PUBLIC SUBMISSION

As of: 11/22/22, 9:45 AM
Received: November 18, 2022
Status: Draft
Tracking No. lam-jysh-ou0d
Comments Due: November 21, 2022
Submission Type: Web

Docket: GSA-GSA-2022-0001
GSA Information Collections -2022

Comment On: GSA-GSA-2022-0001-0024
Agency Information Collection Activities; Proposals, Submissions, and Approvals: Equity Study on Remote Identity Proofing

Document: GSA-GSA-2022-0001-DRAFT-0008
Comment on FR Doc # 2022-20249

Submitter Information

Email: gmonetti@imageware.io
Organization: Imageware Systems

General Comment

Imageware Comments: We have kept our comments within the framework as stated above and in line with your desired comment areas as listed in Section C. Public Comments. Additional comments will follow:
1. Is this collection of information necessary? Yes, we believe it is. However, based on our experience we believe the more recent facial recognition algorithms available on the market offer superior performance to those of the past. This improved level of performance helps improve performance in accurate matching of the facial biometric captured. Having said this, we do believe and concur that this collection of information is necessary to further validate these claims and provide peace of mind for the users. Continual improvement in accuracy will only provide more benefit to public safety and national security without compromising personal information and general privacy concerns. Improvements in document scanning and capture for proofing have also improved over the years such as Real ID drivers' licenses with lots more information embedded in the card. These cards can easily be scanned with cellular devices that can capture the information for enrollment and matching.
2. Will this study have practical utility? We believe it will. The utility comes from the areas for improvement that will be identified to improve the matching accuracy, whether it be in the biometric algorithm or the Biometric match engine. We see value in collecting biometric and non-biometric information in a controlled study environment like this.
3. Will the estimate of public burden of the collection of information be accurate? We think yes. We see little to no burden in collection of the information as it can be captured as simple as taking a selfie or scanning a document like you do at a TSA checkpoint. We don't imply there will not be individuals that do find general burden in sharing their personal information for privacy reasons. Here best practices in securing information as defined by NIST and other cryptographic technologies to include quantum keying can be considered.
4. Will there be ways to enhance the quality, utility, and clarity of the information to be collected based upon valid assumptions and proven methodologies? Yes, we believe there will be. We say this as there are ways to conduct statical analysis to determine how to further refine and fine tune the biometric algorithms

for optimal performance. More recent advances in biometric fusion and neural networking show real promise to enhance the quality, utility, and clarity of the information of the information collected. What's important is to ensure your study team includes vendors like Imageware with decades of R & D experience and practical knowledge.

5. Can the study find ways in which to minimize the burden of the collection of information on those who are to respond through the use of appropriate technology collection techniques or other forms of information technology? Yes, we think the study can. The study can take advantage of the trillion-dollar investment into the commercial cellular markets and 5G. The devices that power these networks are perfect collection devices for multimodal biometric capture with very high-resolution cameras. Other collection platforms include tablets, laptops, desktop PC's and more purpose-built biometric capture devices to include touchless.

6. Other comments and considerations: We believe there are other means of improving accuracy and reducing the burden of identity information collection whether it be facial imagery or documents. This can be achieved by capturing other biometrics from across the diverse demographic study group. You can consider capturing Iris, palm, fingerprint, and voice for example then fusing the scores in a biometric engine to perform matching and scoring. You can also start investigating a Digital ID implementation using digital wallets and blockchain where the credentials are held by the individual in more a decentralized manner. Here you can capture all information from a simple QR code scan to include biometrics.

7. Consider using the recent International Biometric & Identity Association (IBIA) Study on Data Analysis of Facial Recognition Technology for a Diverse Population as reference and to help baseline a means of performance measurement and testing criteria. The full report from August 25th, 2021, can be found here. IBIA Diversity Data Analysis Unabridged FINAL.pdf. We have also uploaded it with our comments

Attachments

Imageware Comments_GSA Equity Study on Remote Identity Proofing

IBIA Diversity Data Analysis Unabridged FINAL

Date: 11/19/2022

To: GSA Equity Study on Remote Identity Proofing Representative
cc: Gerardo Cruz; Tiffany Andrews

From: Gary Monetti, Federal Sales Director Imageware Systems

Subject: Imageware Comments on Information Collection 3090-XXXX; GSA Equity Study on Remote Identity Proofing

Introduction: Imageware Systems hereby submits our comments to the Federal Register with regard to the GSA Equity Study on Remote Identity Proofing (OMB control # 3090-XXXX; Docket Number 2022-001; Sequence No. 16.)

Imageware is a global Biometric Identity Management Software company, and small business enterprise, based in San Diego California. Imageware has been in the Digital Imaging and biometric identity management business for over 35 years. The company brings much biometric and identity experience across several market sectors to include Federal, State and Local, Financial, and commercial. www.imageware.io

General Observations: It is our understanding that GSA wishes to conduct an Equity Study on Remote Identity Proofing to proactively examine the potential burden with the collection of information and technical challenges with achieving very high levels of accurate identification of individuals across several demographics such as race, ethnicity, gender, age, income, educational level, and other demographic data. We also understand you will look at non-biometric proofing checks using credentials such as Driver's License, passport, National ID card, and others. It should be noted facial images are included on these documents and can be used to match an individual facial image captured at enrollment, compared against a particular document provided or another on file remotely such as a credit bureau. We acknowledge the study will use the NIST SP-8000-63-3 Identity Assurance Level 2 (IAL2) standard to help establish a basis to check performance levels across demographic groups participating in the study. We also understand GSA is working with industry to help conduct the study with the vendors identity proofing products that are compatible with the study architecture to include IAL2 certification. This will serve as the general basis for our comments.

Imageware Comments: We have kept our comments within the framework as stated above and in line with your desired comment areas as listed in Section C. Public Comments. Additional comments will follow:

1. **Is this collection of information necessary?** Yes, we believe it is. However, based on our experience we believe the more recent facial recognition algorithms available on the market offer superior performance to those of the past. This improved level of performance helps improve performance in accurate matching of the facial biometric captured. Having said this, we do believe and concur that this collection of information is necessary to further validate these claims and provide peace of mind for the users. Continual improvement in accuracy will only provide more benefit to public safety and national security without compromising personal information and general privacy concerns. Improvements in document scanning and capture for proofing have also improved over the years such as Real ID drivers' licenses with lots more information embedded in the card. These cards can easily be scanned with cellular devices that can capture the information for enrollment and matching.
2. **Will this study have practical utility?** We believe it will. The utility comes from the areas for improvement that will be identified to improve the matching accuracy, whether it be in the biometric algorithm or the Biometric match engine. We see value in collecting biometric and non-biometric information in a controlled study environment like this.
3. **Will the estimate of public burden of the collection of information be accurate?** We think yes. We see little to no burden in collection of the information as it can be captured as simple as taking a selfie or scanning a document like you do at a TSA checkpoint. We don't imply there will not be individuals that do find general burden in sharing their personal

information for privacy reasons. Here best practices in securing information as defined by NIST and other cryptographic technologies to include quantum keying can be considered.

4. **Will there be ways to enhance the quality, utility, and clarity of the information to be collected based upon valid assumptions and proven methodologies?** Yes, we believe there will be. We say this as there are ways to conduct statical analysis to determine how to further refine and fine tune the biometric algorithms for optimal performance. More recent advances in biometric fusion and neural networking show real promise to enhance the quality, utility, and clarity of the information of the information collected. What's important is to ensure your study team includes vendors like Imageware with decades of R & D experience and practical knowledge.
5. **Can the study find ways in which to minimize the burden of the collection of information on those who are to respond through the use of appropriate technology collection techniques or other forms of information technology?** Yes, we think the study can. The study can take advantage of the trillion-dollar investment into the commercial cellular markets and 5G. The devices that power these networks are perfect collection devices for multimodal biometric capture with very high-resolution cameras. Other collection platforms include tablets, laptops, desktop PC's and more purpose-built biometric capture devices to include touchless.
6. **Other comments and considerations:** We believe there are other means of improving accuracy and reducing the burden of identity information collection whether it be facial imagery or documents. This can be achieved by capturing other biometrics from across the diverse demographic study group. You can consider capturing Iris, palm, fingerprint, and voice for example then fusing the scores in a biometric engine to perform matching and scoring. You can also start investigating a Digital ID implementation using digital wallets and blockchain where the credentials are held by the individual in more a decentralized manner. Here you can capture all information from a simple QR code scan to include biometrics.

7. **Consider using the recent International Biometric & Identity Association (IBIA) Study on Data Analysis of Facial Recognition Technology for a Diverse Population as reference and to help baseline a means of performance measurement and testing criteria.** The full report from August 25th, 2021, can be found here. IBIA Diversity Data Analysis Unabridged FINAL.pdf. We have also uploaded it with our comments

Conclusion: In conclusion, Imageware systems concurs with the effort you are about to embark on with the GSA Equity Study on Remote Identity proofing. We trust you will find some very acceptable performance levels with the more recent biometric algorithms on the market today and non-biometric capture and matching technologies. We can be available to contribute to the study should you be interested. We look forward to reviewing your final report and the areas of improvement that your report will specify.

Respectfully,
Gary Monetti

Federal Sales Director
Imageware Systems
301-514-7992
gmonetti@imageware.io

IBIA International
Biometrics+Identity
Association

Data Analysis of Facial Recognition Technology
For a Diverse Population
The International Biometrics + Identity Association (IBIA)
For Publication 8/26/2021

The International Biometrics + Identity Association (IBIA) is a strong advocate for ethical uses of technology, particularly related to identity technology. We appreciate the efforts by privacy, civil liberties, and racial justice organizations to raise important questions about how to use facial recognition technologies in a manner that promotes privacy and social justice, and we are glad that communities around the world are thinking critically about ethical use of biometric technologies. To help ground those broader discussions in a strong understanding of biometric systems and the biometric technology industry, IBIA is publishing a series of papers to help the public and lawmakers better understand how biometric technologies work, factors impacting biometric technology performance, and some of the best practices we have developed to promote ethical use of biometric technologies. In this paper, we provide an explanation of independent testing that NIST, a globally recognized facial recognition algorithm testing authority, has performed on facial recognition vendor algorithms. We seek to help a non-technical audience understand NIST testing results, and we provide an overview of publicly available NIST data on demographic effects for IBIA member company algorithms.

This analysis helps to demonstrate the following facts about facial recognition technologies:

- Modern facial recognition algorithms can often achieve better identification accuracy than humans can,[1] but the best results are often the product of humans and facial recognition technologies working in tandem.[2]
- Variations in algorithm demographic performance naturally exist.
  - Although some algorithms show statistically significant performance differences across demographic groups, NIST has found that top-performing algorithms display "undetectable" false positive error rate differentials across demographic groups.[3]
- It is important to understand variations in algorithm performance for users to make good policy choices for their particular applications.

[1] *See* Meissner, C. A. Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. Psychology, Public Policy, & Law 7, 3–35 (providing information about human face memory across demographic groups and finding that humans remember own-race faces better than faces of people who are members of other, less familiar racial groups).
[2] P. Jonathon Phillips et al., Face *recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms*, 115 PNAS 6171, 6171-76 (2018), https://www.pnas.org/content/pnas/115/24/6171.full.pdf.
[3] NIST, *FRVT Part 3: Demographic Effects*, p. 8.

Reading a DET Curve

Important NIST test results of biometric algorithm performance are shown as DET curves, which characterize the threshold-setting tradeoff between achieving a low false positive identification rate (face matched to the wrong person) vs. achieving a low false negative identification rate (face not matched at all, even when the correct face is presented). NIST builds such curves by testing algorithms at different threshold settings and observing the types of errors against NIST's known dataset. This known dataset can be segmented by demographic, thereby yielding results for variances against each dataset. An example DET curve, showing curves for two different hypothetical demographics, is below:

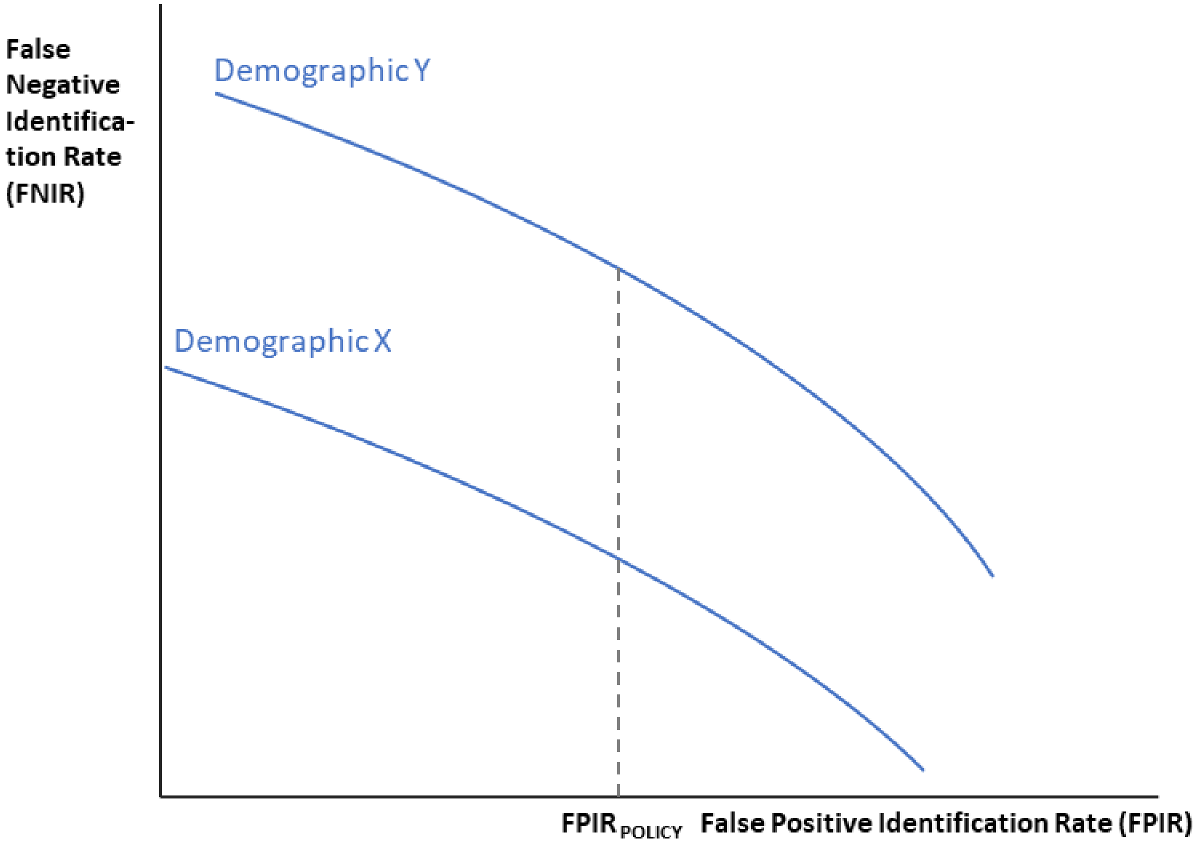

Figure 1. Generic detection error tradeoff curves for an algorithm performing biometric matching on two different demographic data sets. The dashed line illustrates the difference between the two curves at a given false positive identification rate set by policy ($FPIR_{POLICY}$).

Figure 1[4] shows two DET curves for the same algorithm when operating on two different demographic data sets, X and Y. The curves illustrate lower error rates for comparable threshold settings for Demographic X. (Generally, curves that are "down and to the left" demonstrate better results.) This explanation, along with some definitions,[5] sets us up to understand NIST demographic differential results for IBIA company algorithms in the following figures.

[4] NIST report NISTIR 8280 "Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects."

[5] FNIR = "false negative Identification rate", and FPIR = "false positive identification rate". "Identification" is searching a group of ("N") faces to see if any of them match a ("1") face image you provide (hence the term "1:N" identification or matching).

NIST FRVT Part 3 Test Results: Vendor-Specific Demographic Performance

The following NIST DET curves[6] from IBIA members Cogent (Thales), Cognitec, IDEMIA and NEC show excellent demographic differential performance. For this discussion, we show three graphs for each algorithm from left to right, one for white and black populations, one for black males and females, and one for white males and females. NIST used a mugshot dataset for these tests.

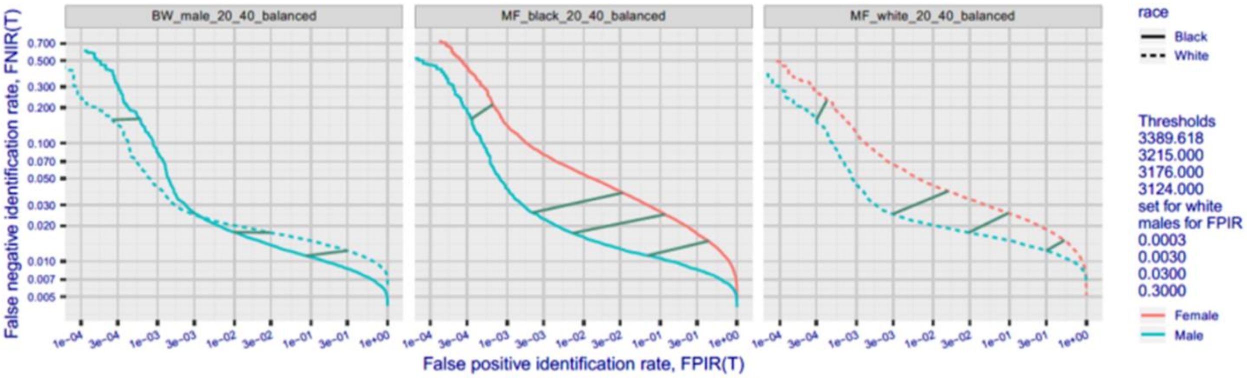

Figure 2. DET curves illustrating demographic differentials for algorithm Cogent-0.

Figure 2 represents DET curves from algorithm Cogent-0. The left-most graph shows performance for black subjects on the solid line, and white subjects on the dashed line. Note that for two of the threshold settings, results are better for black subjects. For the threshold setting that yields a false positive identification rate of 3 in 1000, there is a negligible difference in false negative identification between the demographics. The right-most two graphs show a larger demographic differential (about 0.02) between males and females of both races, with a lower error rate for males at all four selected threshold values. Referencing the left-most graph again, the true identification rate (sometimes called accuracy) is about 97% at the 3 in 1000 false negative identification threshold. For comparison, humans exhibit true identification rates of between 50% and 80% with the average around 60%. Trained face examiners, and people deemed to be super-recognizers (who make up only 2 to 3% of the population) can perform much better than average.[7] Some people, about 2% of the population afflicted with prosopagnosia, cannot recognize faces at all.

[6] NIST report NISTIR 8280 “Annex 16 : Identification error characteristics by race and sex”

[7] Dunn JD, Summersby S, Towler A, Davis JP, White D (2020). UNSW Face Test: A screening tool for super-recognizers. PLOS ONE 15(11): e0241747.

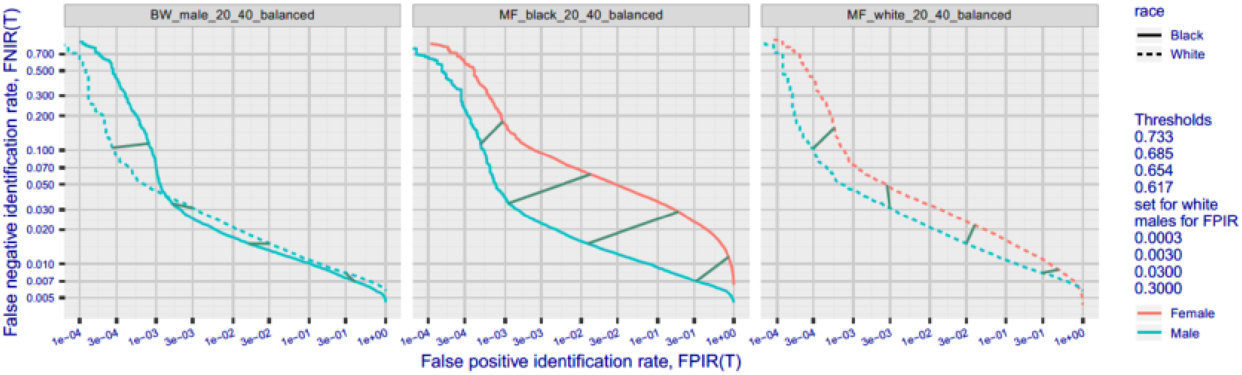

Figure 3.

DET curves illustrating demographic differentials for algorithm Cognitec-2.

Figure 3 represents DET curves from algorithm Cognitec-2. The left-most graph shows performance for black subjects on the solid line, and white subjects on the dashed line. Note that for three of the threshold settings, results are better for black subjects, and for one of the threshold settings it is better for white subjects. For the threshold setting that yields a false positive identification rate of 3 in 1000, as the curve shows, there is a negligible difference in false negative identification between the demographics. As was the case previously, the right-most two graphs show a larger demographic differential between males and females of both races, with lower error rates for males at all four threshold settings. Referencing the left-most graph again, the true identification rate (sometimes called accuracy) is about 96% at the 3 in 1000 false negative identification threshold.

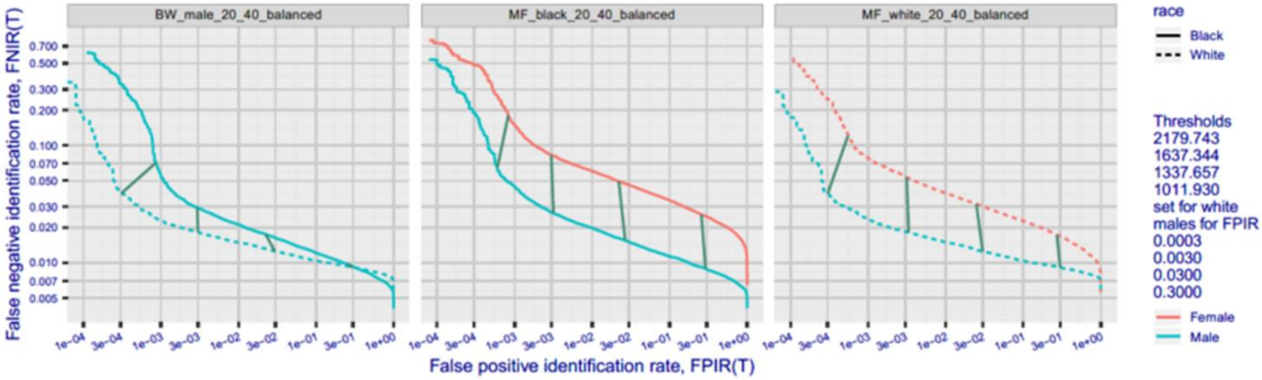

Figure 4.

DET curves illustrating demographic differentials for algorithm Idemia-4.

This set of figures represents DET curves from algorithm Idemia-4. The left-most graph shows performance for black subjects on the solid line, and white subjects on the dashed line. Note that for three threshold settings, results are better for white subjects, and for one of the threshold settings, results were about the same for black subjects and white subjects. For the threshold setting that yields a false positive identification rate of 3 in 1000, as the curve shows, there is about 0.01 difference in false negative identification between the demographics. As was the case previously, the right-most two graphs show a larger demographic differential between males and females of both races, with male error rates lower for all four threshold settings. Referencing the left-most graph again, the true identification rate (sometimes called accuracy) is about 97% at the 3 in 1000 false negative identification threshold.

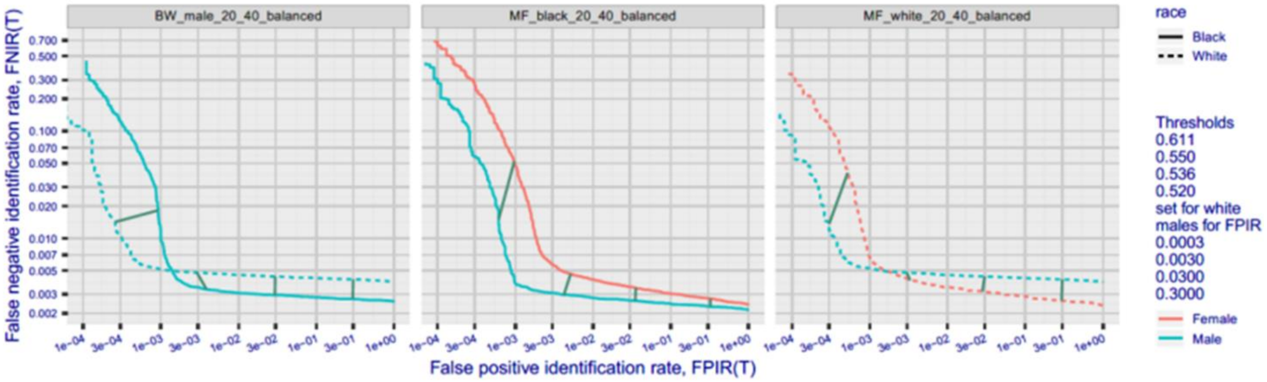

Figure 5.
DET curves illustrating demographic differentials for algorithm NEC-2.

Figure 5 represents DET curves from algorithm NEC-2. The left-most graph shows performance for black subjects on the solid line, and white subjects on the dashed line. For three threshold settings, results are better for black subjects, and for one of the threshold settings, results are better for white subjects. For the threshold setting that yields a false positive identification rate of 3 in 1000, as the curve shows, there is about 0.002 difference in false negative identification between the demographic groups. As was the case previously, the right-most two graphs show a demographic differential between males and females of both races. However, in the right-most graph, white female performance was better than white male performance for three of the four tested threshold settings. Referencing the left-most graph again, the true identification rate (sometimes called accuracy) is about 99% at the 3 in 1000 false negative identification threshold, making this one of the most accurate algorithms in this NIST test.

Summary

These data show that, for the best facial recognition algorithms, including those from IBIA members, demographic differentials are small[8] and, in some cases, actually reveal lower error rates for black subjects than for white subjects. Algorithms from the IBIA companies cited here, including companies that submitted algorithms with "undetectable" false positive error rate differentials across demographic groups,[9] are in use around the world.

The successful implementation of a facial recognition or other biometric systems depends on the choice of high-performing algorithms, best informed by the latest NIST test results, and on environmental factors, the human operators, and system configuration settings (such as threshold settings). Developing and implementing ethical use policies and best practices, including those that IBIA has developed,[10] helps to promote racial justice, privacy, and other civil rights and civil liberties. In addition, using the best algorithms in the right use cases can yield major benefits for society as a whole, including economic efficiency, security, and convenience.

IBIA is dedicated to the ethical use of biometrics and welcomes opportunities to participate in multi-stakeholder dialogues and to serve as a resource to policymakers and media outlets interested in discussing and working to address on these important topics.

For more insights from IBIA, visit www.IBIA.org To contact IBIA, email info@ibia.org.

[8] NIST, *FRVT Part 3: Demographic Effects*, p. 8 (explaining that top-performing algorithms display "undetectable" false positive error rate differentials across demographic groups).

[9] *Id.*

[10] https://www.ibia.org/resources/white-papers