a police vehicle which is moving may result in a live feed of poor quality which may hinder AFR performance.

39. I have been told by the Claimant's lawyers that SWP is currently using a Bosch Mic Starlight 7000 HD (1080x1920p) camera to capture the live feed for its AFR technology. While the image size of this camera (1080x1920p) is large enough to use for AFR, the quality of the image will still be affected by the variables listed above in §38(a) to (f).

40. AFR systems deployed for running checks on two static images taken in a constrained environment – such as passports, driving licences, a police custody database – are generally more accurate than AFR systems deployed for identifying subjects from a live video feed. This is because, in the static context, the two images being compared are acquired in a controlled manner, where the subject is cooperative, the subject's pose and facial expressions are constrained, and the quality of the image is high.

41. On the other hand, where a live feed is used to identify subjects, at least one of the images (the still image obtained from the live feed) is uncontrolled; it is taken in an unconstrained environment where the subject is not necessarily cooperative, not necessarily looking at the camera, and may not even be aware that an image is being taken. These PIE variations and the poor quality of images obtained ((b) and (f) from §[38] above) are the most common causes of inaccuracy in the use of AFR systems in a live deployment.

---

2 D. Deb, L. Best-Rowden, A. K. Jain, "Face Recognition Performance Under Aging", in CVPR, Workshop on Biometrics, 2017 9 AJ 1/2.

Training of AFR

42. As described above, AFR systems rely on representing face images by Features. These Features are used to distinguish between two face images belonging to two different people, as well as match two face images belonging to the same person. Therefore, an AFR system needs to learn the best set of Features that can be extracted from a face image.

43. Typically, all state-of-the-art AFR systems, such as AFR Locate, employ a deep learning network. This is a computer program that automatically learns patterns in input data, and predicts an output. In the case of facial recognition, the "input" is a large database of face images of different individuals, and the "output" is the identity of the person. When given a large number of face images of different individuals under various PIE and occlusion conditions, the network automatically learns patterns in the Features that best represent the person (this is known as best face representation) and distinguishing patterns in face images that differentiate different people; it "learns" by saving these patterns. This learning process is called "training" of AFR.

44. Any state-of-the-art AFR system, such as AFR Locate, is trained using large face datasets. There are two main datasets which are available publicly: (1) CASIA-WebFace,[3] which contains 453,453 images of over 10,575 subjects, and (2) MS-Celeb-1M,[4] which contains around 10,000,000 images of 100,000 subjects. These datasets are created using images of celebrities which have been downloaded from the internet. The images of celebrities are found using face detector software to "crawl" the internet for images. This means that a programme will methodically browse the internet, using public search engines, automatically locating faces in images that it comes across.

45. These publicly-available datasets are not regularly updated. Further, they can be augmented with private datasets, which are considered trade secrets and not released to the public.

46. Once a deep learning network is given images from the training datasets, it will take the images and the associated identity labels, and automatically extract Features in a manner that maximises the "classification accuracy", i.e. the accuracy in predicting the identity given an unknown face image. This is different to the Similarity Score, which is a measure of accuracy of the match.

47. The accuracy of an AFR system depends to a considerable extent on the training dataset. Larger datasets, both in terms of number of subjects and number of images per subject of different ethnicity and skin color are preferred because deep learning networks are data-

[3] Accessed by http://www.cbsr.ia.ac.cn/english/Databases.asp, however at the time of writing I note that the webpage for CASIA-WebFace in particular is currently inoperative. Instead, a paper on CASIA-WebFace can be accessed by: https://arxiv.org/abs/1411.7923

[4] Accessed by: https://www.msceleb.org/

driven and learn better face representations when a large number of training images are present. However, the performance of a deep learning network will eventually plateau once no new information can be learned from additional data. A robust AFR system trained on large datasets will be robust in its ability to recognise faces in a test set, i.e. face images not included in the training set.

48. AFR systems can also suffer from training bias. For example, if only very high quality custody images are provided for training, the AFR system may not work well with poorer quality live feed footage.

49. Training bias can also be caused by any imbalance in the demographic of subjects in the training datasets, resulting in the AFR system having a high False Alarm Rate or a high False Reject Rate for that particular demographic. Therefore, the under or overrepresentation of a particular demographic in any training dataset will affect the accuracy of the system in identifying faces from that demographic. In a recent paper that I co-authored, we found that AFR software performs differently on different demographics, and that this is likely due to the demographic difference in the training dataset.[5]

50. By way of example, a particular AFR system called FaceNet was recently identified as having poor recognition performance for identifying children under the age of 18.[6] By utilising a dataset of children's faces during training, FaceNet's recognition ability has been shown to have improved significantly.[7]

51. I have been told by the Claimant's lawyers that the Claimant asked SWP for information on the dataset used by NEC to train its AFR system, and that SWP replied to say it was in contact with NEC about this request, suggesting that it did not know about the dataset. If SWP was not aware of the dataset used to train the AFR system, it would be difficult for SWP to confirm whether the technology is in fact biased. As a minimum for confirming whether an AFR system is biased, the database statistics such as the number of males to females, and different races considered, would need to be known.

Development of AFR

52. AFR technology is constantly being upgraded to handle the variety of challenges caused by PIE variations, aging, training biases, etc.

53. As AFR technology is updated, old technology becomes so out-of-date that it is no longer considered effective. For example, prior to deep learning networks, face recognition was implemented by selecting certain features to be used for recognition.

---

5 D. Deb, L. Best-Rowden, A. K. Jain, "Face Recognition Performance Under Aging", in CVPR, Workshop on Biometrics, 2017 9 **AJ 1/2.**

6 D. Deb, N. Nain and A.K. Jain, "Longitudinal Study of Child Face Recognition", in ICB, 2018 **AJ 1/3.**

7 D. Deb, N. Nain and A.K. Jain, "Longitudinal Study of Child Face Recognition", in ICB, 2018 **AJ 1/3.**

This could include measuring the distance from the eyes, the distance from the nose to the ears etc. Such technology is now outdated and no longer considered effective.

54. In academia, AFR technology is generally developed, tested and trained on datasets that are publicly available. When academic algorithms are transferred to AFR companies, they will use their resources and capabilities to better engineer and package the software for marketing. This aspect of AFR's development is often kept confidential as a trade secret.

55. All AFR systems need to reduce the likelihood of "edge cases" that occur in particular operational settings. An edge case is a scenario that has not been accounted for while designing an AFR system. This would include obstructions, PIE variations etc. One way of reducing the likelihood of edge cases is to include more variations in the training dataset.

56. The manufacturers of AFR systems are the ones who will understand the learning process of their AFR systems. End users of these systems can only speculate about how the systems are learning and evolving, but they cannot understand this process like the manufacturers.

57. To the best of my knowledge, no AFR system is without some error rates.

Other

58. Manufacturers of AFR systems will provide guidance on how to use their systems, and ongoing technical support to their customers. NEC provides this guidance and support to all its customers using AFR systems.

I confirm that I have made clear which facts and matters referred to in this report are within my own knowledge and which are not. Those that are within my own knowledge I confirm to be true. The opinions I have expressed represent my true and complete professional opinions on the matters to which they refer.

DATED THIS ..... DAY OF SEPTEMBER 2018

SIGNED

DR ANIL JAIN