NISTIR XXXX Draft

Ongoing Face Recognition Vendor Test (FRVT)

Part 1: Verification

Patrick Grother
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing

2020/03/25

National Institute of Standards and Technology
U.S. Department of Commerce

ACKNOWLEDGMENTS

The authors are grateful to staff in the NIST Biometrics Research Laboratory for infrastructure supporting rapid evaluation of algorithms.

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.

INSTITUTIONAL REVIEW BOARD

The National Institute of Standards and Technology's Research Protections Office reviewed the protocol for this project and determined it is not human subjects research as defined in Department of Commerce Regulations, 15 CFR 27, also known as the Common Rule for the Protection of Human Subjects (45 CFR 46, Subpart A).

FRVT STATUS

This report is a draft NIST Interagency Report, and is open for comment. It is the sixteenth edition of the report since the first was published in June 2017. Prior editions of this report are maintained on the FRVT website, and may contain useful information about older algorithms and datasets no longer used in FRVT.

FRVT remains open: All four tracks of the FRVT are closed to new algorithm submissions indefinitely (due to SARS-COV-19). This report will be updated as new algorithms are evaluated, as new datasets are added, and as new analyses are included. Comments and suggestions should be directed to frvt@nist.gov.

Changes since February 27, 2020:

- ▷ The report adds results algorithms from two new developers: Beijing Alleyes Technology, and the Chinese Univeristy of Hong Kong. Results for newly submitted algorithms from two other developers will appear in the next report.
- ▷ The report adds results for algorithms from thirteen returning developers: ASUSTek Computer, Aware, Cyberlink Corp, Gorilla Technology, Innovative Technology, Kakao Enterprise, Lomonosov Moscow State University, Panasonic R+D Center Singapore, Shenzhen AiMall Technology, Shenzhen Intellifusion Technologies, Synology, Tech5 SA, and Via Technologies.
- ▷ Per policy to only list results for two algorithms per developer, we have dropped results for algorithms from Aware, Cyberlink, Gorilla Technology, Kakao Enterprise, Lomonosov Moscow State University, Panasonic R+D Center Singapore, and Tech5 SA.

Changes since January 20, 2020:

- ▷ The report adds results for five new developers: Ability Enterprise (Andro Video), Chosun University, Fujitsu Research and Development Center, University of Coimbra, and Xforward AI Technology.
- ▷ The report adds results for algorithms from six returning developers: AlphaSSTG, Incode Technologies, Kneron. Shanghai Jiao Tong University, Vocord, and X-Laboratory.
- ▷ We have corrected template comparison timing numbers for algorithms submitted September 2019 to January 2020. The values reported previously were slower due to a software bug.
- ▷ We have dropped results for algorithms from Vocord and Incode per policy to only list results for two algorithms per developer.
- ▷ The FRVT 1:1 homepage has been updated with latest accuracy results.
- ▷ The FRVT 1:N homepage now includes an update to the September 2019 NIST Interagency Report 8271. The new report adds results for one-to-many search algorithms submitted to NIST from June 2019 to January 2020.

Changes since January 6, 2020:

- ▷ Section 2 has been updated to better describe the Visa and Border images. The caption for Table 10 has been updated to better relate the accuracy values to particular image comparisons.
- ▷ The report adds results for five new developers: Acer, Advance.AI, Expasoft, Netbridge Technology, and Videmo Intelligente Videoanalyse.
- ▷ The report adds results for algorithms from 7 returning developers: China Electronics Import-Export Corp, Intel Research Group, ITMO University, Neurotechnology, N-Tech Lab, Rokid, and VisionLabs.

- ▷ We have dropped results from this edition of the report per policy to only list results for two algorithms per developer: N-Tech Lab, Neurotechnology, ITMO, Visionlabs, and CEIEC.
- ▷ The FRVT homepage has been updated with latest accuracy results.

Changes since November 11, 2019:

- ▷ Table 5 has been updated to include runtime memory usage. This is the first time such a quantity has been reported. The value is the peak size of the resident set size logged during enrollment of single images.
- ▷ We have migrated summary results table to a new platform that supports sortable tables: https://pages.nist.gov/frvt/html/frvt11.html
- ▷ The report adds results for four new developers: Antheus Technologia, BioID Technologies SA, Canon Information Tech. (Beijing), Samsung S1 (listed in the tables as S1), and Taiwan AI Labs.
- ▷ The report adds results for algorithms from 13 returning developers: Anke Investments, Chunghwa Telecom, Deepglint, Institute of Information Technologies, iQIYI, Kneron, Ping An Technology, Paravision, KanKan Ai, Rokid Corporation, Shanghai Universiy - Shanghai Film Academy, Veridas Digital Authentication Solutions, and Videonetics Technology.
- ▷ We have dropped results from this edition of the report per policy to only list results for two algorithms per developer: remarkai-000, veridas-001, sensetime-001, iit-000, anke-003, and everai-002. Results for these are available in prior editions of this report linked from the FRVT page.
- ▷ We issued NIST Interagency Report 8280: FRVT Part 3: Demographics on 2019-12-19. It includes results for many of the algorithms covered by this report.

Changes since October 16, 2019:

- ▷ The report adds results for ten new developers: Ai-Union Technology, ASUSTek Computer, DiDi ChuXing Technology, Innovative Technology, Luxand, MVision, Pyramid Cyber Security + Forensic, Scanovate, Shenzhen AiMall Tech, and TUPU Technology.
- ▷ The report adds results for 12 returning developers: CTBC Bank Glory Gorilla Technology Guangzhou Pixel Solutions Imagus Technology Incode Technologies Lomonosov Moscow State University Rank One Computing Samtech InfoNet Shanghai Ulucu Electronics Technology Synesis, and Winsense.
- ▷ We have dropped results from this edition of the report per policy to only list results for two algorithms per developer: glory-000, gorilla-002, incode-003, rankone-006, and synesis-004.
- ▷ Results for five recently submitted algorithms will appear in the next report.

Changes since September 11, 2019:

- ▷ The report adds results for five new participants: Awidit Systems (Awiros), Momemtum Digital (Sertis), Trueface AI, Shanghai Jiao Tong University, and X-Laboratory.
- ▷ The reports adds results for five new algorithms from returning developers: Cyberlink, Hengrui AI Technology, Idemia, Panasonic R+D Singapore, and Tevian. This causes three algorithm, to be de-listed from the report per policy to list results for two algorithms per developer.

Changes since July 31 2019:

- ▷ The HTML table on the FRVT 1:1 homepage has been updated to include a column for cross-domain Visa-Border verification. Results for this new dataset appeared in the July 29 report under the name "CrossEV" - these are now renamed "Visa-Border".
- ▷ The FRVT 1:1 homepage lists algorithms according to lowest mean rank accuracy:
  Rank($FNMR_{VISA}$ at FMR = 0.000001) +
  Rank($FNMR_{VISA\text{-}BORDER}$ at FMR = 0.000001) +
  Rank($FNMR_{MUGSHOT}$ at FMR = 0.00001 after 14 years) +
  Rank($FNMR_{WILD}$ at FMR = 0.00001)
  This ordering rewards high accuracy across all datasets.
- ▷ The main results in Table 10 is now in landscape format to accomodate extra columns for the Visa-Border set, and mugshot comparisons after at least 12 years.
- ▷ The report adds results for nine new participants: Alpha SSTG, Intel Research, ULSee, Chungwa Telecon, iSAP Solution, Rokid, Shenzhen EI Networks, CSA Intellicloud, Shenzhen Intellifusion Technologies.
- ▷ The reports adds results for six new algorithms from returning developers: Innovatrics, Dahua Technology, Tech5 SA, Intellivision, Nodeflux and Imperial College, London. One algorithm, from Imperial has been retired, per policy to list results for two algorithms per developer.
- ▷ The cross-country false match rate heatmaps starting from Figure **??** have been replotted to reveal more structure by listing countries by region instead of alphabetically.
- ▷ The next version of this report will be posted around October 18, 2019.

Changes since July 3 2019:

- ▷ The HTML table on the FRVT 1:1 homepage has been updated to list the 20 most accurate developers rather than algorithms, choosing the most accurate algorithm from each developer based on visa and mugshot results. Also, the algorithms are ordered in terms of lowest mean rank across mugshot, visa and wild datasets, rewarding broad accuracy over a good result on one particular dataset.
- ▷ This report includes results for a new dataset - see the column labelled "visa-border" in Table 5. It compares a new set of high quality visa-like portraits with a set webcam border-crossing photos that exhibit moderately poor pose variations and background illumination. The two new sets are described in sections 2.3 and 2.4. The comparisons are "cross-domain" in that the algorithm must compare "visa" and "wild" images. Results for other algorithms will be added in future reports as they become available.
- ▷ This report adds results for algorithms from 9 developers submitted in early July 2019. These are from 3DiVi, Camvi, EverAI-Paravision, Facesoft, Farbar (F8), Institute of Information Technologies, Shanghai U. Film Academy, Via Technologies, and Ulucu Electronics Tech. Six of these are new participants.
- ▷ Several other algorithms have been submitted and are being evaluated. Results will be released in the next report, scheduled for September 5. That report will include results for new datasets.
- ▷ Older algorithms from Everai, Camvi and 3DiVi, have been retired, per the policy to list only two algorithms per developer.

Changes since June 20 2019:

- ▷ This report adds results for algorithms from 18 developers submitted in early June 2019. These are from CTBC Bank, Deep Glint, Thales Cogent, Ever AI Paravision, Gorilla Technology, Imagus, Incode, Kneron, N-Tech Lab, Neurotechnology, Notiontag Technologies, Star Hybrid, Videonetics, Vigilant Solutions, Winsense, Anke Investments, CEIEC, and DSK. Nine of these are new participants.
- ▷ Several other algorithms have been submitted and are being evaluated. Results will be released in the next report, scheduled for August 1.
- ▷ Older algorithms from Everai, Thales Cogent, Gorilla Technology, Incode, Neurotechnology, N-Tech Lab and Vigilant Solutions have been retired, per the policy to list only two algorithms per developer.

Changes since April 2019:

- ▷ This report adds results for nine algorithms from nine developers submitted in early June 2019. These are from Tencent Deepsea, Hengrui, Kedacom, Moontime, Guangzhou Pixel, Rank One Computing, Synesis, Sensetime and Vocord.
- ▷ Another 23 algorithms have been submitted and are being evaluated. Results will be released in the next report, scheduled for July 3.
- ▷ Older algorithms for Rank One, Synesis, and Vocord have been retired, per the policy to list only two algorithms per developer.

Changes since February 2019:

- ▷ This report adds results for 49 algorithms from 42 developers submitted in early March 2019.
- ▷ This report omits results for algorithms that we retired. We retired for three reasons: 1. The developer submitted a new algorithm, and we only list two. 2. The algorithm needs a GPU, and we no longer allow GPU-based algorithms. 3. Inoperable algorithms.
- ▷ Previous results for retired algorithms are available in older editions of this report linked here.
- ▷ The mugshot database used from February 2017 to January 2019 has been replaced with an extract of the mugshot database documented in NIST Interagency Report 8238, November 2018. The new mugshot set is described in section 2.5 and is adopted because:
  - ▷▷ It has much better identity label integrity, so that false non-match rates are substantially lower than those reported in FRVT 1:1 reports to date - see Figure 44.
  - ▷▷ It includes images collected over a 17 year period such that ageing can be much better characterized - - see Figure 166.
- ▷ Using the new mugshot database, Figure 166 shows accuracy for four demographic groups identified in the biographic metadata that accompanies the data: black females, black males, white females and white males.
- ▷ The report adds Figure 12 with results for the twenty human-difficult pairs used in the May 2018 paper *Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms* by Phillips et al. [1].
- ▷ The report uses an update to the wild image database that corrects some ground truth labels.
- ▷ Some results for the child exploitation database are not complete. They are typically updated less frequently than for other image sets.

Contents

List of Tables

List of Figures

	Developer	Short	Seq.	Validation	Config[1]	Template			Comparison Time (ns)[4]	
	Name	Name	Num.	Date	Data (KB)	Memory (MB)[2]	Size (B)	Time (ms)[3]	Genuine	Impostor
1	3Divi	3divi	003	2018-10-09	191636	^{39}358	^{197}4096 ± 0	^{128}650 ± 90	^{22}627 ± 11	^{26}623 ± 32
2	3Divi	3divi	004	2019-07-22	263670	^{50}430	^{102}2048 ± 0	^{205}984 ± 131	^{40}794 ± 35	^{42}801 ± 40
3	ADVANCE.AI	advance	002	2019-12-19	257173	^{31}295	^{65}2048 ± 0	^{169}811 ± 2	^{54}987 ± 10	^{53}988 ± 45
4	ASUSTek Computer Inc	asusaics	000	2019-10-24	257418	^{89}605	^{89}2048 ± 0	^{76}484 ± 13	^{157}5455 ± 78	^{158}5422 ± 112
5	ASUSTek Computer Inc	asusaics	001	2020-02-25	257418	^{84}595	^{199}4096 ± 0	^{181}842 ± 17	^{171}8618 ± 42	^{171}8638 ± 136
6	Ability Enterprise Co. Ltd - Andro Video	andro	000	2020-02-03	128502	^{41}360	^{83}2048 ± 0	^{46}308 ± 1	^{144}3580 ± 32	^{144}3609 ± 60
7	Acer Incorporated	acer	000	2020-01-08	109735	^{62}478	^{99}2048 ± 0	^{24}222 ± 0	^{61}1065 ± 40	^{70}1109 ± 35
8	Adera Global PTE Ltd	adera	001	2019-06-17	0	^{24}190	^{180}2560 ± 0	^{9}97 ± 0	^{96}1604 ± 71	^{96}1649 ± 56
9	Ai First	aifirst	001	2019-11-21	224157	^{63}485	^{64}2048 ± 0	^{105}587 ± 2	^{67}1099 ± 14	^{69}1087 ± 45
10	AiUnion Technology Co Ltd	aiunionface	000	2019-10-22	241642	^{45}402	^{104}2048 ± 0	^{124}637 ± 13	^{62}1072 ± 19	^{68}1080 ± 47
11	Alchera Inc	alchera	000	2019-03-01	258450	^{91}614	^{86}2048 ± 0	^{106}587 ± 13	^{135}3189 ± 32	^{135}3031 ± 142
12	Alchera Inc	alchera	000	2019-03-01	174013	^{61}473	^{79}2048 ± 0	^{119}627 ± 11	^{137}3342 ± 81	^{137}3243 ± 47
13	Alivia / Innovation Sys	isystems	001	2018-06-12	274621	^{143}1091	^{63}2048 ± 0	^{39}291 ± 9	^{14}557 ± 16	^{16}564 ± 22
14	Alivia / Innovation Sys	isystems	002	2018-10-18	358984	^{174}1595	^{115}2048 ± 0	^{174}822 ± 8	^{37}749 ± 31	^{29}632 ± 28
15	AllGoVision	allgovision	000	2019-03-01	172509	^{81}561	^{125}2048 ± 0	^{58}384 ± 8	^{192}29903 ± 406	^{193}29735 ± 194
16	AlphaSSTG	alphaface	001	2019-09-03	259849	^{77}527	^{92}2048 ± 0	^{114}612 ± 1	^{55}1008 ± 10	^{56}1002 ± 19
17	AlphaSSTG	alphaface	002	2020-02-20	768995	^{165}1434	^{112}2048 ± 0	^{120}628 ± 2	^{46}945 ± 25	^{48}935 ± 17
18	Amplified Group	amplifiedgroup	001	2019-03-01	0	^{8}81	^{32}866 ± 2	^{6}93 ± 0	^{197}57803 ± 4210	^{197}56365 ± 1196
19	Anke Investments	anke	004	2019-06-27	349388	^{98}706	^{157}2056 ± 0	^{118}625 ± 1	^{23}633 ± 22	^{28}632 ± 34
20	Anke Investments	anke	005	2019-11-21	328553	^{147}1134	^{159}2056 ± 0	^{107}590 ± 2	^{28}685 ± 19	^{32}687 ± 26
21	Antheus Technologia Ltda	antheus	000	2019-12-05	119453	^{15}116	^{26}520 ± 0	^{10}109 ± 1	^{165}6901 ± 268	^{164}6936 ± 103
22	AnyVision	anyvision	002	2018-01-31	662659	^{59}468	^{39}1024 ± 0	^{28}248 ± 0	^{199}74069 ± 188	^{199}74019 ± 198
23	AnyVision	anyvision	004	2018-06-15	401001	^{145}1102	^{36}1024 ± 0	^{53}355 ± 1	^{110}1891 ± 51	^{105}1829 ± 85
24	Aware	aware	004	2019-03-01	427829	^{180}1820	^{172}2084 ± 0	^{191}900 ± 10	^{84}1279 ± 50	^{86}1287 ± 100
25	Aware	aware	005	2020-02-27	300017	^{153}1265	^{56}1572 ± 0	^{188}886 ± 23	^{92}1475 ± 63	^{90}1427 ± 115
26	Awidit Systems	awiros	001	2019-09-23	15499	^{11}88	^{17}512 ± 0	^{8}97 ± 6	^{64}1079 ± 44	^{62}1050 ± 45
27	Ayonix	ayonix	000	2017-06-22	58505	^{5}69	^{41}1036 ± 0	^{2}18 ± 2	^{21}621 ± 23	^{25}620 ± 26
28	Beijing Alleyes Technology Co Ltd	alleyes	000	2020-03-09	507636	^{120}857	^{111}2048 ± 0	^{164}784 ± 1	^{86}1298 ± 34	^{87}1303 ± 51
29	Beijing Vion Technology Inc	vion	000	2018-10-19	228219	^{66}498	^{148}2052 ± 0	^{51}333 ± 1	^{194}39839 ± 3561	^{191}26830 ± 2241
30	BioID Technologies SA	bioidtechswiss	000	2019-11-15	758466	^{134}1039	^{10}256 ± 0	^{122}630 ± 2	^{193}34416 ± 137	^{194}34403 ± 126
31	Bitmain	bitmain	001	2018-10-17	287734	^{17}148	^{1}64 ± 0	^{71}444 ± 88	^{109}1887 ± 31	^{107}1877 ± 26
32	CSA IntelliCloud Technology	intellicloudai	001	2019-08-13	220831	^{95}655	^{138}2048 ± 0	^{74}468 ± 2	^{59}1056 ± 4	^{63}1051 ± 72
33	CTBC Bank Co Ltd	ctbcbank	000	2019-06-28	257208	^{82}570	^{87}2048 ± 0	^{98}568 ± 43	^{142}3551 ± 87	^{153}4805 ± 209
34	CTBC Bank Co Ltd	ctbcbank	001	2019-10-28	275511	^{86}603	^{130}2048 ± 0	^{130}652 ± 35	^{149}3926 ± 45	^{150}3924 ± 56
35	Camvi Technologies	camvitech	002	2018-10-19	236278	^{103}737	^{38}1024 ± 0	^{142}677 ± 7	^{19}612 ± 26	^{20}603 ± 20
36	Camvi Technologies	camvitech	004	2019-07-12	280733	^{127}919	^{116}2048 ± 0	^{159}759 ± 10	^{47}948 ± 40	^{50}963 ± 31
37	Canon Information Technology (Beijing) Co Ltd	cib	000	2019-12-11	340288	^{80}557	^{110}2048 ± 0	^{207}993 ± 40	^{189}24340 ± 60	^{189}25972 ± 97
38	China Electronics Import-Export Corp	ceiec	002	2019-06-12	269063	^{47}426	^{80}2048 ± 0	^{115}612 ± 17	^{116}2188 ± 57	^{121}2301 ± 56
39	China Electronics Import-Export Corp	ceiec	003	2020-01-06	260371	^{49}430	^{81}2048 ± 0	^{171}817 ± 4	^{118}2256 ± 38	^{118}2241 ± 54
40	China University of Petroleum	upc	001	2019-06-05	0	^{139}1077	^{43}1052 ± 0	^{95}551 ± 15	^{134}3114 ± 44	^{136}3165 ± 97
41	Chinese Univeristy of Hong Kong	cuhkee	001	2020-03-18	787853	^{191}2515	^{146}2052 ± 0	^{203}977 ± 31	^{131}2719 ± 60	^{131}2783 ± 56
42	Chosun University	chosun	000	2020-02-12	167093	^{16}136	^{74}2048 ± 0	^{70}441 ± 1	^{52}983 ± 20	^{52}983 ± 29
43	Chunghwa Telecom Co. Ltd	chtface	001	2019-08-06	94088	^{36}355	^{119}2048 ± 0	^{23}212 ± 10	^{119}2256 ± 26	^{119}2252 ± 91
44	Chunghwa Telecom Co. Ltd	chtface	002	2019-12-07	363153	^{144}1100	^{98}2048 ± 0	^{103}584 ± 14	^{120}2264 ± 26	^{117}2234 ± 103

	Notes
1	The configuration size does not capture static data included in libraries. We do not count these because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The memory usage is the peak resident set size reported by the ps system call during template generation.
3	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, for GPU-enabled implementations, NVidia Tesla K40.
4	The comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts 1ns clock ticks. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 1: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

	Developer	Short	Seq.	Validation	Config[1]	Template			Comparison Time (ns)[4]	
	Name	Name	Num.	Date	Data (KB)	Memory (MB)[2]	Size (B)	Time (ms)[3]	Genuine	Impostor
45	Cognitec Systems GmbH	cognitec	000	2018-10-19	474759	[65]495	[145]2052 ± 0	[25]224 ± 1	[148]3835 ± 108	[146]3782 ± 83
46	Cognitec Systems GmbH	cognitec	001	2019-03-01	476809	[67]498	[151]2052 ± 0	[40]293 ± 17	[152]4253 ± 59	[151]4102 ± 167
47	Cyberextruder	cyberex	001	2017-08-02	121211	[21]178	[9]256 ± 0	[190]893 ± 25	[65]1083 ± 16	[67]1079 ± 19
48	Cyberextruder	cyberex	002	2018-01-30	168909	[26]194	[66]2048 ± 0	[85]532 ± 6	[106]1803 ± 14	[103]1779 ± 22
49	Cyberlink Corp	cyberlink	003	2019-10-07	470949	[114]824	[144]2052 ± 0	[67]423 ± 1	[122]2366 ± 38	[123]2373 ± 45
50	Cyberlink Corp	cyberlink	004	2020-02-27	340894	[135]1046	[203]4140 ± 0	[150]712 ± 1	[147]3693 ± 51	[148]3898 ± 71
51	DSK	dsk	000	2019-06-28	11967	[27]252	[19]512 ± 0	[44]304 ± 47	[167]7152 ± 115	[165]7134 ± 111
52	Dahua Technology Co Ltd	dahua	003	2019-08-14	605337	[197]3739	[118]2048 ± 0	[86]535 ± 2	[25]648 ± 20	[27]631 ± 34
53	Dahua Technology Co Ltd	dahua	004	2019-12-18	832455	[200]4885	[93]2048 ± 0	[153]735 ± 3	[35]730 ± 25	[34]707 ± 44
54	Deepglint	deepglint	001	2019-06-21	569802	[126]917	[194]4096 ± 0	[151]721 ± 4	[146]3680 ± 35	[143]3517 ± 182
55	Deepglint	deepglint	002	2019-11-15	459642	[175]1614	[191]4096 ± 0	[141]677 ± 2	[178]13633 ± 87	[176]12905 ± 440
56	Dermalog	dermalog	005	2018-02-02	0	[38]357	[2]128 ± 0	[13]130 ± 11	[9]499 ± 22	[10]500 ± 22
57	Dermalog	dermalog	006	2018-10-18	0	[132]970	[3]128 ± 0	[84]532 ± 12	[10]506 ± 23	[7]459 ± 23
58	DiDi ChuXing Technology Co	didiglobalface	001	2019-10-23	259849	[76]527	[91]2048 ± 0	[113]612 ± 1	[51]973 ± 20	[54]988 ± 20
59	Digital Barriers	barriers	002	2019-03-01	83002	[184]1930	[161]2056 ± 0	[21]209 ± 11	[176]13409 ± 228	[177]13267 ± 206
60	Expasoft LLC	expasoft	000	2020-01-06	15341	[12]100	[135]2048 ± 0	[4]68 ± 0	[104]1779 ± 26	[102]1757 ± 97
61	FaceSoft Ltd	facesoft	000	2019-07-10	370120	[110]796	[140]2048 ± 0	[140]675 ± 18	[117]2239 ± 28	[120]2277 ± 96
62	FarBar Inc	f8	001	2019-07-11	272977	[154]1276	[77]2048 ± 0	[175]822 ± 39	[179]15262 ± 139	[179]15277 ± 212
63	Fujitsu Research and Development Center Co Ltd	fujitsulab	000	2020-02-04	0	[55]453	[21]512 ± 0	[66]419 ± 1	[145]3613 ± 37	[145]3621 ± 29
64	Gemalto Cogent	cogent	003	2019-03-01	698290	[166]1445	[33]973 ± 0	[199]952 ± 0	[175]12496 ± 75	[174]11822 ± 163
65	Gemalto Cogent	cogent	004	2019-06-14	722919	[137]1059	[61]1983 ± 0	[206]987 ± 50	[180]15536 ± 75	[180]15964 ± 708
66	Glory Ltd	glory	001	2018-06-08	0	[157]1331	[58]1726 ± 0	[61]393 ± 2	[173]9607 ± 128	[173]9539 ± 182
67	Glory Ltd	glory	002	2019-11-12	0	[133]982	[175]2106 ± 0	[109]594 ± 3	[164]6787 ± 85	[161]6551 ± 249
68	Gorilla Technology	gorilla	004	2019-11-04	186952	[118]853	[179]2192 ± 0	[59]389 ± 3	[132]2768 ± 44	[129]2745 ± 44
69	Gorilla Technology	gorilla	005	2020-03-11	100684	[93]629	[178]2192 ± 0	[63]407 ± 3	[129]2678 ± 42	[130]2770 ± 112
70	Guangzhou Pixel Solutions Co Ltd	pixelall	002	2019-06-06	0	[33]342	[181]2560 ± 0	[20]191 ± 1	[80]1223 ± 56	[81]1230 ± 47
71	Guangzhou Pixel Solutions Co Ltd	pixelall	003	2019-10-15	0	[121]865	[182]2560 ± 0	[147]699 ± 8	[76]1174 ± 28	[73]1139 ± 68
72	Hengrui AI Technology Ltd	hr	001	2019-06-04	346156	[176]1682	[168]2057 ± 0	[135]665 ± 3	[181]17816 ± 260	[181]17878 ± 464
73	Hengrui AI Technology Ltd	hr	001	2019-10-08	390059	[158]1337	[169]2057 ± 0	[194]908 ± 3	[187]22530 ± 416	[187]21651 ± 533
74	Hikvision Research Institute	hik	001	2019-03-01	667866	[202]6597	[49]1408 ± 0	[129]651 ± 0	[8]488 ± 19	[8]477 ± 22
75	ID3 Technology	id3	003	2018-10-05	265951	[35]354	[12]264 ± 0	[47]316 ± 19	[87]1330 ± 25	[88]1354 ± 28
76	ID3 Technology	id3	004	2019-03-01	171526	[90]613	[11]264 ± 0	[90]541 ± 11	[71]1135 ± 23	[78]1156 ± 32
77	ITMO University	itmo	006	2019-03-01	599187	[168]1489	[177]2121 ± 0	[170]814 ± 1	[191]26154 ± 148	[190]26217 ± 260
78	ITMO University	itmo	007	2020-01-06	415979	[187]2199	[122]2048 ± 0	[154]741 ± 2	[125]2551 ± 50	[127]2529 ± 80
79	Idemia	Idemia	004	2019-03-01	406924	[68]501	[15]352 ± 0	[45]306 ± 5	[159]5592 ± 518	[159]5533 ± 426
80	Idemia	Idemia	005	2019-10-11	509824	[92]618	[31]588 ± 0	[81]514 ± 15	[163]6657 ± 54	[162]6616 ± 53
81	Imagus Technology Pty Ltd	imagus	000	2019-06-19	183453	[58]466	[132]2048 ± 0	[68]425 ± 24	[72]1145 ± 25	[100]1718 ± 63
82	Imagus Technology Pty Ltd	imagus	001	2019-10-22	282680	[115]826	[103]2048 ± 0	[167]807 ± 29	[57]1045 ± 22	[47]934 ± 45
83	Imperial College London	imperial	000	2019-03-01	370120	[111]796	[101]2048 ± 0	[137]669 ± 1	[114]2130 ± 32	[113]2052 ± 100
84	Imperial College London	imperial	002	2019-08-28	472327	[181]1826	[68]2048 ± 0	[99]569 ± 1	[121]2278 ± 90	[114]2131 ± 44
85	Incode Technologies Inc	incode	005	2019-10-17	256242	[142]1089	[97]2048 ± 0	[97]566 ± 1	[100]1695 ± 38	[97]1663 ± 85
86	Incode Technologies Inc	incode	006	2020-02-20	266095	[113]814	[71]2048 ± 0	[75]472 ± 1	[105]1788 ± 41	[104]1798 ± 59
87	Innovative Technology Ltd	innovativetechnologyltd	001	2019-10-22	177232	[32]341	[108]2048 ± 0	[69]433 ± 7	[108]1877 ± 42	[108]1924 ± 97
88	Innovative Technology Ltd	innovativetechnologyltd	002	2020-02-26	173939	[124]912	[117]2048 ± 0	[133]661 ± 2	[107]1841 ± 50	[106]1857 ± 59

Notes	
1	The configuration size does not capture static data included in libraries. We do not count these because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The memory usage is the peak resident set size reported by the ps system call during template generation.
3	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, for GPU-enabled implementations, NVidia Tesla K40.
4	The comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts 1ns clock ticks. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 2: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

	Developer	Short	Seq.	Validation	Config[1]	Template			Comparison Time (ns)[4]	
	Name	Name	Num.	Date	Data (KB)	Memory (MB)[2]	Size (B)	Time (ms)[3]	Genuine	Impostor
89	Innovatrics	innovatrics	004	2018-10-19	0	[161] 1367	[44] 1076 ± 0	[60] 391 ± 0	[170] 8573 ± 274	[169] 7929 ± 244
90	Innovatrics	innovatrics	006	2019-08-13	0	[146] 1107	[27] 538 ± 0	[173] 820 ± 5	[160] 5855 ± 204	[156] 5266 ± 118
91	Institute of Information Technologies	iitvision	001	2019-07-05	269176	[108] 788	[106] 2048 ± 0	[146] 699 ± 4	[60] 1060 ± 48	[65] 1074 ± 54
92	Institute of Information Technologies	iitvision	002	2019-12-04	259579	[102] 731	[128] 2048 ± 0	[82] 514 ± 1	[56] 1023 ± 7	[57] 1011 ± 66
93	Intel Research Group	intelresearch	000	2019-07-08	388229	[189] 2406	[69] 2048 ± 0	[192] 902 ± 6	[153] 4800 ± 152	[152] 4561 ± 97
94	Intel Research Group	intelresearch	001	2020-01-14	353997	[164] 1433	[114] 2048 ± 0	[144] 682 ± 4	[143] 3553 ± 57	[141] 3462 ± 161
95	Intellivision	intellivision	001	2017-10-10	43692	[7] 74	[164] 2056 ± 0	[3] 62 ± 2	[126] 2573 ± 91	[128] 2544 ± 38
96	Intellivision	intellivision	002	2019-08-23	43692	[9] 81	[158] 2056 ± 0	[49] 322 ± 1	[177] 13525 ± 134	[175] 12782 ± 278
97	Is It You	isityou	000	2017-06-26	48010	[13] 110	[206] 19200 ± 0	[11] 113 ± 5	[201] 237517 ± 1318	[201] 237374 ± 1279
98	Kakao Enterprise	kakao	002	2019-06-19	479406	[87] 603	[141] 2048 ± 0	[155] 747 ± 6	[102] 1720 ± 62	[99] 1715 ± 83
99	Kakao Enterprise	kakao	003	2020-02-26	414379	[179] 1754	[96] 2048 ± 0	[186] 878 ± 3	[113] 2128 ± 34	[115] 2134 ± 60
100	Kedacom International Pte	kedacom	000	2019-06-03	245292	[205] 23574	[13] 292 ± 0	[79] 506 ± 3	[27] 684 ± 14	[31] 682 ± 16
101	Kneron Inc	kenron	003	2019-07-01	58366	[22] 188	[94] 2048 ± 0	[36] 281 ± 3	[156] 5237 ± 63	[157] 5274 ± 99
102	Kneron Inc	kenron	005	2020-02-21	375374	[56] 457	[72] 2048 ± 0	[83] 518 ± 2	[111] 1922 ± 11	[109] 1926 ± 20
103	Lomonosov Moscow State University	intsysmsu	001	2019-10-22	384409	[109] 789	[95] 2048 ± 0	[116] 614 ± 2	[20] 621 ± 8	[22] 611 ± 31
104	Lomonosov Moscow State University	intsysmsu	002	2020-03-12	765921	[107] 786	[40] 1024 ± 0	[108] 593 ± 1	[13] 549 ± 25	[15] 548 ± 29
105	Lookman Electroplast Industries	lookman	002	2018-06-13	138200	[203] 16518	[29] 548 ± 0	[16] 173 ± 1	[18] 610 ± 19	[23] 612 ± 22
106	Lookman Electroplast Industries	lookman	004	2019-06-03	244775	[204] 23548	[28] 548 ± 0	[80] 507 ± 5	[44] 871 ± 29	[46] 878 ± 29
107	Luxand Inc	luxand	000	2019-11-07	0	[160] 1366	[42] 1040 ± 0	[65] 407 ± 23	[43] 828 ± 28	[44] 828 ± 32
108	MVision	mvision	001	2019-11-12	227502	[100] 723	[23] 512 ± 0	[145] 691 ± 21	[70] 1123 ± 40	[76] 1154 ± 38
109	Megvii/Face++	megvii	001	2018-06-15	1361523	[163] 1426	[137] 2048 ± 0	[91] 543 ± 0	[155] 5228 ± 32	[155] 5252 ± 60
110	Megvii/Face++	megvii	002	2018-10-19	1809564	[183] 1879	[201] 4100 ± 0	[125] 644 ± 0	[196] 50630 ± 183	[196] 47591 ± 716
111	MicroFocus	microfocus	001	2018-06-13	104524	[23] 190	[7] 256 ± 0	[32] 264 ± 18	[1] 215 ± 8	[1] 217 ± 10
112	MicroFocus	microfocus	002	2018-10-17	96288	[19] 176	[6] 256 ± 0	[30] 259 ± 18	[2] 337 ± 34	[2] 230 ± 25
113	Momentum Digital Co Ltd	sertis	000	2019-10-07	265572	[48] 427	[70] 2048 ± 0	[157] 754 ± 0	[93] 1497 ± 29	[94] 1582 ± 38
114	Moontime Smart Technology	mt	000	2019-06-03	372169	[136] 1056	[142] 2049 ± 0	[152] 724 ± 12	[99] 1678 ± 47	[95] 1614 ± 85
115	N-Tech Lab	ntech	007	2019-06-25	2509686	[201] 5070	[186] 3348 ± 0	[165] 792 ± 3	[78] 1209 ± 59	[84] 1267 ± 65
116	N-Tech Lab	ntech	008	2020-01-06	1138002	[193] 2707	[48] 1300 ± 0	[96] 556 ± 1	[66] 1095 ± 45	[61] 1049 ± 51
117	Netbridge Technology Incoporation	netbridgetech	001	2020-01-08	133108	[72] 508	[192] 4096 ± 0	[5] 85 ± 1	[172] 9280 ± 74	[172] 9446 ± 512
118	Neurotechnology	neurotech	006	2019-06-26	525541	[170] 1538	[20] 512 ± 0	[143] 678 ± 56	[11] 513 ± 14	[14] 535 ± 26
119	Neurotechnology	neurotech	008	2020-01-08	119718	[96] 694	[8] 256 ± 0	[52] 339 ± 0	[5] 467 ± 19	[9] 486 ± 26
120	Nodeflux	nodeflux	001	2019-03-01	262553	[42] 363	[113] 2048 ± 0	[27] 247 ± 1	[136] 3242 ± 81	[138] 3255 ± 93
121	Nodeflux	nodeflux	002	2019-08-13	774668	[57] 466	[139] 2048 ± 0	[149] 708 ± 4	[140] 3475 ± 62	[140] 3408 ± 143
122	NotionTag Technologies Private Limited	notiontag	000	2019-06-12	92753	[75] 525	[30] 584 ± 0	[93] 548 ± 64	[195] 44672 ± 269	[195] 44593 ± 358
123	Panasonic R+D Center Singapore	psl	003	2019-10-01	1159643	[199] 3960	[176] 2120 ± 0	[183] 865 ± 3	[58] 1052 ± 14	[59] 1025 ± 51
124	Panasonic R+D Center Singapore	psl	004	2020-03-03	771727	[188] 2257	[185] 3144 ± 0	[179] 839 ± 3	[26] 680 ± 34	[30] 670 ± 22
125	Paravision (EverAI)	everai paravision	003	2019-07-01	539802	[150] 1225	[193] 4096 ± 0	[139] 674 ± 4	[32] 699 ± 20	[36] 713 ± 47
126	Paravision (EverAI)	everai paravision	004	2019-12-11	556670	[172] 1572	[190] 4096 ± 0	[176] 829 ± 2	[36] 737 ± 31	[38] 718 ± 38
127	Pyramid Cyber Security + Forensic (P) Ltd	pyramid	000	2019-11-04	372608	[112] 804	[155] 2056 ± 0	[102] 583 ± 2	[166] 7147 ± 59	[167] 7586 ± 425
128	Rank One Computing	rankone	007	2019-06-03	0	[4] 67	[5] 165 ± 0	[26] 245 ± 5	[29] 688 ± 20	[19] 601 ± 16
129	Rank One Computing	rankone	008	2019-11-12	0	[6] 70	[4] 165 ± 0	[34] 272 ± 3	[38] 750 ± 20	[24] 613 ± 23
130	Realnetworks Inc	realnetworks	002	2019-02-28	95328	[43] 370	[60] 1848 ± 0	[29] 250 ± 2	[85] 1285 ± 17	[82] 1247 ± 42
131	Realnetworks Inc	realnetworks	003	2019-06-12	95334	[34] 345	[59] 1848 ± 0	[18] 177 ± 10	[94] 1516 ± 29	[92] 1522 ± 60
132	Remark Holdings	remarkai	001	2019-03-01	241857	[101] 730	[150] 2052 ± 0	[177] 831 ± 6	[81] 1229 ± 20	[43] 805 ± 56

Notes	
1	The configuration size does not capture static data included in libraries. We do not count these because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The memory usage is the peak resident set size reported by the ps system call during template generation.
3	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, for GPU-enabled implementations, NVidia Tesla K40.
4	The comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts 1ns clock ticks. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 3: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

	Developer	Short	Seq.	Validation	Config[1]	Template			Comparison Time (ns)[4]	
	Name	Name	Num.	Date	Data (KB)	Memory (MB)[2]	Size (B)	Time (ms)[3]	Genuine	Impostor
133	Remark Holdings	remarkai	001	2019-11-21	224157	[52]443	[85]2048 ± 0	[198]950 ± 6	[69]1115 ± 25	[64]1068 ± 54
134	Rokid Corporation Ltd	rokid	000	2019-08-01	258612	[149]1218	[165]2056 ± 0	[92]546 ± 3	[139]3457 ± 62	[142]3463 ± 77
135	Rokid Corporation Ltd	rokid	001	2019-12-13	641223	[138]1071	[170]2060 ± 0	[195]911 ± 2	[138]3345 ± 50	[139]3346 ± 149
136	Saffe Ltd	saffe	001	2018-10-19	85973	[18]168	[47]1280 ± 0	[35]281 ± 1	[83]1274 ± 19	[85]1277 ± 26
137	Saffe Ltd	saffe	002	2019-03-01	260622	[119]855	[127]2048 ± 0	[172]817 ± 11	[33]717 ± 7	[37]714 ± 29
138	Samsung S1 Corp	s1	001	2019-12-06	435491	[105]772	[174]2092 ± 0	[111]605 ± 24	[91]1428 ± 34	[89]1415 ± 85
139	Samtech InfoNet Limited	samtech	001	2019-10-15	288082	[88]605	[160]2056 ± 0	[41]294 ± 3	[168]7694 ± 59	[168]7678 ± 91
140	Scanovate Ltd	scanovate	001	2019-11-12	257083	[85]601	[75]2048 ± 0	[100]577 ± 24	[174]12054 ± 699	[178]13795 ± 705
141	Sensetime Group Ltd	sensetime	002	2018-10-19	531783	[185]2094	[153]2052 ± 0	[166]797 ± 3	[130]2713 ± 90	[122]2301 ± 25
142	Sensetime Group Ltd	sensetime	003	2019-06-04	787853	[192]2519	[152]2052 ± 0	[193]908 ± 4	[124]2527 ± 65	[134]3004 ± 174
143	Shaman Software	shaman	000	2017-12-05	0	[70]507	[196]4096 ± 0	[131]653 ± 16	[3]380 ± 25	[4]379 ± 31
144	Shaman Software	shaman	001	2018-01-13	0	[73]511	[189]4096 ± 0	[42]294 ± 2	[24]635 ± 19	[5]441 ± 25
145	Shanghai Jiao Tong University	sjtu	001	2019-09-27	347115	[51]438	[67]2048 ± 0	[127]650 ± 2	[206]2243804 ± 12751	[206]2249915 ± 19380
146	Shanghai Jiao Tong University	sjtu	002	2020-02-12	446215	[79]538	[129]2048 ± 0	[178]835 ± 2	[205]2211198 ± 26052	[205]2210428 ± 21877
147	Shanghai Ulucu Electronics Technology Co. Ltd	uluface	002	2019-07-10	0	[141]1088	[105]2048 ± 0	[185]873 ± 42	[185]19207 ± 1114	[183]18501 ± 274
148	Shanghai Ulucu Electronics Technology Co. Ltd	uluface	003	2019-11-12	97357	[152]1264	[184]3072 ± 0	[202]965 ± 11	[190]26057 ± 195	[192]26865 ± 566
149	Shanghai Universiy - Shanghai Film Academy	shu	001	2019-06-17	329513	[46]402	[134]2048 ± 0	[112]612 ± 5	[128]2619 ± 19	[132]2987 ± 143
150	Shanghai Universiy - Shanghai Film Academy	shu	002	2019-12-10	731250	[123]890	[188]4096 ± 0	[156]751 ± 2	[207]2930763 ± 47355	[207]2929759 ± 39149
151	Shanghai Yitu Technology	yitu	003	2019-03-01	1525719	[196]3737	[171]2082 ± 0	[182]860 ± 0	[182]18305 ± 71	[182]18286 ± 62
152	Shenzhen AiMall Tech Ltd	aimall	001	2019-11-12	128820	[20]178	[37]1024 ± 0	[50]326 ± 1	[151]4127 ± 72	[149]3909 ± 166
153	Shenzhen AiMall Tech Ltd	aimall	002	2020-03-12	370156	[173]1576	[107]2048 ± 0	[162]776 ± 4	[198]72811 ± 7399	[198]71216 ± 6286
154	Shenzhen EI Networks Limited	einetworks	000	2019-08-13	372608	[122]880	[166]2056 ± 0	[126]645 ± 3	[154]4876 ± 66	[154]5156 ± 77
155	Shenzhen Inst Adv Integrated Tech CAS	SIAT	002	2018-06-13	486842	[190]2434	[154]2052 ± 0	[101]579 ± 0	[39]769 ± 13	[40]750 ± 13
156	Shenzhen Inst Adv Integrated Tech CAS	SIAT	004	2019-03-01	940063	[198]3860	[200]4100 ± 0	[138]670 ± 0	[150]4013 ± 45	[147]3782 ± 173
157	Shenzhen Intellifusion Technologies Co Ltd	intellifusion	001	2019-08-22	271872	[104]762	[73]2048 ± 0	[160]764 ± 38	[68]1112 ± 28	[71]1128 ± 41
158	Shenzhen Intellifusion Technologies Co Ltd	intellifusion	002	2020-03-18	762731	[130]941	[195]4096 ± 0	[197]950 ± 2	[101]1713 ± 57	[98]1665 ± 87
159	Smilart	smilart	002	2018-02-06	111826	[28]263	[35]1024 ± 0	[17]176 ± 16	[183]18784 ± 136	[184]18795 ± 151
160	Smilart	smilart	003	2018-06-18	67339	[25]192	[18]512 ± 0	[19]180 ± 12	[88]1395 ± 74	[60]1027 ± 66
161	Star Hybrid Limited	starhybrid	001	2019-06-19	100509	[117]845	[88]2048 ± 0	[55]358 ± 82	[63]1075 ± 51	[66]1078 ± 53
162	Synesis	synesis	005	2019-06-06	146509	[94]638	[109]2048 ± 0	[22]211 ± 9	[17]599 ± 23	[17]581 ± 32
163	Synesis	synesis	006	2019-10-10	731941	[167]1472	[202]4104 ± 0	[94]549 ± 1	[31]697 ± 32	[33]688 ± 31
164	Synology Inc	synology	000	2019-10-23	221021	[54]453	[76]2048 ± 0	[64]407 ± 14	[186]19720 ± 203	[185]19767 ± 379
165	Synology Inc	synology	001	2020-02-26	256895	[78]528	[124]2048 ± 0	[196]942 ± 21	[90]1402 ± 29	[101]1730 ± 220
166	TUPU Technology Co Ltd	tuputech	000	2019-10-11	11476	[2]33	[84]2048 ± 0	[12]122 ± 4	[188]23893 ± 406	[188]25279 ± 406
167	Taiwan AI Labs	ailabs	001	2019-12-18	1054663	[151]1252	[78]2048 ± 0	[134]664 ± 4	[200]104034 ± 661	[200]103415 ± 7722
168	Tech5 SA	tech5	003	2019-08-19	1427464	[178]1745	[50]1536 ± 0	[187]885 ± 20	[73]1151 ± 37	[72]1128 ± 32
169	Tech5 SA	tech5	004	2020-03-09	2410272	[194]2733	[14]321 ± 0	[184]872 ± 2	[16]597 ± 13	[18]592 ± 16
170	Tencent Deepsea Lab	deepsea	001	2019-06-03	147497	[40]358	[34]1024 ± 0	[121]630 ± 7	[89]1401 ± 37	[91]1467 ± 50
171	Tevian	tevian	004	2019-03-01	863474	[106]774	[90]2048 ± 0	[78]506 ± 30	[6]474 ± 31	[3]326 ± 20
172	Tevian	tevian	005	2019-09-21	921043	[140]1083	[123]2048 ± 0	[123]633 ± 21	[15]568 ± 22	[21]607 ± 35
173	TigerIT Americas LLC	tiger	002	2018-06-13	341638	[74]522	[162]2056 ± 0	[62]393 ± 20	[115]2135 ± 29	[116]2137 ± 38
174	TigerIT Americas LLC	tiger	003	2018-10-16	426164	[99]708	[156]2056 ± 0	[72]458 ± 21	[112]2031 ± 35	[112]2029 ± 38
175	TongYi Transportation Technology	tongyi	005	2019-06-12	1140701	[186]2121	[173]2089 ± 0	[15]165 ± 1	[184]18924 ± 65	[186]20158 ± 103
176	Toshiba	toshiba	002	2018-10-19	813606	-	[54]1560 ± 0	[89]541 ± 0	[141]3521 ± 369	[125]2449 ± 124

Notes	
1	The configuration size does not capture static data included in libraries. We do not count these because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The memory usage is the peak resident set size reported by the ps system call during template generation.
3	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, for GPU-enabled implementations, NVidia Tesla K40.
4	The comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts 1ns clock ticks. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 4: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

	Developer Name	Short Name	Seq. Num.	Validation Date	Config[1] Data (KB)	Template Memory (MB)[2]	Template Size (B)	Template Time (ms)[3]	Comparison Time (ns)[4] Genuine	Comparison Time (ns)[4] Impostor
177	Toshiba	toshiba	003	2019-03-01	984125	^{148}1197	^{55}1560 ± 0	^{88}540 ± 0	^{123}2390 ± 41	^{124}2407 ± 81
178	Trueface.ai	trueface	000	2019-10-08	255123	^{64}493	^{131}2048 ± 0	^{56}367 ± 8	^{7}482 ± 13	^{13}528 ± 20
179	ULSee Inc	ulsee	001	2019-07-31	370519	-	^{126}2048 ± 0	^{132}654 ± 2	^{161}6065 ± 94	^{160}6228 ± 77
180	Universidade de Coimbra	visteam	000	2020-01-14	32729	^{10}83	^{51}1536 ± 0	^{7}96 ± 7	^{162}6361 ± 87	^{163}6668 ± 277
181	VCognition	vcog	002	2017-06-12	3229434	^{195}3666	^{207}61504 ± 5	^{54}357 ± 25	^{202}296154 ± 3077	^{202}296436 ± 4183
182	Veridas Digital Authentication Solutions S.L.	veridas	002	2019-03-01	193466	^{69}505	^{25}512 ± 0	^{136}669 ± 20	^{103}1733 ± 81	^{110}1934 ± 44
183	Veridas Digital Authentication Solutions S.L.	veridas	003	2019-11-27	293109	^{60}469	^{100}2048 ± 0	^{189}890 ± 33	^{158}5484 ± 42	^{166}7306 ± 410
184	Via Technologies Inc	via	000	2019-07-08	124422	^{131}964	^{120}2048 ± 0	^{148}707 ± 8	^{50}966 ± 28	^{58}1021 ± 44
185	Via Technologies Inc	via	001	2020-01-08	370255	^{177}1697	^{82}2048 ± 0	^{201}964 ± 3	^{53}983 ± 31	^{55}989 ± 40
186	Videmo Intelligente Videoanalyse	videmo	000	2019-12-19	139643	^{44}390	^{136}2048 ± 0	^{14}142 ± 5	^{12}513 ± 16	^{11}523 ± 38
187	Videonetics Technology Pvt Ltd	videonetics	001	2019-06-19	30875	^{3}61	^{16}512 ± 0	^{31}262 ± 3	^{74}1153 ± 38	^{74}1142 ± 65
188	Videonetics Technology Pvt Ltd	videonetics	002	2019-11-21	121981	^{14}115	^{149}2052 ± 0	^{37}282 ± 5	^{79}1219 ± 57	^{83}1262 ± 56
189	Vigilant Solutions	vigilant	006	2019-03-01	343048	^{156}1316	^{53}1548 ± 0	^{180}841 ± 8	^{45}939 ± 32	^{35}711 ± 37
190	Vigilant Solutions	vigilant	007	2019-06-27	255600	^{125}912	^{52}1548 ± 0	^{77}493 ± 6	^{41}803 ± 35	^{41}800 ± 40
191	Visidon	visidon	001	2019-02-26	170262	^{30}281	^{147}2052 ± 0	^{48}316 ± 6	^{82}1258 ± 38	^{75}1148 ± 109
192	Vision-Box	visionbox	000	2019-02-26	176501	^{37}355	^{133}2048 ± 0	^{43}304 ± 7	^{98}1648 ± 57	^{80}1192 ± 42
193	Vision-Box	visionbox	001	2019-03-01	256869	^{83}579	^{121}2048 ± 0	^{204}983 ± 7	^{75}1161 ± 22	^{77}1154 ± 20
194	VisionLabs	visionlabs	007	2019-06-12	357204	^{29}266	^{22}512 ± 0	^{33}272 ± 0	^{49}965 ± 41	^{51}972 ± 31
195	VisionLabs	visionlabs	008	2020-01-06	706099	^{53}446	^{24}512 ± 0	^{73}467 ± 1	^{48}955 ± 23	^{49}962 ± 25
196	Vocord	vocord	007	2019-06-06	587489	^{159}1352	^{57}1664 ± 0	^{163}780 ± 2	^{127}2593 ± 83	^{126}2526 ± 59
197	Vocord	vocord	008	2020-01-031	603867	^{171}1559	^{183}2688 ± 0	^{200}962 ± 2	^{133}3015 ± 50	^{133}2988 ± 62
198	Winsense Co Ltd	winsense	000	2019-06-17	270819	^{116}833	^{45}1280 ± 0	^{38}283 ± 1	^{95}1551 ± 31	^{93}1532 ± 42
199	Winsense Co Ltd	winsense	001	2019-10-16	264428	^{128}922	^{46}1280 ± 0	^{161}766 ± 7	^{97}1631 ± 28	^{111}1964 ± 171
200	X-Laboratory	x-laboratory	000	2019-09-03	520020	^{169}1524	^{167}2056 ± 0	^{168}808 ± 7	^{34}725 ± 19	^{39}749 ± 34
201	X-Laboratory	x-laboratory	001	2020-01-21	625140	^{182}1844	^{163}2056 ± 0	^{104}586 ± 2	^{42}813 ± 28	^{45}872 ± 32
202	Xforward AI Technology Co LTD	xforwardai	000	2020-02-06	242457	^{162}1392	^{62}2048 ± 0	^{158}757 ± 6	^{77}1185 ± 44	^{79}1157 ± 44
203	Xiamen Meiya Pico Information Co. Ltd	meiya	001	2019-03-01	280055	^{71}507	^{143}2049 ± 0	^{117}622 ± 12	^{169}8356 ± 615	^{170}8134 ± 97
204	Zhuhai Yisheng Electronics Technology	yisheng	004	2018-06-12	486351	^{155}1279	^{187}3704 ± 0	^{57}378 ± 12	^{30}693 ± 137	^{12}526 ± 34
205	iQIYI Inc	iqface	000	2019-06-04	268819	^{97}704	^{204}4750 ± 32	^{87}538 ± 26	^{204}636433 ± 38446	^{204}632654 ± 85615
206	iQIYI Inc	iqface	001	2019-12-11	1145590	^{129}938	^{205}4752 ± 35	^{110}603 ± 1	^{203}568761 ± 3659	^{203}570335 ± 5122
207	iSAP Solution Corporation	isap	001	2019-08-07	99049	^{1}18	^{198}4096 ± 0	^{1}1 ± 0	^{4}459 ± 17	^{6}456 ± 11

Notes	
1	The configuration size does not capture static data included in libraries. We do not count these because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The memory usage is the peak resident set size reported by the ps system call during template generation.
3	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, for GPU-enabled implementations, NVidia Tesla K40.
4	The comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts 1ns clock ticks. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 5: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

	Algorithm	FALSE NON-MATCH RATE (FNMR)															
		CONSTRAINED, COOPERATIVE												LESS CONSTRAINED, NON-COOP.			
	Name	VISAMC		VISA		VISA		MUGSHOT		MUGSHOT12+YRS		VISABORDER		WILD		CHILDEXP	
	FMR	0.0001		1E-06		0.0001		1E-05		1E-05		1E-06		0.0001		0.01	
1	3divi-003	0.0318	152	0.0588	149	0.0097	139	0.0389	156	0.0639	153	0.0619	150	0.0867	147	0.5365	33
2	3divi-004	0.0095	61	0.0153	63	0.0049	85	0.0097	67	0.0145	65	0.0175	73	0.0665	140	0.5025	26
3	acer-000	0.1393	173	0.9075	198	0.0650	175	0.9981	199	-		1.0000	194	0.9841	194	-	
4	adera-001	0.1021	170	0.1757	167	0.0368	169	0.1823	177	0.2967	175	0.1714	164	0.1965	168	0.7202	65
5	advance-002	0.0089	55	0.0137	55	0.0034	54	0.0073	38	0.0115	44	0.0400	131	0.0498	125	-	
6	aifirst-001	0.0119	80	0.0170	70	0.0067	114	0.0084	52	0.0127	50	0.0131	49	0.0432	110	0.4301	17
7	ailabs-001	0.0158	112	0.0276	119	0.0065	106	0.0192	128	0.0317	124	0.0352	126	0.0338	89	-	
8	aimall-001	0.0176	121	0.0324	127	0.0074	127	0.0225	135	0.0407	135	-		0.0344	91	0.6587	53
9	aimall-002	0.0119	79	0.0167	68	0.0052	90	0.0224	133	0.0411	136	0.0233	96	0.0327	80	-	
10	aiunionface-000	0.0104	68	0.0154	65	0.0051	89	0.0082	51	0.0122	48	0.0141	50	0.0306	63	-	
11	alchera-000	0.0165	116	0.0243	106	0.0086	132	0.0125	96	0.0186	90	0.0204	86	0.0370	97	-	
12	alchera-001	0.0183	122	0.0299	121	0.0078	128	0.0142	105	0.0234	106	0.0239	101	0.0372	98	-	
13	alleyes-000	0.0058	27	0.0090	30	0.0019	23	0.0055	22	0.0087	29	0.0068	11	0.0282	21	-	
14	allgovision-000	0.0346	153	0.0527	145	0.0210	158	0.0232	137	0.0339	126	0.0372	129	0.0607	136	-	
15	alphaface-001	0.0065	34	0.0097	34	0.0025	39	0.0039	13	0.0063	17	0.0083	18	0.0280	13	-	
16	alphaface-002	0.0052	19	0.0075	20	0.0019	22	0.0030	2	0.0044	5	1.0000	197	0.0279	10	-	
17	amplifiedgroup-001	0.5034	192	0.5848	190	0.2999	194	0.6973	191	0.8316	189	0.7807	184	0.4250	185	-	
18	androvideo-000	0.0243	138	0.0438	142	0.0098	140	0.0239	140	0.0365	132	0.0483	140	0.1162	157	-	
19	anke-004	0.0080	49	0.0154	64	0.0031	48	0.0073	37	0.0112	42	0.0102	35	0.0288	34	0.3577	8
20	anke-005	0.0070	40	0.0109	44	0.0022	31	0.0059	29	0.0094	31	0.0105	36	0.0289	36	0.3337	6
21	antheus-000	0.2564	181	0.3776	181	0.1059	182	0.7240	192	0.8699	192	0.8899	188	0.7668	188	0.9233	88
22	anyvision-002	0.0660	165	0.0898	160	0.0387	170	0.0928	171	0.1512	168	0.0899	157	0.2227	172	0.6960	59
23	anyvision-004	0.0267	145	0.0385	139	0.0081	130	0.0258	144	0.0487	146	0.0234	98	0.0470	120	0.4633	21
24	asusaics-000	0.0125	87	0.0209	86	0.0043	68	0.0085	53	0.0134	56	0.0143	52	0.0295	46	-	
25	asusaics-001	0.0125	88	0.0210	87	0.0044	69	0.0085	55	0.0134	57	0.0143	53	0.0295	45	-	
26	aware-004	0.0690	166	0.0949	163	0.0257	163	0.0837	169	0.1436	167	0.1171	161	0.0516	127	-	
27	aware-005	0.0457	158	0.0643	153	0.0126	149	0.0603	164	0.1094	161	0.0613	148	0.0314	71	-	
28	awiros-001	0.4044	186	0.4622	184	0.2880	193	0.5530	186	0.6518	184	0.2008	165	0.5584	187	-	
29	ayonix-000	0.4351	189	0.4872	185	0.2299	188	0.6150	189	0.7510	187	0.6557	180	0.3635	182	0.8434	77
30	bioidtechswiss-000	0.0066	36	0.0082	28	0.0015	14	0.0113	84	0.0225	101	0.0078	16	0.0278	8	-	
31	bm-001	0.7431	198	0.9494	200	0.6188	199	0.9586	196	0.9843	194	0.9049	189	0.9935	195	0.8845	83
32	camvi-002	0.0125	89	0.0221	93	0.0049	87	0.0089	59	0.0145	67	0.0142	51	0.0288	32	0.5760	41
33	camvi-004	0.0171	119	0.0316	126	0.0049	84	0.0042	15	0.0049	10	0.0097	33	0.0284	23	0.5788	43
34	ceiec-002	0.0161	115	0.0193	82	0.0124	148	0.0122	93	0.0164	77	0.0270	108	0.0465	119	0.5156	30
35	ceiec-003	0.0071	43	0.0107	42	0.0024	38	0.0061	31	0.0079	24	0.0160	60	0.0308	68	-	
36	chosun-000	0.8481	200	1.0000	204	0.7301	200	1.0000	205	-		1.0000	204	1.0000	204	-	
37	chtface-001	0.9993	204	0.9994	203	0.9993	204	0.9999	201	-		1.0000	193	0.9980	196	-	
38	chtface-002	0.0150	105	0.0268	115	0.0054	93	0.0096	66	0.0140	61	0.0186	77	0.0306	64	-	
39	cib-000	0.0084	52	0.0156	66	0.0053	91	0.0049	18	0.0051	12	0.0275	111	0.0294	44	-	
40	cogent-003	0.0091	56	0.0188	79	0.0032	49	0.0098	69	0.0132	54	0.0187	78	0.0406	105	-	
41	cogent-004	0.0064	33	0.0116	47	0.0024	36	0.0096	65	0.0134	58	0.0157	56	0.0379	100	0.7177	64
42	cognitec-000	0.0116	75	0.0177	71	0.0036	56	0.0118	89	0.0167	79	0.0285	115	0.0953	151	0.8365	76
43	cognitec-001	0.0126	90	0.0185	77	0.0047	80	0.0120	91	0.0168	80	0.0270	107	0.0598	135	-	
44	ctbcbank-000	0.0168	117	0.0250	111	0.0064	104	0.0146	109	0.0224	100	0.0211	89	1.0000	203	0.8803	81

Table 6: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. The light grey values give rank over all algorithms in that column. The pink columns use only same-sex impostors; others are selected regardless of demographics. The exception, in the green column, uses "matched-covariates" i.e. impostors of the same sex, age group, and country of birth. The second pink column includes effects of extended ageing. Missing entries for border, visa, mugshot and wild images generally mean the algorithm did not run to completion. For child exploitation, missing entries arise because NIST executes those runs only infrequently. The VISA columns compare images described in section 2.2. The MUGSHOT columns compare images described in section 2.5. The VISA-BORDER column compare images described in section 2.3 with those described in section 2.4. The WILD columns compare images described in section 2.6. The CHILD-EXPLOITATION columns compare images described in section 2.1.

	Algorithm Name	FALSE NON-MATCH RATE (FNMR) CONSTRAINED, COOPERATIVE VISAMC		VISA		VISA		MUGSHOT		MUGSHOT12+YRS		VISABORDER		LESS CONSTRAINED, NON-COOP. WILD		CHILDEXP	
	FMR	0.0001		1E-06		0.0001		1E-05		1E-05		1E-06		0.0001		0.01	
45	ctbcbank-001	0.0155	108	0.0235	103	0.0060	101	0.0148	113	0.0243	109	0.0207	87	1.0000	206	-	
46	cuhkee-001	0.0036	8	0.0045	6	0.0007	5	0.0031	5	0.0046	7	0.0051	7	0.1492	162	-	
47	cyberextruder-001	0.1972	177	0.2547	174	0.0755	177	0.4686	185	0.6387	183	0.3807	176	0.1747	165	0.7804	74
48	cyberextruder-002	0.0811	167	0.1336	166	0.0265	165	0.1465	175	0.2266	174	0.2086	168	0.1000	154	0.6105	47
49	cyberlink-003	0.0118	77	0.0192	81	0.0042	64	0.0098	70	0.0161	72	0.0153	55	0.0303	57	-	
50	cyberlink-004	0.0074	48	0.0105	40	0.0030	47	0.0068	35	0.0089	30	0.0094	29	0.0283	22	-	
51	dahua-003	0.0052	20	0.0068	16	0.0023	35	0.0056	25	0.0062	16	0.0113	41	0.0285	26	0.6253	49
52	dahua-004	0.0045	16	0.0058	11	0.0019	25	0.0036	9	0.0048	9	0.0051	5	0.0281	15	-	
53	deepglint-001	0.0040	11	0.0062	15	0.0014	13	0.0047	17	0.0067	18	0.0069	12	0.0278	7	0.4006	12
54	deepglint-002	0.0016	2	0.0027	3	0.0004	2	0.0032	7	0.0033	2	0.0043	2	0.0280	12	0.3422	7
55	deepsea-001	0.0136	96	0.0215	90	0.0071	122	0.0142	106	0.0214	97	0.0163	64	0.0347	92	0.5606	38
56	dermalog-005	0.1526	175	0.1823	170	0.0658	176	0.2580	179	0.4018	177	-		0.0855	145	0.6842	56
57	dermalog-006	0.0253	143	0.0369	137	0.0172	155	0.0171	122	0.0283	118	0.0217	92	0.0623	138	0.5852	44
58	didiglobalface-001	0.0055	24	0.0092	31	0.0016	15	0.0030	3	0.0045	6	0.0088	24	0.0282	19	0.4270	15
59	digitalbarriers-002	0.3360	184	0.3690	179	0.0968	181	0.0877	170	0.1557	169	0.0971	160	0.0436	112	-	
60	dsk-000	0.1526	174	0.2169	171	0.0765	178	0.3787	182	0.5426	181	0.3115	171	0.2201	170	0.7313	67
61	einetworks-000	0.0099	64	0.0180	74	0.0047	79	0.0088	58	0.0140	63	0.0130	47	0.0293	41	-	
62	everai-paravision-003	0.0034	7	0.0050	9	0.0011	8	0.0036	12	0.0052	14	0.0092	27	0.0278	9	0.2669	2
63	expasoft-000	0.0427	156	0.0655	155	0.0233	162	0.0239	139	0.0393	133	0.0673	152	0.0565	133	-	
64	f8-001	0.0249	142	0.0336	129	0.0182	156	0.0178	125	0.0232	105	0.0303	121	0.0475	122	0.5272	31
65	facesoft-000	0.0085	53	0.0112	46	0.0032	51	0.0064	33	0.0107	37	0.0091	25	0.0275	4	0.4992	25
66	fujitsulab-000	0.0123	83	0.0212	88	0.0057	97	0.0091	60	0.0133	55	0.0251	104	0.0445	114	-	
67	glory-001	0.0902	168	0.1082	165	0.0410	171	0.1642	176	0.2065	173	0.2186	169	0.4261	186	0.8831	82
68	glory-002	0.0241	135	0.0311	125	0.0188	157	0.0116	86	0.0151	69	0.0157	58	0.1265	159	-	
69	gorilla-004	0.0138	97	0.0239	105	0.0049	86	0.0243	141	0.0444	139	0.0275	110	0.0287	31	-	
70	gorilla-005	-		-		-		0.0142	107	0.0267	116	0.0228	95	0.0307	65	-	
71	hik-001	0.0096	62	0.0125	51	0.0036	58	0.0093	63	0.0164	76	0.0108	38	0.0271	1	-	
72	hr-001	0.0044	15	0.0072	19	0.0019	21	0.0073	40	0.0108	40	0.0125	44	0.0303	56	0.5499	35
73	hr-002	0.0043	13	0.0059	12	0.0017	16	0.0054	19	0.0076	21	0.0076	15	0.0338	88	-	
74	id3-003	0.0361	154	0.0757	157	0.0104	146	0.0292	150	0.0476	143	0.0447	139	0.0848	144	0.5865	45
75	id3-004	0.0198	129	0.0344	132	0.0084	131	0.0238	138	0.0423	138	0.0289	117	-		-	
76	idemia-004	0.0160	114	0.0244	108	0.0065	107	0.0199	130	0.0354	128	0.0268	106	0.0309	69	-	
77	idemia-005	0.0132	93	0.0216	91	0.0057	98	0.0121	92	0.0218	99	0.0215	91	0.0294	43	0.4343	18
78	iit-001	0.0104	69	0.0179	73	0.0048	82	0.0099	72	0.0142	64	0.1222	162	0.3092	181	0.4836	23
79	iit-002	0.0111	73	0.0177	72	0.0054	92	0.0085	54	0.0140	62	0.0193	82	0.1373	160	-	
80	imagus-000	0.0642	163	0.0882	158	0.0330	168	0.0497	159	0.0905	157	0.0848	155	0.1158	156	0.6936	58
81	imagus-001	0.0245	139	0.0407	140	0.0091	135	0.0257	143	0.0497	148	0.0514	143	0.0971	152	-	
82	imperial-000	0.0067	38	0.0108	43	0.0022	32	0.0080	48	0.0134	59	0.0087	23	0.0281	16	-	
83	imperial-002	0.0058	26	0.0081	26	0.0027	40	0.0055	23	0.0085	26	0.0083	19	0.0273	2	0.5151	27
84	incode-005	0.0111	74	0.0189	80	0.0042	65	0.0156	119	0.0304	122	0.0195	83	0.0562	132	-	
85	incode-006	0.0156	109	0.0273	117	0.0051	88	0.0117	87	0.0238	107	0.0202	85	0.0325	78	-	
86	innovativetechnologyltd-001	0.0578	161	0.0938	162	0.0258	164	0.0501	160	0.0981	158	-		0.0449	115	-	
87	innovativetechnologyltd-002	0.0451	157	0.0716	156	0.0172	154	0.0541	162	0.1009	160	0.0506	142	0.0804	143	-	
88	innovatrics-004	0.0194	126	0.0292	120	0.0068	116	0.0344	153	0.0617	152	0.0562	146	0.0454	118	0.4650	22

Table 7: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. The light grey values give rank over all algorithms in that column. The pink columns use only same-sex impostors; others are selected regardless of demographics. The exception, in the green column, uses "matched-covariates" i.e. impostors of the same sex, age group, and country of birth. The second pink column includes effects of extended ageing. Missing entries for border, visa, mugshot and wild images generally mean the algorithm did not run to completion. For child exploitation, missing entries arise because NIST executes those runs only infrequently. The VISA columns compare images described in section 2.2. The MUGSHOT columns compare images described in section 2.5. The VISA-BORDER column compare images described in section 2.3 with those described in section 2.4. The WILD columns compare images described in section 2.6. The CHILD-EXPLOITATION columns compare images described in section 2.1.

	Algorithm Name	FALSE NON-MATCH RATE (FNMR)															
		CONSTRAINED, COOPERATIVE												LESS CONSTRAINED, NON-COOP.			
		VISAMC		VISA		VISA		MUGSHOT		MUGSHOT12+YRS		VISABORDER		WILD		CHILDEXP	
	FMR	0.0001		1E-06		0.0001		1E-05		1E-05		1E-06		0.0001		0.01	
89	innovatrics-006	0.0058	28	0.0089	29	0.0021	27	0.0061	32	0.0096	34	0.0096	32	0.0281	14	0.3056	4
90	intellicloudai-001	0.0142	99	0.0234	101	0.0064	105	0.0092	62	0.0145	66	0.0162	62	0.0409	106	-	
91	intellifusion-001	0.0072	44	0.0094	32	0.0028	45	0.0056	26	0.0085	27	0.0111	40	0.0289	35	0.5454	34
92	intellifusion-002	0.0059	29	0.0077	22	0.0022	28	0.0040	14	0.0074	19	0.0085	22	0.0305	60	-	
93	intellivision-001	0.1335	172	0.2205	172	0.0417	172	0.1090	173	0.1670	170	0.1385	163	0.2445	173	0.7766	73
94	intellivision-002	0.1000	169	0.1775	168	0.0265	166	0.0610	165	0.1009	159	0.0805	154	0.0768	142	-	
95	intelresearch-000	0.0307	149	0.0578	148	0.0093	136	0.0385	155	0.0751	156	0.0409	133	0.0324	75	-	
96	intelresearch-001	0.0242	136	0.0595	151	0.0097	138	0.0129	100	0.0292	120	0.0351	125	0.0919	150	-	
97	intsysmsu-001	0.9543	203	0.9888	202	0.9165	203	0.9923	197	-		0.9977	190	0.7871	189	-	
98	intsysmsu-002	0.0130	91	0.0254	112	0.0040	63	0.0137	102	0.0267	117	0.0160	59	0.0289	37	-	
99	iqface-000	0.0091	58	0.0143	58	0.0043	67	0.0075	43	0.0110	41	0.0171	71	0.0381	101	0.6490	51
100	iqface-001	0.0088	54	0.0117	48	0.0047	78	0.0177	123	0.0230	104	0.0165	67	0.0597	134	-	
101	isap-001	0.5092	193	0.6588	192	0.2338	190	0.6899	190	0.7978	188	0.7200	181	0.1931	167	-	
102	isityou-000	0.5682	195	0.7033	194	0.4145	197	1.0000	202	-		1.0000	198	1.0000	199	1.0000	93
103	isystems-001	0.0149	104	0.0245	109	0.0067	113	0.0138	104	0.0210	95	0.0209	88	0.0524	130	0.5152	29
104	isystems-002	0.0118	76	0.0182	75	0.0066	108	0.0111	79	0.0162	74	0.0166	68	0.0516	128	0.4876	24
105	itmo-006	0.0125	86	0.0220	92	0.0046	74	0.0149	114	0.0266	115	0.0233	97	0.0329	82	-	
106	itmo-007	0.0080	50	0.0125	52	0.0033	52	0.0107	75	0.0185	88	0.0167	69	0.0300	52	-	
107	kakao-002	0.0625	162	0.1779	169	0.0168	153	0.0791	168	0.1381	166	0.0636	151	1.0000	201	1.0000	101
108	kakao-003	0.0130	92	0.0185	78	0.0049	83	0.0261	146	0.0464	142	0.0252	105	0.0274	3	-	
109	kedacom-000	0.0055	23	0.0081	27	0.0027	41	0.0111	81	0.0120	47	0.0415	134	0.2511	175	0.7650	71
110	kneron-003	0.0542	159	0.0902	161	0.0218	160	0.0346	154	0.0562	151	0.0919	158	0.3053	180	0.6962	60
111	kneron-005	0.0157	110	0.0259	114	0.0072	124	0.0126	98	0.0212	96	0.0406	132	0.0471	121	-	
112	lookman-002	0.0297	147	0.0547	147	0.0102	145	0.0339	152	0.0562	150	0.0614	149	0.2640	178	-	
113	lookman-004	0.0074	46	0.0099	36	0.0037	59	0.0124	95	0.0149	68	0.0430	137	0.2516	176	0.7664	72
114	luxand-000	0.2056	179	0.2814	175	0.0895	180	0.4053	183	0.5365	180	0.3497	173	0.2222	171	-	
115	megvii-001	0.0157	111	0.0244	107	0.0045	73	0.0392	157	0.0671	154	0.0168	70	0.0916	149	0.4418	19
116	megvii-002	0.0104	70	0.0145	60	0.0036	57	0.0225	134	0.0345	127	0.0099	34	0.0692	141	0.3013	3
117	meiya-001	0.0171	118	0.0275	118	0.0066	110	0.0159	120	0.0261	114	0.0311	122	0.0363	96	-	
118	microfocus-001	0.4482	190	0.5524	189	0.2309	189	0.7256	193	0.8416	190	0.7301	182	0.2567	177	0.6890	57
119	microfocus-002	0.3605	185	0.5057	186	0.1566	185	0.5783	187	0.7223	185	0.5909	179	0.1582	164	0.6517	52
120	mt-000	0.0100	65	0.0170	69	0.0047	77	0.0074	42	0.0118	46	0.0127	45	0.0326	79	0.3773	11
121	mvision-001	0.0191	124	0.0233	99	0.0131	150	0.0204	131	0.0356	129	0.0198	84	0.0431	109	-	
122	netbridgetech-001	0.4749	191	0.6599	193	0.1800	186	0.4438	184	0.5676	182	0.4491	177	0.1098	155	-	
123	neurotechnology-006	0.0098	63	0.0136	54	0.0040	62	0.0105	74	0.0182	86	0.0164	65	0.0303	58	0.5911	46
124	neurotechnology-008	0.0091	57	0.0184	76	0.0023	33	0.0076	45	0.0116	45	0.0157	57	0.0301	54	-	
125	nodeflux-001	1.0000	206	1.0000	205	1.0000	205	1.0000	207	-		0.5169	178	1.0000	207	-	
126	nodeflux-002	0.0186	123	0.0340	130	0.0070	119	0.0261	145	0.0451	141	0.0548	145	0.0299	51	-	
127	notiontag-000	0.6669	196	0.7885	195	0.3222	196	0.3715	181	0.4978	179	0.8571	186	0.1807	166	0.6479	50
128	ntechlab-007	0.0056	25	0.0076	21	0.0018	18	0.0073	41	0.0128	52	0.0079	17	0.0276	6	0.3316	5
129	ntechlab-008	0.0041	12	0.0061	13	0.0011	7	0.0056	24	0.0108	38	0.0042	1	0.0289	39	-	
130	paravision-004	0.0030	6	0.0046	7	0.0012	9	0.0030	4	0.0036	3	0.0091	26	0.0288	33	0.2467	1
131	pixelall-002	0.0193	125	0.0340	131	0.0066	109	0.0127	99	0.0209	93	0.0234	99	0.0342	90	0.8963	85
132	pixelall-003	0.0074	47	0.0118	49	0.0032	50	0.0057	27	0.0079	23	0.0121	43	0.0285	28	-	

Table 8: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. The light grey values give rank over all algorithms in that column. The pink columns use only same-sex impostors; others are selected regardless of demographics. The exception, in the green column, uses "matched-covariates" i.e. impostors of the same sex, age group, and country of birth. The second pink column includes effects of extended ageing. Missing entries for border, visa, mugshot and wild images generally mean the algorithm did not run to completion. For child exploitation, missing entries arise because NIST executes those runs only infrequently. The VISA columns compare images described in section 2.2. The MUGSHOT columns compare images described in section 2.5. The VISA-BORDER column compare images described in section 2.3 with those described in section 2.4. The WILD columns compare images described in section 2.6. The CHILD-EXPLOITATION columns compare images described in section 2.1.

	Algorithm Name	FALSE NON-MATCH RATE (FNMR) CONSTRAINED, COOPERATIVE VISAMC		VISA		VISA		MUGSHOT		MUGSHOT12+YRS		VISABORDER		LESS CONSTRAINED, NON-COOP. WILD		CHILDEXP	
	FMR	0.0001		1E-06		0.0001		1E-05		1E-05		1E-06		0.0001		0.01	
133	psl-003	0.0065	35	0.0099	37	0.0028	43	0.0055	21	0.0075	20	1.0000	196	0.0296	48	-	
134	psl-004	0.0064	32	-		-		0.0035	8	0.0050	11	0.0084	20	0.0301	55	-	
135	pyramid-000	0.0136	95	0.0233	100	0.0056	95	0.0117	88	0.0192	92	0.0185	76	0.0304	59	-	
136	rankone-007	0.0197	128	0.0366	136	0.0057	99	0.0113	83	0.0177	83	0.0286	116	0.0450	116	0.6686	54
137	rankone-008	0.0124	84	0.0232	98	0.0045	72	0.0082	50	0.0107	36	0.0188	80	0.0420	107	-	
138	realnetworks-002	0.0248	140	0.0358	133	0.0099	142	0.0513	161	0.1127	162	0.0371	128	0.0334	84	-	
139	realnetworks-003	0.0259	144	0.0372	138	0.0100	143	0.0541	163	0.1208	164	0.0378	130	0.0335	86	0.5152	28
140	remarkai-001	0.0144	100	0.0256	113	0.0061	102	0.0102	73	0.0159	71	0.0162	63	0.0308	67	-	
141	remarkai-002	0.0151	107	0.0197	83	0.0102	144	0.0075	44	0.0108	39	0.0119	42	0.0426	108	-	
142	rokid-000	0.0093	60	0.0145	59	0.0038	60	0.0073	39	0.0102	35	0.0164	66	0.0857	146	-	
143	rokid-001	0.0105	71	0.0162	67	0.0042	66	0.0094	64	0.0163	75	0.0181	74	0.0325	76	-	
144	s1-001	0.0314	151	0.0651	154	0.0099	141	0.0252	142	0.0357	130	0.0444	138	0.8493	190	-	
145	saffe-001	0.4339	188	0.5261	187	0.2340	191	0.7539	195	0.8736	193	0.7977	185	0.3887	183	0.8973	86
146	saffe-002	0.0119	81	0.0206	84	0.0054	94	0.0107	77	0.0177	82	0.0244	102	0.0308	66	-	
147	samtech-001	0.0197	127	0.0365	134	0.0066	111	0.0146	111	0.0241	108	0.0238	100	0.0337	87	-	
148	scanovate-001	0.0175	120	0.0331	128	0.0061	103	0.0163	121	0.0248	110	0.2476	170	0.4060	184	-	
149	sensetime-002	0.0068	39	0.0098	35	0.0035	55	0.0143	108	-		0.0278	113	0.9999	197	0.5309	32
150	sensetime-003	0.0021	3	0.0027	2	0.0005	3	0.0027	1	0.0027	1	0.0051	6	0.0329	81	0.3683	9
151	sertis-000	0.0118	78	0.0208	85	0.0047	75	0.0080	47	0.0127	49	0.0110	39	0.0285	27	-	
152	shaman-000	0.9297	202	0.9774	201	0.9128	202	0.9990	200	-		0.9999	192	0.9575	193	0.9618	89
153	shaman-001	0.3346	183	0.4616	183	0.1360	184	0.2368	178	0.3723	176	0.3574	174	0.1498	163	0.8990	87
154	shu-001	0.0103	67	0.0140	57	0.0044	70	0.0293	151	0.0688	155	0.0172	72	0.0986	153	0.4590	20
155	shu-002	-		0.0079	24	0.0017	17	0.0146	110	0.0308	123	1.0000	195	0.0284	24	-	
156	siat-002	0.0091	59	0.0126	53	0.0039	61	0.0109	78	0.0190	91	0.0276	112	0.0520	129	0.4277	16
157	siat-004	0.0067	37	0.0099	38	0.0028	44	0.0152	116	-		0.0275	109	1.0000	198	-	
158	sjtu-001	0.0051	18	0.0080	25	0.0019	20	0.0211	132	0.0446	140	0.0131	48	0.0289	38	-	
159	sjtu-002	0.0053	22	0.0078	23	0.0027	42	0.0138	103	0.0296	121	0.0084	21	0.0284	25	-	
160	smilart-002	0.2440	180	0.3532	178	0.0821	179	-		-		0.3785	175	-		0.6999	61
161	smilart-003	0.6944	197	0.8836	196	0.1088	183	0.0695	166	0.1193	163	0.0894	156	0.1190	158	-	
162	starhybrid-001	0.0108	72	0.0138	56	0.0058	100	0.0081	49	0.0113	43	0.0152	54	0.0350	94	0.5584	36
163	synesis-005	0.0147	101	0.0226	95	0.0073	125	0.0153	117	0.0226	102	0.0213	90	0.0334	85	0.5601	37
164	synesis-006	0.0070	41	0.0096	33	0.0023	34	0.0107	76	0.0166	78	-		0.0292	40	-	
165	synology-000	0.0149	103	0.0238	104	0.0067	115	0.0148	112	0.0261	113	0.0221	93	0.0330	83	-	
166	synology-001	0.0214	131	0.0307	123	0.0073	126	0.0266	147	0.0513	149	0.0283	114	0.0296	49	-	
167	tech5-003	0.0053	21	0.0070	17	0.0014	12	0.0099	71	0.0185	87	0.0095	31	0.0306	61	0.8963	84
168	tech5-004	0.0123	82	0.0234	102	0.0024	37	0.0086	57	0.0162	73	0.0065	10	0.0281	17	-	
169	tevian-004	0.0228	133	0.0304	122	0.0069	117	0.0226	136	0.0478	144	0.0128	46	0.0394	103	-	
170	tevian-005	0.0043	14	0.0062	14	0.0020	26	0.0057	28	0.0085	28	0.0070	13	0.0300	53	0.5625	40
171	tiger-002	0.0658	164	0.0889	159	0.0227	161	0.1083	172	0.1766	171	0.0952	159	0.0512	126	0.7862	75
172	tiger-003	0.0313	150	0.0602	152	0.0087	133	0.0188	127	0.0359	131	0.0344	123	0.0482	124	0.5610	39
173	tongyi-005	0.0073	45	0.0146	61	0.0019	24	0.0187	126	0.0421	137	0.0161	61	0.0399	104	0.6195	48
174	toshiba-002	0.0134	94	0.0222	94	0.0048	81	0.0097	68	0.0154	70	-		0.0434	111	0.7103	62
175	toshiba-003	0.0125	85	0.0214	89	0.0047	76	0.0085	56	0.0131	53	-		0.0282	18	-	
176	trueface-000	0.0249	141	0.4321	182	0.0069	118	0.0119	90	0.0180	85	0.0297	120	0.0614	137	-	

Table 9: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. The light grey values give rank over all algorithms in that column. The pink columns use only same-sex impostors; others are selected regardless of demographics. The exception, in the green column, uses "matched-covariates" i.e. impostors of the same sex, age group, and country of birth. The second pink column includes effects of extended ageing. Missing entries for border, visa, mugshot and wild images generally mean the algorithm did not run to completion. For child exploitation, missing entries arise because NIST executes those runs only infrequently. The VISA columns compare images described in section 2.2. The MUGSHOT columns compare images described in section 2.5. The VISA-BORDER column compare images described in section 2.3 with those described in section 2.4. The WILD columns compare images described in section 2.6. The CHILD-EXPLOITATION columns compare images described in section 2.1.

	Algorithm	FALSE NON-MATCH RATE (FNMR)															
	Algorithm	CONSTRAINED, COOPERATIVE												LESS CONSTRAINED, NON-COOP.			
	Name	VISAMC		VISA		VISA		MUGSHOT		MUGSHOT12+YRS		VISABORDER		WILD		CHILDEXP	
	FMR	0.0001		1E-06		0.0001		1E-05		1E-05		1E-06		0.0001		0.01	
177	tuputech-000	0.3218	182	0.3696	180	0.2779	192	-		-		0.3237	172	0.9415	192	-	
178	ulsee-001	0.0151	106	0.0246	110	0.0080	129	0.0113	82	0.0185	89	0.0187	79	0.0316	72	-	
179	uluface-002	0.0081	51	0.0123	50	0.0033	53	0.0071	36	0.0095	33	0.0107	37	0.0444	113	0.6729	55
180	uluface-003	0.0100	66	0.0150	62	0.0044	71	0.0079	46	0.0128	51	-		0.0635	139	-	
181	upc-001	0.0234	134	0.0519	144	0.0071	121	0.0291	149	0.0490	147	0.0294	118	0.0314	70	0.4224	14
182	vcog-002	0.7522	199	0.9033	197	0.5040	198	-		-		-		-		0.7523	69
183	vd-001	0.0243	137	0.0452	143	0.0093	137	0.0271	148	0.0402	134	0.0424	136	0.1389	161	-	
184	veridas-002	0.1733	176	0.2257	173	0.0528	174	0.2617	180	0.4147	178	0.2073	167	0.0450	117	-	
185	veridas-003	0.0557	160	0.0983	164	0.0154	152	0.0734	167	0.1267	165	0.0694	153	0.0299	50	0.5785	42
186	via-000	0.0216	132	0.0365	135	0.0088	134	0.0177	124	0.0287	119	0.0296	119	0.0349	93	0.7638	70
187	via-001	0.0149	102	0.0229	97	0.0067	112	0.0114	85	0.0177	84	0.0183	75	0.0373	99	-	
188	videmo-000	0.0298	148	0.0423	141	0.0211	159	0.0155	118	0.0260	112	0.0246	103	0.0541	131	-	
189	videonetics-001	0.5483	194	0.6446	191	0.3063	195	0.7517	194	0.8607	191	0.8664	187	0.2986	179	0.7297	66
190	videonetics-002	0.4274	187	0.5329	188	0.2168	187	0.6081	188	0.7438	186	0.7775	183	0.1976	169	0.7435	68
191	vigilantsolutions-006	0.1264	171	0.3221	176	0.0136	151	0.0150	115	0.0254	111	0.0493	141	0.0321	74	-	
192	vigilantsolutions-007	0.0202	130	0.0307	124	0.0070	120	0.0136	101	0.0227	103	0.0356	127	0.0306	62	1.0000	193
193	vion-000	0.0419	155	0.0590	150	0.0288	167	0.0422	158	0.0478	145	0.0581	147	0.2479	174	0.8765	80
194	visionbox-000	0.0293	146	0.0541	146	0.0110	147	0.0197	129	0.0339	125	0.0349	124	0.0476	123	-	
195	visionbox-001	0.0159	113	0.0270	116	0.0072	123	0.0111	80	0.0173	81	0.0190	81	0.0389	102	-	
196	visionlabs-007	0.0038	9	0.0048	8	0.0012	11	0.0036	10	0.0048	8	0.0057	8	0.0286	30	0.3708	10
197	visionlabs-008	0.0026	4	0.0036	4	0.0007	4	0.0031	6	0.0040	4	0.0045	3	0.0282	20	-	
198	visteam-000	0.9200	201	0.9489	199	0.8616	201	0.9959	198	-		0.9994	191	0.8783	191	-	
199	vocord-007	0.0039	10	0.0053	10	0.0012	10	0.0061	30	0.0094	32	0.0520	144	0.0280	11	0.8468	78
200	vocord-008	0.0029	5	0.0038	5	0.0008	6	0.0042	16	0.0055	15	0.0045	4	0.0286	29	-	
201	winsense-000	0.0140	98	0.0228	96	0.0056	96	0.0125	97	0.0215	98	0.0226	94	0.0352	95	0.8600	79
202	winsense-001	0.0062	31	0.0099	39	0.0022	30	0.0092	61	0.0210	94	0.0093	28	0.0320	73	0.4155	13
203	x-laboratory-000	0.0071	42	0.0106	41	0.0030	46	0.0123	94	0.0138	60	0.0419	135	0.0295	47	0.9686	90
204	x-laboratory-001	0.0059	30	0.0110	45	0.0022	29	0.0054	20	0.0078	22	0.0094	30	0.0294	42	-	
205	xforwardai-000	0.0050	17	0.0072	18	0.0018	19	0.0036	11	0.0051	13	0.0074	14	0.0276	5	-	
206	yisheng-004	0.1988	178	0.3329	177	0.0475	173	0.1147	174	0.1849	172	0.2044	166	0.0908	148	0.7152	63
207	yitu-003	0.0015	1	0.0026	1	0.0003	1	0.0066	34	0.0085	25	0.0064	9	0.0325	77	-	

Table 10: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. The light grey values give rank over all algorithms in that column. The pink columns use only same-sex impostors; others are selected regardless of demographics. The exception, in the green column, uses "matched-covariates" i.e. impostors of the same sex, age group, and country of birth. The second pink column includes effects of extended ageing. Missing entries for border, visa, mugshot and wild images generally mean the algorithm did not run to completion. For child exploitation, missing entries arise because NIST executes those runs only infrequently. The VISA columns compare images described in section 2.2. The MUGSHOT columns compare images described in section 2.5. The VISA-BORDER column compare images described in section 2.3 with those described in section 2.4. The WILD columns compare images described in section 2.6. The CHILD-EXPLOITATION columns compare images described in section 2.1.

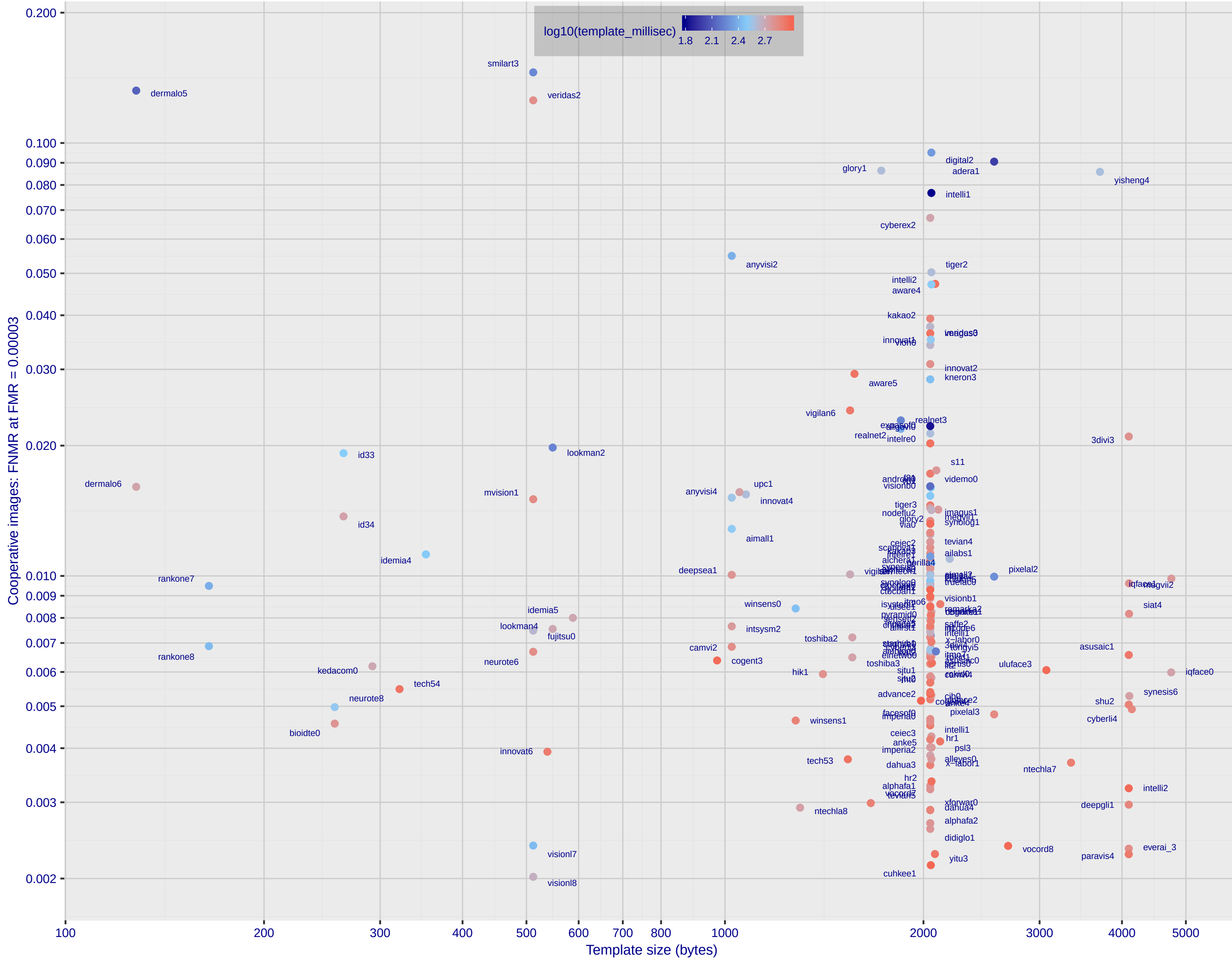


Figure 1: The points show false non-match rates (FNMR) versus the size of the encoded template. FNMR is the geometric mean of FNMR values for visa and mugshot images (from Figs. 33 and 44) at a false match rate (FMR) of 0.0001. The color of the points encodes template generation time - which spans at least one order of magnitude. Durations are measured on a single core of a c. 2016 Intel Xeon CPU E5-2630 v4 running at 2.20GHz. Algorithms with poor FNMR are omitted.

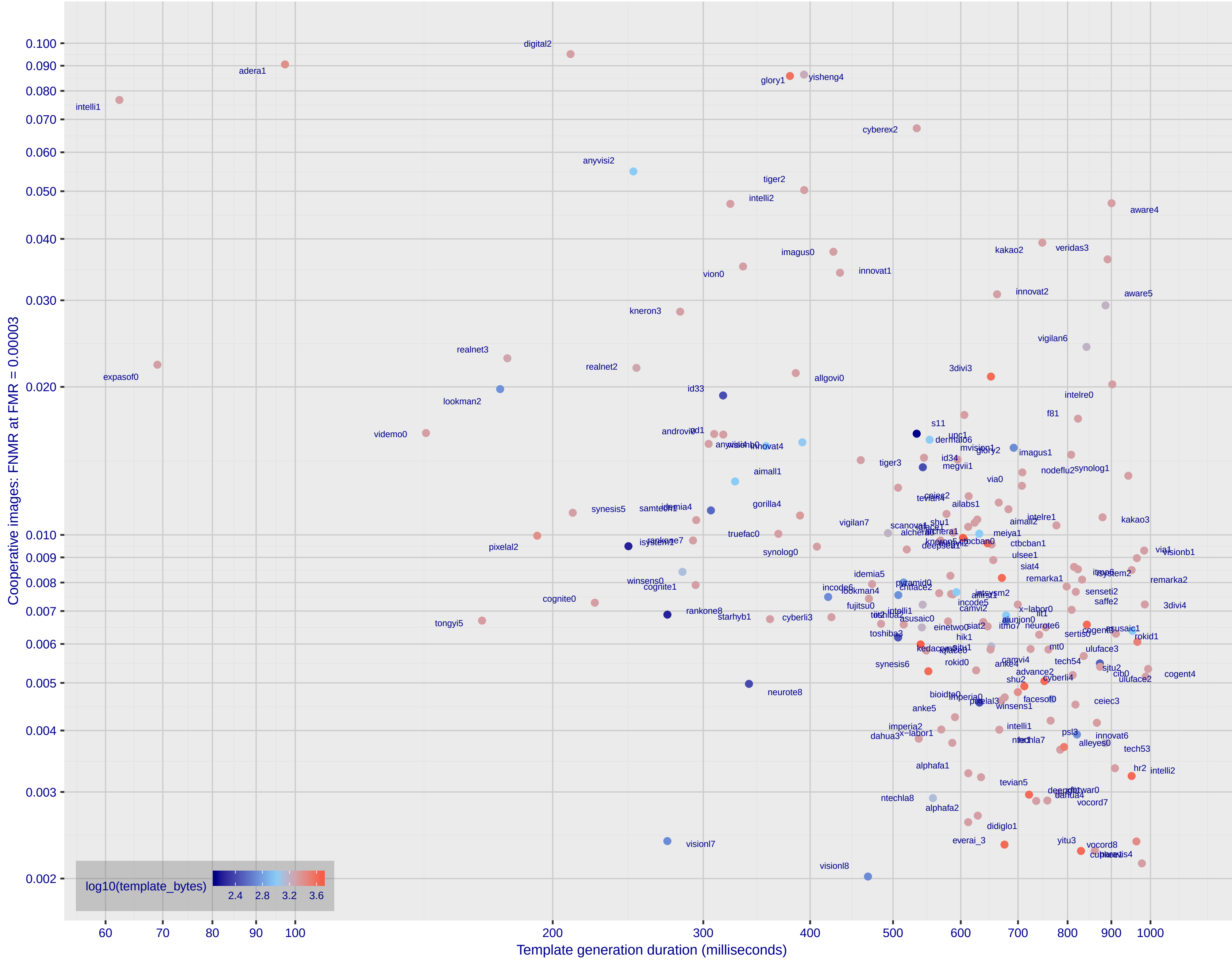


Figure 2: The points show false non-match rates (FNMR) versus the duration of the template generation operation. FNMR is the geometric mean of FNMR values for visa and mugshot images (from Figs. 33 and 44) at a false match rate (FMR) of 0.0001. Template generation time is a median estimated over 640 x 480 pixel portraits. It is measured on a single core of a c. 2016 Intel Xeon CPU E5-2630 v4 running at 2.20GHz. The color of the points encodes template size - which span two orders of magnitude. Algorithms with poor FNMR are omitted.

1 Metrics

1.1 Core accuracy

Given a vector of N genuine scores, u, the false non-match rate (FNMR) is computed as the proportion below some threshold, T:

$$\text{FNMR}(T) = 1 - \frac{1}{N} \sum_{i=1}^{N} H(u_i - T) \tag{1}$$

where $H(x)$ is the unit step function, and $H(0)$ taken to be 1.

Similarly, given a vector of N impostor scores, v, the false match rate (FMR) is computed as the proportion above T:

$$\text{FMR}(T) = \frac{1}{N} \sum_{i=1}^{N} H(v_i - T) \tag{2}$$

The threshold, T, can take on any value. We typically generate a set of thresholds from quantiles of the observed impostor scores, v, as follows. Given some interesting false match rate range, $[\text{FMR}_L, \text{FMR}_U]$, we form a vector of K thresholds corresponding to FMR measurements evenly spaced on a logarithmic scale

$$T_k = Q_v(1 - \text{FMR}_k) \tag{3}$$

where Q is the quantile function, and FMR_k comes from

$$\log_{10} \text{FMR}_k = \log_{10} \text{FMR}_L + \frac{k}{K} \left[\log_{10} \text{FMR}_U - \log_{10} \text{FMR}_L\right] \tag{4}$$

Error tradeoff characteristics are plots of FNMR(T) vs. FMR(T). These are plotted with $\text{FMR}_U \rightarrow 1$ and FMR_L as low as is sustained by the number of impostor comparisons, N. This is somewhat higher than the "rule of three" limit $3/N$ because samples are not independent, due to re-use of images.

2 Datasets

2.1 Child exploitation images

- ▷ The number of images is on the order of 10^4.
- ▷ The number of subjects is on the order of 10^3.
- ▷ The number of subjects with two images on the order of 10^3.
- ▷ The images are operational. They are taken from ongoing investigations of child exploitation crimes. The images are arbitrarily unconstrained. Pose varies considerably around all three axes, including subject lying down. Resolution varies very widely. Faces can be occluded by other objects, including hair and hands. Lighting varies, although the images are intended for human viewing. Mis-focus is rare. Images are given to the algorithm without any cropping; faces may occupy widely varying areas.
- ▷ The images are usually large from contemporary cameras. The mean interocular distance (IOD) is 70 pixels.
- ▷ The images are of subjects from several countries, due to the global production of this imagery.
- ▷ The images are of children, from infancy to late adolescence.
- ▷ All of the images are live capture, none are scanned. Many have been cropped.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”EXPLOITATION” - see Table 4 of the FRVT API.

2.2 Visa images

- ▷ The number of images is on the order of 10^5.
- ▷ The number of subjects is on the order of 10^5.
- ▷ The number of subjects with two images is on the order of 10^4.
- ▷ The images have geometry in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type. Pose is generally excellent.
- ▷ The images are of size 252x300 pixels. The mean interocular distance (IOD) is 69 pixels.
- ▷ The images are of subjects from greater than 100 countries, with significant imbalance due to visa issuance patterns.
- ▷ The images are of subjects of all ages, including children, again with imbalance due to visa issuance demand.
- ▷ Many of the images are live capture. A substantial number of the images are photographs of paper photographs.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”ISO” - see Table 4 of the FRVT API.

2.3 Application images

- ▷ The number of images is on the order of 10^6.
- ▷ The number of subjects is on the order of 10^6.
- ▷ The number of subjects with two images is on the order of 10^6.

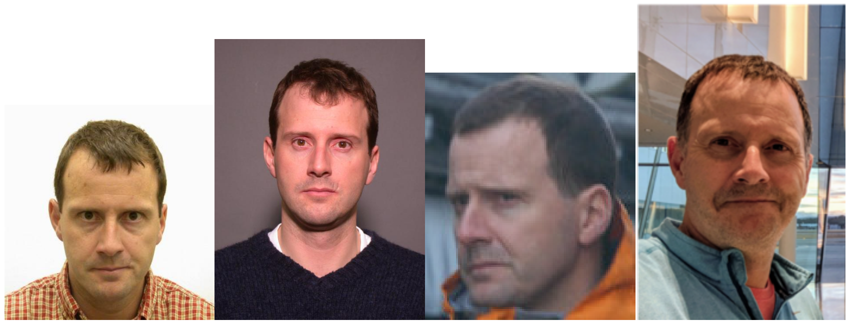
- ▷ The images have geometry in good conformance with the ISO/IEC 19794-5 Full Frontal image type. Pose is generally excellent.
- ▷ The images are of size 300x300 pixels. The mean interocular distance (IOD) is 61 pixels.
- ▷ The images are of subjects from greater than 100 countries, with significant imbalance due to population and immigration patterns.
- ▷ The images are of subjects of adults with imbalance due due to population and immigration patterns and demand.
- ▷ All of the images are live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”ISO” - see Table 4 of the FRVT API.

2.4 Border crossing images

- ▷ The number of images is on the order of 10^6.
- ▷ The number of subjects is on the order of 10^6.
- ▷ The number of subjects with two images is on the order of 10^6.
- ▷ The images are taken with at camera oriented by an attendant toward a cooperating subject. This is done under time constraints so there are role, pitch and yaw angle variations. Also background illumination is sometimes strong, so the face is under-exposed. There is some perspective distortion due to close range images. Some faces are partially cropped.
- ▷ The images are of subjects from greater than 100 countries, with significant imbalance due to population and immigration patterns.
- ▷ The images are of subjects of adults with imbalance due due to population and immigration patterns and demand.
- ▷ The images have mean IOD of 38 pixels.
- ▷ The images are all live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”WILD” - see Table 4 of the FRVT API.

2.5 Mugshot images

- ▷ The number of images is on the order of 10^6.
- ▷ The number of subjects is on the order of 10^6.
- ▷ The number of subjects with two images is on the order of 10^6.
- ▷ The images have geometry in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type.
- ▷ The images are of variable sizes. The median IOD is 105 pixels. The mean IOD is 113 pixels. The 1-st, 5-th, 10-th, 25-th, 75-th, 90-th and 99-th percentiles are 34, 58, 70, 87, 121, 161 and 297 pixels.
- ▷ The images are of subjects from the United States.
- ▷ The images are of adults.
- ▷ The images are all live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”mugshot” - see Table 4 of the FRVT API.



(a) Visa (b) Mugshot (c) Wild (d) Border (from VisaBorder)

Figure 3: The figure gives simulated samples of image types used in this report.

2.6 Wild images

- ▷ The number of images is on the order of 10^5.
- ▷ The number of subjects is on the order of 10^3.
- ▷ The number of subjects with two images on the order of 10^3.
- ▷ The images include many photojournalism-style images. Images are given to the algorithm using a variable but generally tight crop of the head. Resolution varies very widely. The images are very unconstrained, with wide yaw and pitch pose variation. Faces can be occluded, including hair and hands.
- ▷ The images are of adults.
- ▷ All of the images are live capture, none are scanned.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”WILD” - see Table 4 of the FRVT API.

3 Results

3.1 Test goals

- ▷ To state overall accuracy.
- ▷ To compare algorithms.

3.2 Test design

Method: For visa images:

- ▷ The comparisons are of visa photos against visa photos.
- ▷ The number of genuine comparisons is on the order of 10^4.
- ▷ The number of impostor comparisons is on the order of 10^{10}.

- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates. However, later analysis is conducted on subsets.
- ▷ The number of persons is on the order of 10^5.
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For mugshot images:

- ▷ The comparisons are of mugshot photos against mugshot photos.
- ▷ The number of genuine comparisons is on the order of 10^6.
- ▷ The number of impostor comparisons is on the order of 10^8.
- ▷ The impostors are paired by sex, but not by age or other covariates.
- ▷ The number of persons is on the order of 10^6.
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

Method: For wild images:

- ▷ The comparisons are of wild photos against wild photos.
- ▷ The number of genuine comparisons is on the order of 10^6.
- ▷ The number of impostor comparisons is on the order of 10^7.
- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.
- ▷ The number of persons is on the order of 10^4.
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For child exploitation images:

- ▷ The comparisons are of unconstrained child exploitation photos against others of the same type.
- ▷ The number of genuine comparisons is on the order of 10^4.
- ▷ The number of impostor comparisons is on the order of 10^7.
- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.
- ▷ The number of persons is on the order of 10^3.
- ▷ The number of images used to make 1 template is 1.

- The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.
- We produce two performance statements. First, is a DET as used for visa and mugshot images. The second is a cumulative match characteristic (CMC) summarizing a simulated one-to-many search process. This is done as follows.
  - We regard M enrollment templates as items in a gallery.
  - These M templates come from $M > N$ individuals, because multiple images of a subject are present in the gallery under separate identifiers.
  - We regard the verification templates as search templates.
  - For each search we compute the rank of the highest scoring mate.
  - This process should properly be conducted with a 1:N algorithm, such as those tested in NIST IR 8009. We use the 1:1 algorithms in a simulated 1:N mode here to a) better reflect what a child exploitation analyst does, and b) to show algorithm efficacy is better than that revealed in the verification DETs.

3.3 Failure to enroll

	Algorithm Name	Failure to Enrol Rate[1] CHILD-EXPLOIT		MUGSHOT		VISA		WILD	
1	3divi-003	0.1806	68	0.0007	160	0.0006	155	0.0294	177
2	3divi-004	0.2302	78	0.0008	164	0.0006	157	0.0222	173
3	acer-000	-	207	0.0002	114	0.0004	130	0.0008	111
4	adera-001	0.1928	72	0.0003	128	0.0005	149	0.0505	189
5	advance-002	-	207	0.0000	71	0.0004	111	0.0009	119
6	aifirst-001	0.0000	4	0.0000	3	0.0000	3	0.0000	45
7	ailabs-001	-	207	0.0007	163	0.0005	137	0.0016	130
8	aimall-001	0.0000	20	0.0000	21	0.0000	22	0.0001	53
9	aimall-002	-	207	0.0012	179	0.0005	147	0.0005	96
10	aiunionface-000	-	207	0.0000	29	0.0000	32	0.0000	47
11	alchera-000	-	207	0.0004	141	0.0014	187	0.0038	140
12	alchera-001	-	207	0.0004	140	0.0014	186	0.0038	139
13	alleyes-000	-	207	0.0002	109	0.0004	119	0.0004	89
14	allgovision-000	-	207	0.0026	193	0.0052	203	0.0131	161
15	alphaface-001	-	207	0.0000	74	0.0004	118	0.0004	76
16	alphaface-002	-	207	0.0000	75	0.0004	120	0.0004	78
17	amplifiedgroup-001	-	207	0.0189	206	0.0279	209	0.1390	203
18	androvideo-000	-	207	0.0000	18	0.0000	18	0.0002	57
19	anke-004	0.0944	55	0.0001	92	0.0004	121	0.0006	102
20	anke-005	0.1228	62	0.0001	103	0.0004	129	0.0007	105
21	antheus-000	0.0000	18	0.0000	47	0.0000	52	0.0242	174
22	anyvision-002	0.4866	95	0.0070	203	0.0090	206	0.1146	198
23	anyvision-004	0.1660	65	0.0001	102	0.0004	108	0.0080	149
24	asusaics-000	-	207	0.0000	23	0.0000	24	0.0000	16
25	asusaics-001	-	207	0.0000	52	0.0000	57	0.0000	42
26	aware-004	-	207	0.0002	112	0.0005	135	0.0014	128
27	aware-005	-	207	0.0001	108	0.0004	124	0.0011	123
28	awiros-001	-	207	0.0386	207	0.0872	210	0.3415	207
29	ayonix-000	0.0000	3	0.0113	204	0.0137	208	0.1194	199
30	bioidtechswiss-000	-	207	0.0003	126	0.0004	133	0.0006	97
31	bm-001	0.0000	22	0.0000	55	0.0000	25	0.0000	17
32	camvi-002	0.0000	5	0.0000	35	0.0000	38	0.0000	27
33	camvi-004	0.0000	10	0.0000	39	0.0000	43	0.0000	32
34	ceiec-002	0.2482	82	0.0036	197	0.0031	199	0.0081	150
35	ceiec-003	-	207	0.0001	80	0.0004	116	0.0004	73
36	chosun-000	-	207	0.0000	11	0.0000	11	0.0000	9
37	chtface-001	-	207	0.0000	42	0.0000	46	0.0000	35
38	chtface-002	-	207	0.0002	123	0.0007	160	0.0014	127
39	cib-000	-	207	0.0001	105	0.0000	40	0.0000	29
40	cogent-003	-	207	0.0001	86	0.0004	105	0.0009	121
41	cogent-004	0.0000	7	0.0000	5	0.0000	5	0.0000	3
42	cognitec-000	0.6342	99	0.0007	161	0.0007	164	0.0388	184
43	cognitec-001	-	207	0.0008	168	0.0010	167	0.0185	169
44	ctbcbank-000	0.3285	87	0.0011	176	0.0019	191	0.0868	197
45	ctbcbank-001	-	207	0.0005	153	0.0010	168	0.0844	196
46	cuhkee-001	-	207	0.0000	58	0.0004	101	0.1278	200
47	cyberextruder-001	0.5338	97	0.0024	191	0.0029	197	0.0597	193
48	cyberextruder-002	0.2672	85	0.0027	194	0.0028	196	0.0335	181
49	cyberlink-003	-	207	0.0001	78	0.0004	90	0.0008	113
50	cyberlink-004	-	207	0.0004	137	0.0003	78	0.0003	66
51	dahua-003	0.2222	77	0.0002	119	0.0003	68	0.0002	59
52	dahua-004	-	207	0.0001	101	0.0003	65	0.0002	60
53	deepglint-001	0.0000	23	0.0000	24	0.0000	26	0.0000	18
54	deepglint-002	0.0669	49	0.0002	120	0.0004	92	0.0003	65
55	deepsea-001	0.0000	11	0.0000	9	0.0000	9	0.0000	7
56	dermalog-005	0.1796	66	0.0013	183	0.0041	200	0.0163	167
57	dermalog-006	0.1797	67	0.0013	182	0.0041	201	0.0163	168
58	didiglobalface-001	0.2175	75	0.0000	73	0.0004	117	0.0004	75
59	digitalbarriers-002	-	207	0.0028	195	0.0027	195	0.0071	148
60	dsk-000	0.0000	17	0.0000	16	0.0000	16	0.0000	13

Table 11: FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a "small" template. This is defined as one whose size is less than 0.3 times the median template size for that algorithm. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails.

[1] *The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR.*

	Algorithm Name	Failure to Enrol Rate[1] CHILD-EXPLOIT		MUGSHOT		VISA		WILD	
61	einetworks-000	-	207	0.0002	117	0.0005	146	0.0008	114
62	everai-paravision-003	0.0705	51	0.0002	111	0.0004	91	0.0004	87
63	expasoft-000	-	207	0.0000	50	0.0000	55	0.0000	41
64	f8-001	0.2026	74	0.0035	196	0.0030	198	0.0087	153
65	facesoft-000	0.0000	29	0.0000	53	0.0000	58	0.0000	43
66	fujitsulab-000	-	207	0.0005	149	0.0002	61	0.0099	155
67	glory-001	0.0000	14	0.0051	200	0.0010	170	0.1651	204
68	glory-002	-	207	0.0015	185	0.0011	181	0.0557	191
69	gorilla-004	-	207	0.0000	61	0.0003	73	0.0004	83
70	gorilla-005	-	207	0.0000	60	0.0003	71	0.0004	81
71	hik-001	-	207	0.0000	41	0.0000	45	0.0000	34
72	hr-001	0.1198	61	0.0001	77	0.0004	95	0.0003	70
73	hr-002	-	207	0.0002	116	0.0004	122	0.0004	74
74	id3-003	0.3032	86	0.0016	188	0.0011	180	0.0317	179
75	id3-004	-	207	0.0015	187	0.0011	179	-	207
76	idemia-004	-	207	0.0000	63	0.0004	87	0.0003	71
77	idemia-005	0.0239	34	0.0000	59	0.0003	74	0.0003	64
78	iit-001	0.0843	54	0.0001	107	0.0004	113	0.0104	156
79	iit-002	-	207	0.0009	171	0.0005	152	0.0443	187
80	imagus-000	0.1107	59	0.0010	175	0.0012	182	0.0347	182
81	imagus-001	-	207	0.0001	93	0.0004	114	0.0396	185
82	imperial-000	-	207	0.0000	28	0.0000	31	0.0000	22
83	imperial-002	0.0000	9	0.0000	8	0.0000	8	0.0000	6
84	incode-005	-	207	0.0001	87	0.0004	112	0.0007	109
85	incode-006	-	207	0.0001	104	0.0004	80	0.0004	80
86	innovativetechnologyltd-001	-	207	0.0024	192	0.0025	194	0.0055	146
87	innovativetechnologyltd-002	-	207	0.0057	202	0.0005	144	0.0247	176
88	innovatrics-004	0.1170	60	0.0000	72	0.0004	115	0.0041	142
89	innovatrics-006	0.0350	37	0.0000	65	0.0004	84	0.0003	72
90	intellicloudai-001	-	207	0.0000	51	0.0000	56	0.0001	55
91	intellifusion-001	0.0949	56	0.0001	84	0.0003	77	0.0005	95
92	intellifusion-002	-	207	0.0000	56	0.0000	29	0.0001	54
93	intellivision-001	0.5495	98	0.0048	199	0.0042	202	0.1358	201
94	intellivision-002	-	207	0.0012	178	0.0005	153	0.0146	162
95	intelresearch-000	-	207	0.0000	67	0.0003	76	0.0001	56
96	intelresearch-001	-	207	0.0005	152	0.0010	169	0.0407	186
97	intsysmsu-001	-	207	0.0001	90	0.0004	109	0.0004	84
98	intsysmsu-002	-	207	0.0001	91	0.0004	110	0.0004	86
99	iqface-000	0.0000	12	0.0000	40	0.0000	44	0.0000	33
100	iqface-001	-	207	0.0000	54	0.0000	19	0.0000	14
101	isap-001	-	207	0.0000	38	0.0000	42	0.0000	31
102	isityou-000	0.4714	93	0.0023	189	0.0010	172	0.0663	195
103	isystems-001	0.1421	63	0.0010	173	0.0007	161	0.0128	159
104	isystems-002	0.1421	64	0.0010	174	0.0007	162	0.0128	160
105	itmo-006	-	207	0.0004	144	0.0004	107	0.0006	101
106	itmo-007	-	207	0.0003	134	0.0000	48	0.0004	82
107	kakao-002	0.2494	83	0.0002	121	0.0005	141	0.0310	178
108	kakao-003	-	207	0.0000	25	0.0000	27	0.0000	19
109	kedacom-000	0.0000	21	0.0000	22	0.0000	23	0.0000	15
110	kneron-003	0.4883	96	0.0044	198	0.0016	190	0.1823	205
111	kneron-005	-	207	0.0006	157	0.0005	142	0.0097	154
112	lookman-002	-	207	0.0000	37	0.0000	41	0.0000	30
113	lookman-004	0.0000	1	0.0000	33	0.0000	36	0.0000	25
114	luxand-000	-	207	0.0000	10	0.0000	10	0.0000	8
115	megvii-001	0.0274	36	0.0007	162	0.0004	89	0.0152	165
116	megvii-002	0.0274	35	0.0054	201	0.0004	88	0.0126	158
117	meiya-001	-	207	0.0004	145	0.0010	173	0.0025	135
118	microfocus-001	0.0791	53	0.0008	167	0.0016	189	0.0220	172
119	microfocus-002	0.0791	52	0.0008	166	0.0016	188	0.0220	171
120	mt-000	0.1043	57	0.0002	118	0.0004	123	0.0004	77

Table 12: FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a "small" template. This is defined as one whose size is less than 0.3 times the median template size for that algorithm. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails.

[1] *The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR.*

	Algorithm Name	Failure to Enrol Rate[1]							
		CHILD-EXPLOIT		MUGSHOT		VISA		WILD	
121	mvision-001	-	207	0.0000	34	0.0000	37	0.0000	26
122	netbridgetech-001	-	207	0.0000	7	0.0000	7	0.0000	5
123	neurotechnology-006	0.1068	58	0.0004	139	0.0004	97	0.0018	131
124	neurotechnology-008	-	207	0.0004	138	0.0004	94	0.0018	132
125	nodeflux-001	-	207	0.0001	94	0.0002	64	0.0003	63
126	nodeflux-002	-	207	0.0008	165	0.0005	145	0.0008	117
127	notiontag-000	0.0000	27	0.0000	30	0.0000	33	0.0000	23
128	ntechlab-007	0.0682	50	0.0001	79	0.0004	85	0.0005	94
129	ntechlab-008	-	207	0.0001	81	0.0004	82	0.0005	91
130	paravision-004	0.0570	44	0.0002	115	0.0004	100	0.0008	112
131	pixelall-002	0.0001	30	0.0000	15	0.0000	15	0.0001	49
132	pixelall-003	-	207	0.0000	44	0.0000	49	0.0000	36
133	psl-003	-	207	0.0000	69	0.0004	106	0.0003	69
134	psl-004	-	207	0.0000	68	0.0004	104	0.0003	67
135	pyramid-000	-	207	0.0005	150	0.0007	163	0.0015	129
136	rankone-007	0.3518	89	0.0003	129	0.0004	126	0.0043	143
137	rankone-008	-	207	0.0003	124	0.0004	79	0.0040	141
138	realnetworks-002	-	207	0.0004	136	0.0003	70	0.0004	85
139	realnetworks-003	0.0076	32	0.0004	135	0.0003	69	0.0004	88
140	remarkai-001	-	207	0.0000	32	0.0000	35	0.0000	48
141	remarkai-002	-	207	0.0000	20	0.0000	21	0.0000	46
142	rokid-000	-	207	0.0001	89	0.0005	143	0.0354	183
143	rokid-001	-	207	0.0000	19	0.0000	20	0.0007	110
144	s1-001	-	207	0.0013	180	0.0007	159	0.0600	194
145	saffe-001	0.0000	26	0.0000	27	0.0000	30	0.0000	21
146	saffe-002	-	207	0.0000	48	0.0000	53	0.0000	38
147	samtech-001	-	207	0.0004	143	0.0008	165	0.0013	126
148	scanovate-001	-	207	0.0024	190	0.0014	185	0.2751	206
149	sensetime-002	0.3345	88	0.0011	177	0.0005	151	0.0218	170
150	sensetime-003	0.0554	42	0.0000	57	0.0004	102	0.0004	90
151	sertis-000	-	207	0.0000	76	0.0004	93	0.0004	79
152	shaman-000	0.0000	6	0.0000	36	0.0000	39	0.0000	28
153	shaman-001	0.0000	2	0.0000	2	0.0000	2	0.0000	44
154	shu-001	0.1822	69	0.0010	172	0.0006	154	0.0499	188
155	shu-002	-	207	0.0005	146	0.0004	125	0.0007	106
156	siat-002	0.0616	46	0.0000	66	0.0004	103	0.0048	145
157	siat-004	-	207	0.0000	64	0.0004	99	0.0003	68
158	sjtu-001	-	207	0.0005	148	0.0004	132	0.0008	115
159	sjtu-002	-	207	0.0005	147	0.0004	128	0.0007	108
160	smilart-002	0.2422	80	0.0003	133	0.0011	176	0.0575	192
161	smilart-003	-	207	0.0014	184	0.0013	184	0.0555	190
162	starhybrid-001	0.2340	79	0.0009	170	0.0023	193	0.0044	144
163	synesis-005	0.1862	70	0.0001	88	0.0005	136	0.0021	133
164	synesis-006	-	207	0.0000	70	0.0003	66	0.0002	62
165	synology-000	-	207	0.0000	13	0.0000	13	0.0000	11
166	synology-001	-	207	0.0000	45	0.0000	50	0.0000	37
167	tech5-003	0.0016	31	0.0001	85	0.0003	67	0.0002	58
168	tech5-004	-	207	0.0003	127	0.0004	131	0.0006	98
169	tevian-004	-	207	0.0002	113	0.0005	148	0.0057	147
170	tevian-005	0.3606	90	0.0006	158	0.0006	158	0.0012	124
171	tiger-002	0.0619	47	0.0001	97	0.0004	98	0.0082	152
172	tiger-003	0.0619	48	0.0001	95	0.0004	96	0.0082	151
173	tongyi-005	0.0000	8	0.0000	6	0.0000	6	0.0000	4
174	toshiba-002	0.0000	16	0.0000	14	0.0000	14	0.0000	12
175	toshiba-003	-	207	0.0001	98	0.0001	60	0.0002	61
176	trueface-000	-	207	0.0000	49	0.0000	54	0.0000	39
177	tuputech-000	-	207	0.0632	208	0.0081	205	0.6383	208
178	ulsee-001	-	207	0.0000	46	0.0000	51	0.0001	50
179	uluface-002	0.0000	28	0.0000	31	0.0000	34	0.0000	24
180	uluface-003	-	207	0.0002	110	0.0002	62	0.0244	175

Table 13: FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a "small" template. This is defined as one whose size is less than 0.3 times the median template size for that algorithm. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails.

[1] *The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR.*

	Algorithm	Failure to Enrol Rate[1]							
	Name	CHILD-EXPLOIT		MUGSHOT		VISA		WILD	
181	upc-001	0.0450	39	0.0003	125	0.0003	75	0.0011	122
182	vd-001	-	207	0.0004	142	0.0009	166	0.0024	134
183	veridas-002	-	207	0.0001	99	0.0005	140	0.0006	100
184	veridas-003	0.1893	71	0.0001	96	0.0005	139	0.0006	99
185	via-000	0.0000	13	0.0000	43	0.0000	47	0.0001	52
186	via-001	-	207	0.0000	17	0.0000	17	0.0001	51
187	videmo-000	-	207	0.0003	130	0.0012	183	0.0158	166
188	videonetics-001	0.4799	94	0.0015	186	0.0010	171	0.0112	157
189	videonetics-002	0.4598	92	0.0006	159	0.0005	150	0.0013	125
190	vigilantsolutions-006	-	207	0.0001	83	0.0004	86	0.0005	93
191	vigilantsolutions-007	0.2538	84	0.0001	82	0.0004	83	0.0005	92
192	vion-000	0.6388	100	0.0130	205	0.0078	204	0.1389	202
193	visionbox-000	-	207	0.0005	156	0.0011	178	0.0028	138
194	visionbox-001	-	207	0.0005	155	0.0011	177	0.0028	137
195	visionlabs-007	0.1939	73	0.0003	131	0.0005	138	0.0008	116
196	visionlabs-008	-	207	0.0002	122	0.0004	134	0.0009	118
197	visteam-000	-	207	0.0005	151	0.0011	174	0.0026	136
198	vocord-007	0.0000	19	0.0001	106	0.0004	81	0.0009	120
199	vocord-008	-	207	0.0003	132	0.0001	59	0.0007	107
200	winsense-000	0.0000	15	0.0000	12	0.0000	12	0.0000	10
201	winsense-001	0.0000	24	0.0000	26	0.0000	28	0.0000	20
202	x-laboratory-000	0.0000	25	0.0005	154	0.0002	63	0.0000	40
203	x-laboratory-001	-	207	0.0001	100	0.0004	127	0.0007	103
204	xforwardai-000	-	207	0.0000	1	0.0000	1	0.0000	1
205	yisheng-004	0.4279	91	0.0013	181	0.0006	156	0.0321	180
206	yitu-003	-	207	0.0009	169	0.0000	4	0.0000	2

Table 14: FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a "small" template. This is defined as one whose size is less than 0.3 times the median template size for that algorithm. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails.

[1] *The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR.*

3.4 Recognition accuracy

Core algorithm accuracy is stated via:

- **Cooperative subjects**
  - The summary table of Figure 10;
  - The visa image DETs of Figure 33;
  - The mugshot DETs of Figure 44;
  - The mugshot ageing profiles of Figure 166;
  - The human-difficult pairs of Figure 12
- **Non-cooperative subjects**
  - The photojournalism DET of Figure 53
  - The child-exploitation DET of Figure 57;
  - The child-exploitation CMC of Figure 61.

Figure 135 shows dependence of false match rate on algorithm score threshold. This allows a deployer to set a threshold to target a particular false match rate appropriate to the security objectives of the application.

Figure 113 likewise shows FMR(T) but for mugshots, and specially four subsets of the population.

Note that in both the mugshot and visa sets false match rates vary with the ethnicity, age, and sex, of the enrollee and impostor - see section 3.6. For example figure 73 summarizes FMR for impostors paired from four groups black females, black males, white females, white males.

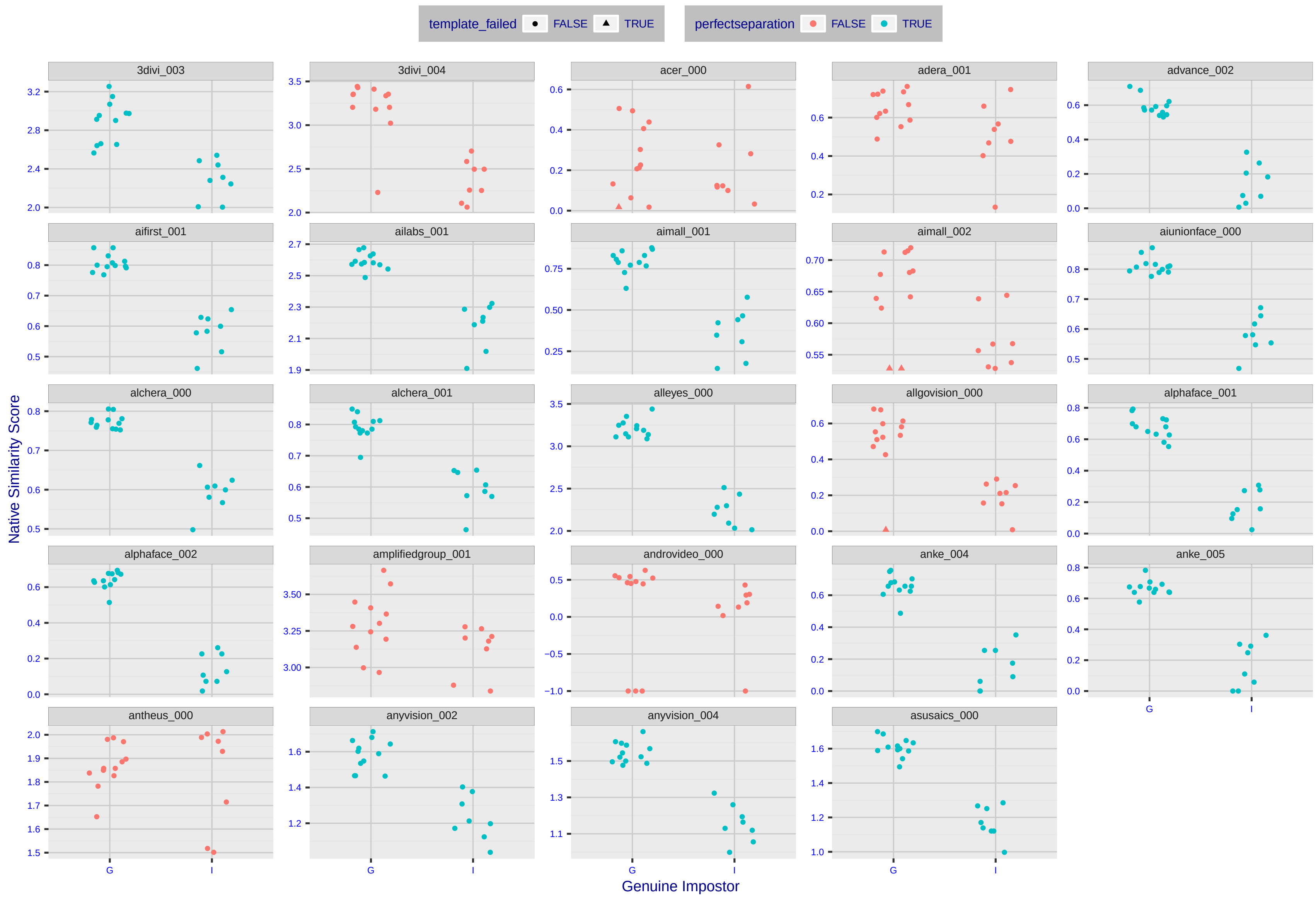


Figure 4: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

Figure 5: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

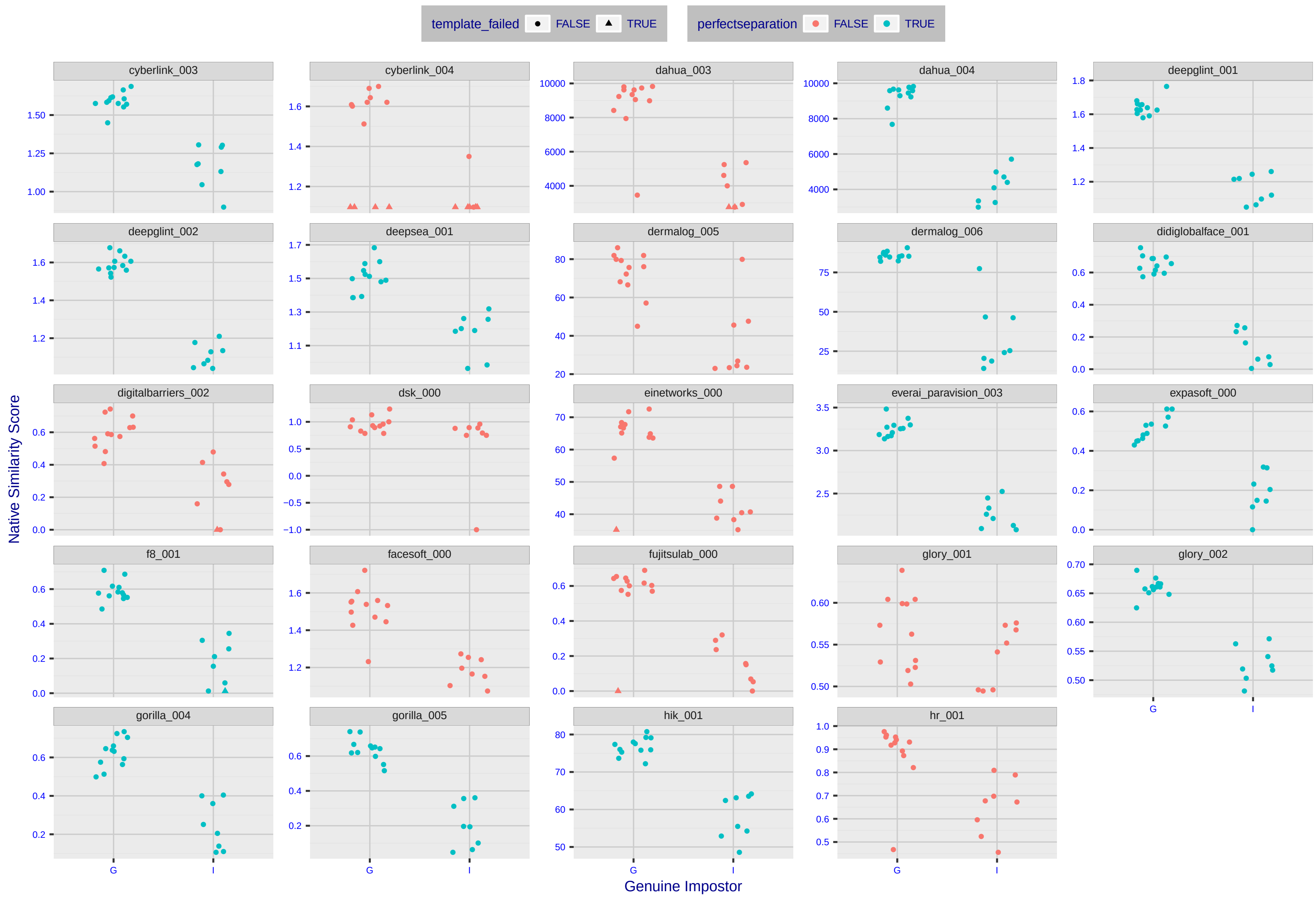


Figure 6: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

Figure 7: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

Figure 8: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

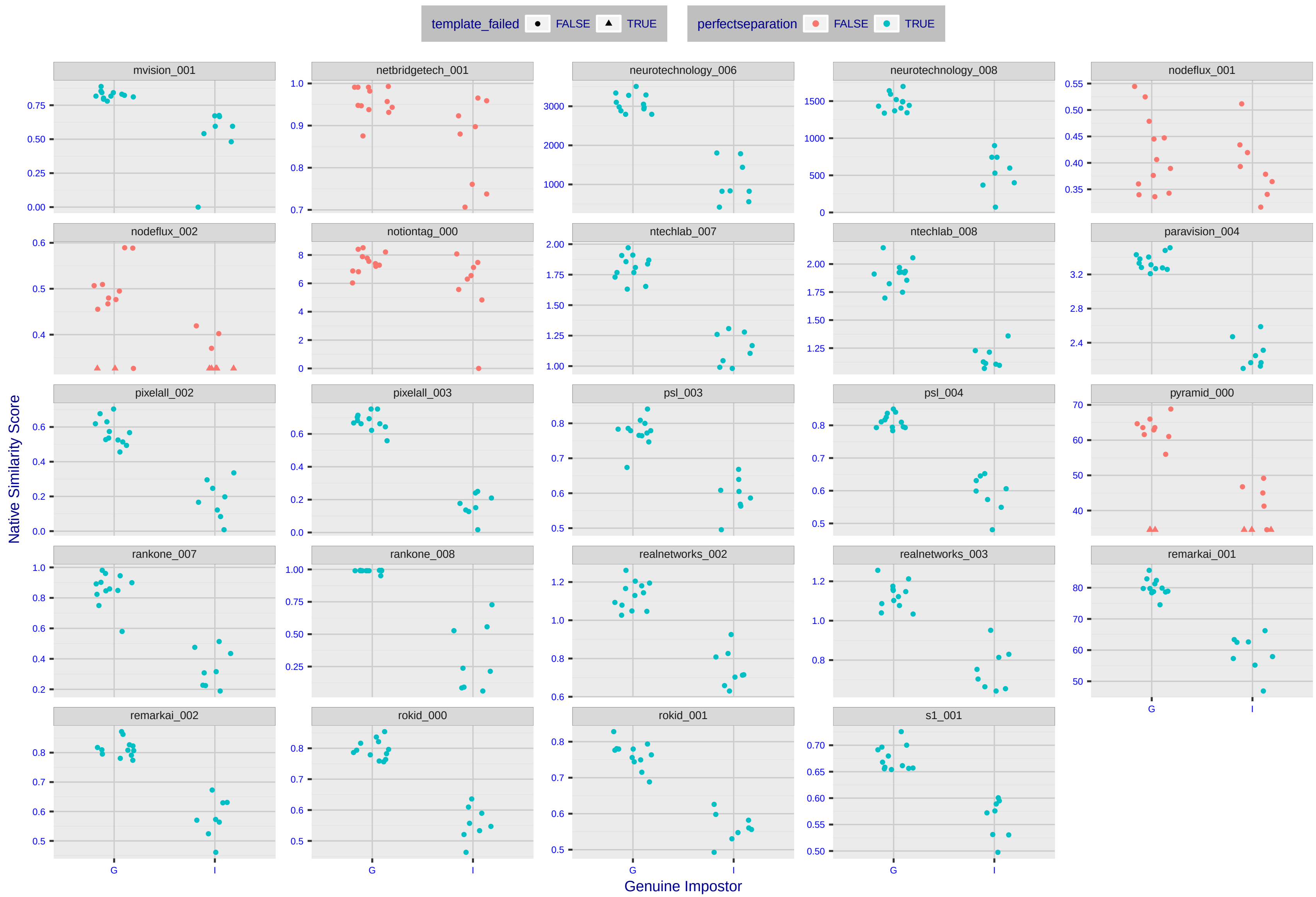


Figure 9: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

Figure 10: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

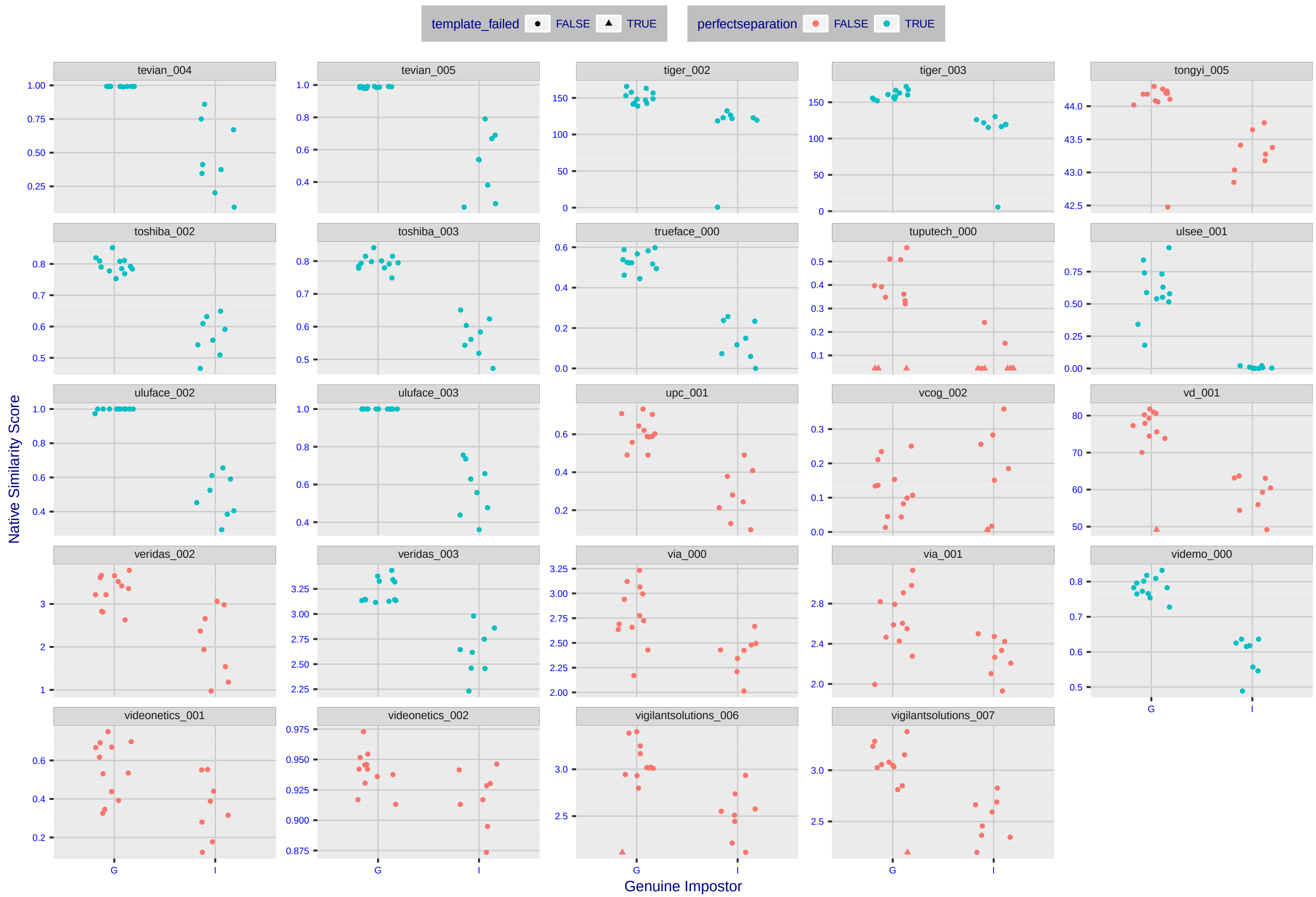


Figure 11: The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.

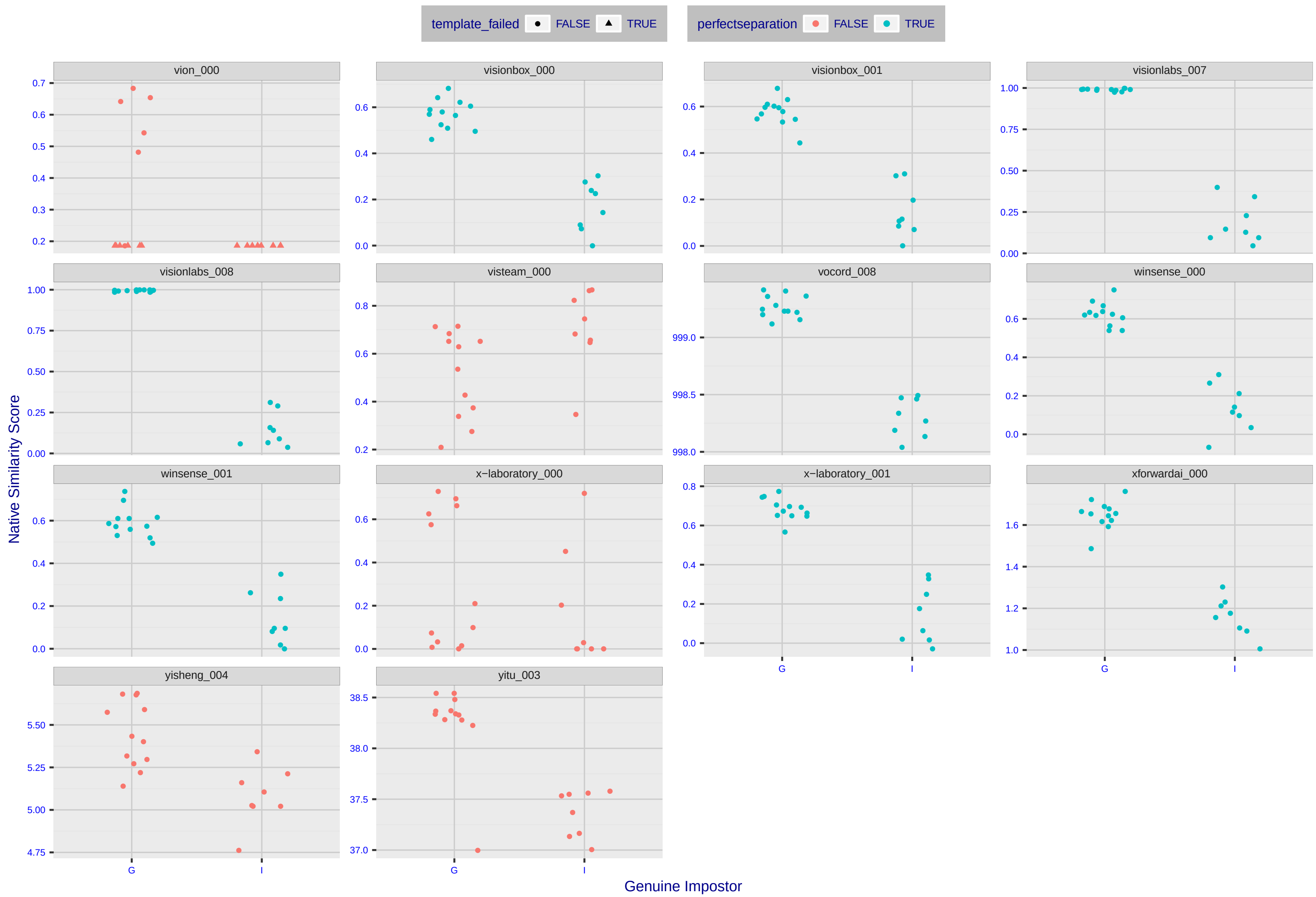


Figure 12: *The Figure shows, in blue, algorithms that correctly separate the 12 genuine and 8 impostor pairs used in the May 2018 paper Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms (Phillips et al. [1]). In red are algorithms that are imperfect. Some algorithms fail only because they failed to make a template e.g. due to face detection failure (shown as a triangle). Others fail because the pairs were selected for that study because they had been difficult for three leading algorithms used in FRVT 2006. Caution: Given the small sample size (n=20) the figure may change substantially if larger or different sets were used. The images can be downloaded from the Supplemental Information page provided with that publication.*

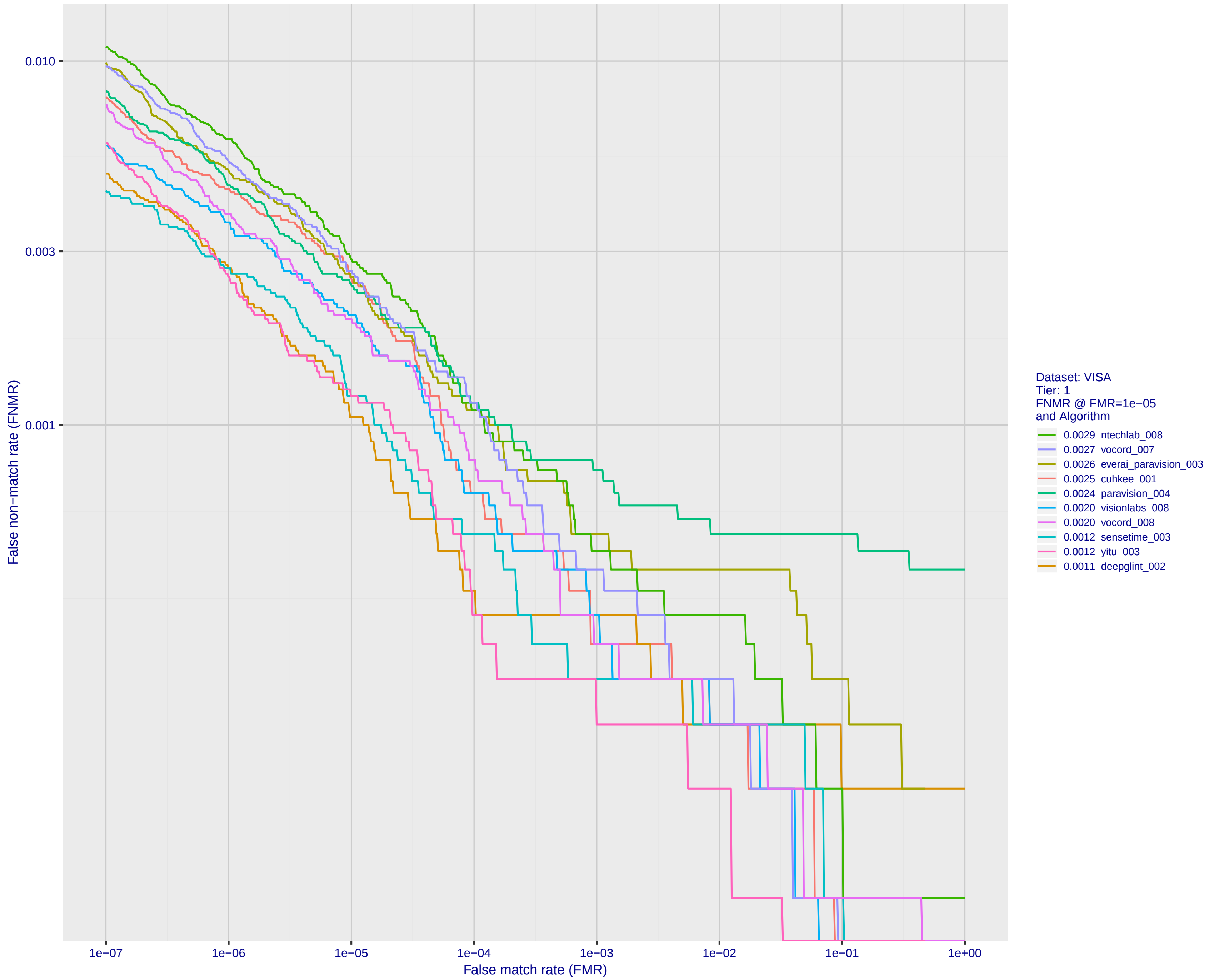


Figure 13: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

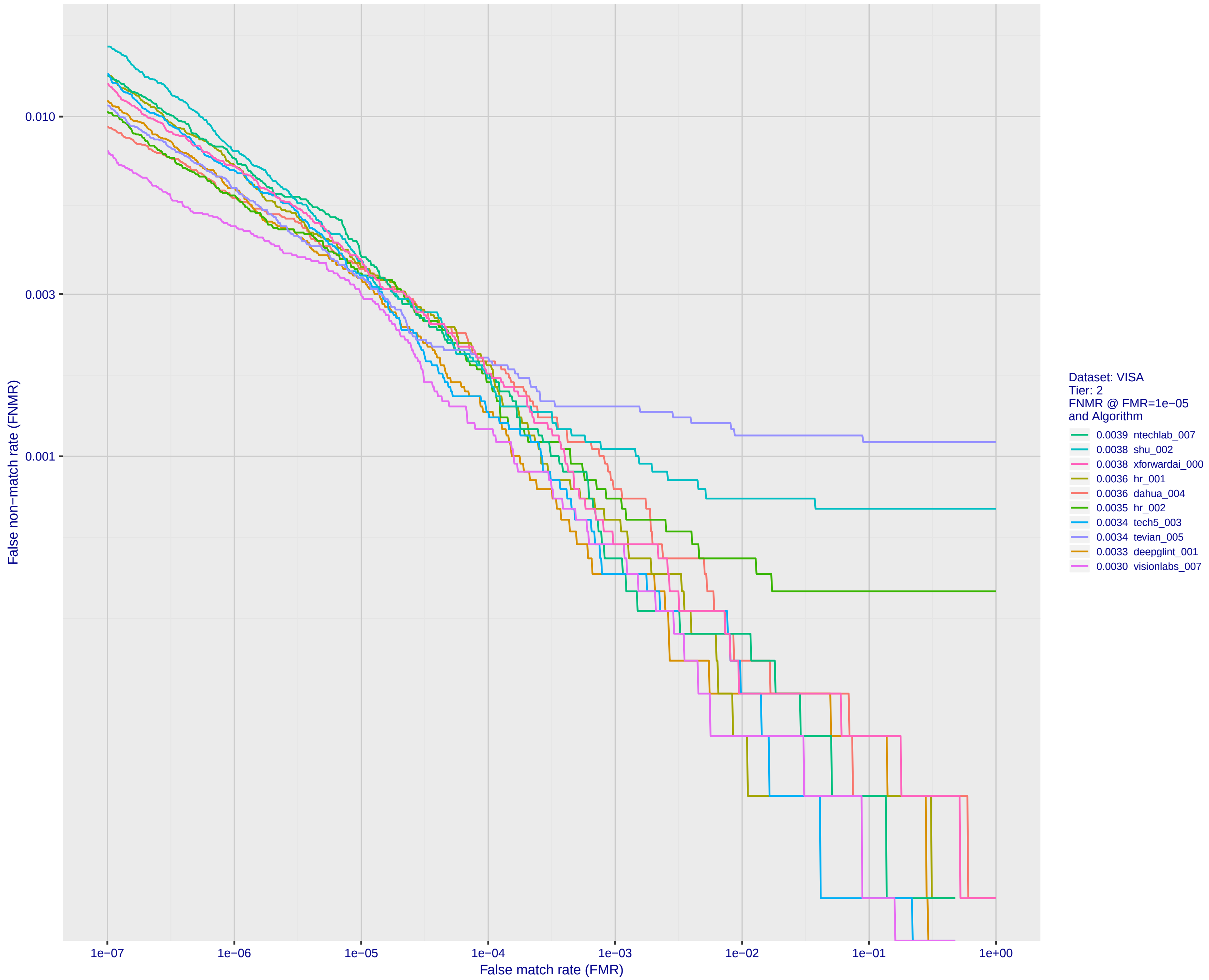


Figure 14: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

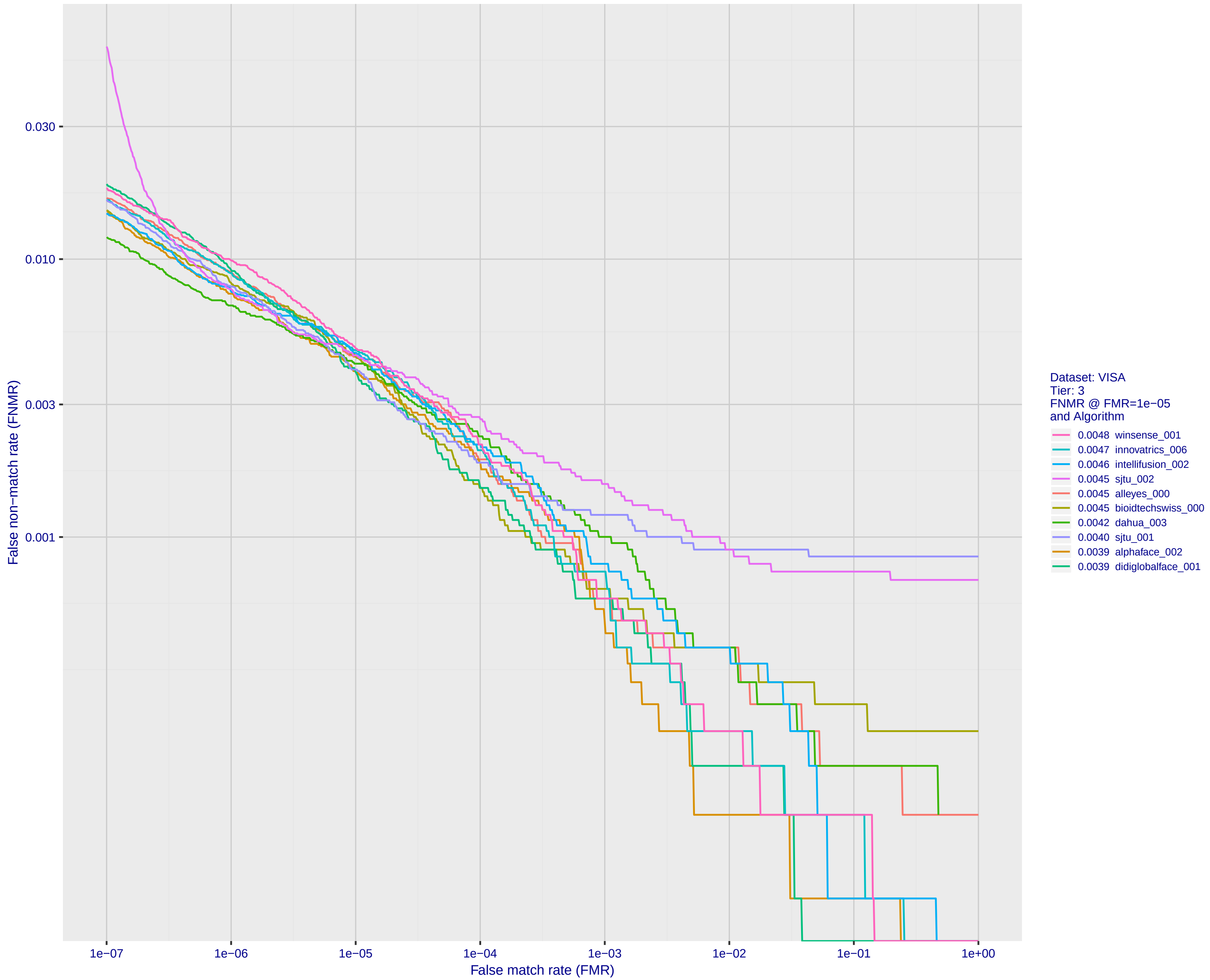


Figure 15: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

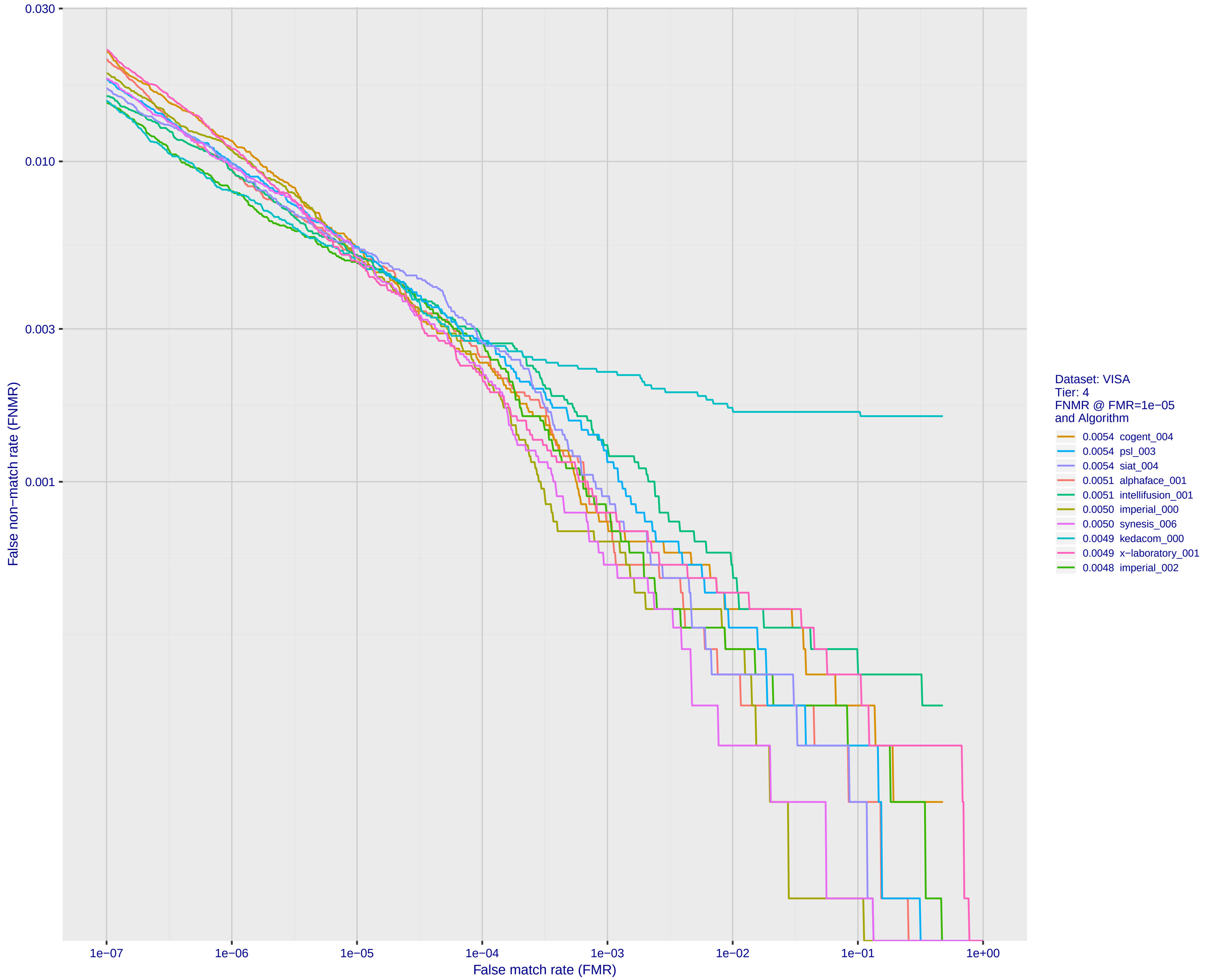


Figure 16: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

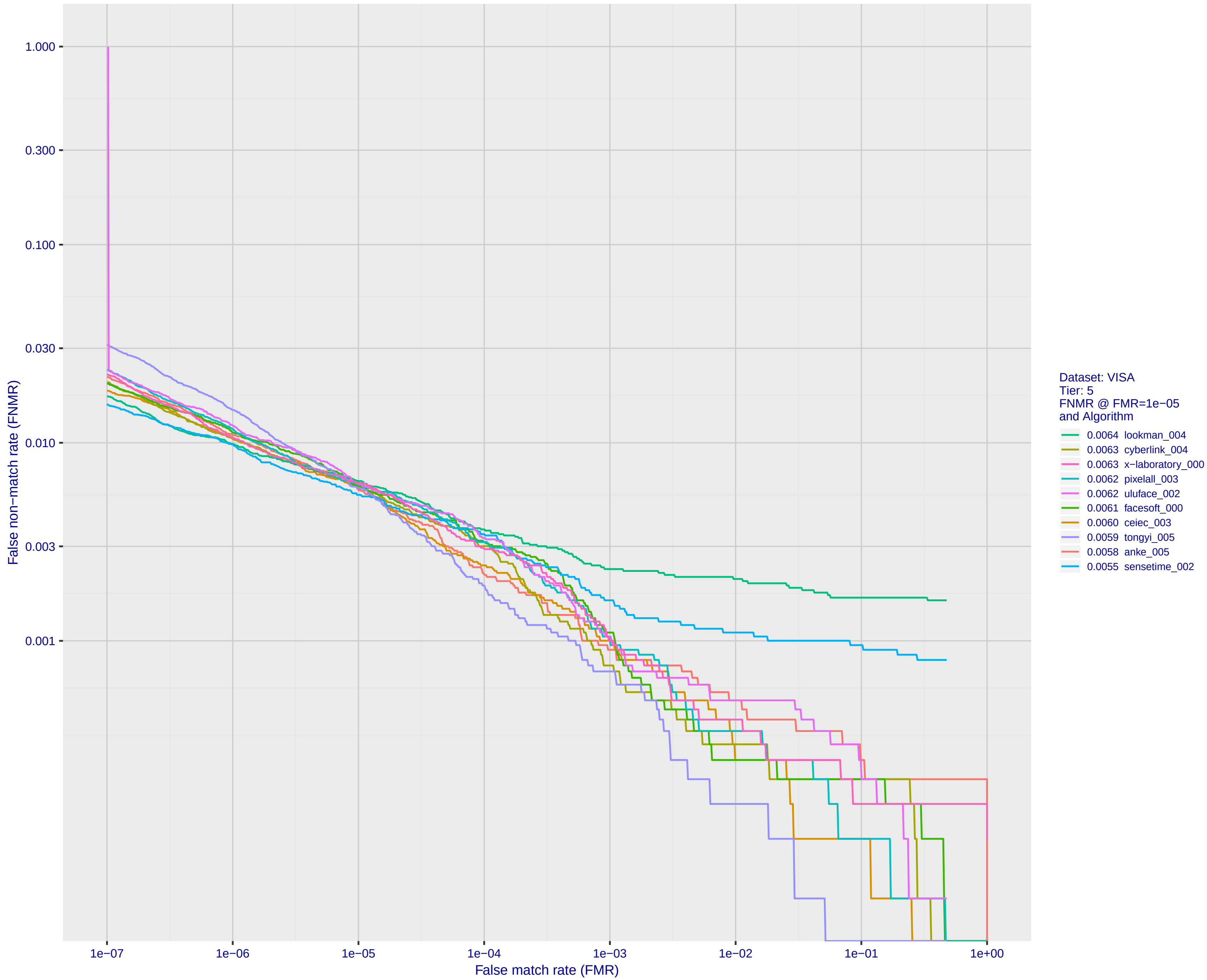


Figure 17: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

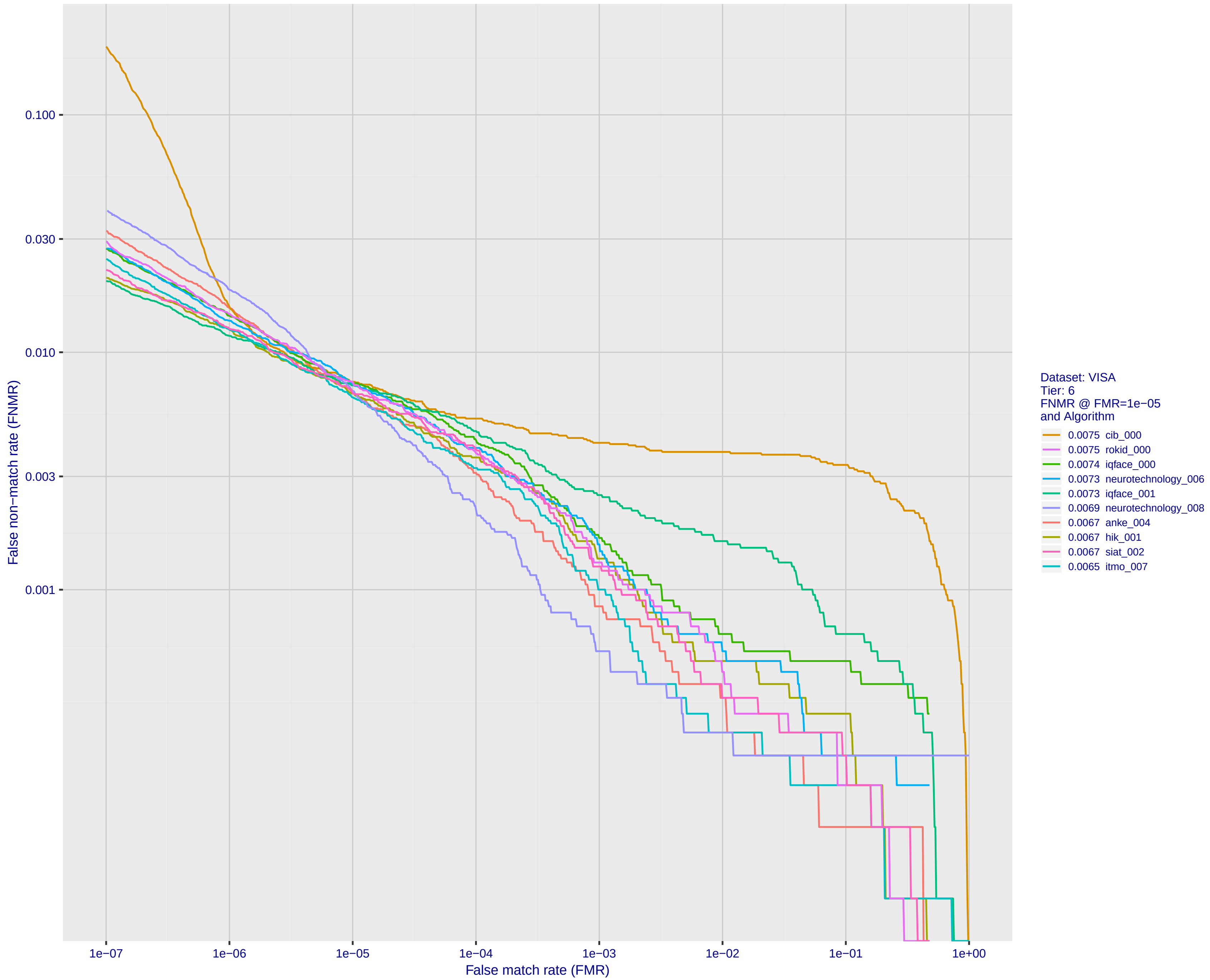


Figure 18: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

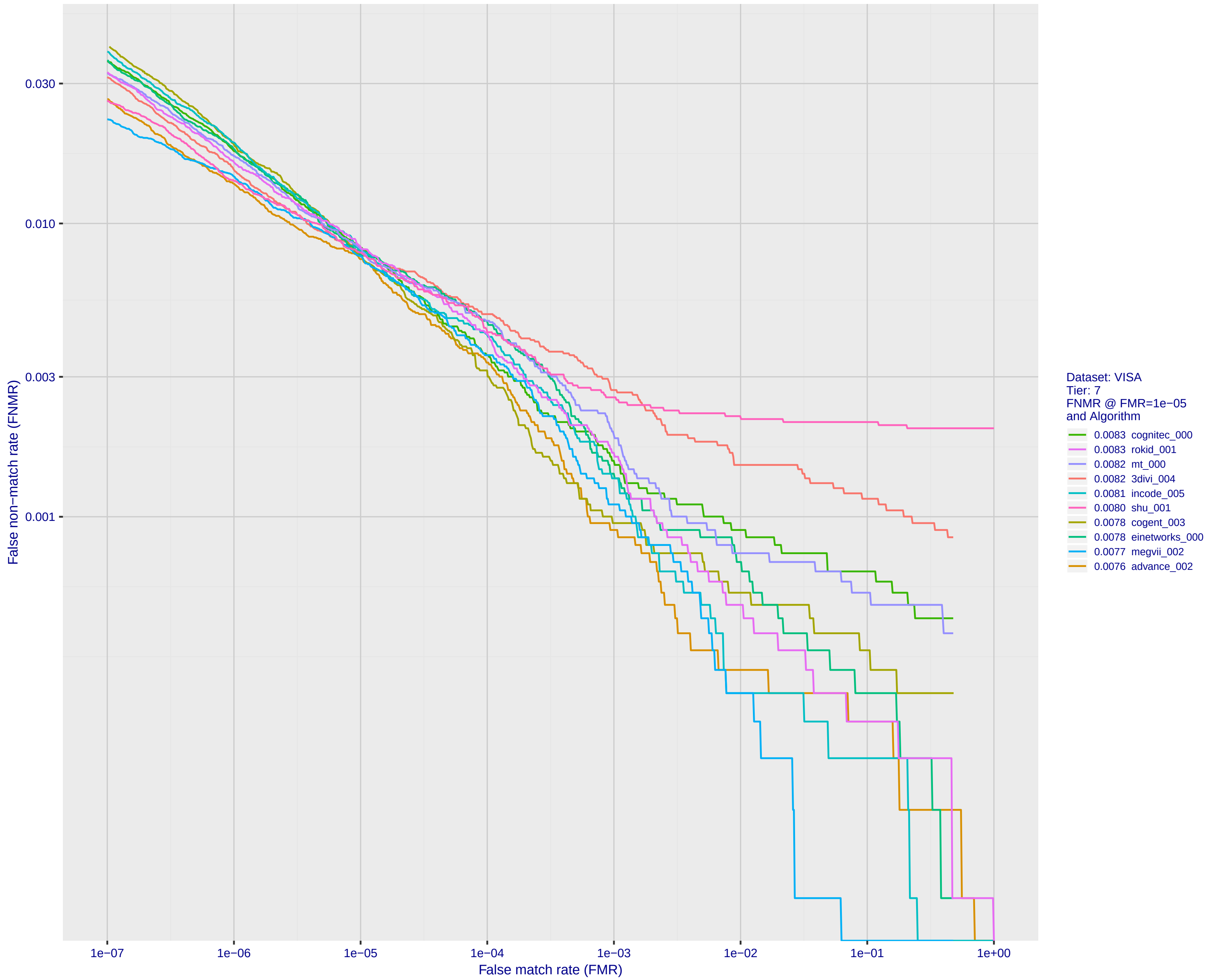


Figure 19: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

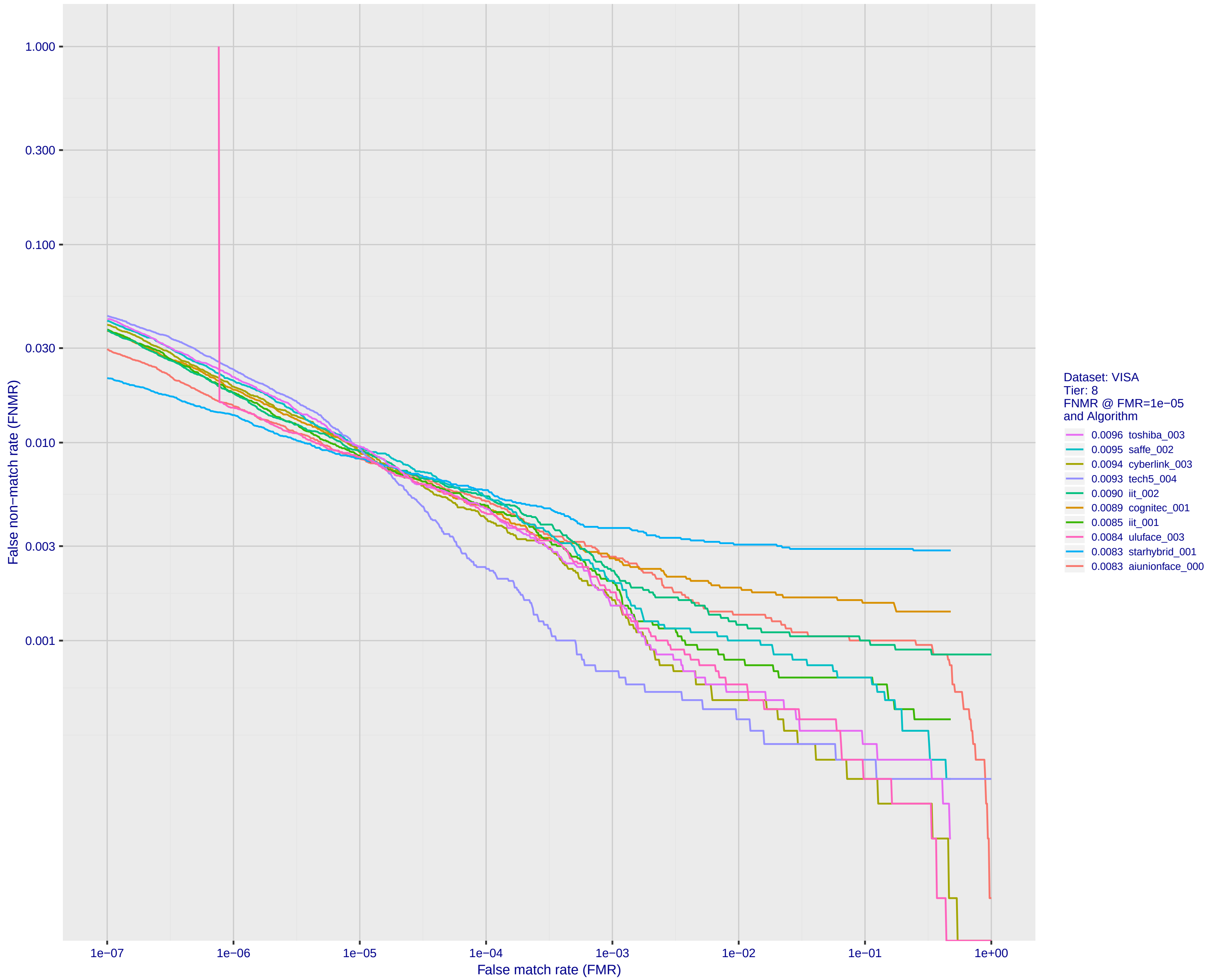


Figure 20: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

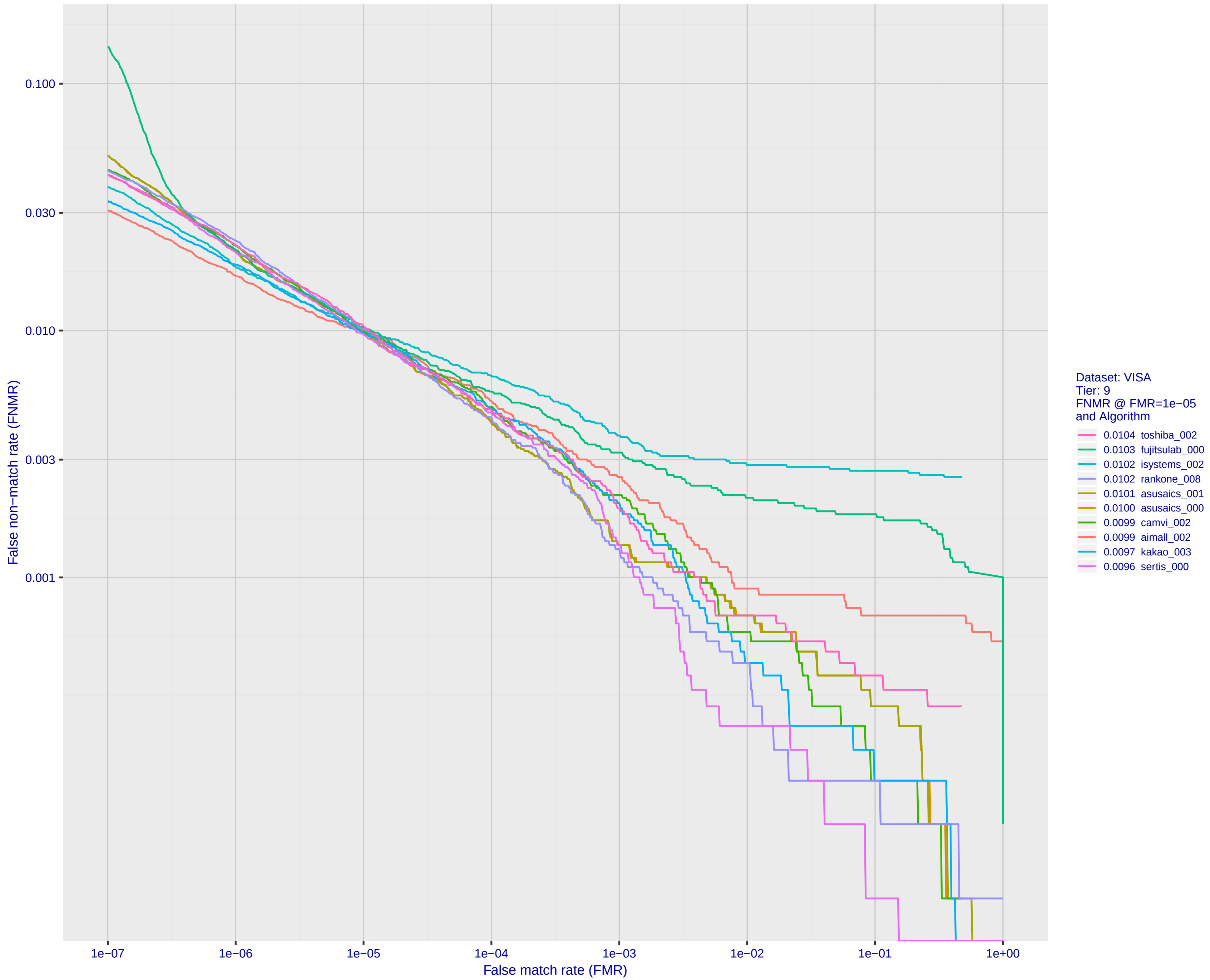


Figure 21: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

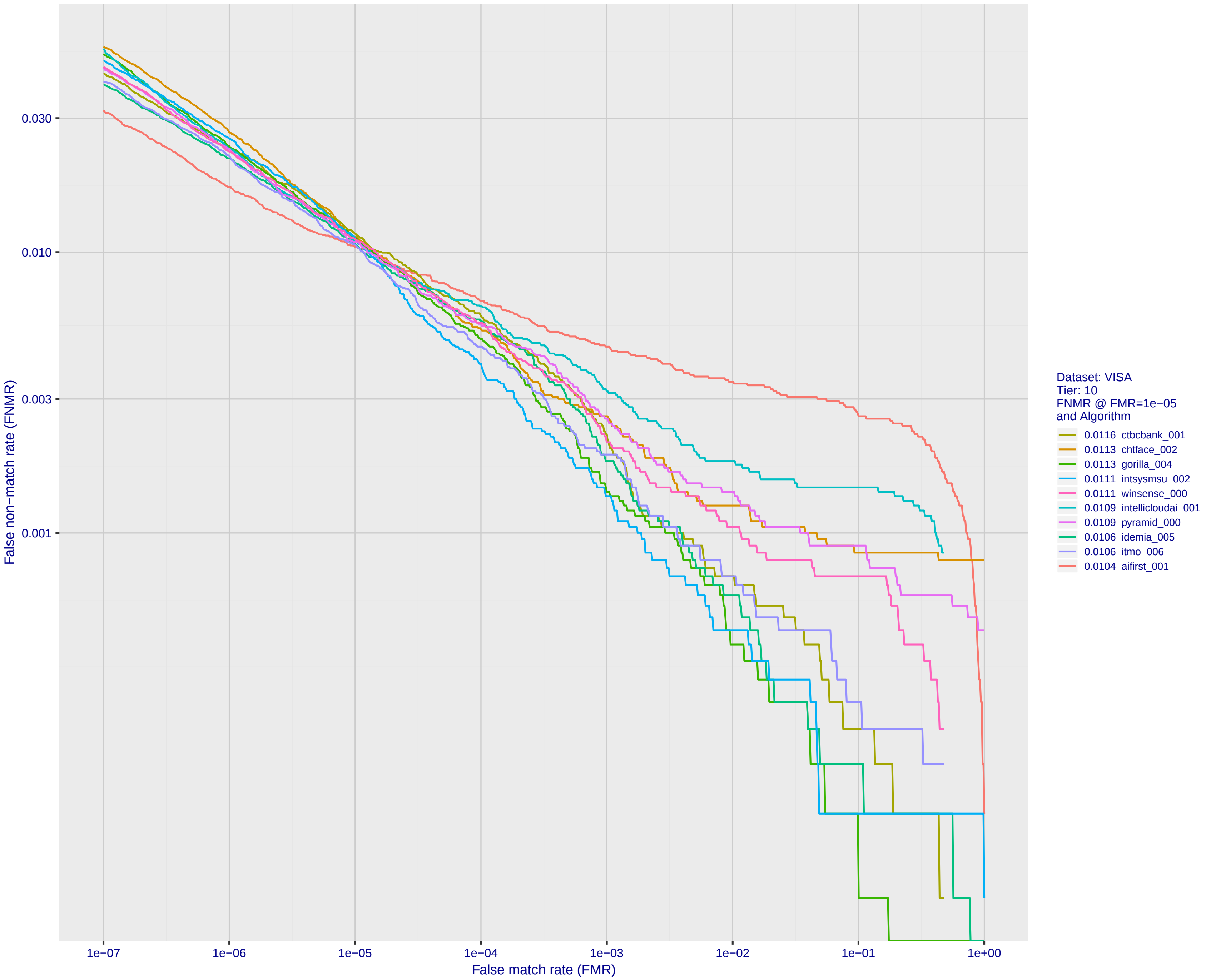


Figure 22: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

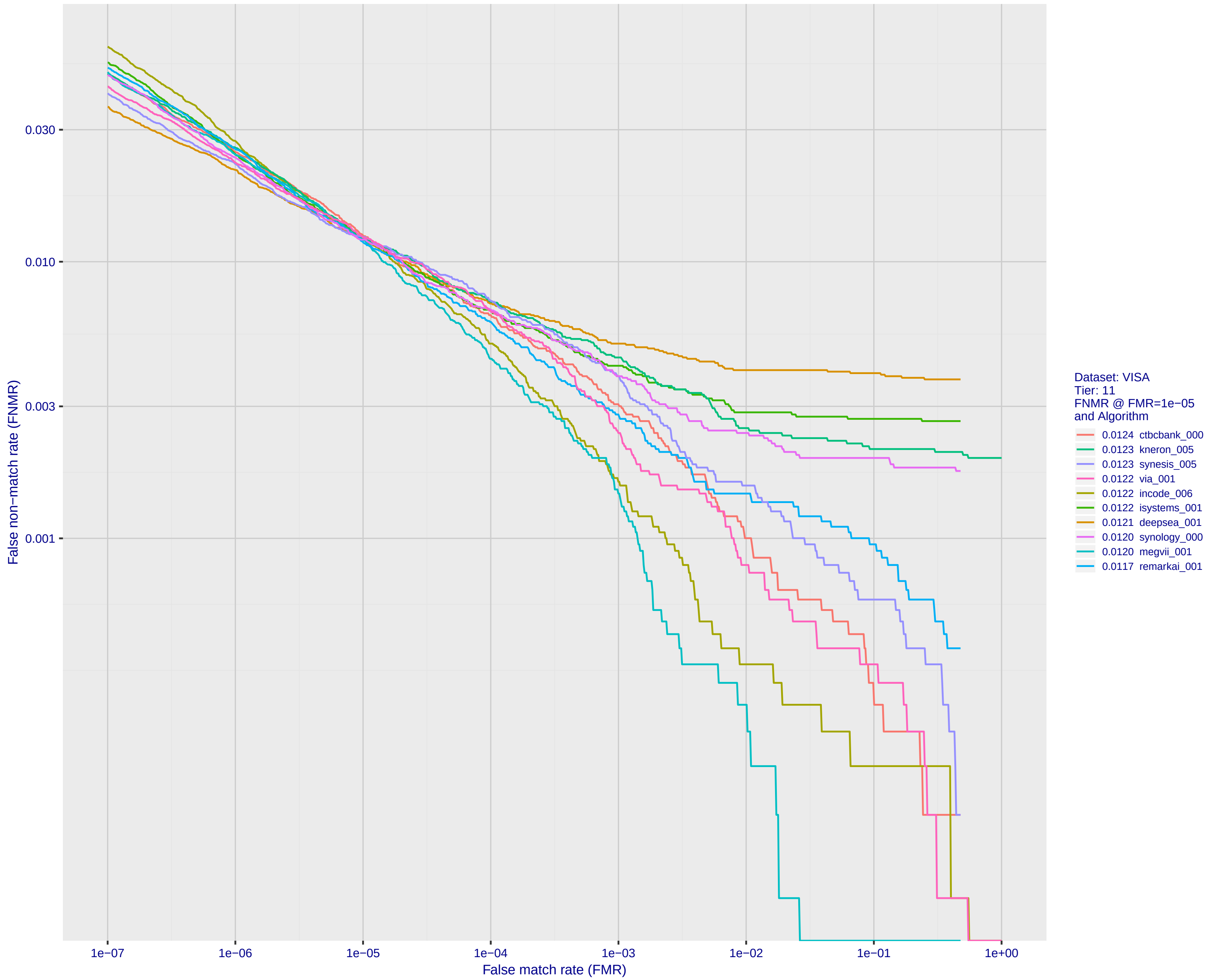


Figure 23: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

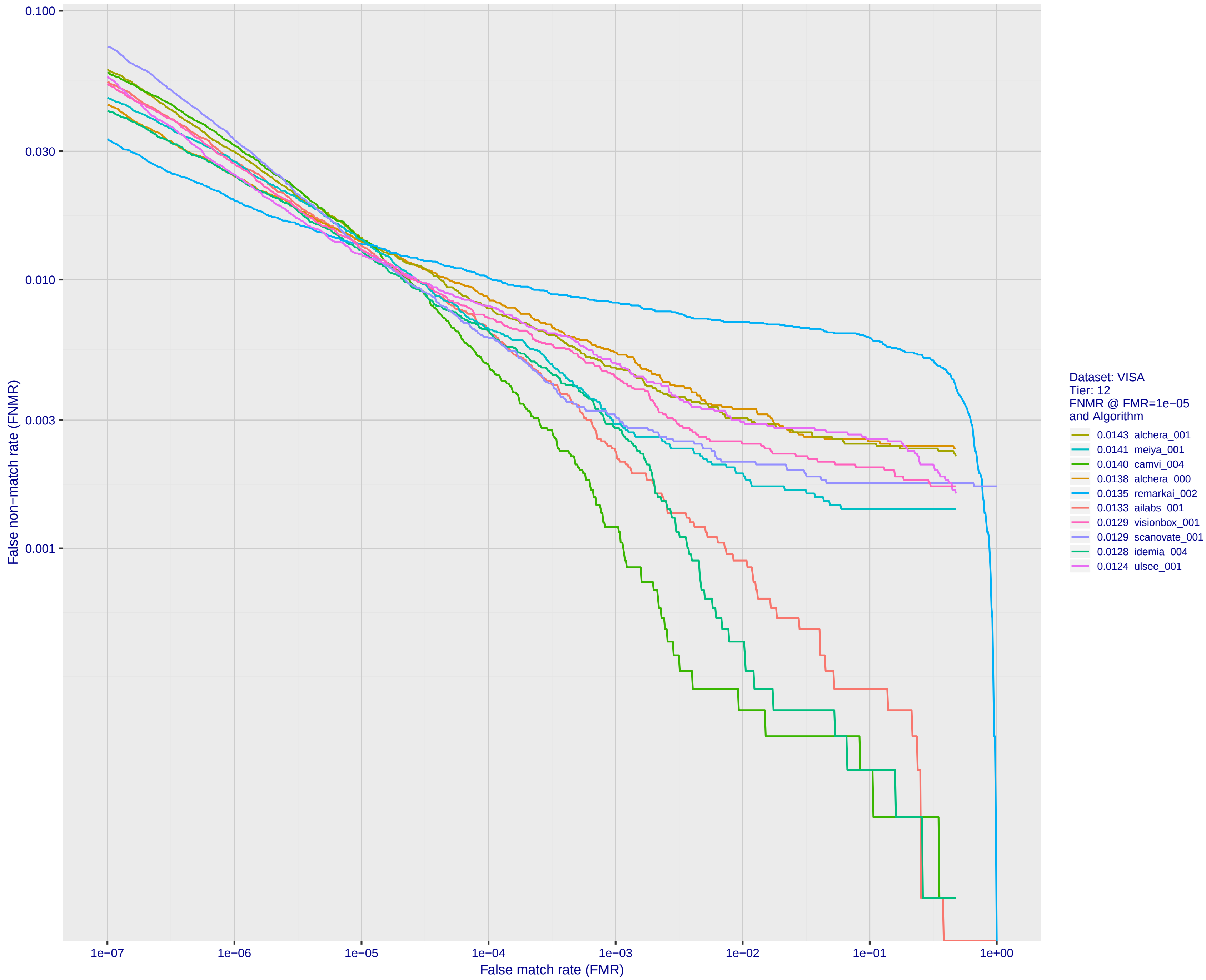


Figure 24: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

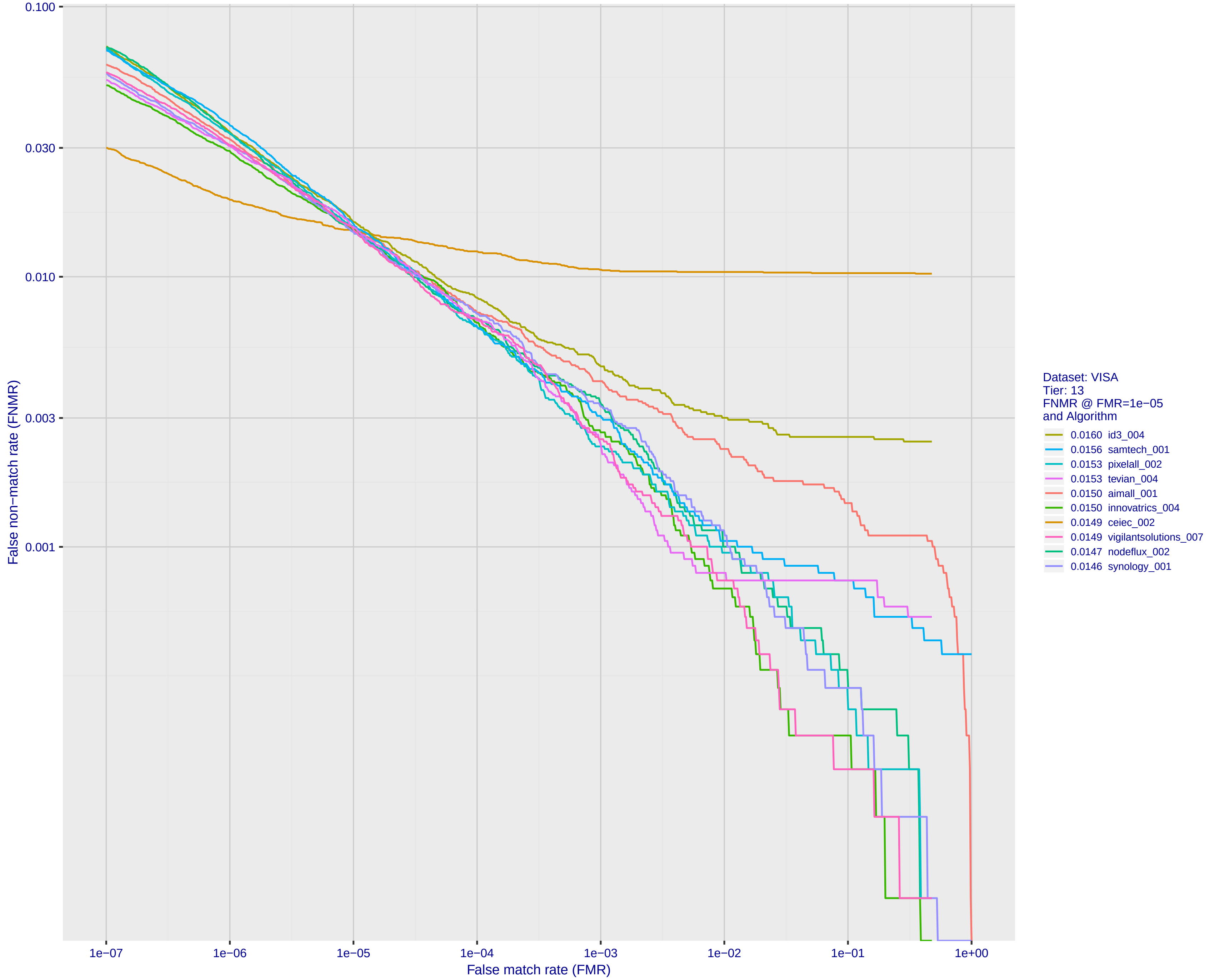


Figure 25: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

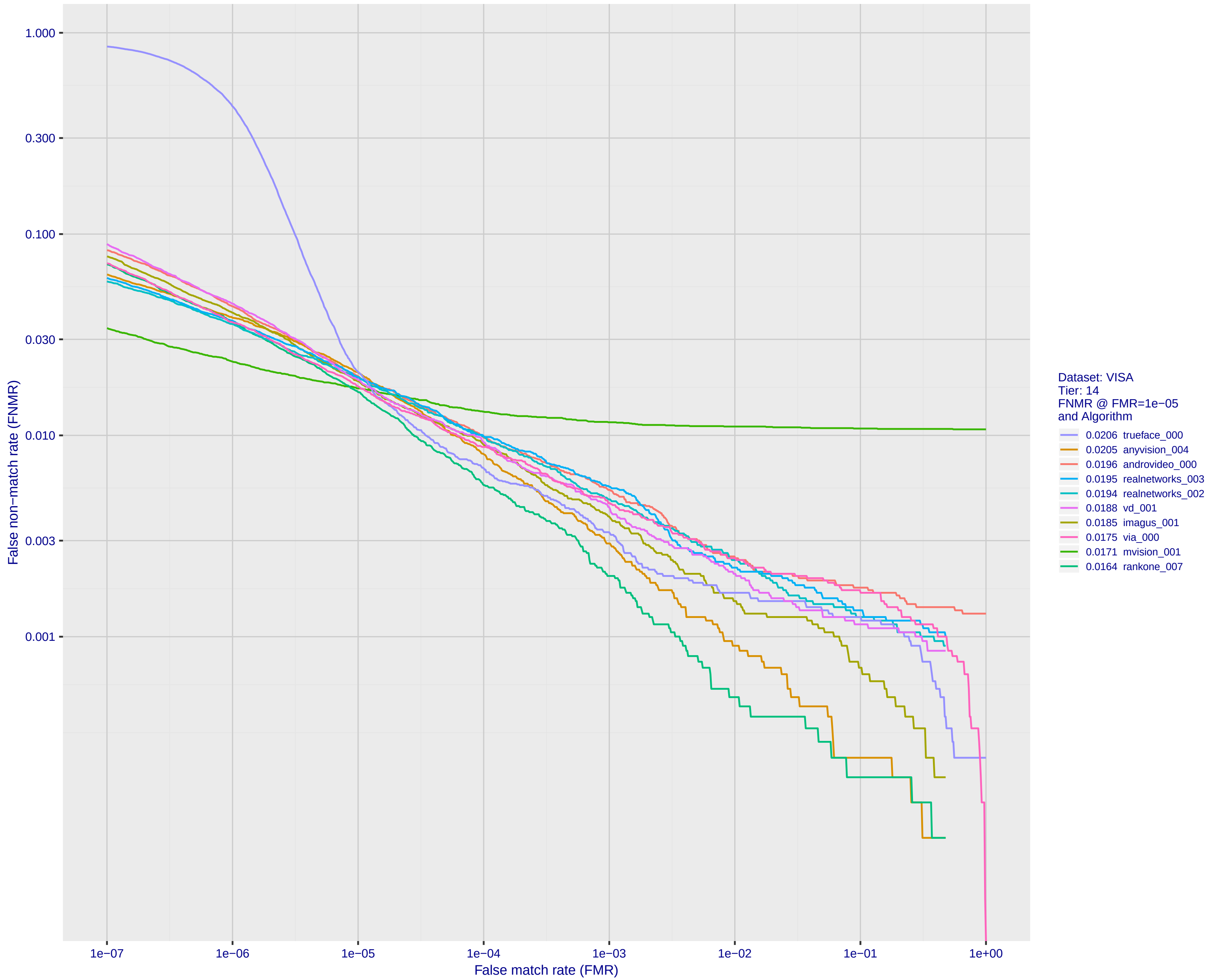


Figure 26: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

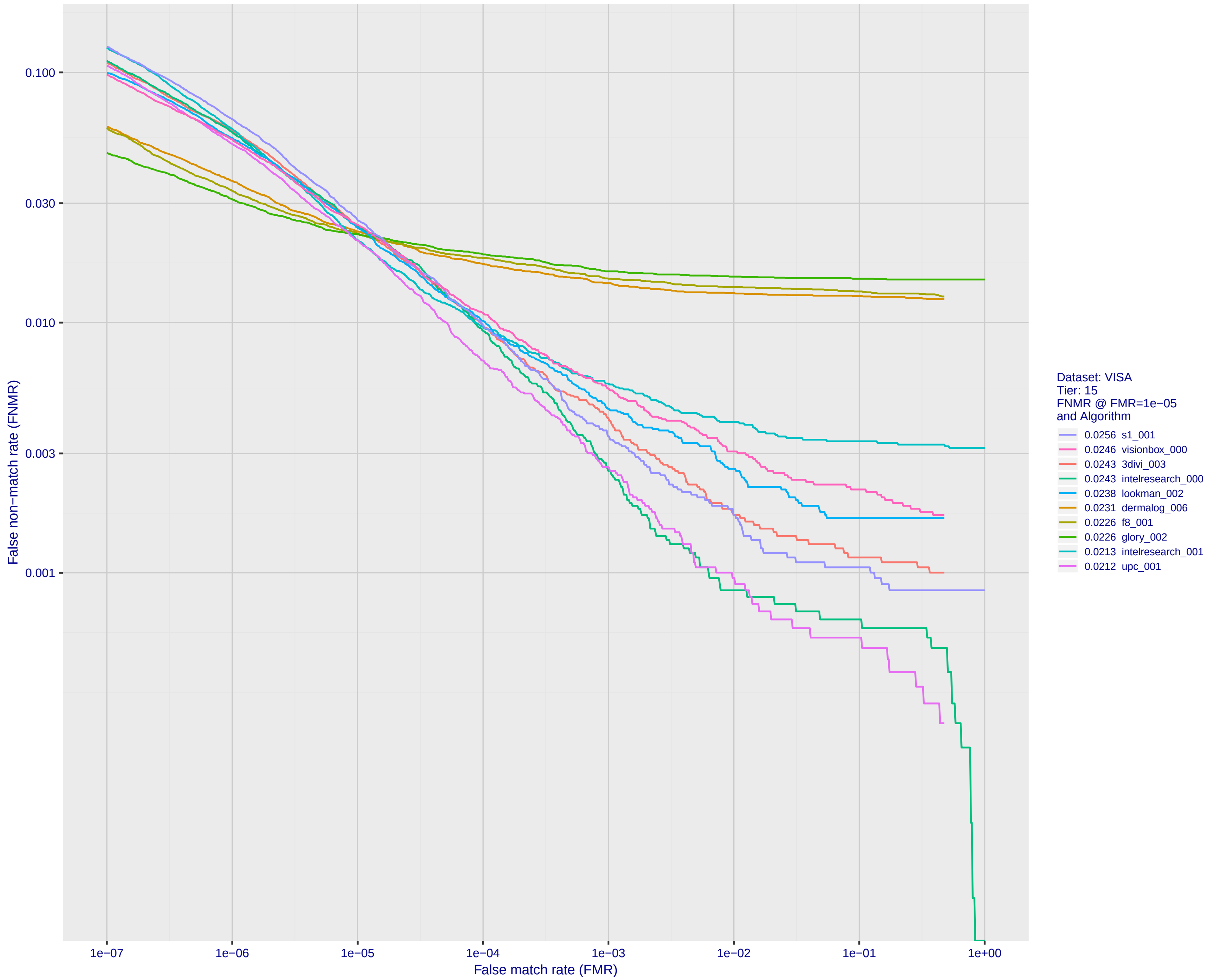


Figure 27: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

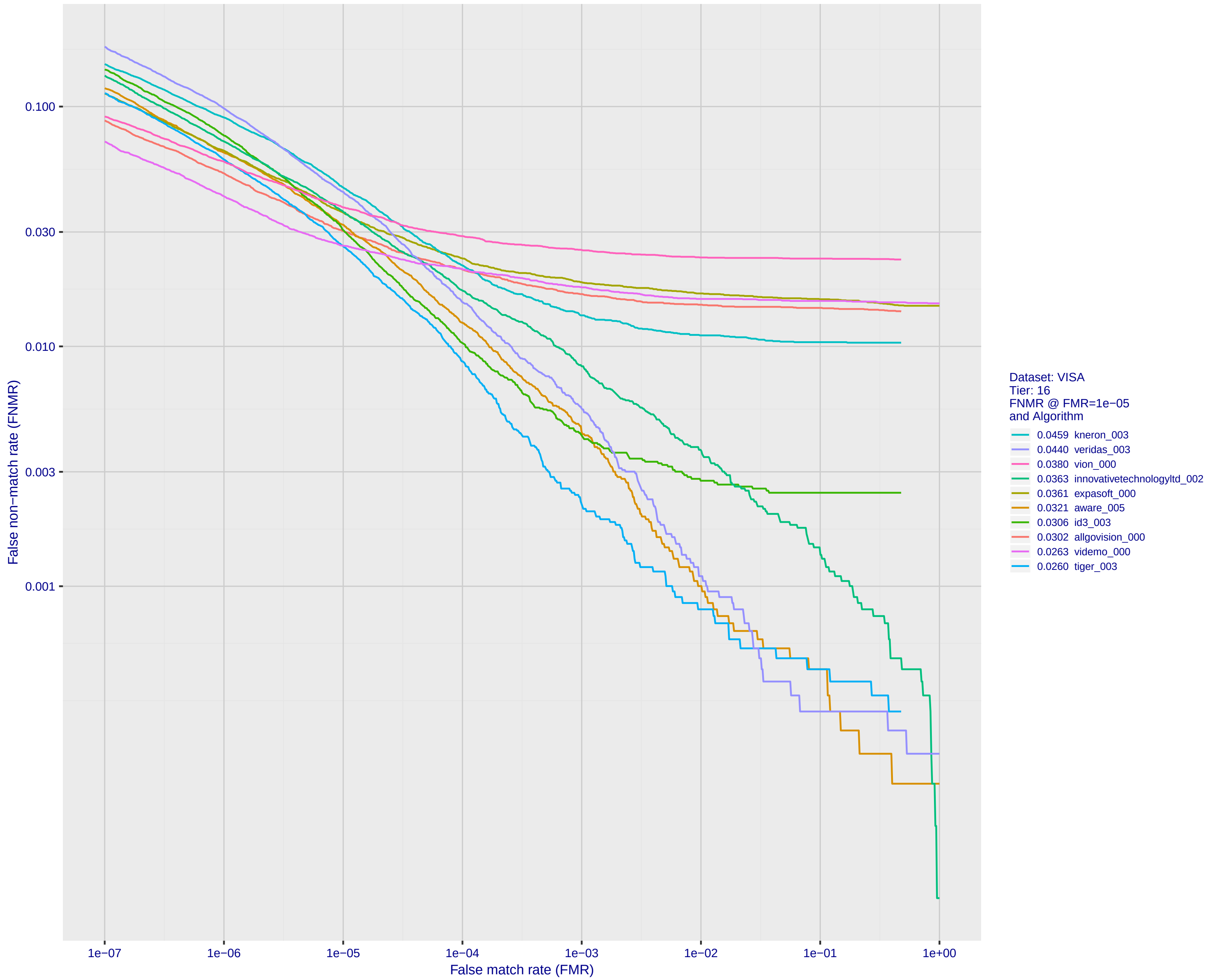


Figure 28: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

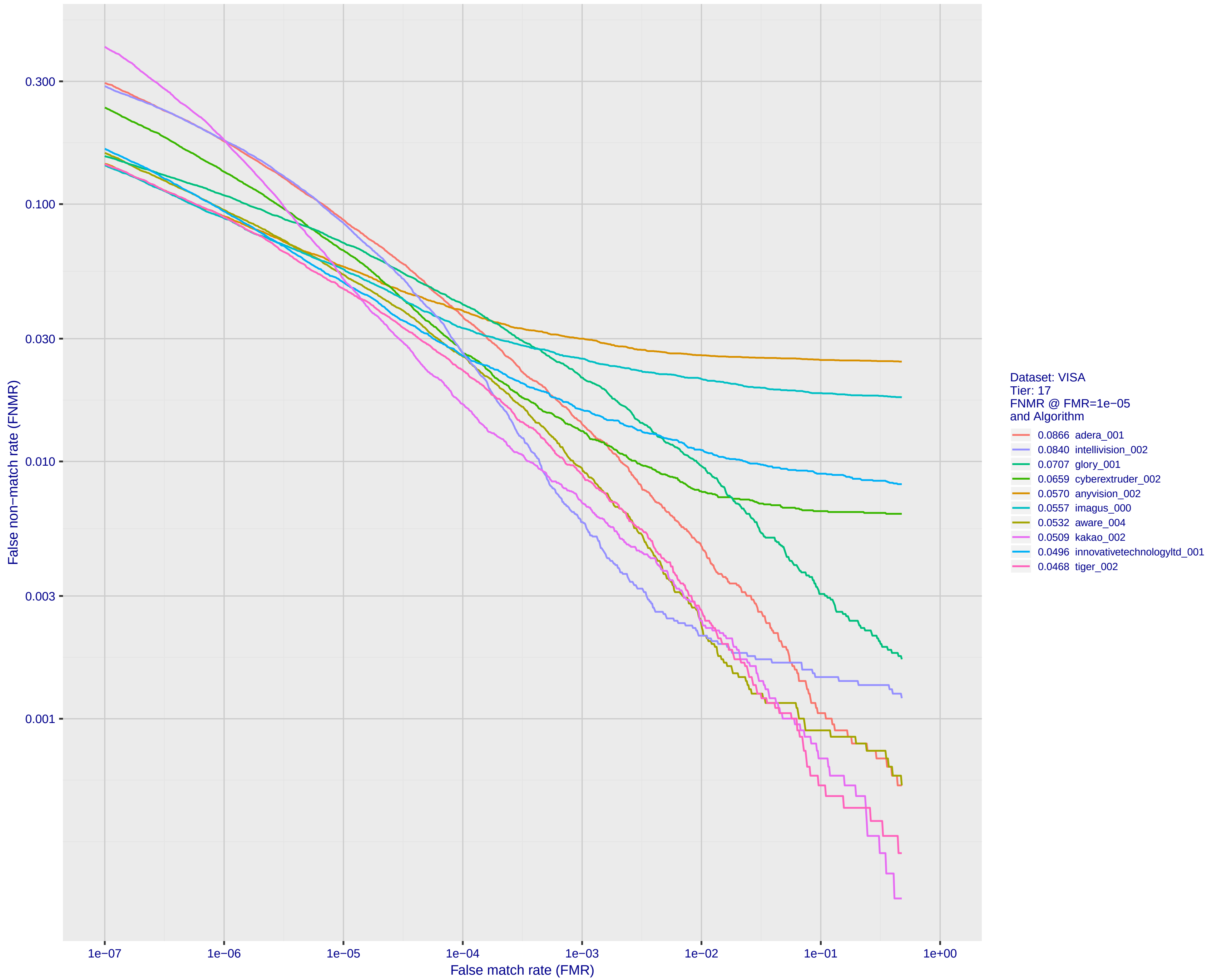


Figure 29: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

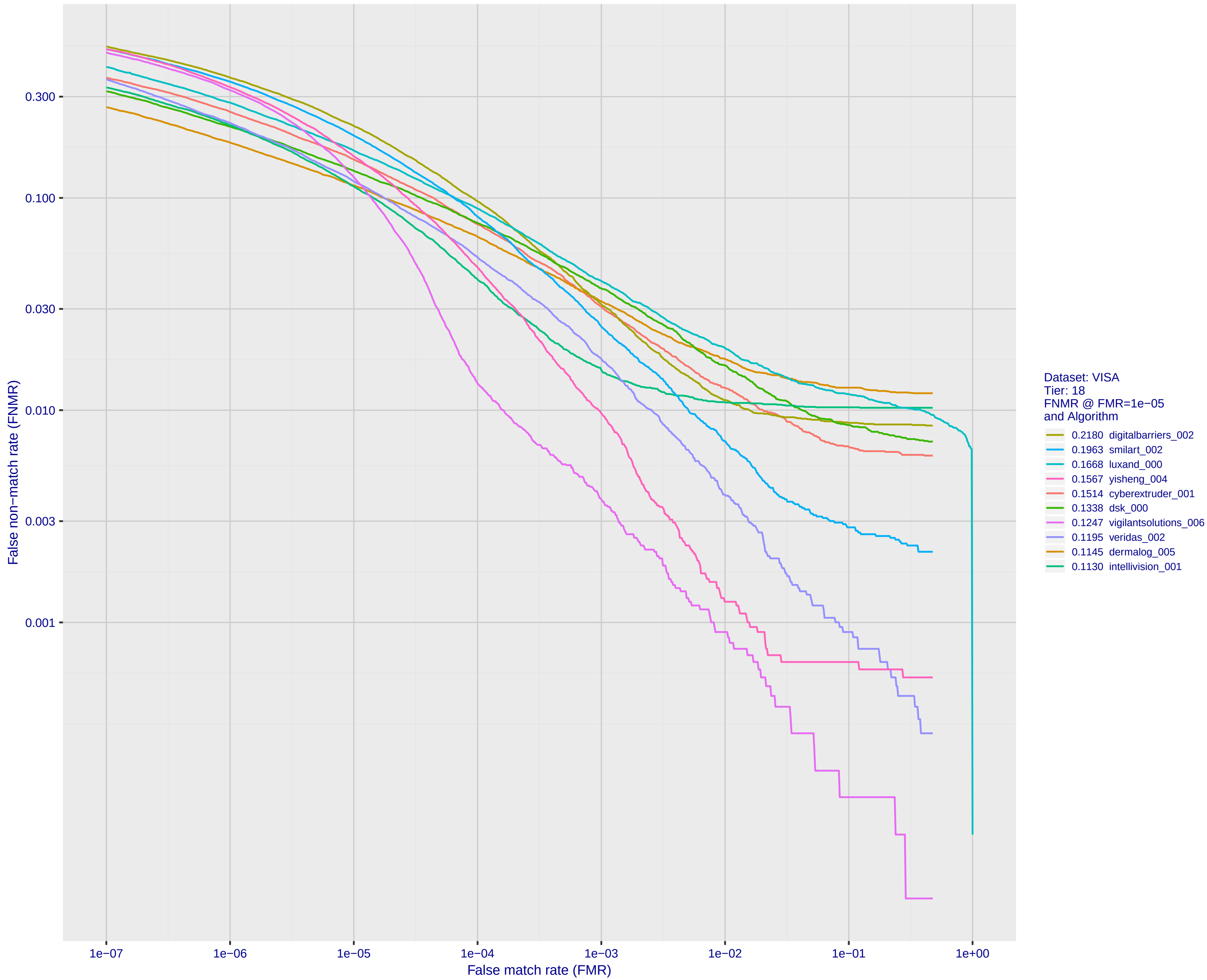


Figure 30: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

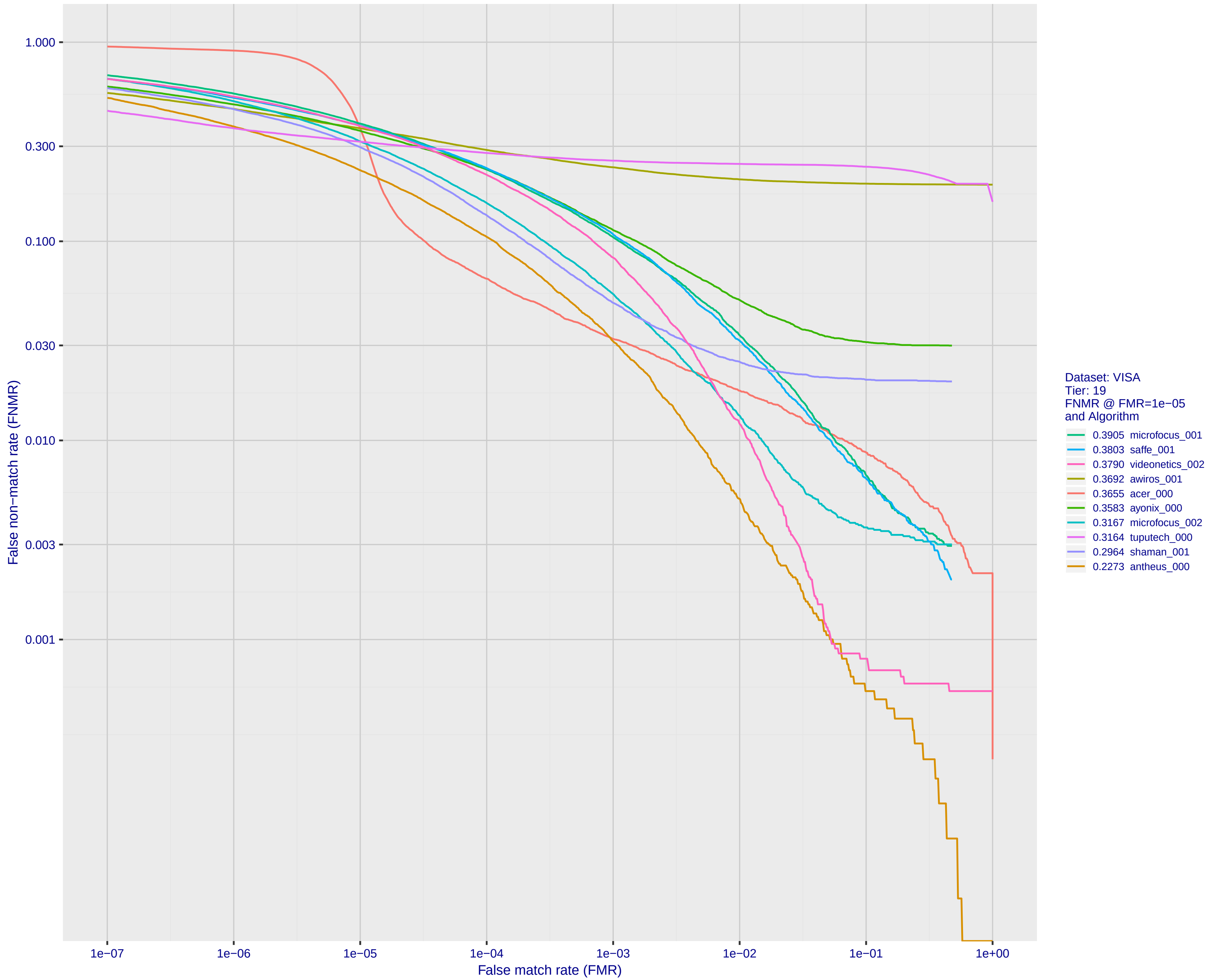


Figure 31: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

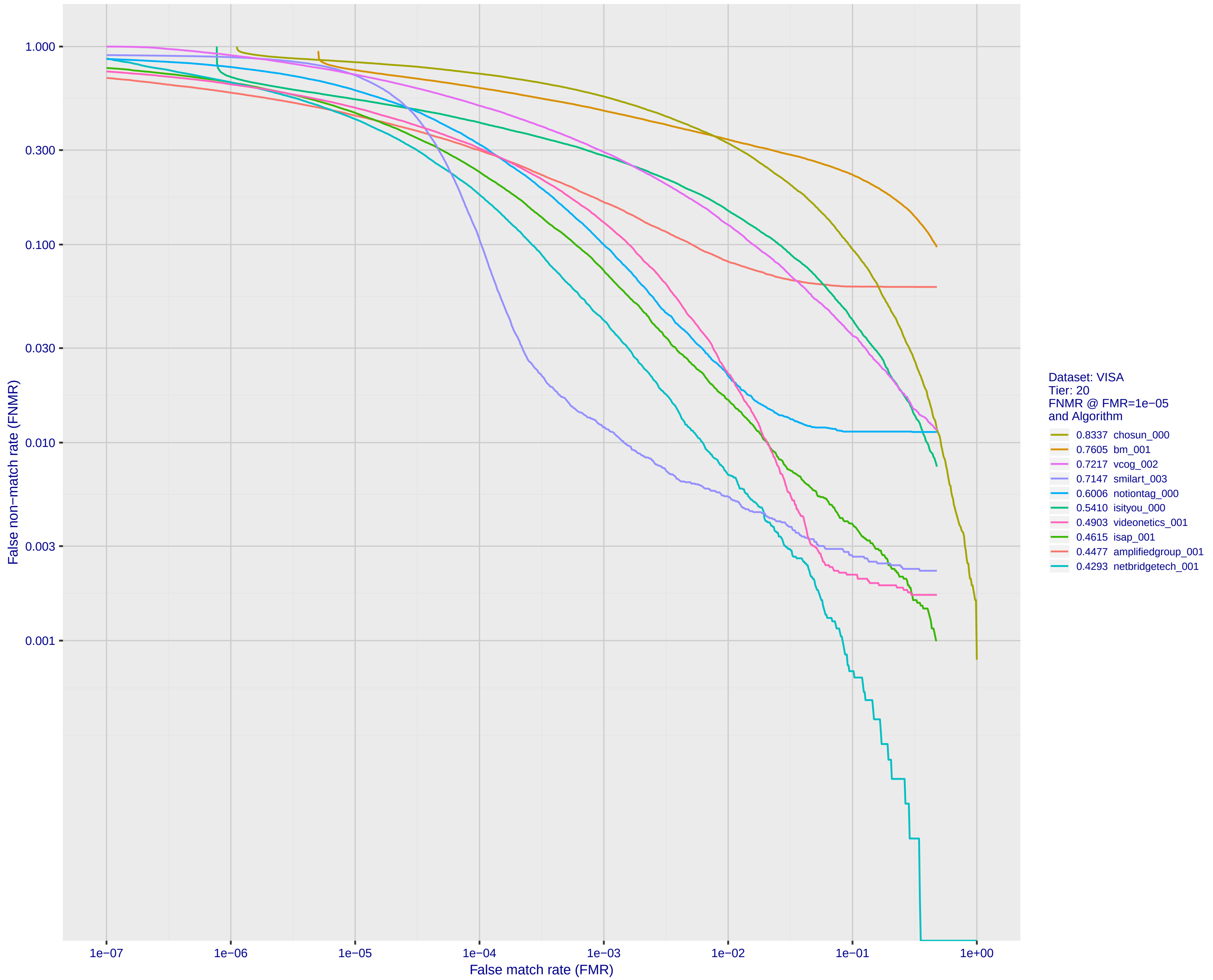


Figure 32: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

2020/03/25 16:36:23

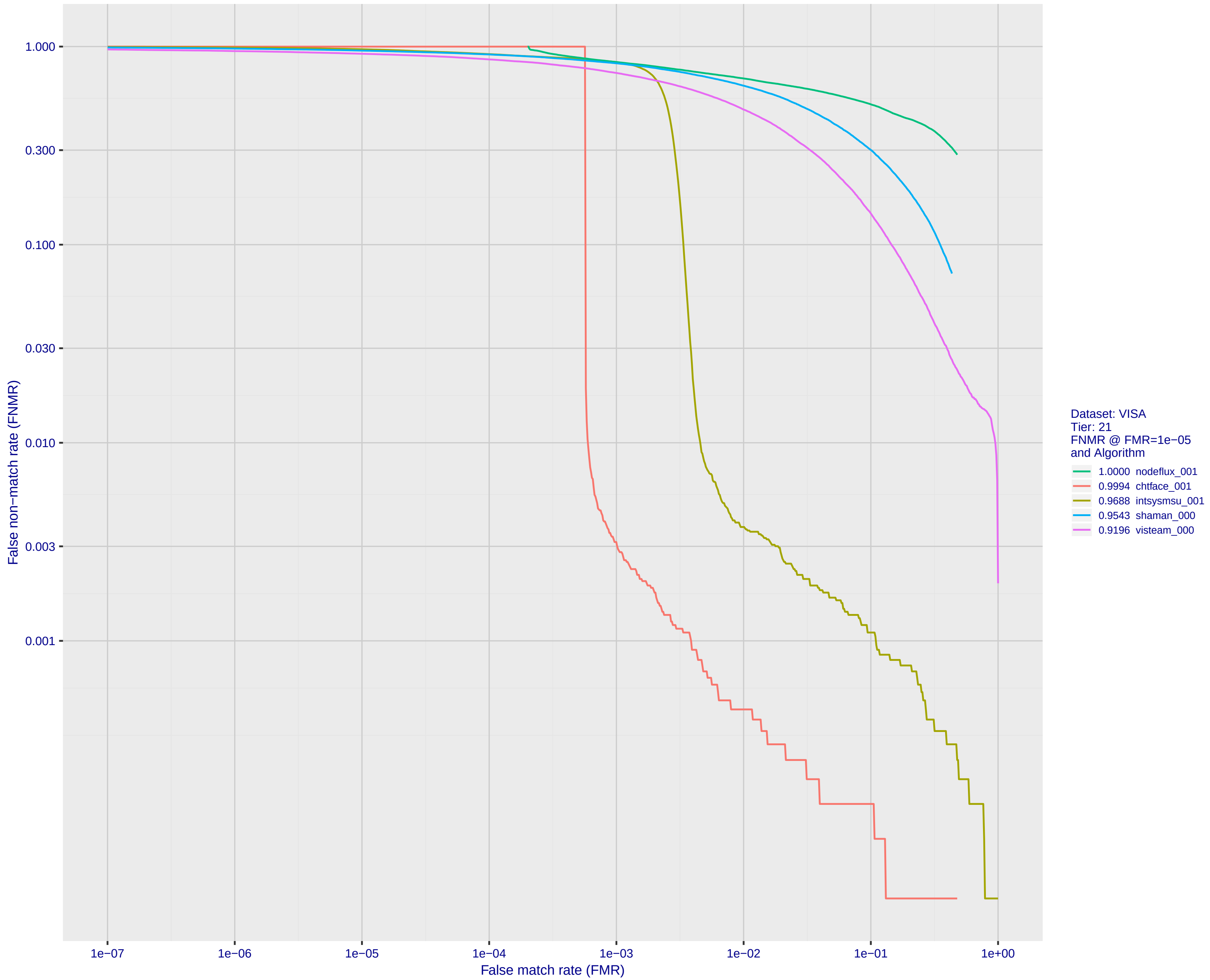


Figure 33: For the visa images, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

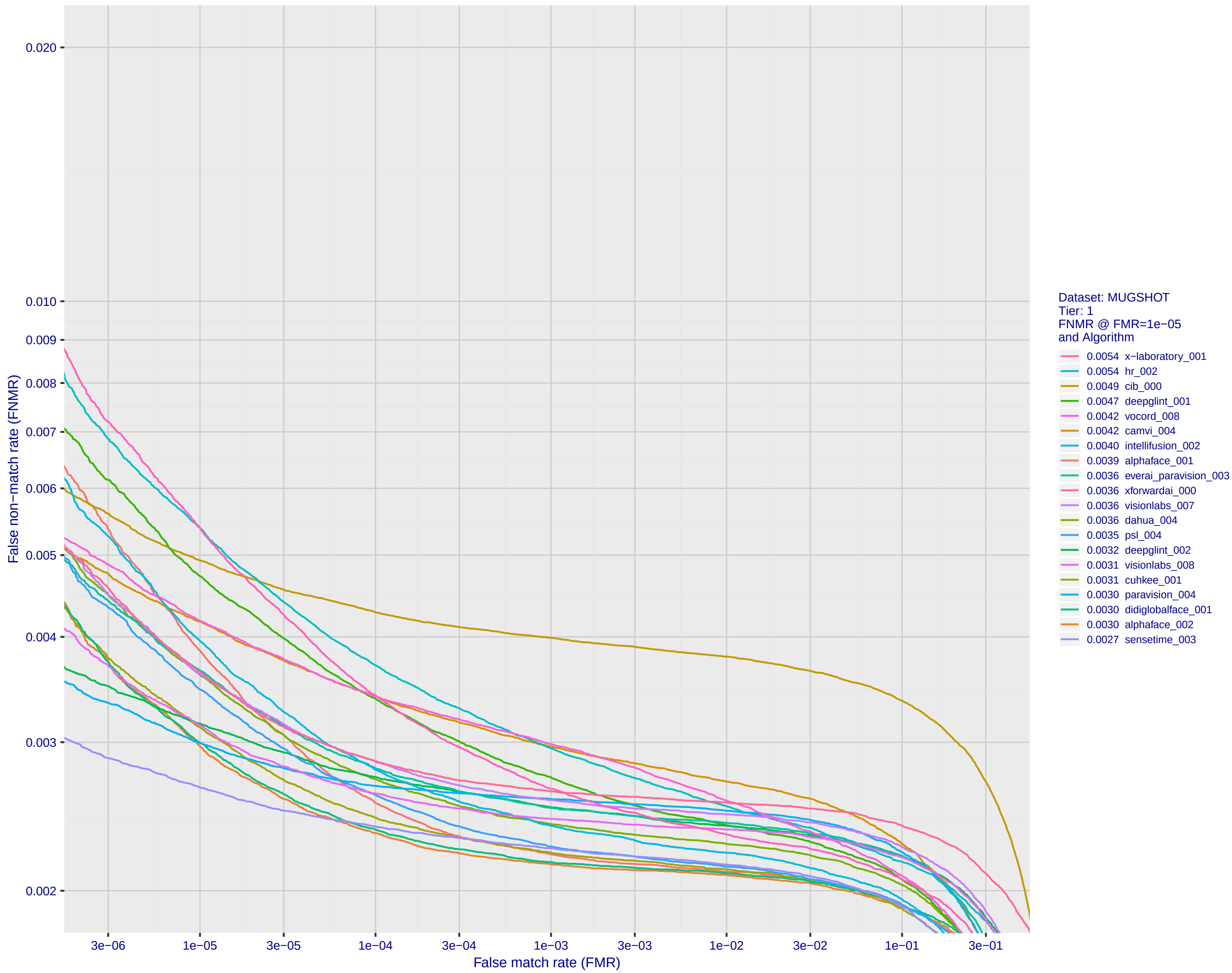


Figure 34: For the mugshot images, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

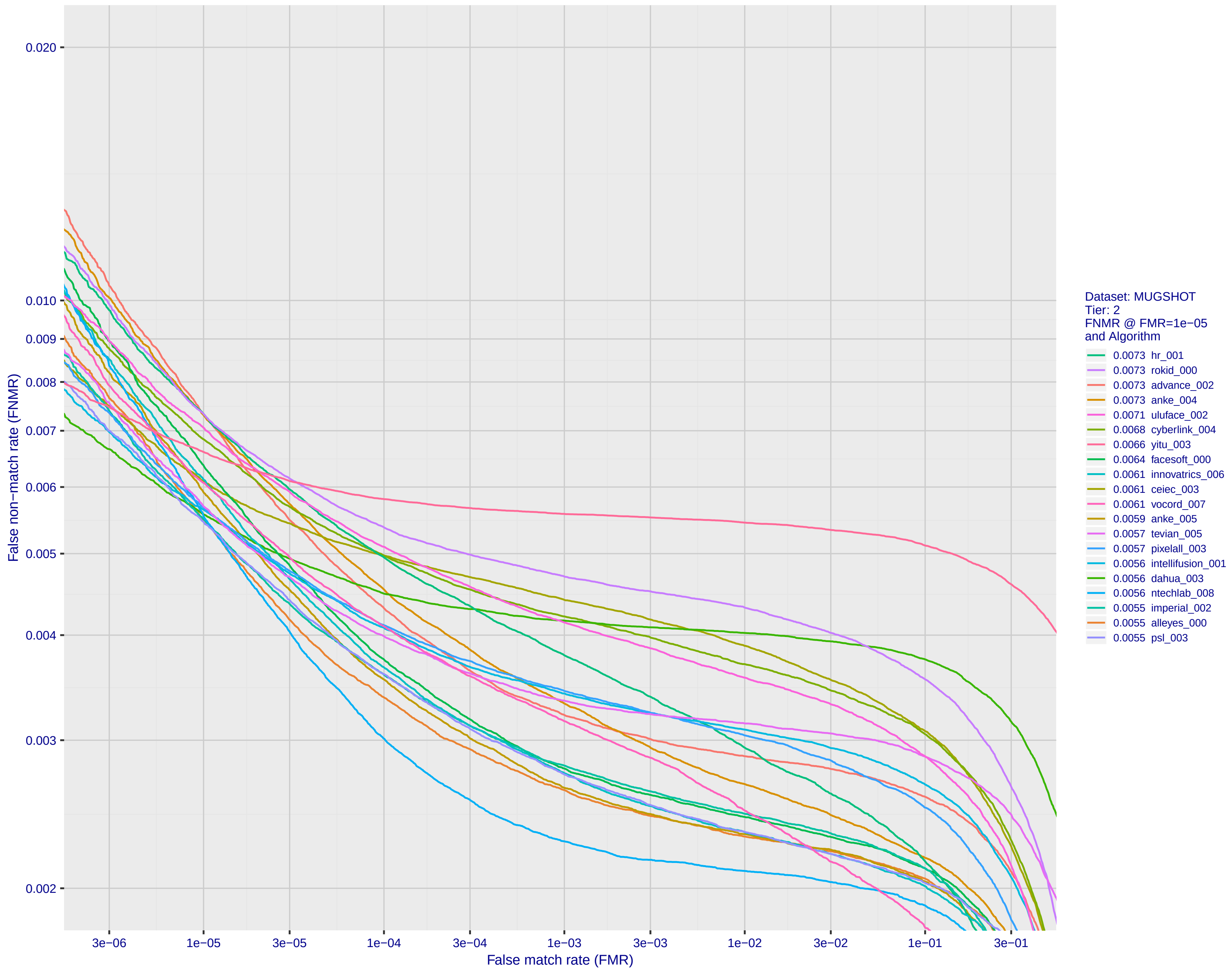


Figure 35: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

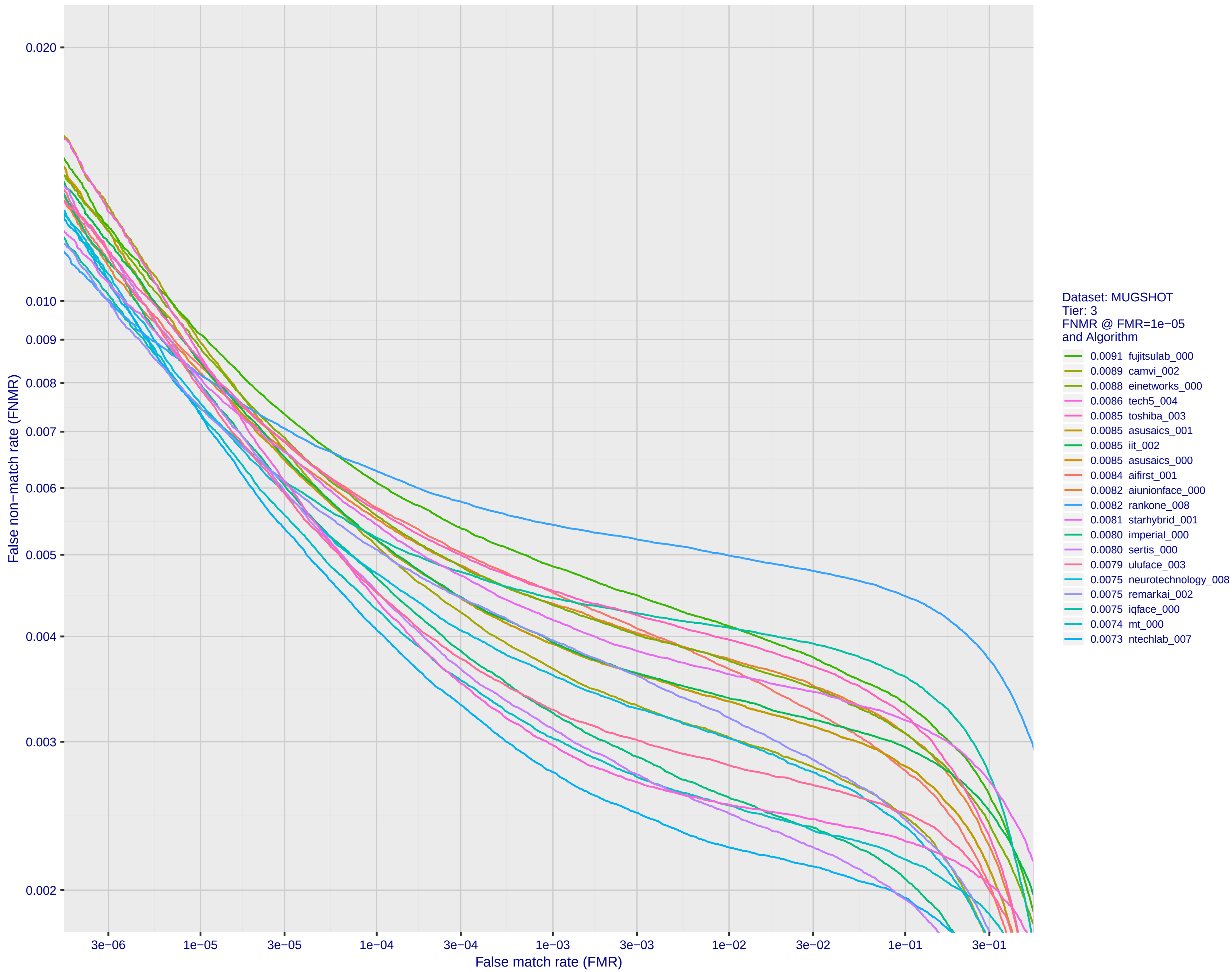


Figure 36: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

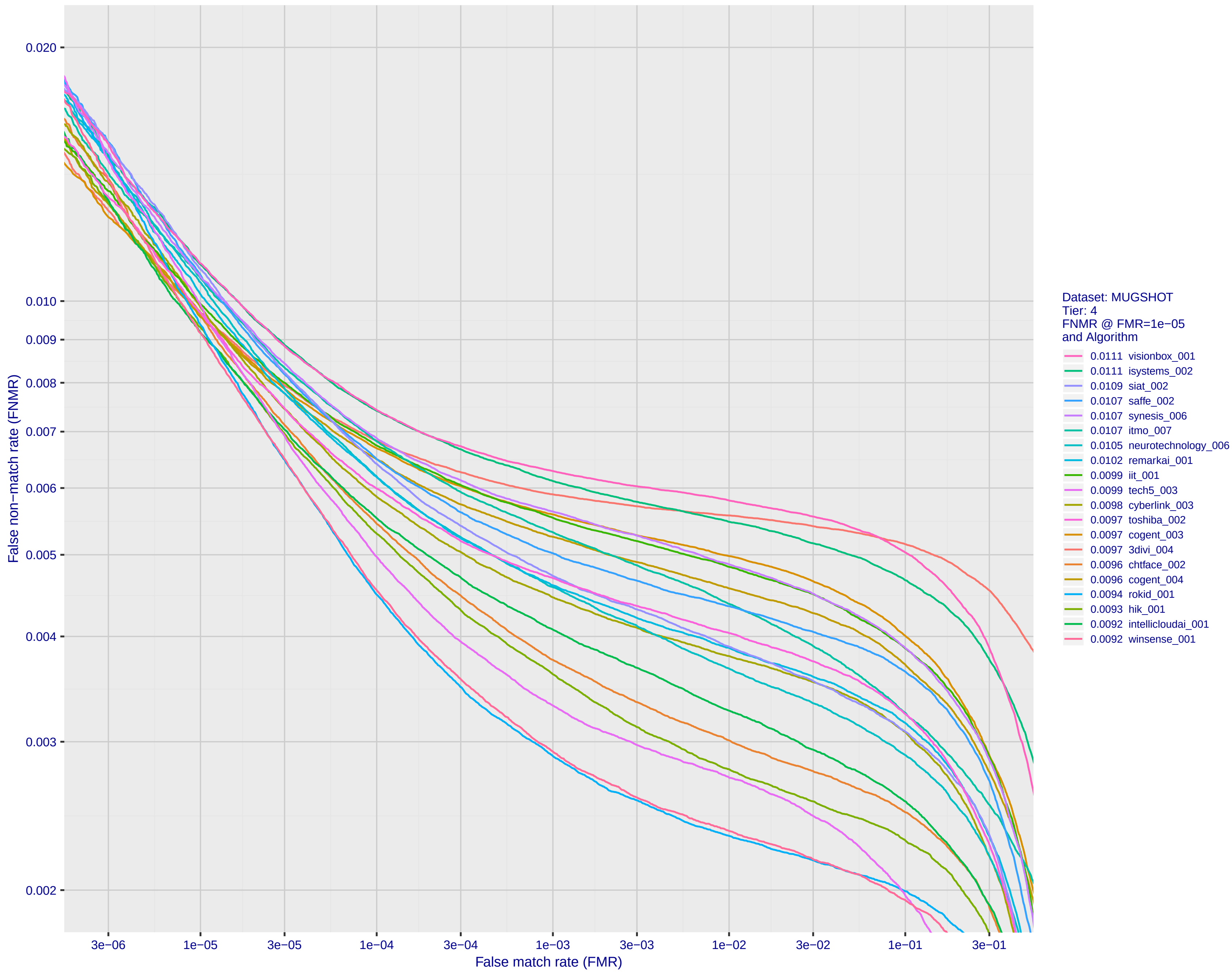


Figure 37: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

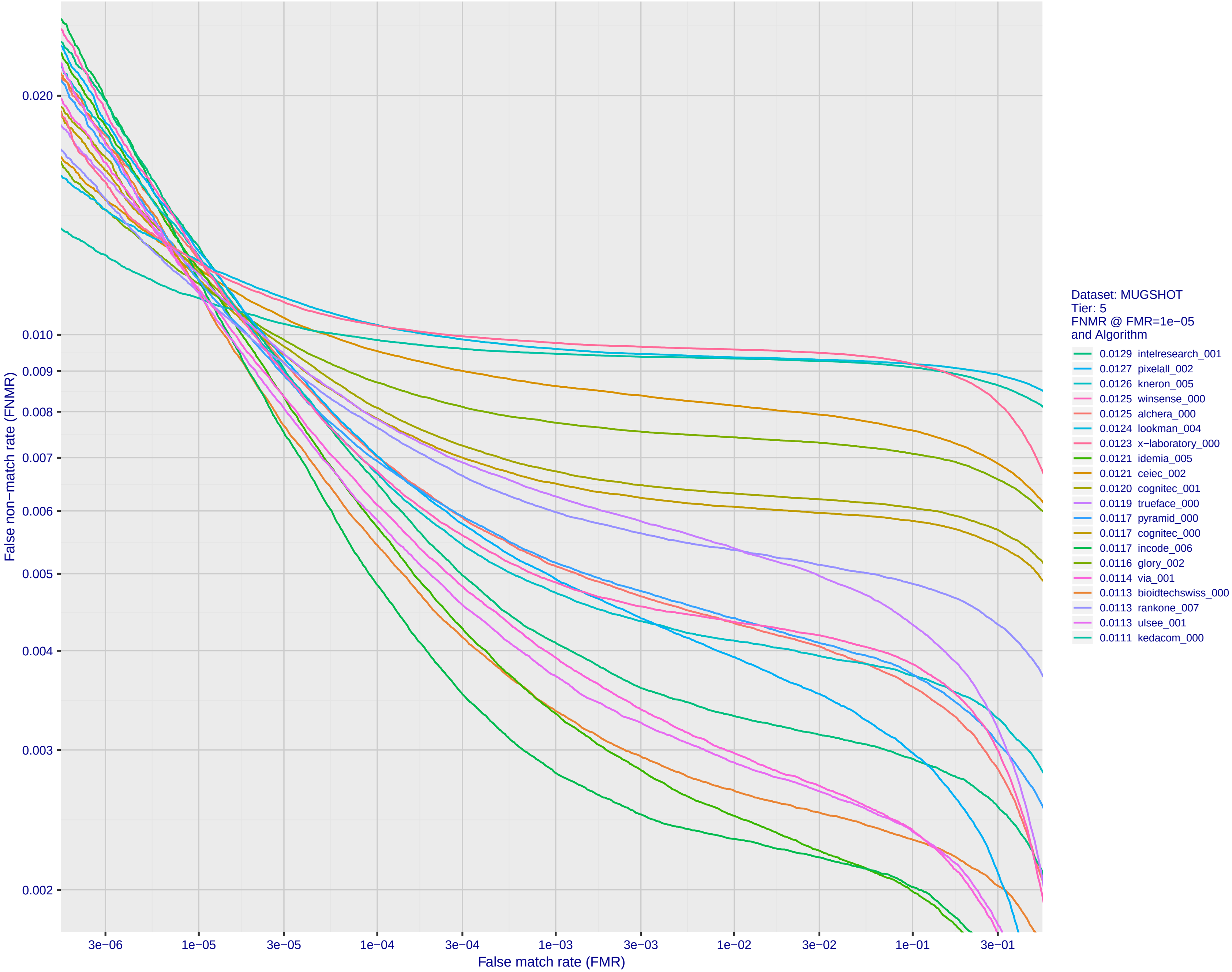


Figure 38: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

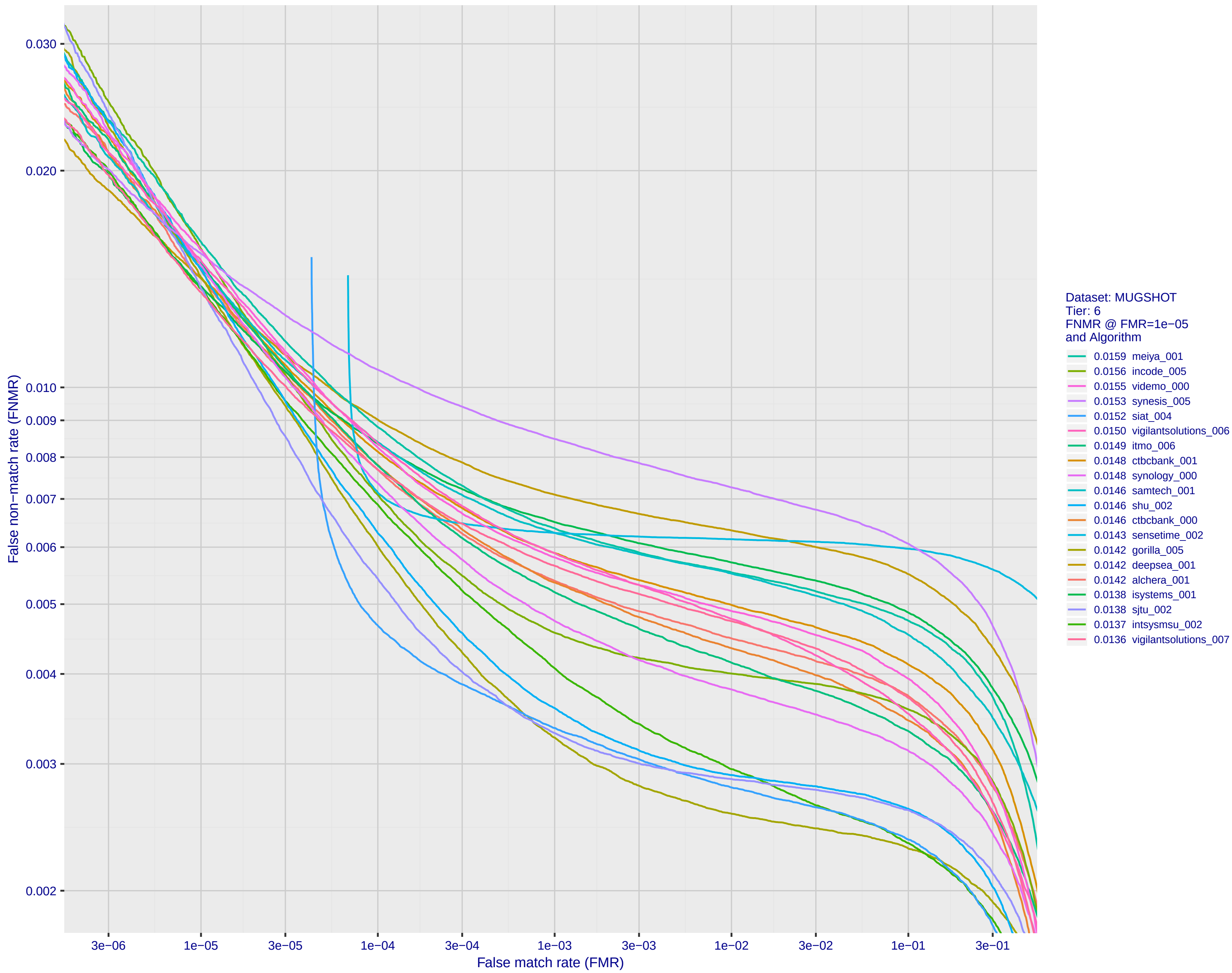


Figure 39: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

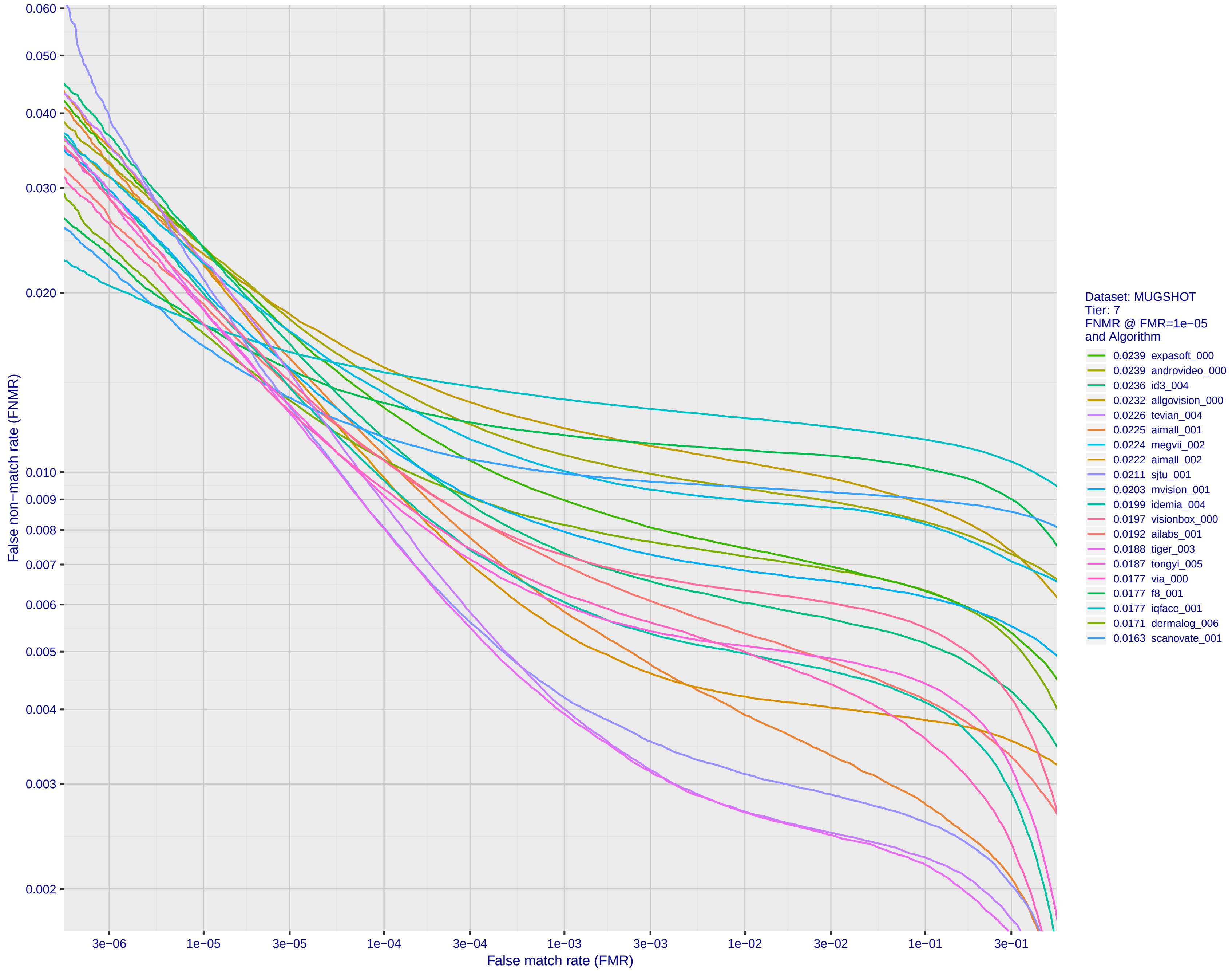


Figure 40: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

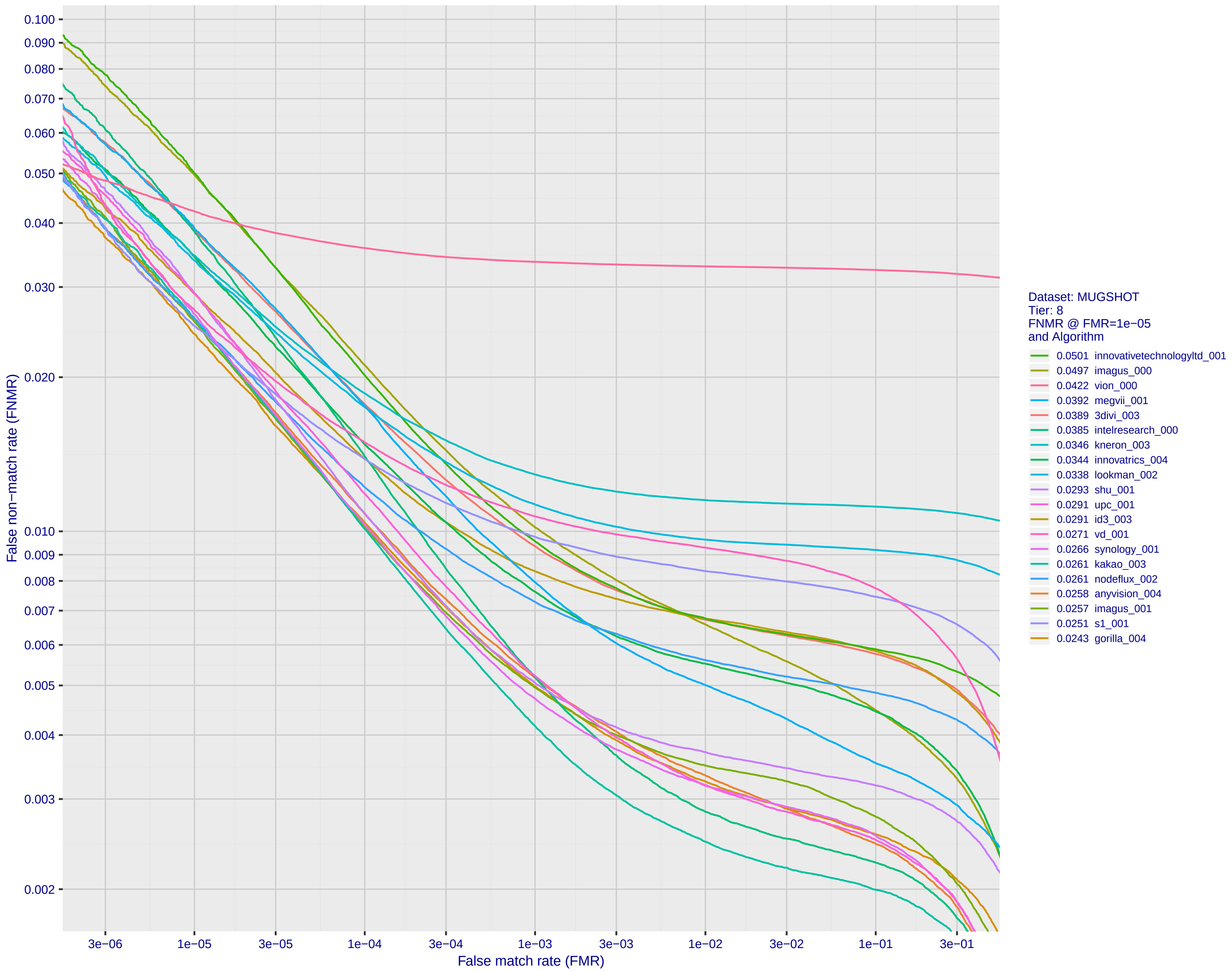


Figure 41: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

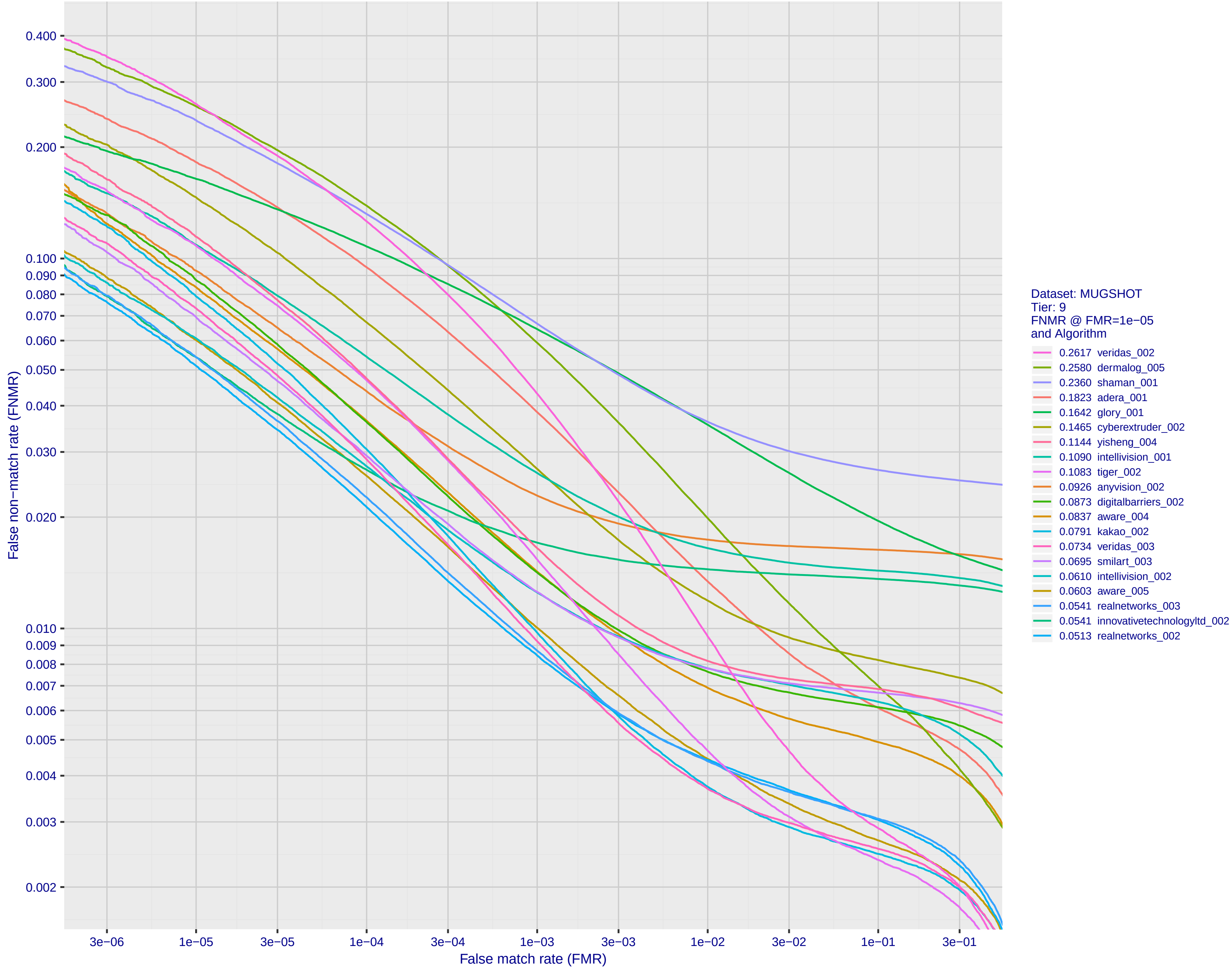


Figure 42: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

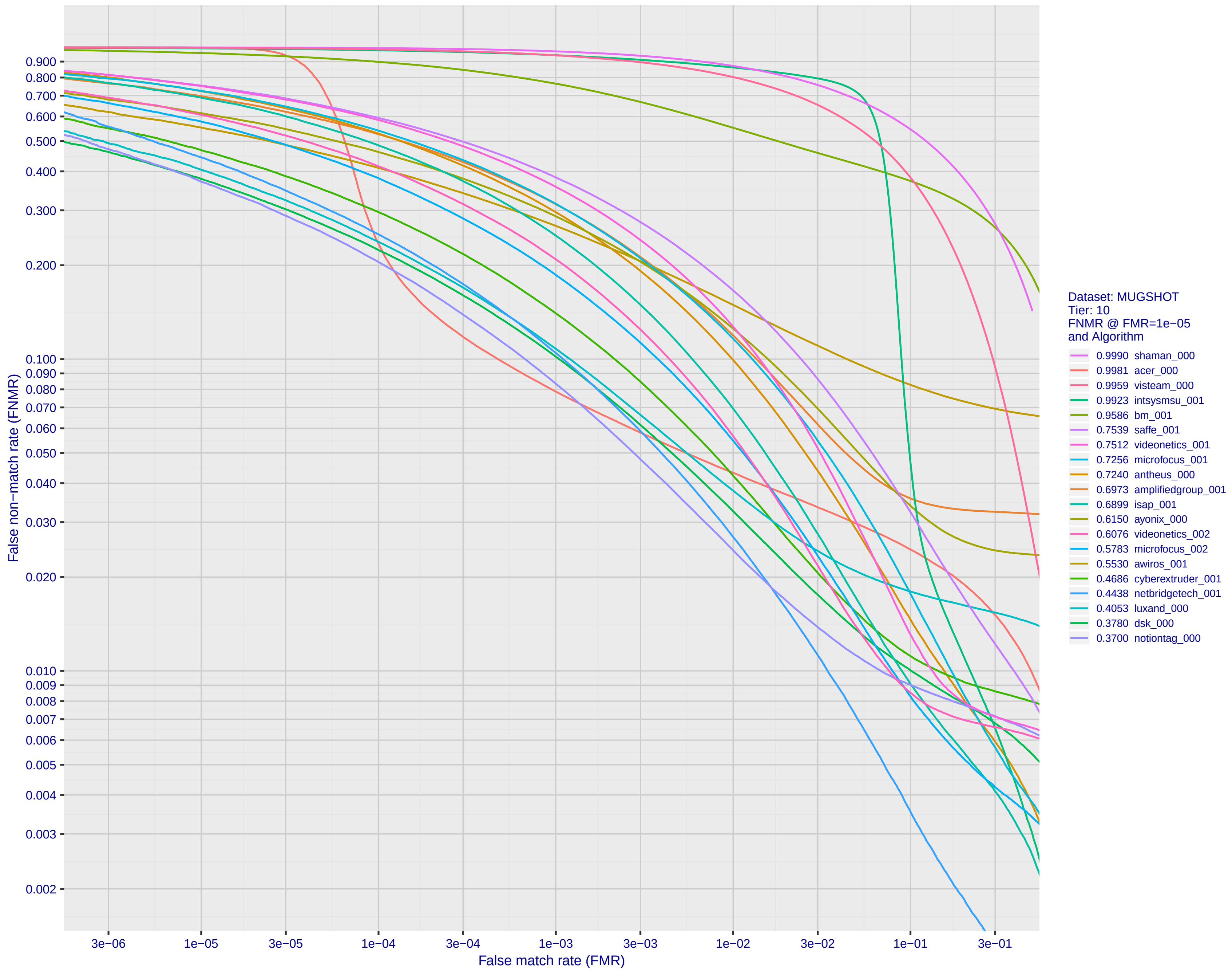


Figure 43: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

2020/03/25 16:36:23

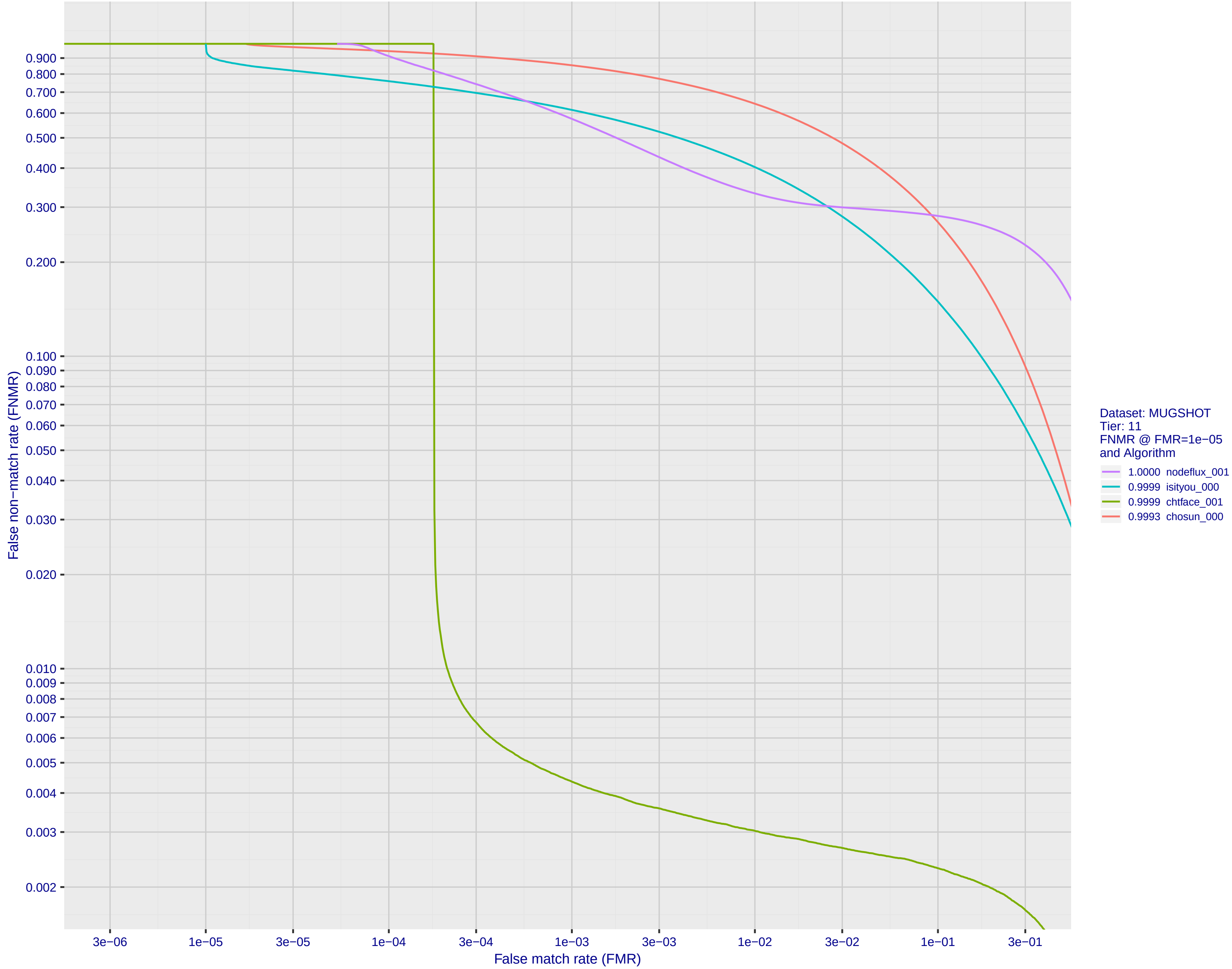


Figure 44: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

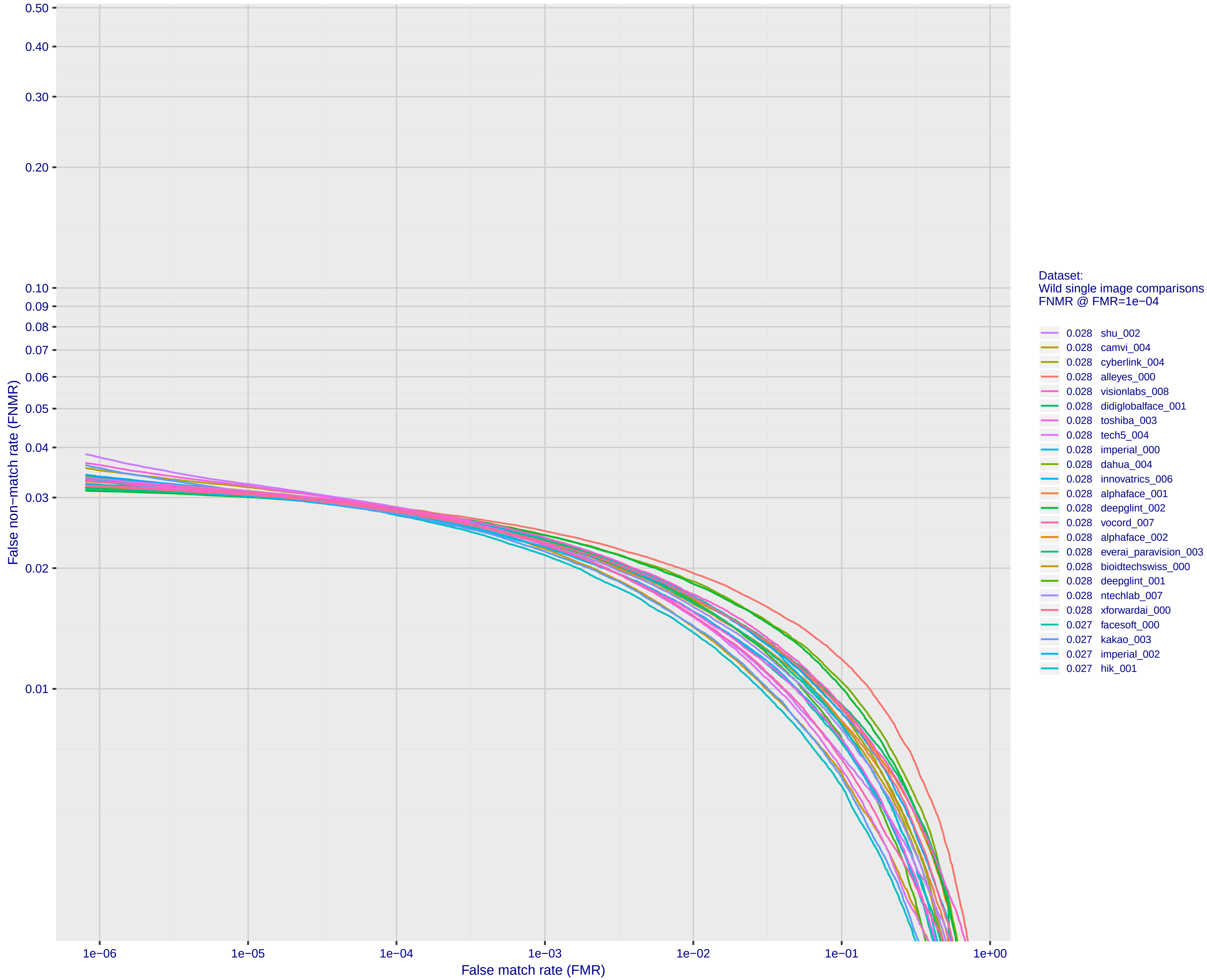


Figure 45: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

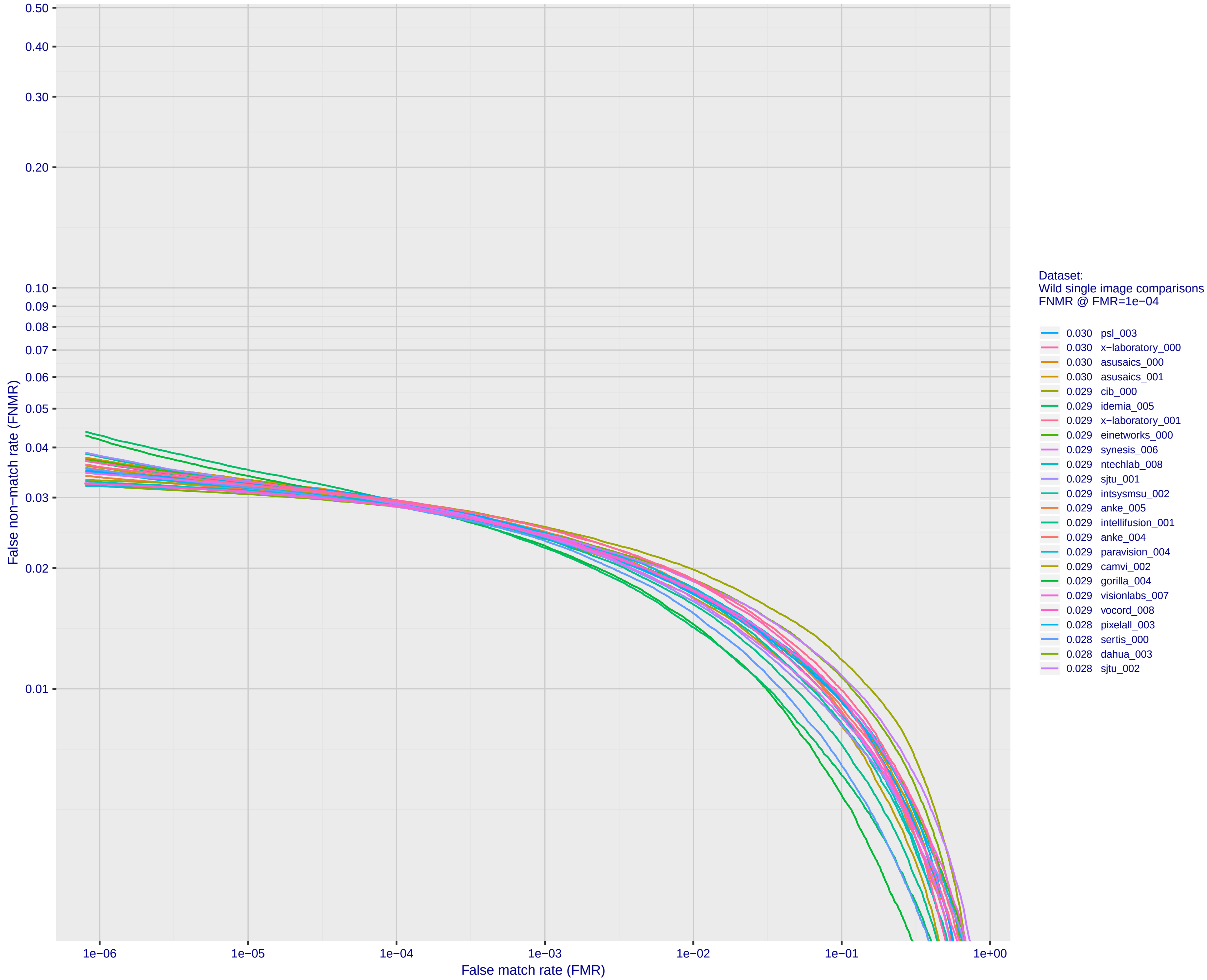


Figure 46: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

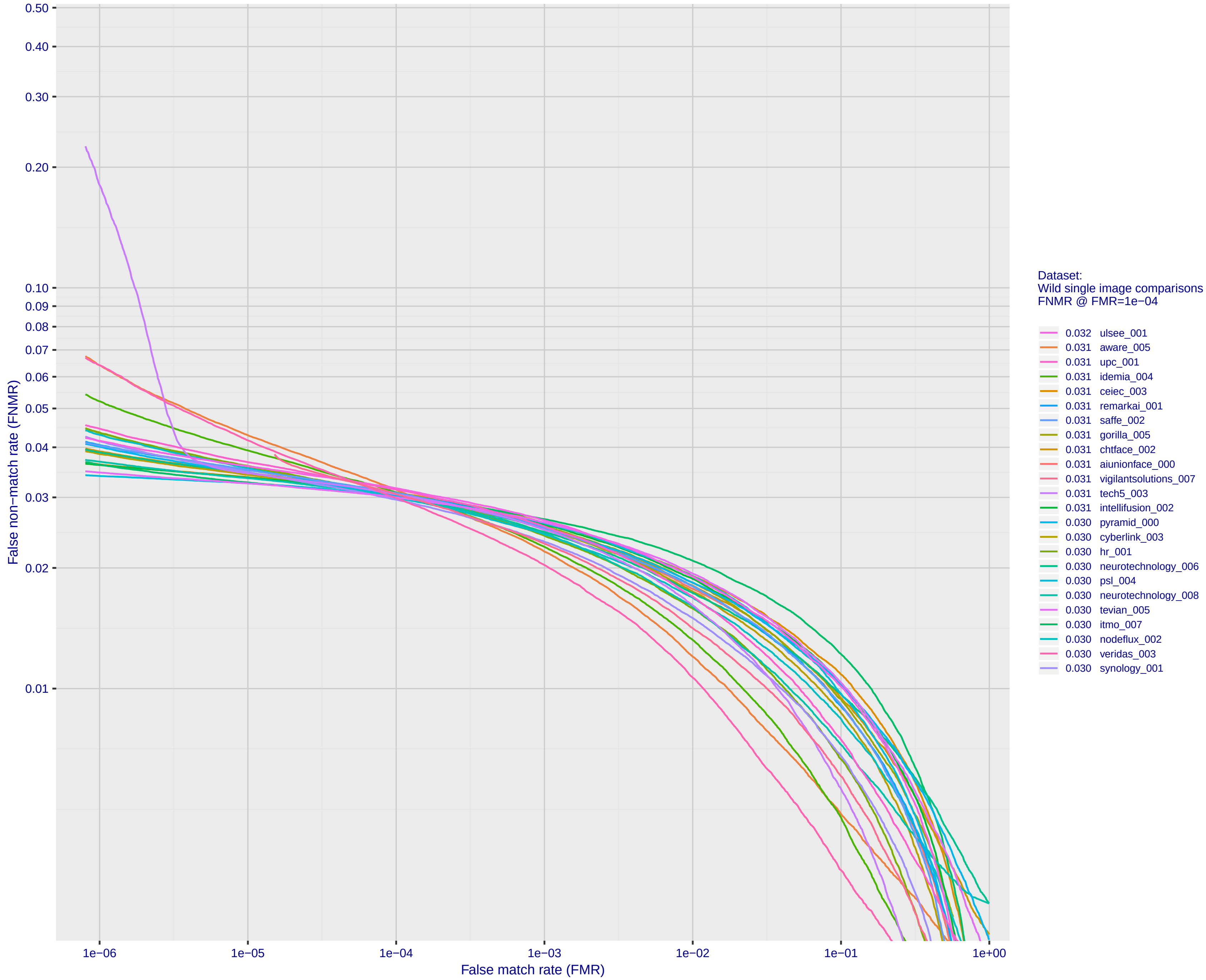


Figure 47: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

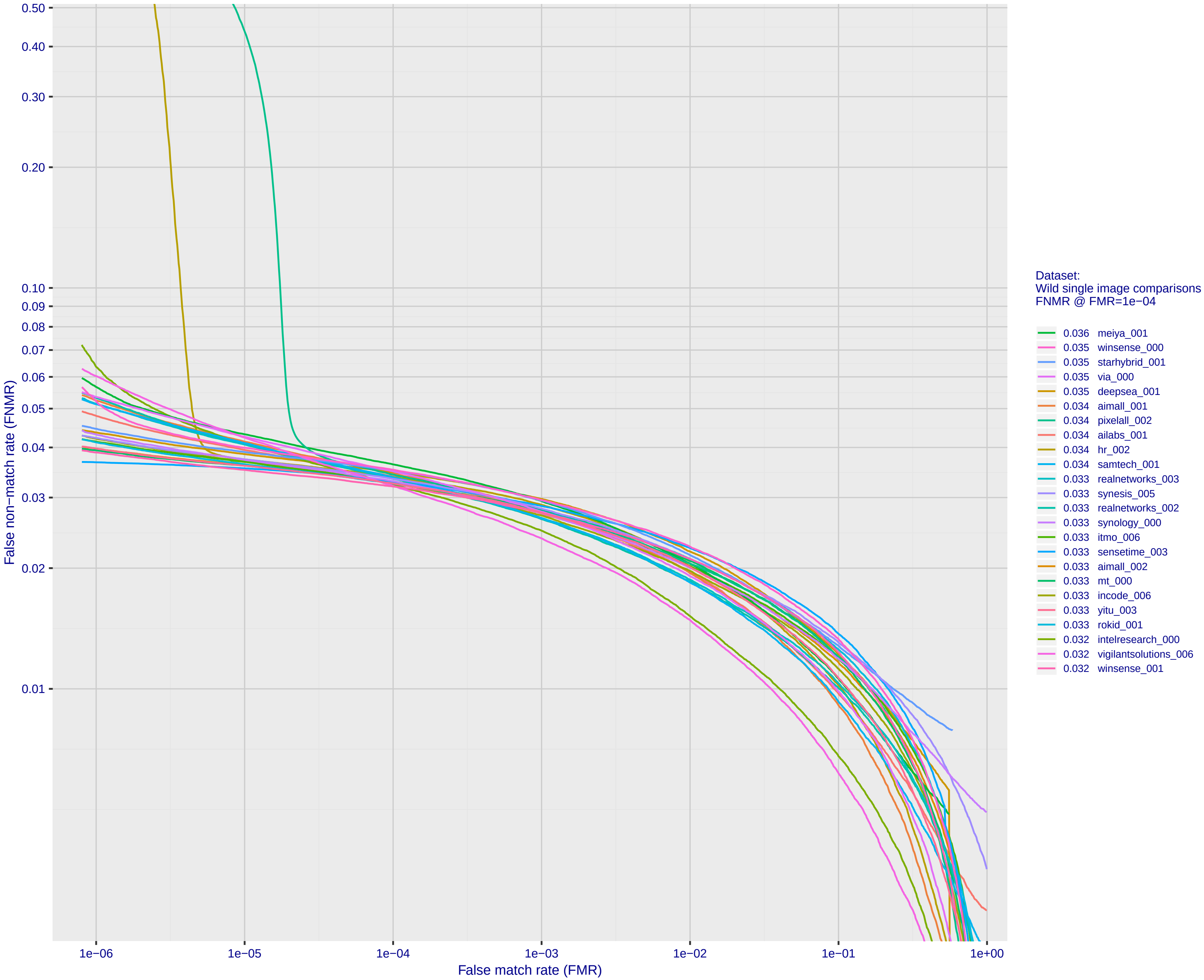


Figure 48: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

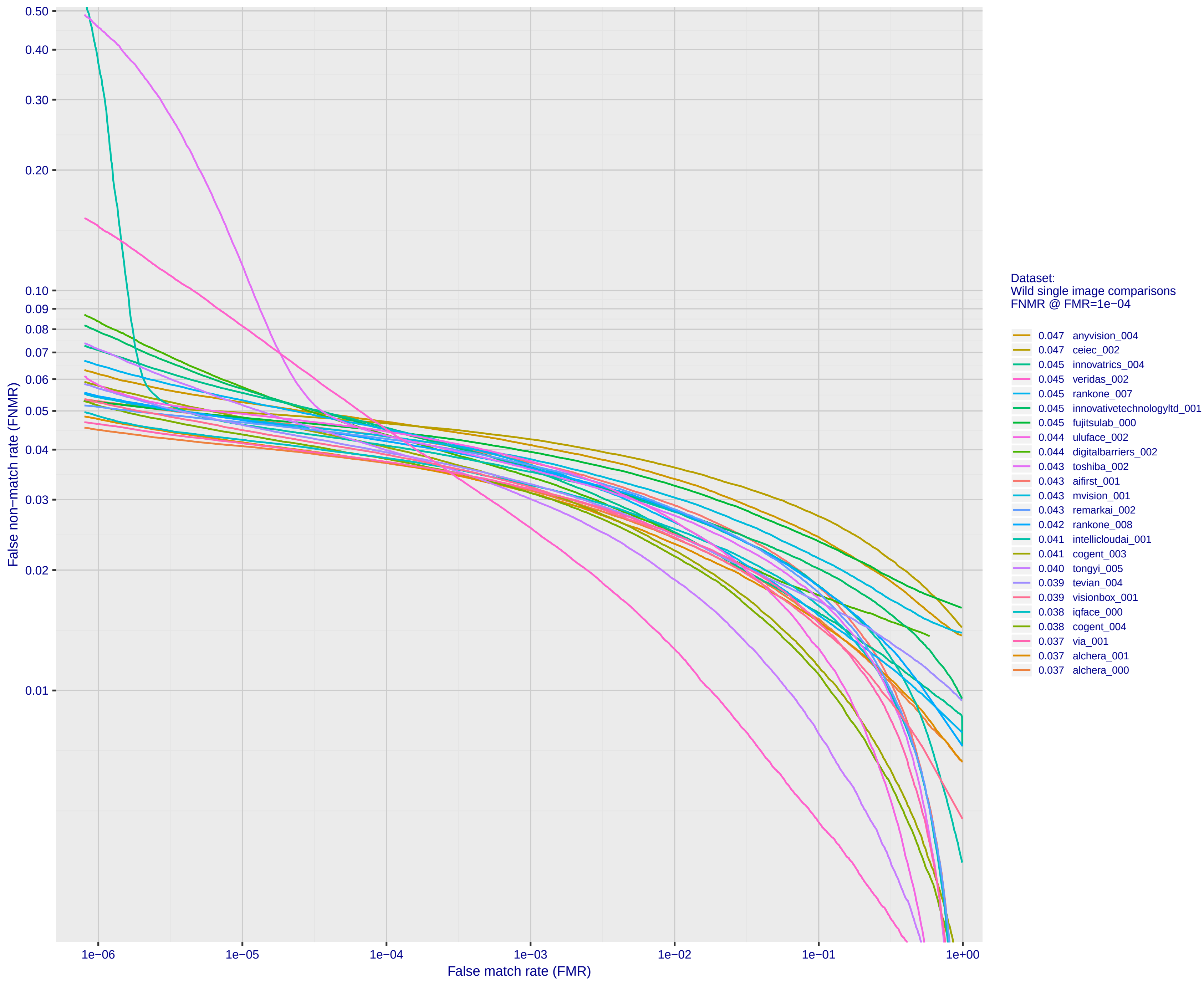


Figure 49: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

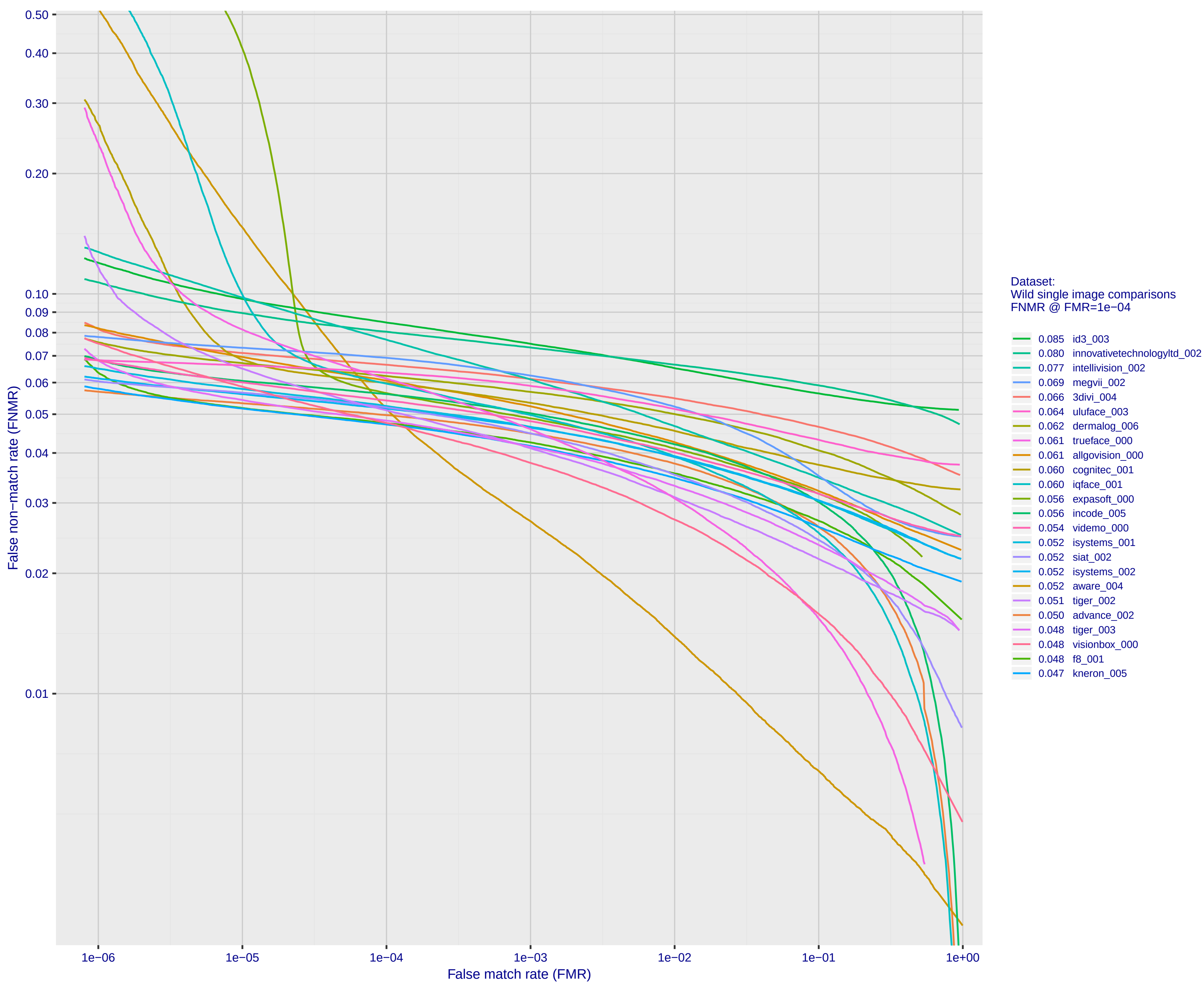


Figure 50: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

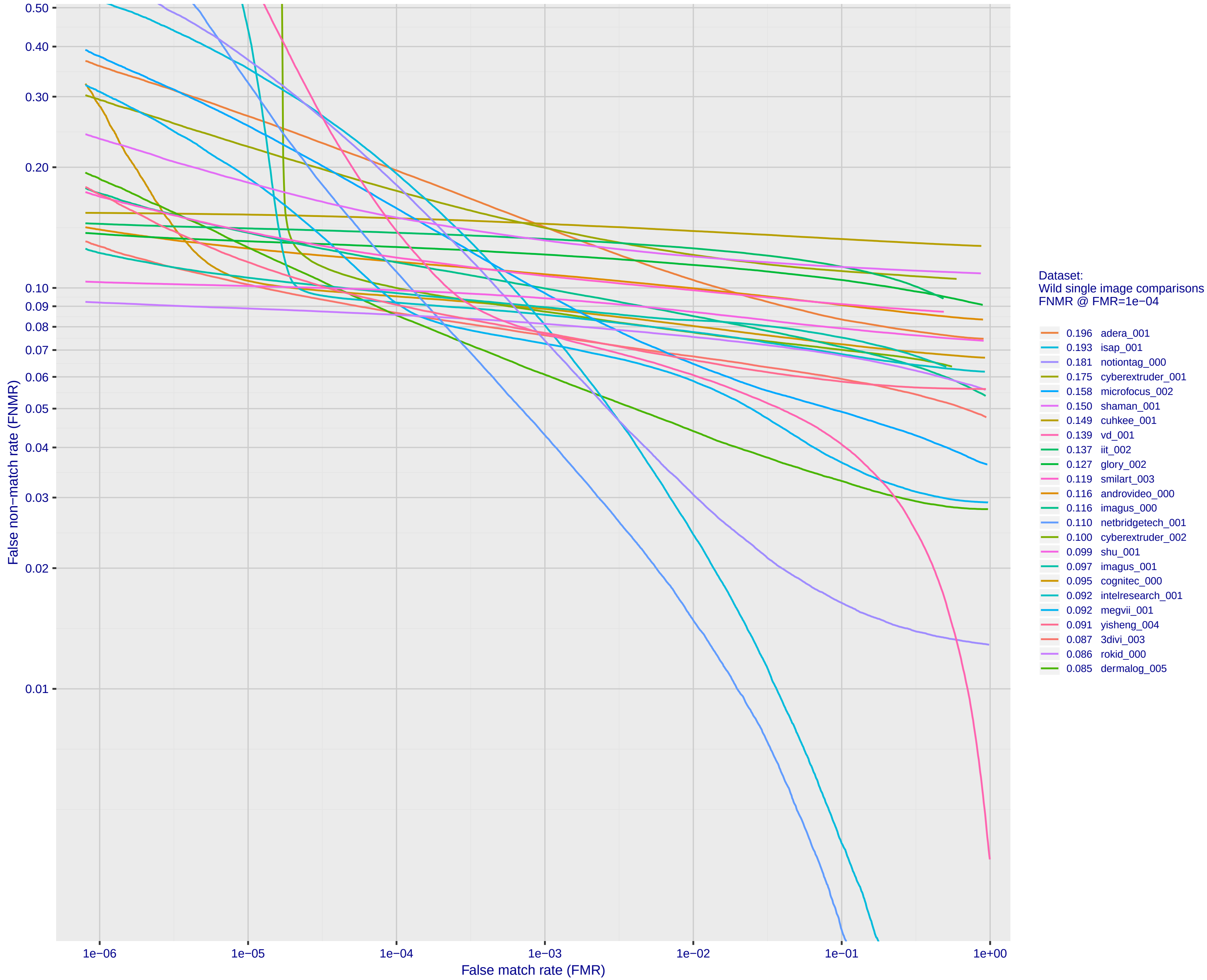


Figure 51: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

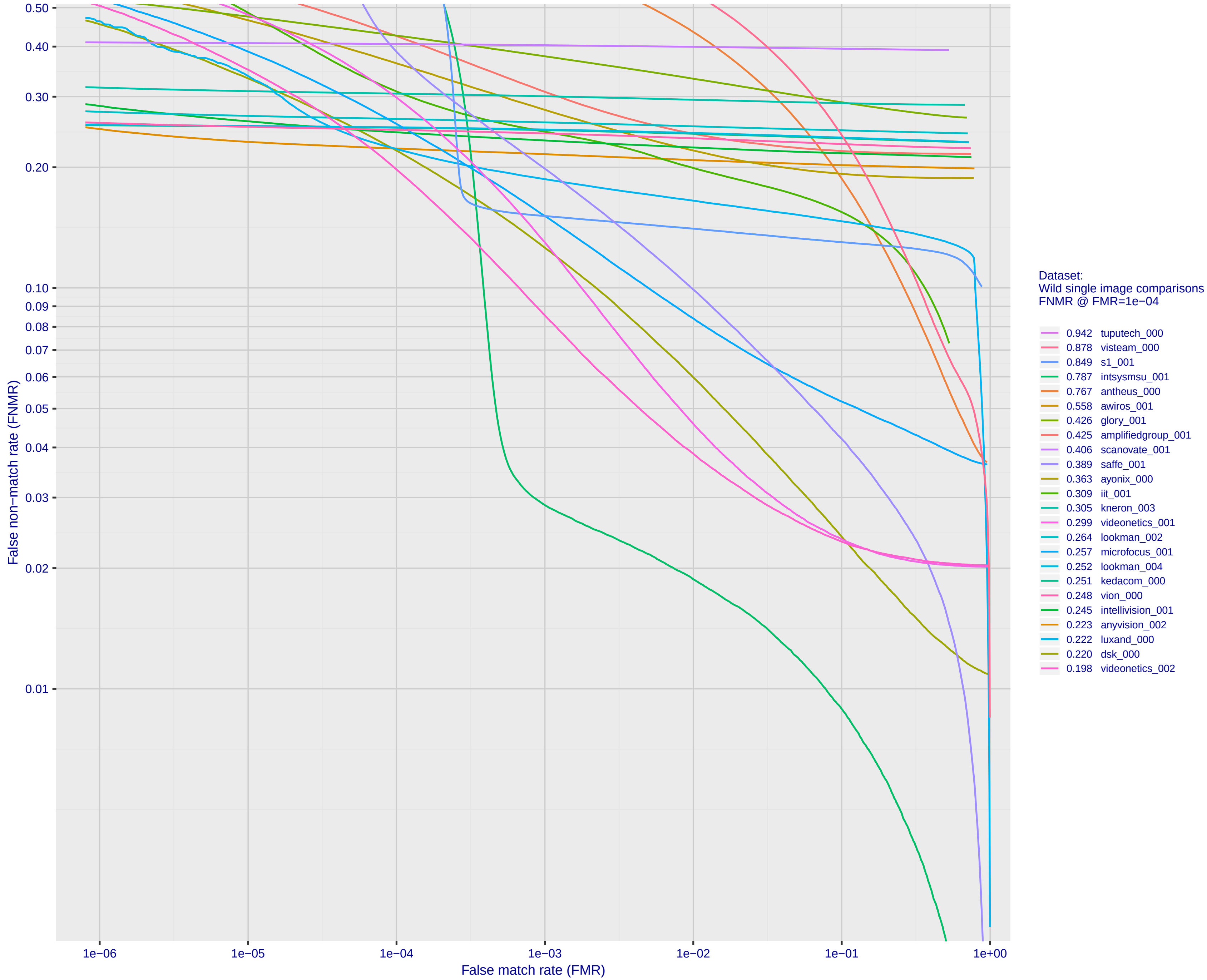


Figure 52: For the 2018 wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

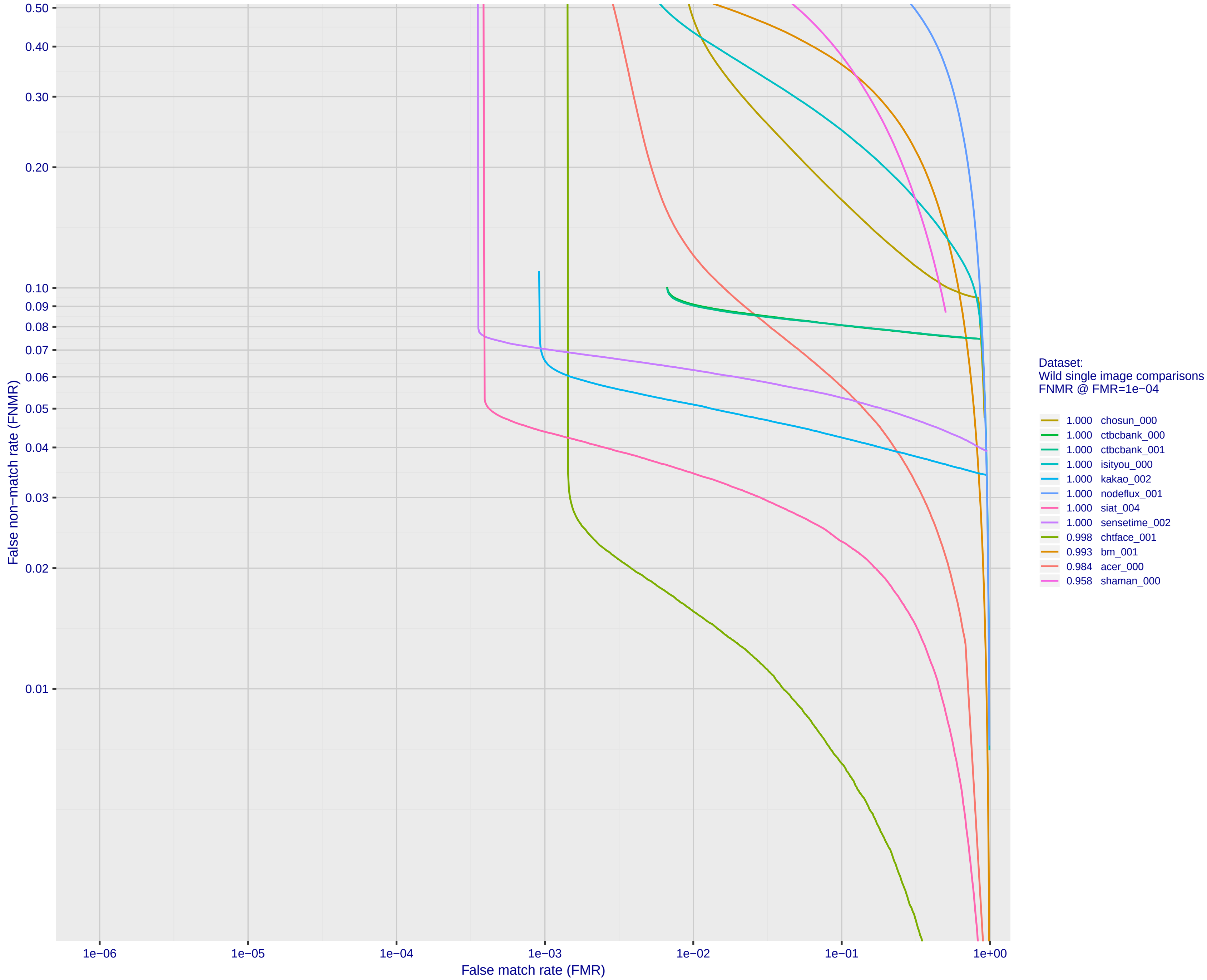


*Figure 53: For the 2018 wild image comparisons, detection error tradeoff (*DET*) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

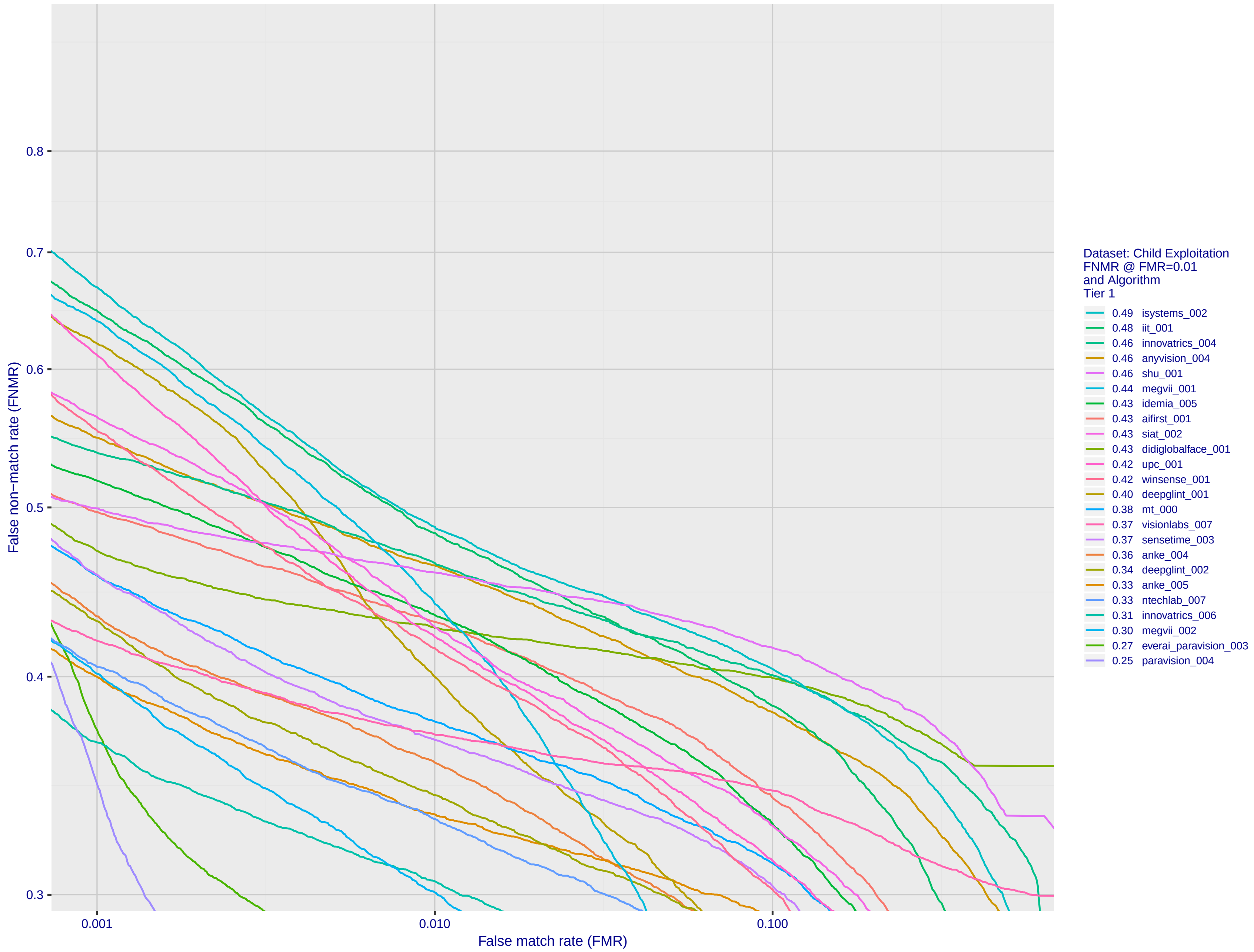


Figure 54: For child exploitation images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR. Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

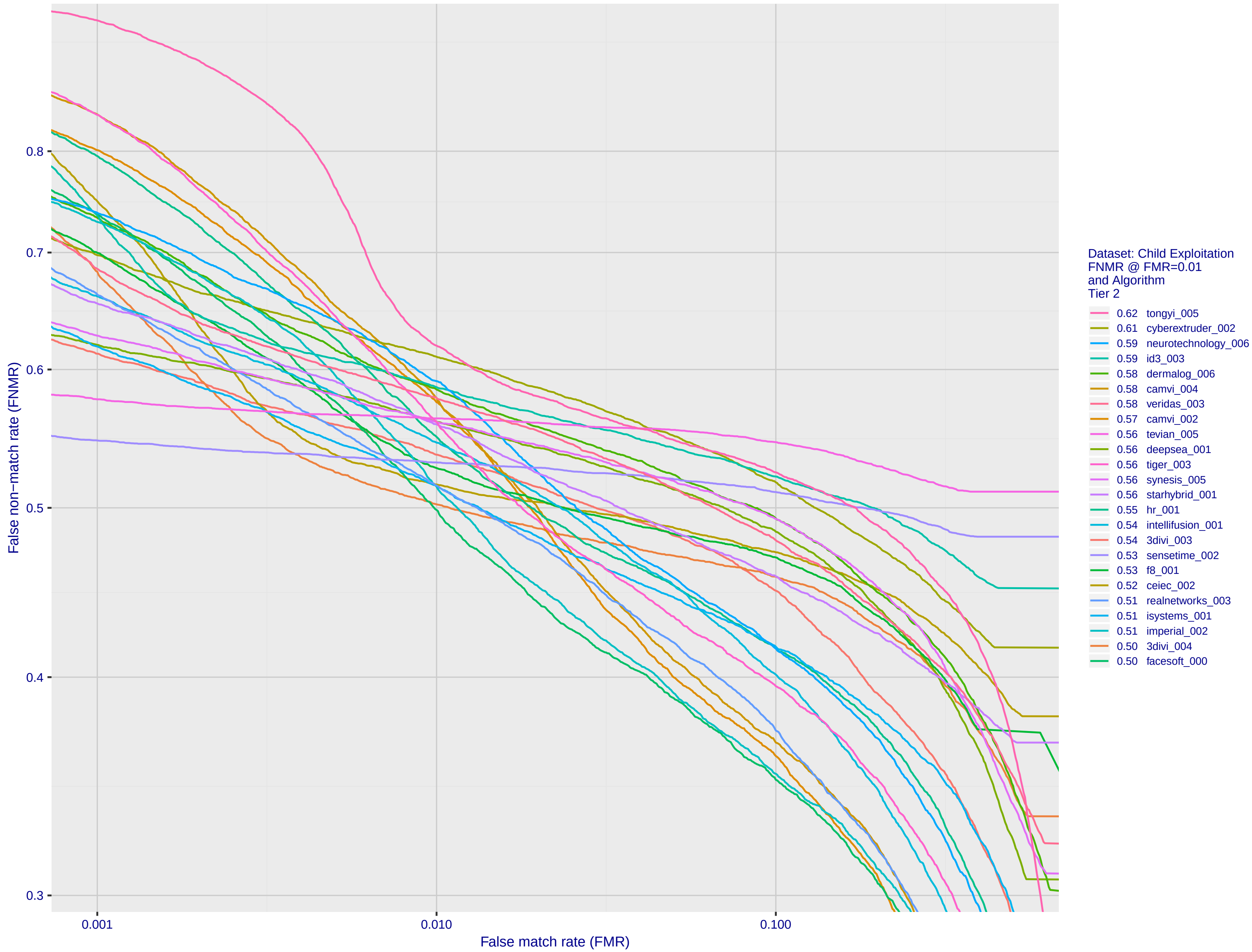


Figure 55: For child exploitation images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR. Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.

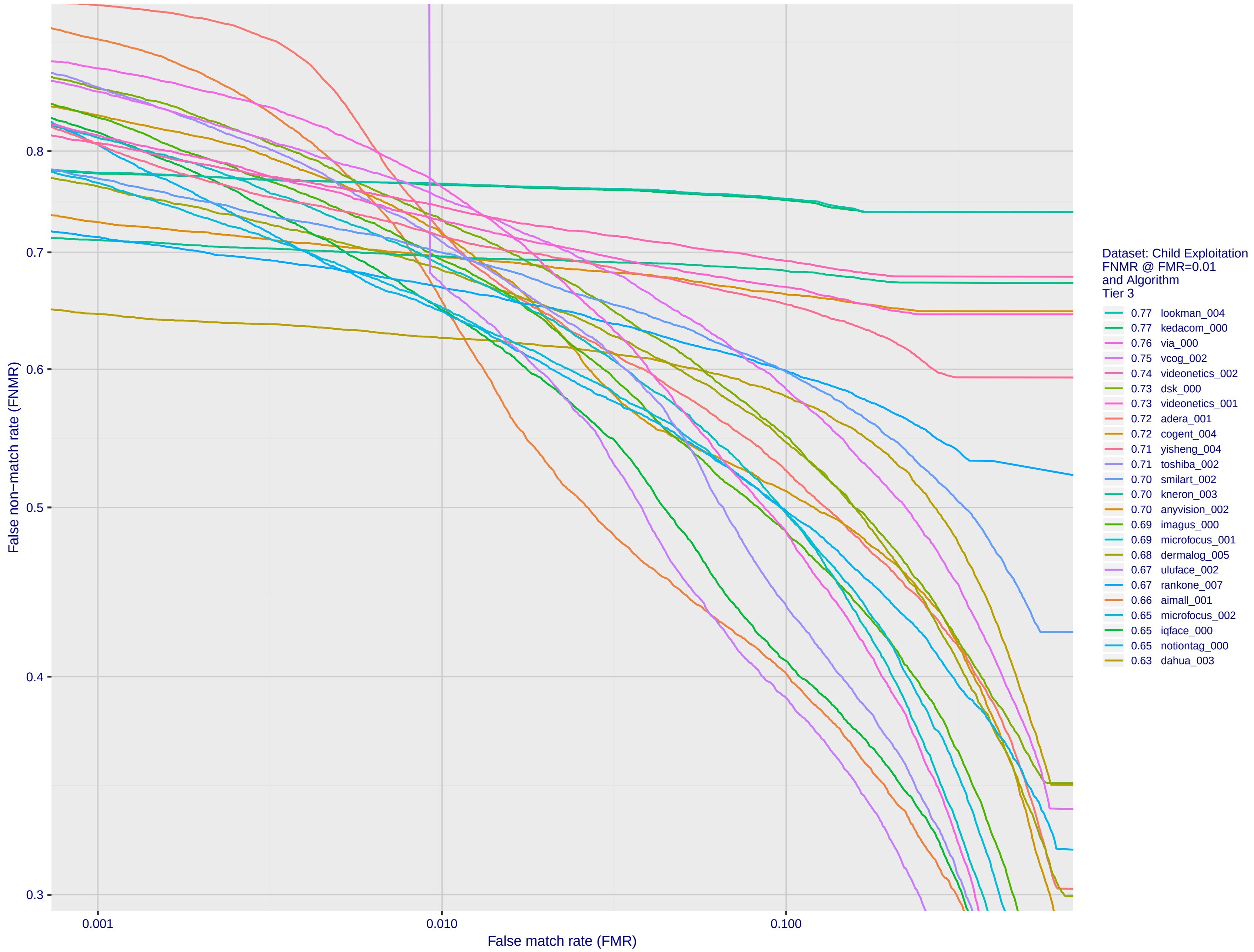


Figure 56: For child exploitation images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR. Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.

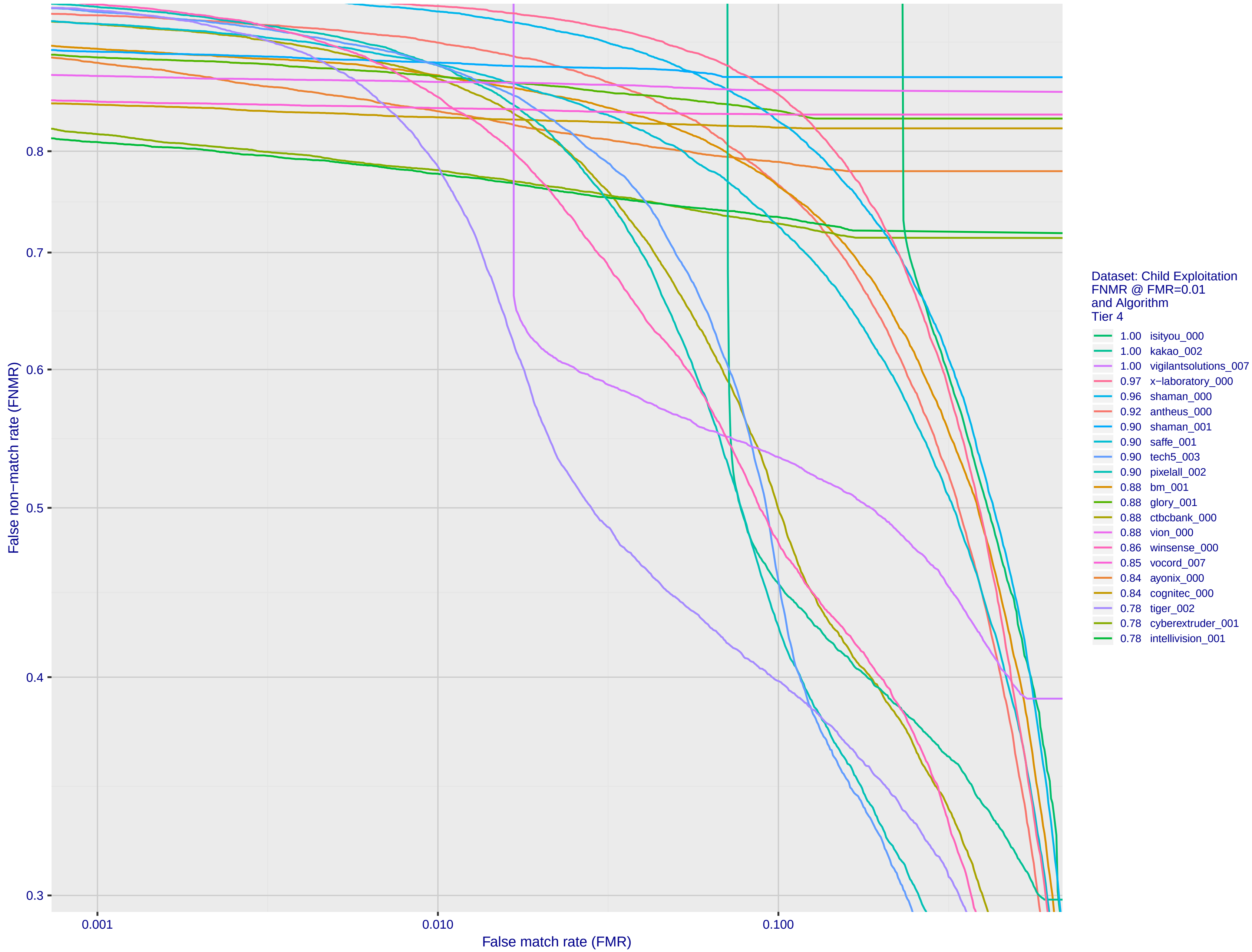


Figure 57: For child exploitation images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR. Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

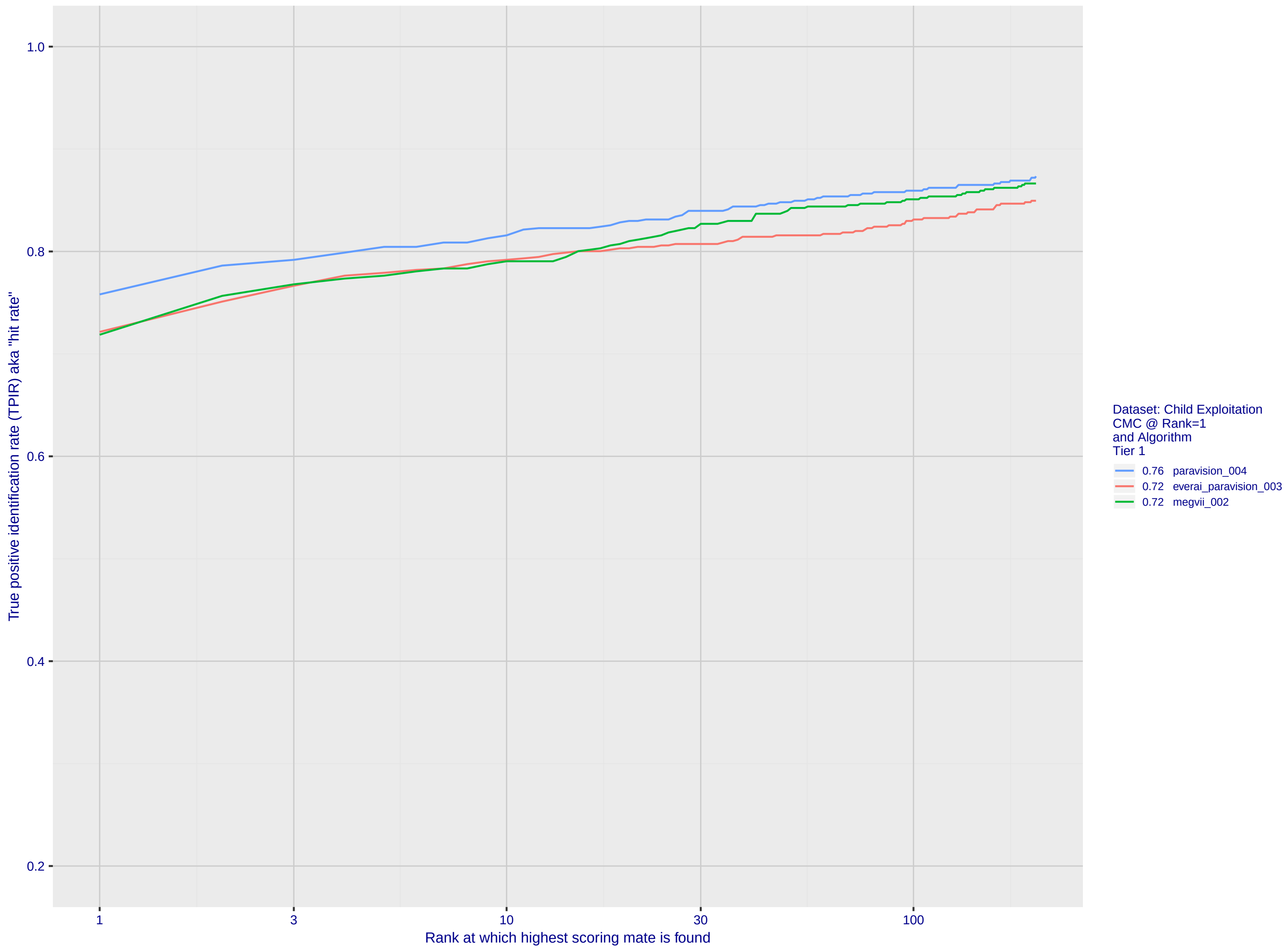


Figure 58: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 3.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 57 because a search can succeed if any *of a subject's several enrolled images matches the search image with a high score.*

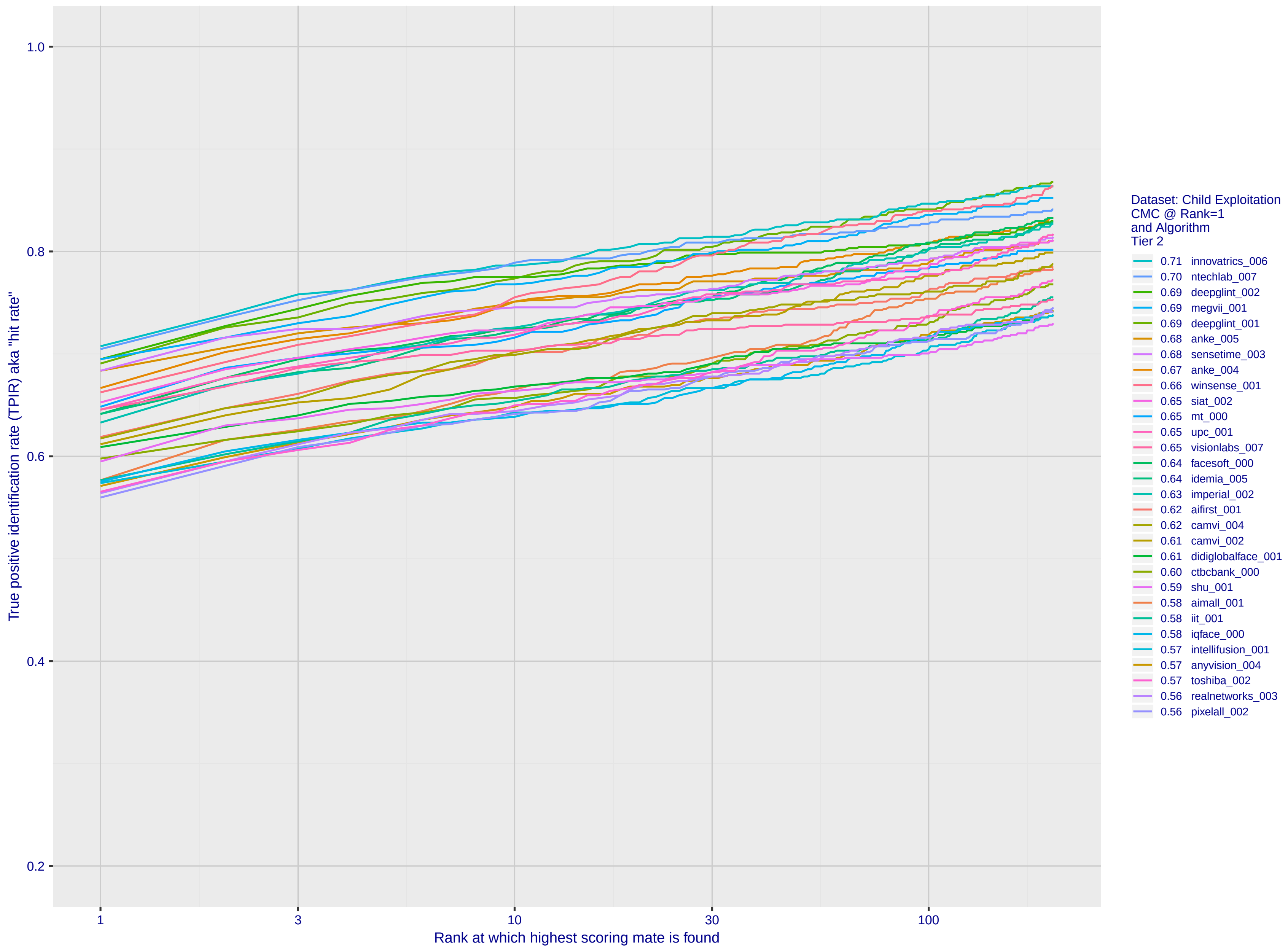


Figure 59: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 3.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 57 because a search can succeed if any *of a subject's several enrolled images matches the search image with a high score.*

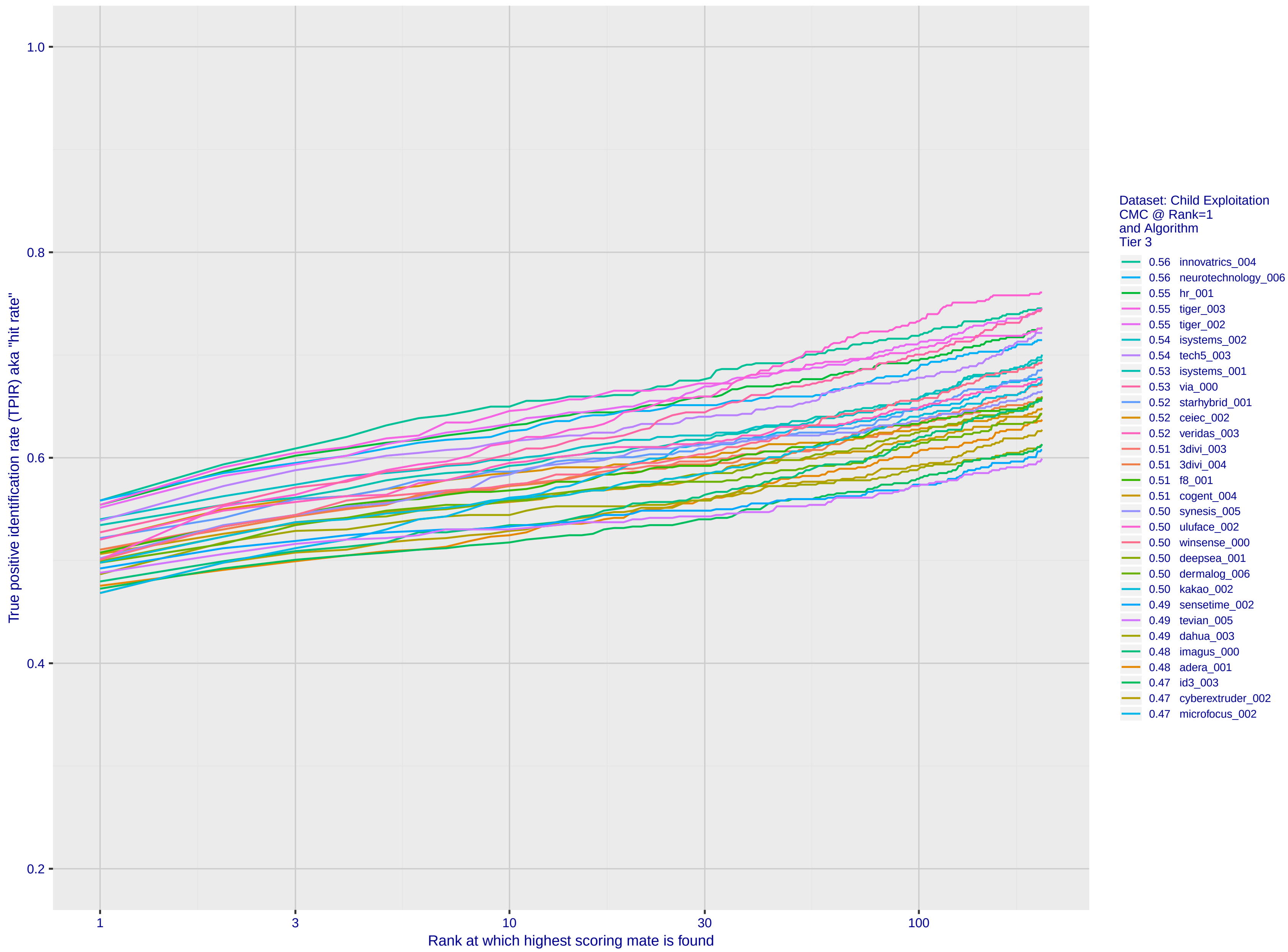


Figure 60: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 3.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 57 because a search can succeed if any *of a subject's several enrolled images matches the search image with a high score.*

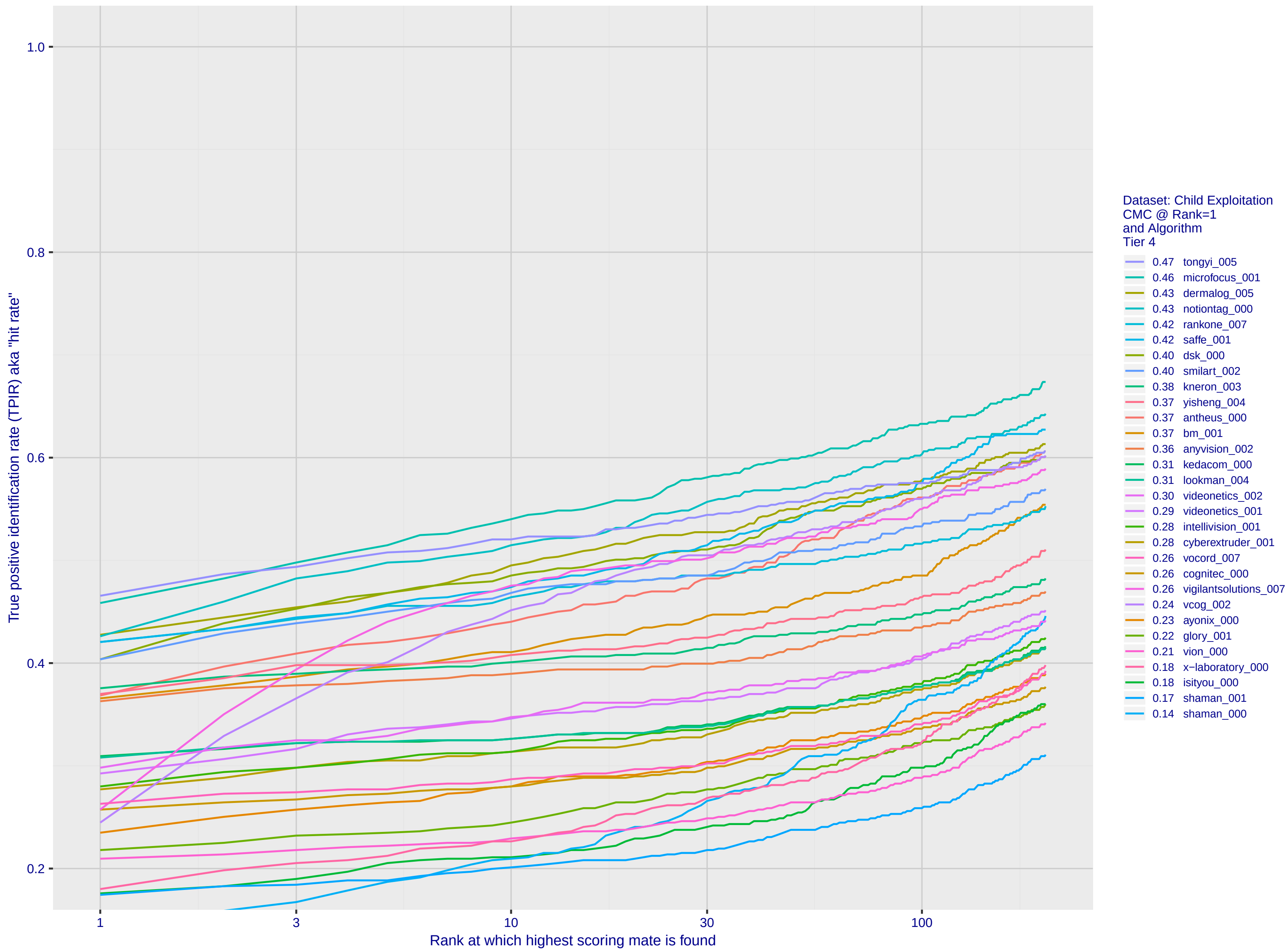


Figure 61: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 3.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 57 because a search can succeed if any *of a subject's several enrolled images matches the search image with a high score.*

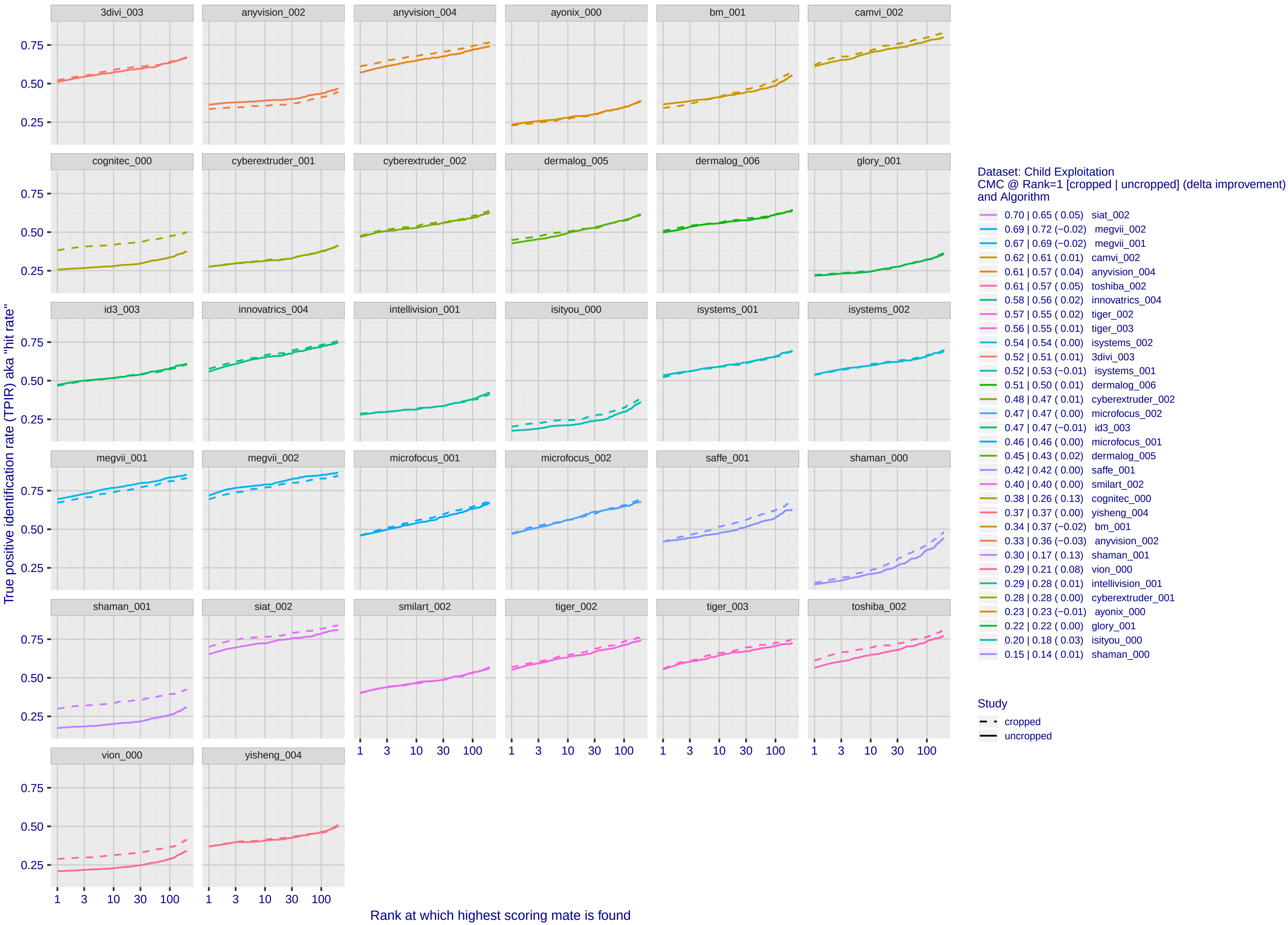


Figure 62: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank for two cases: 1. Whole image provided to the algorithm; 2. Human annotated rectangular region, cropped and provided to the algorithm. The difference between the traces is associated with detection of difficult faces, and fine localization.

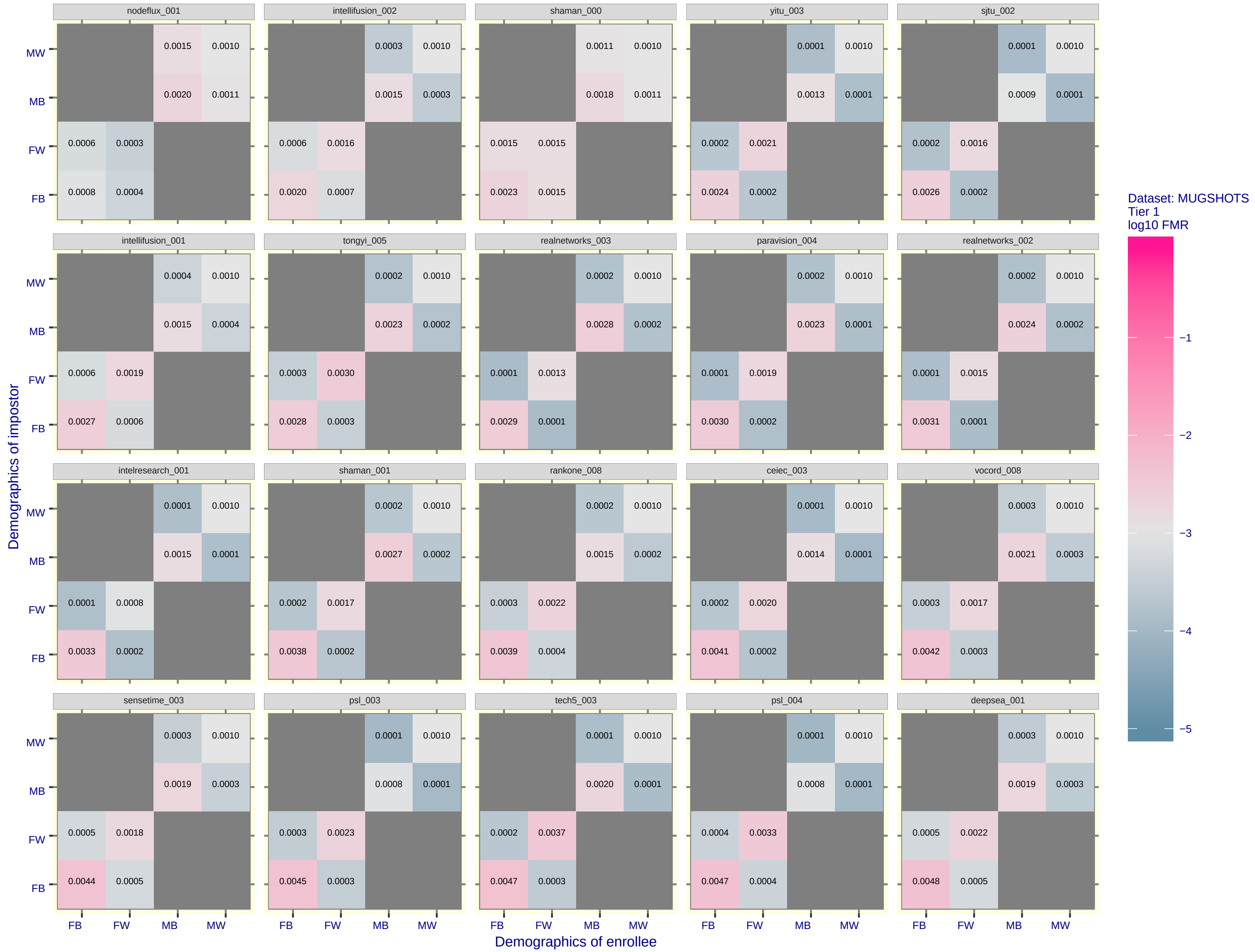


Figure 63: For the mugshot images, FMR for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give FMR = 0.001 for white males which is the demographc that usually gives the lowest FMR. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of FMR on black females, which is the demographc that usually gives the highest FMR.

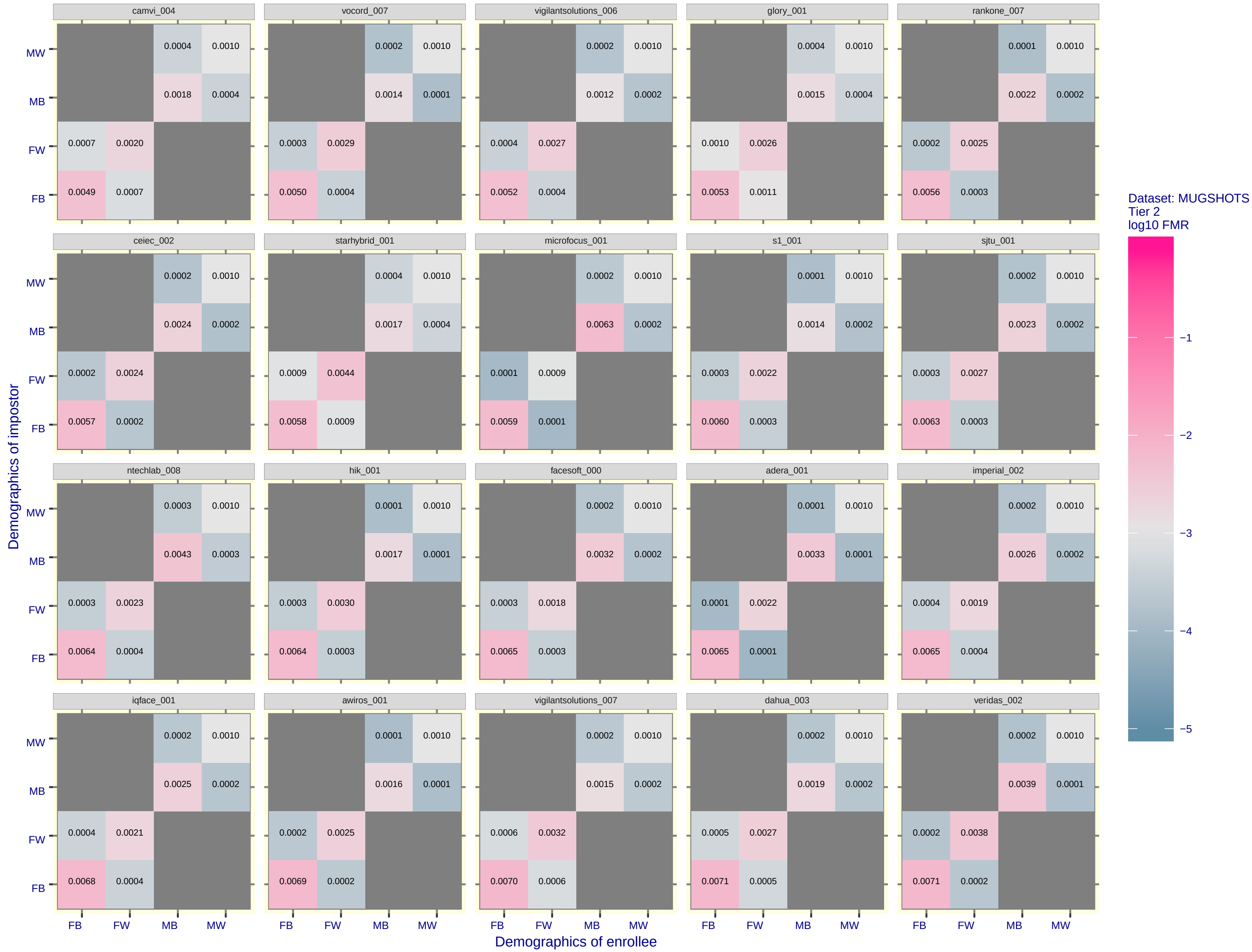


Figure 64: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

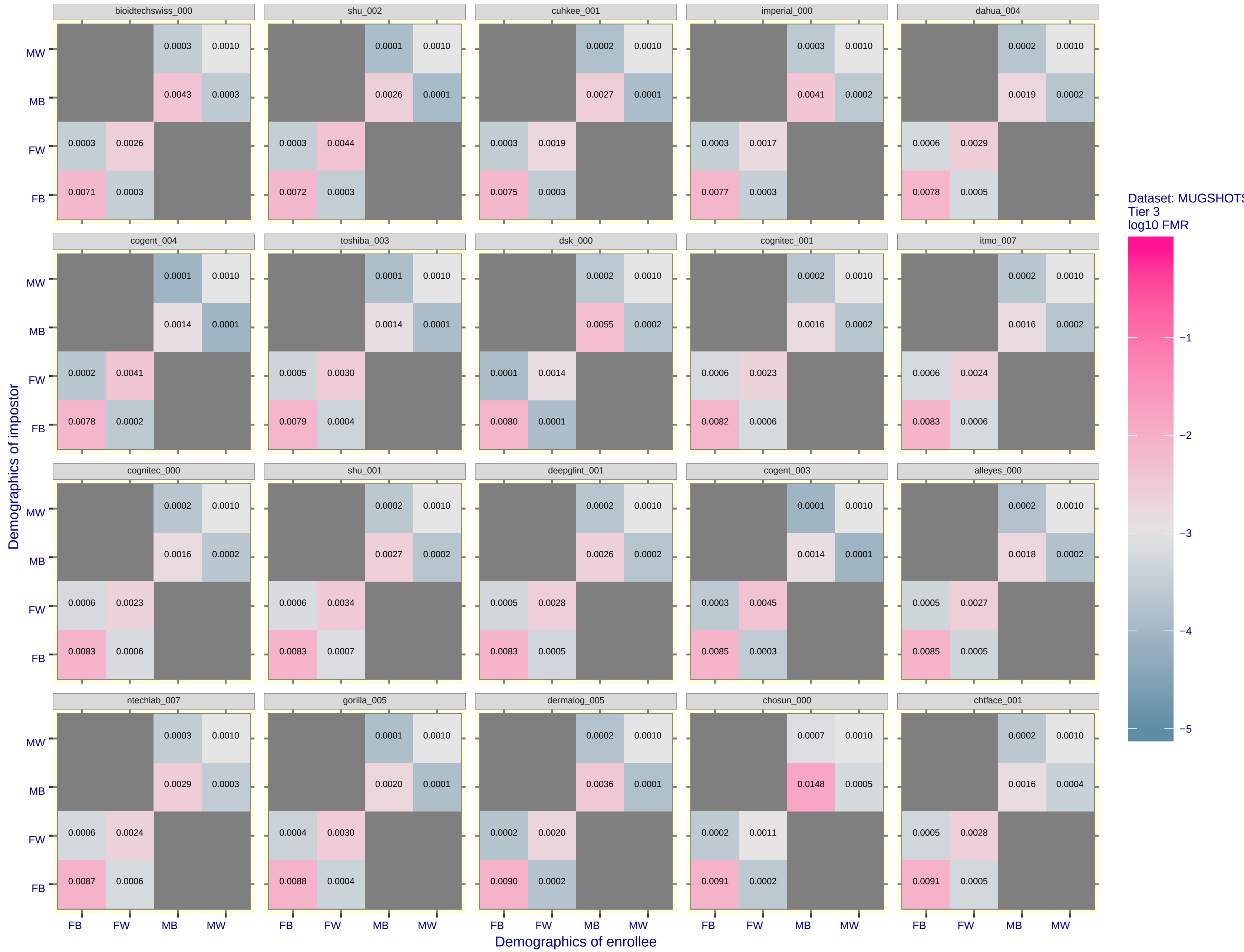


Figure 65: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

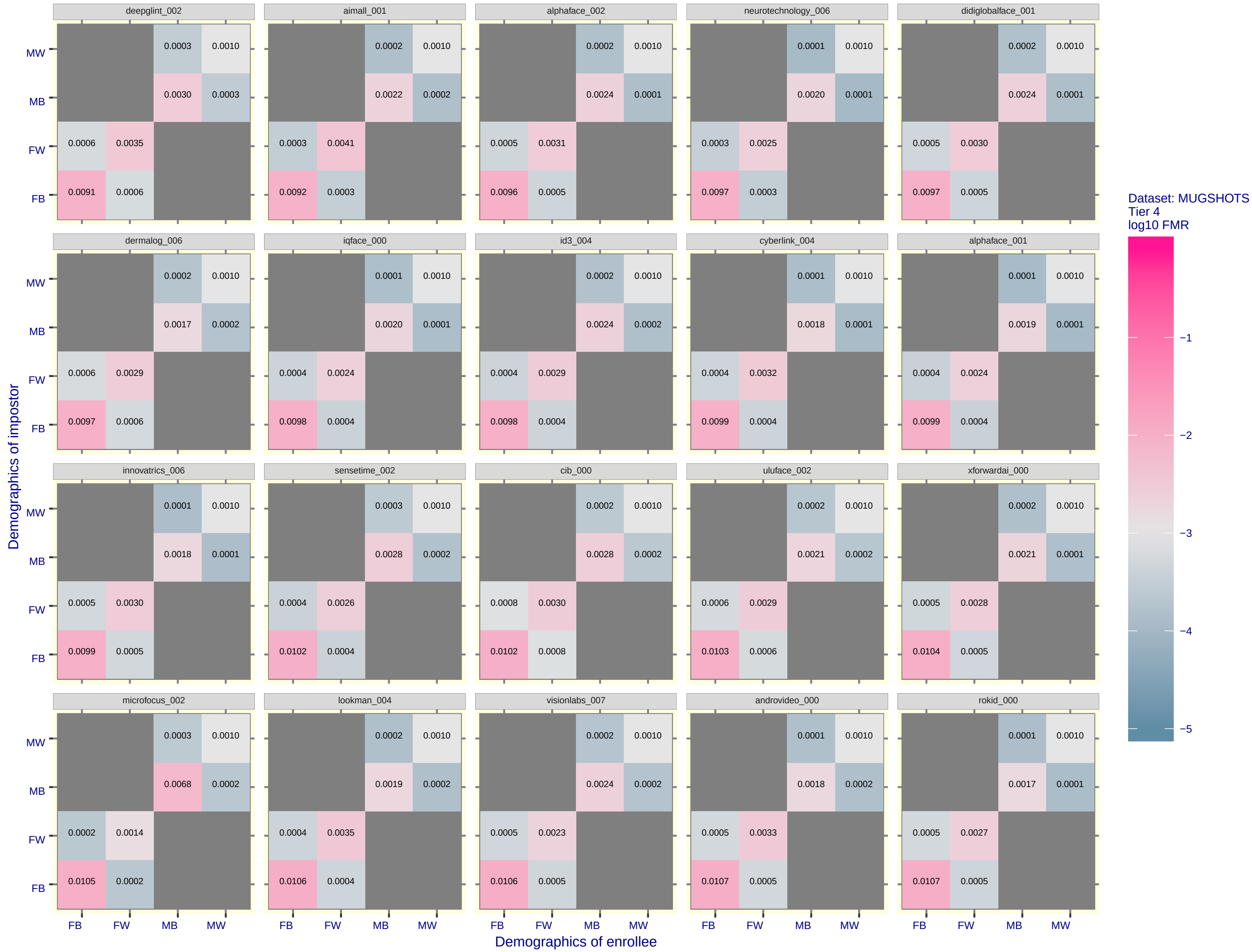


Figure 66: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

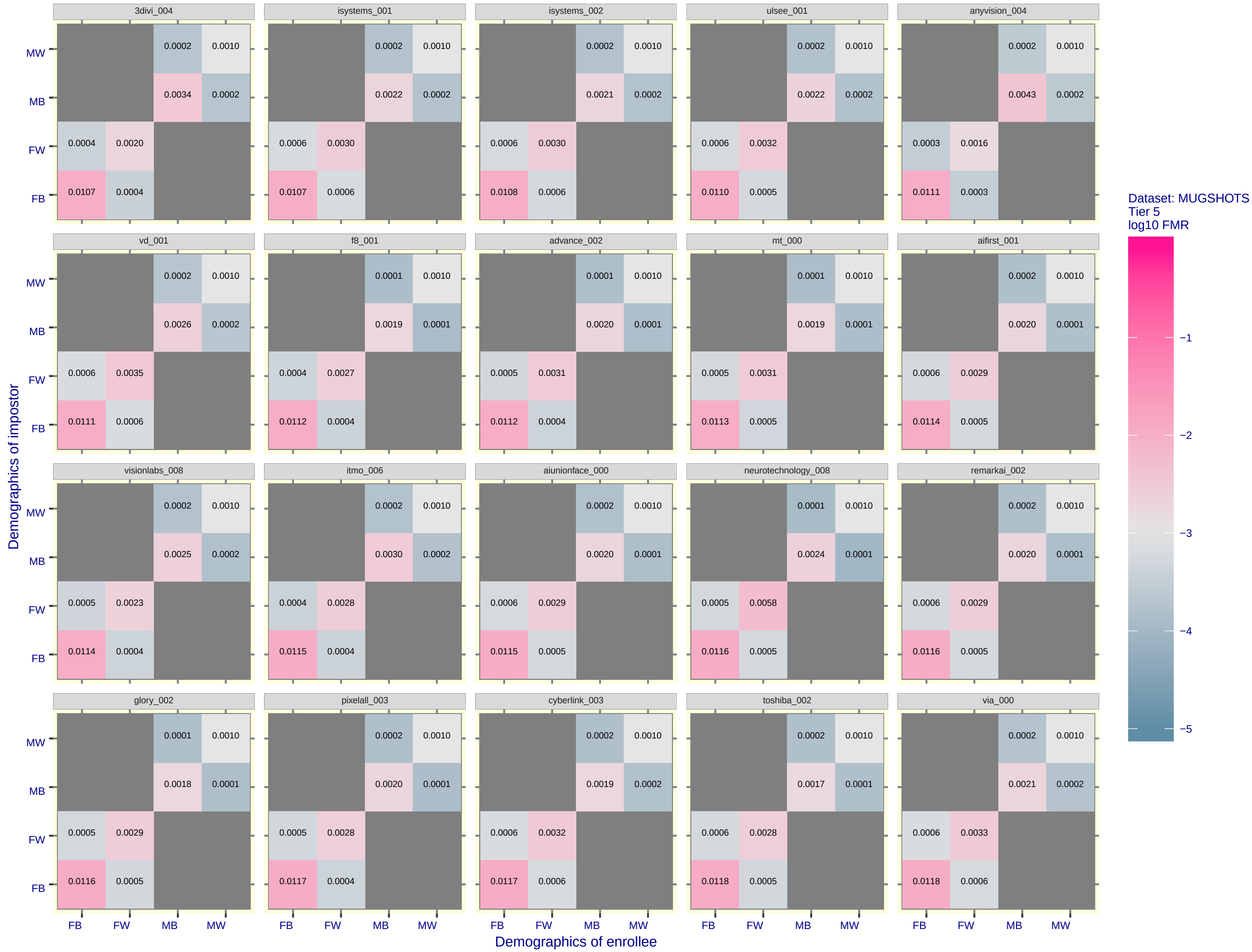


Figure 67: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

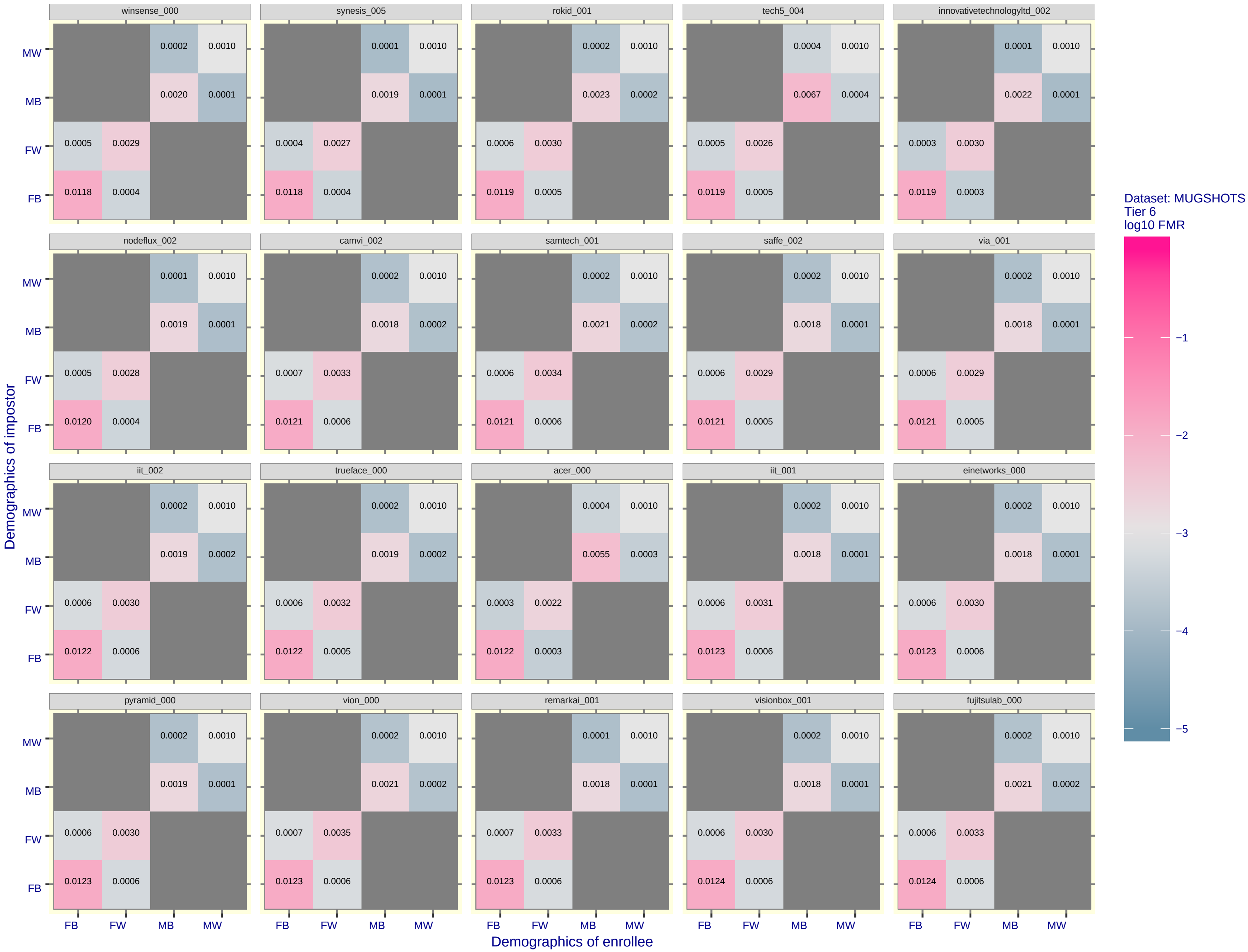


Figure 68: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

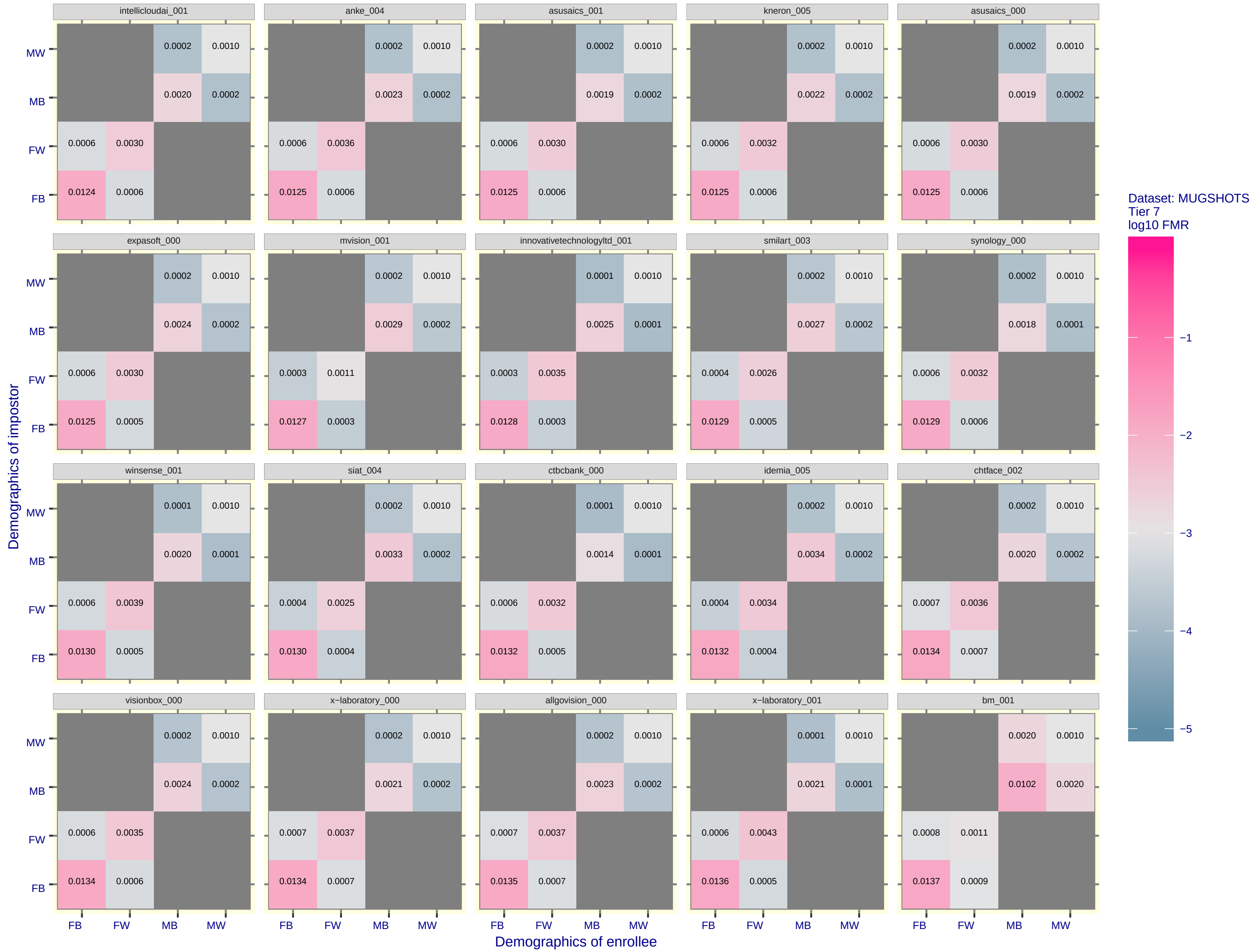


Figure 69: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

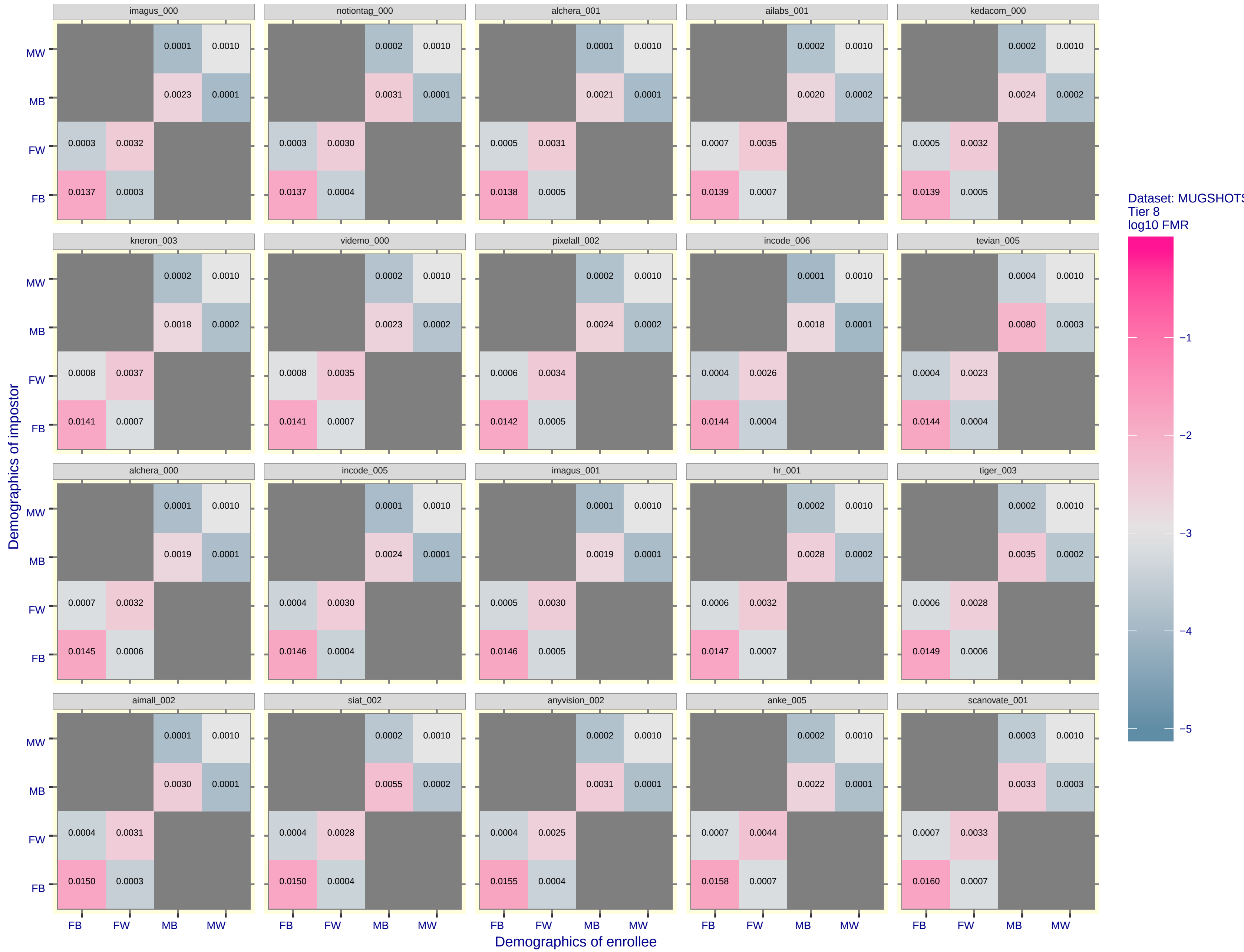


Figure 70: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

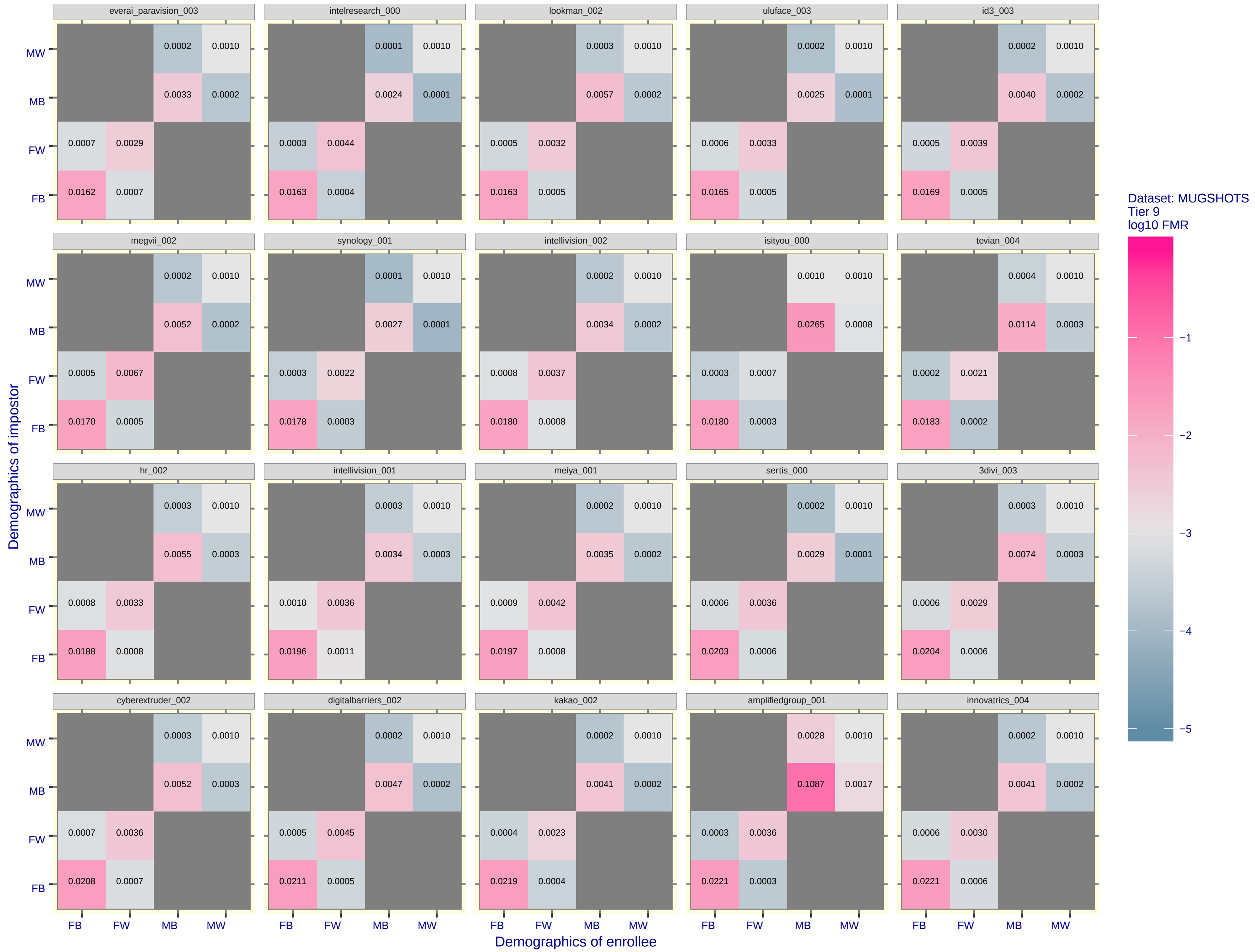


Figure 71: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

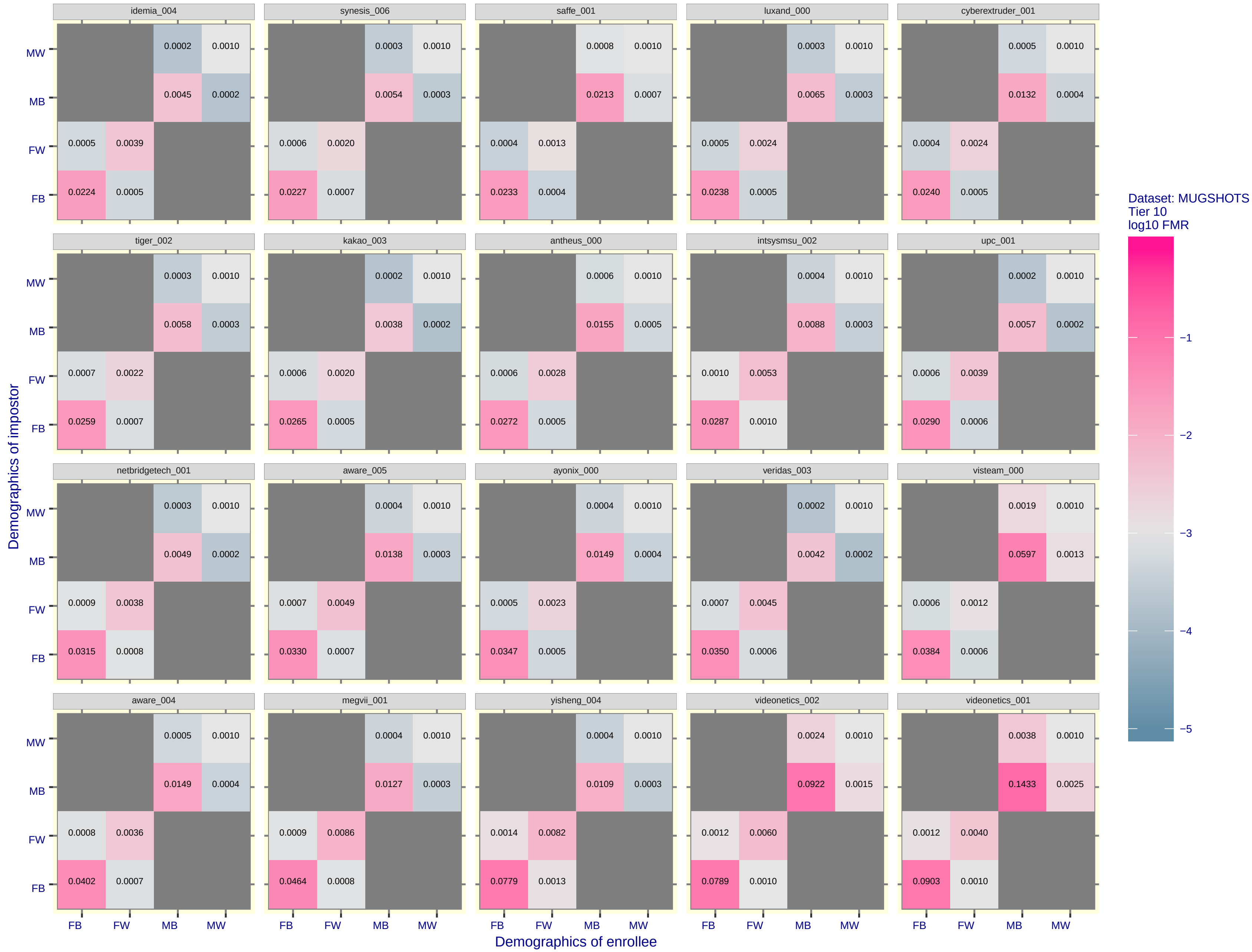


Figure 72: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

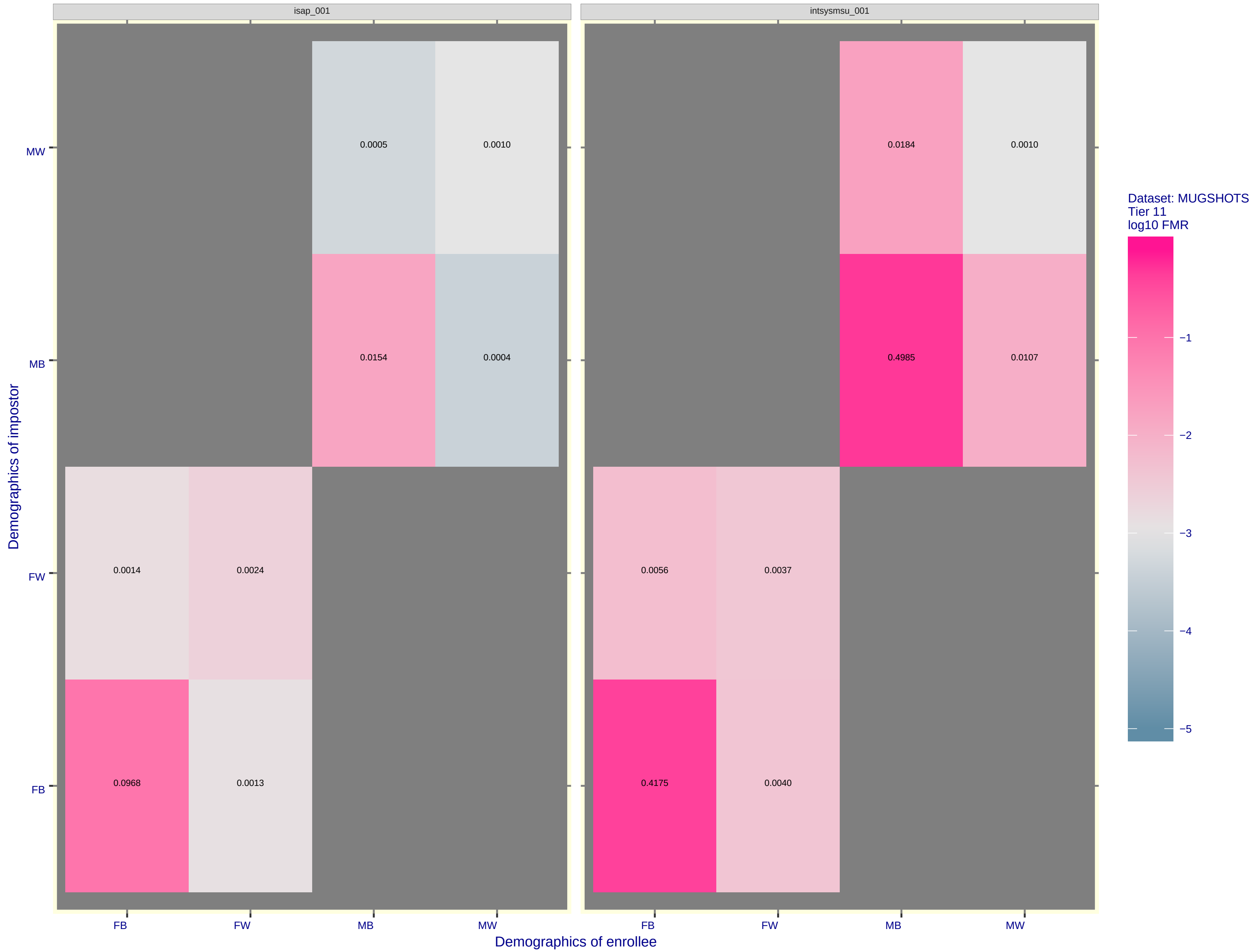


Figure 73: For the mugshot images, FMR *for same-sex impostor pairs of images annotated with codes for black female, black male, white female, white male. The threshold is set for each algorithm to give* FMR *= 0.001 for white males which is the demographc that usually gives the lowest* FMR*. This means the top right box is the same color in all panels. The panels are sorted over multiple pages in order of* FMR *on black females, which is the demographc that usually gives the highest* FMR*.*

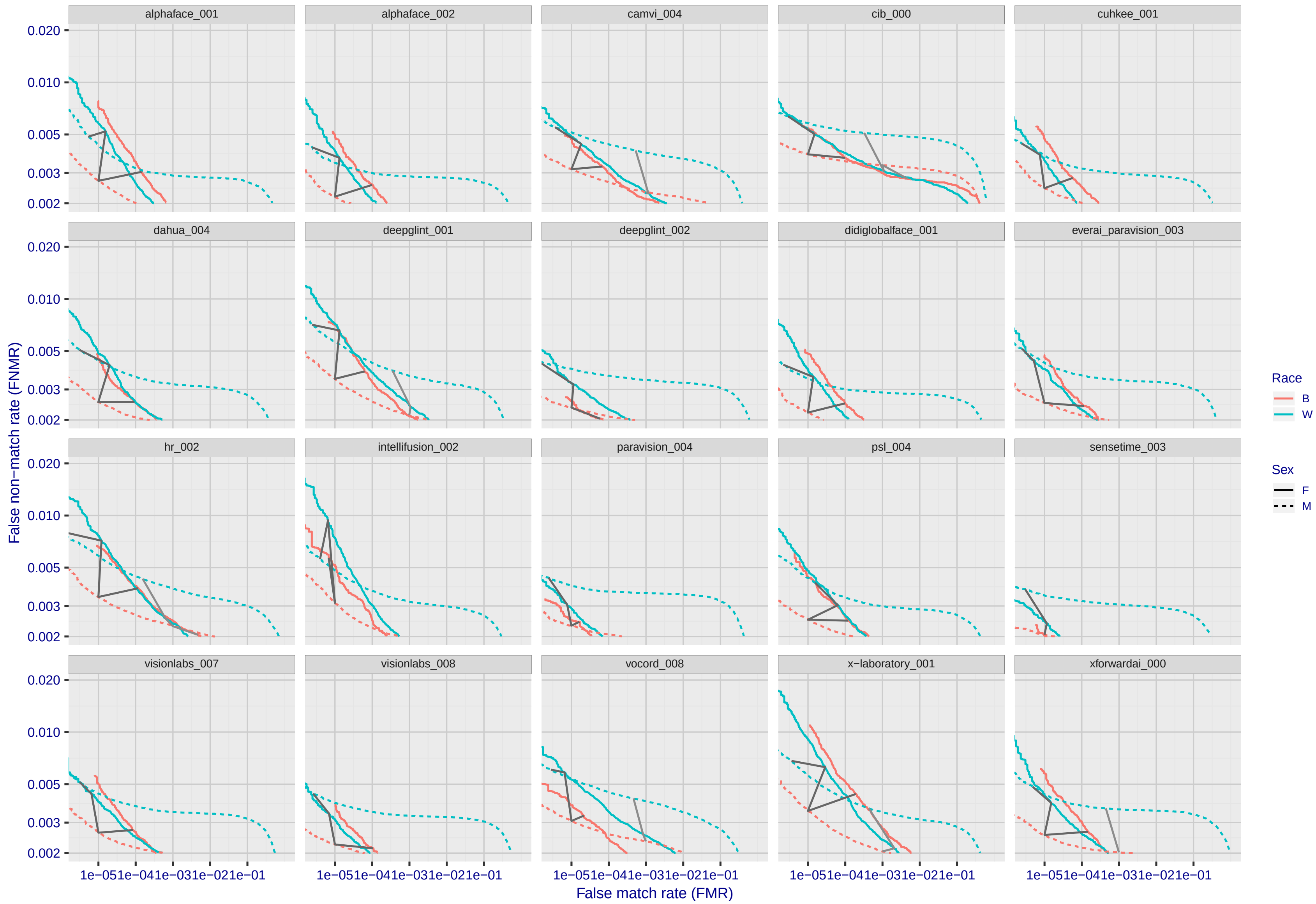


Figure 74: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

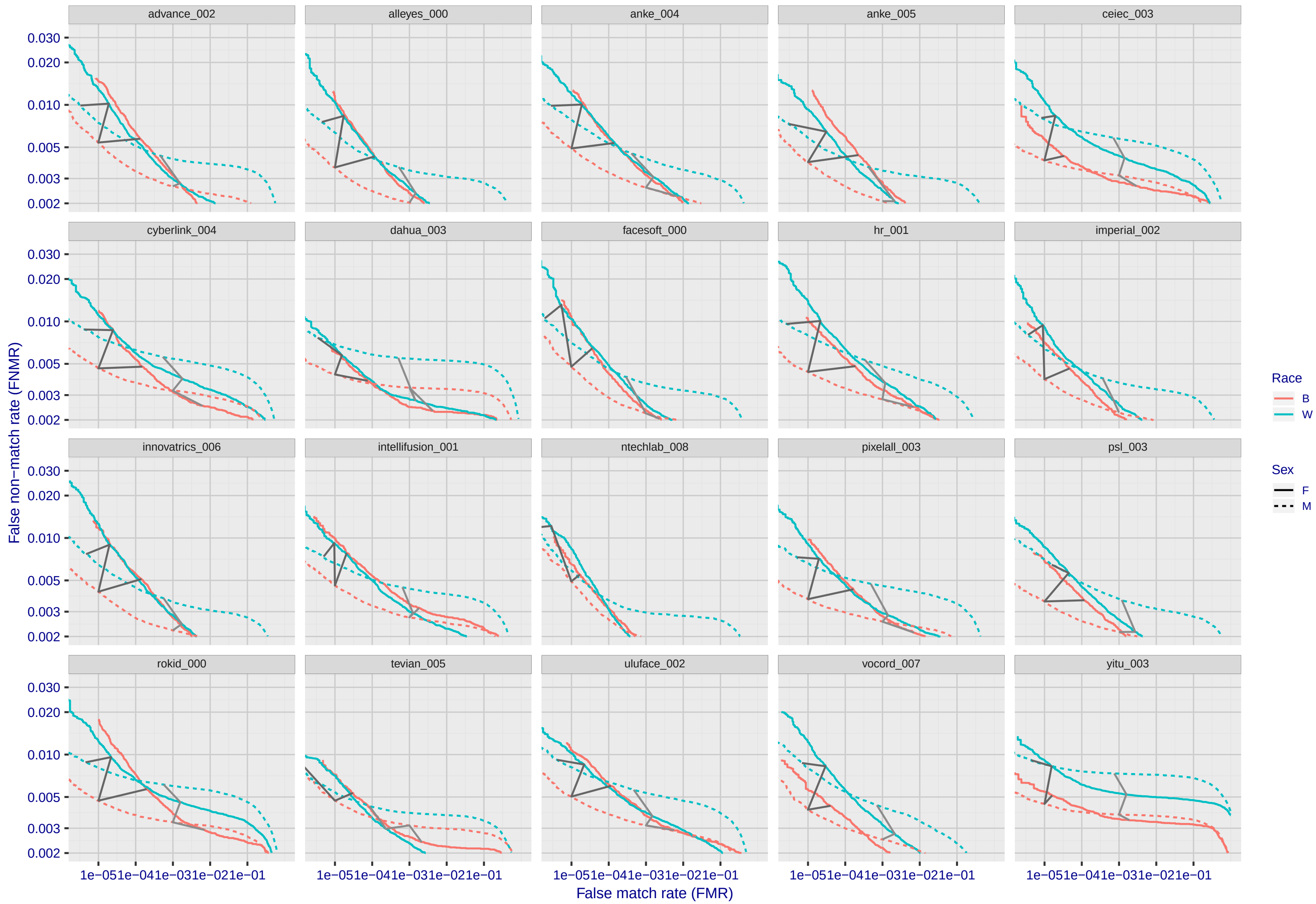


Figure 75: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

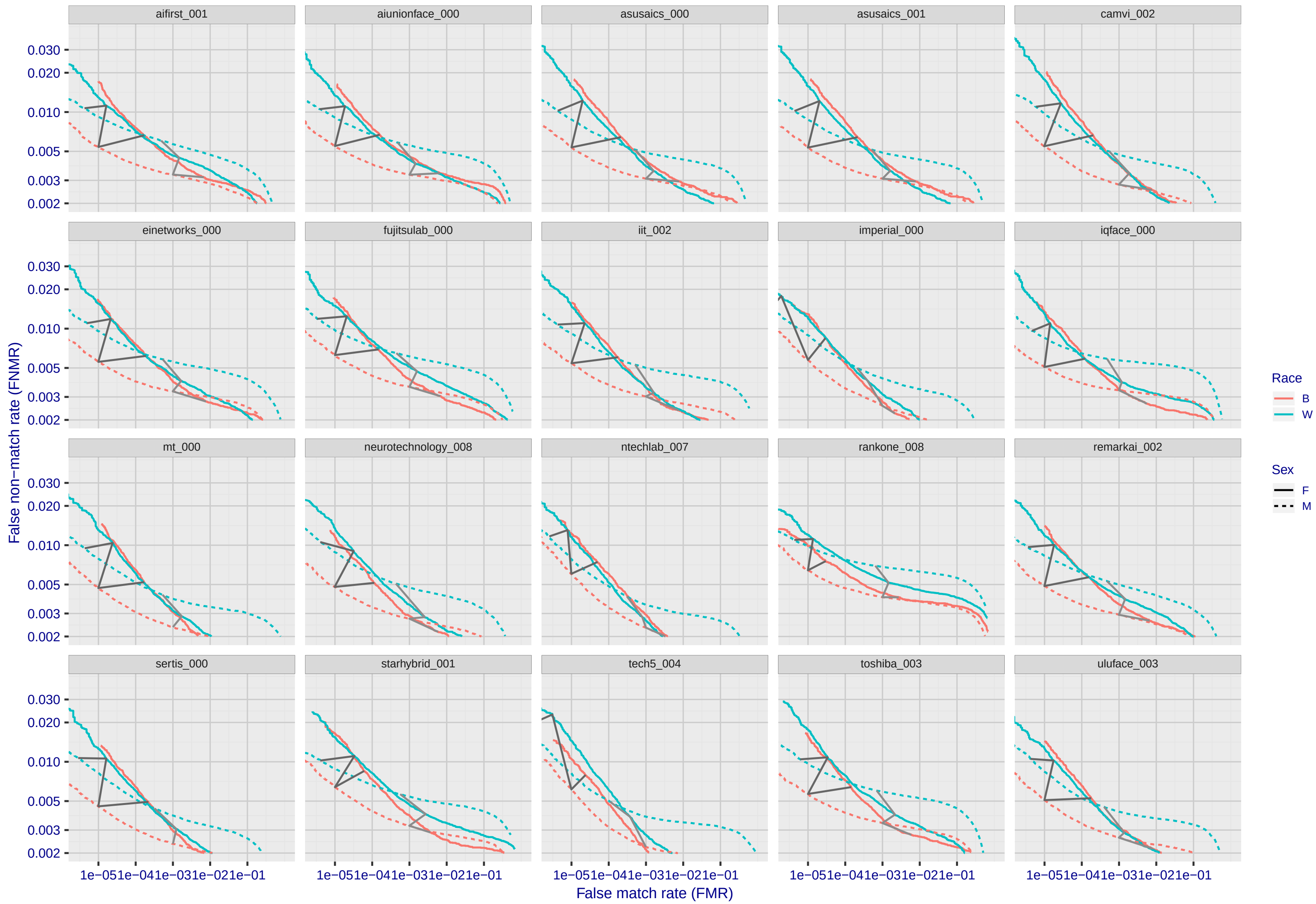


Figure 76: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

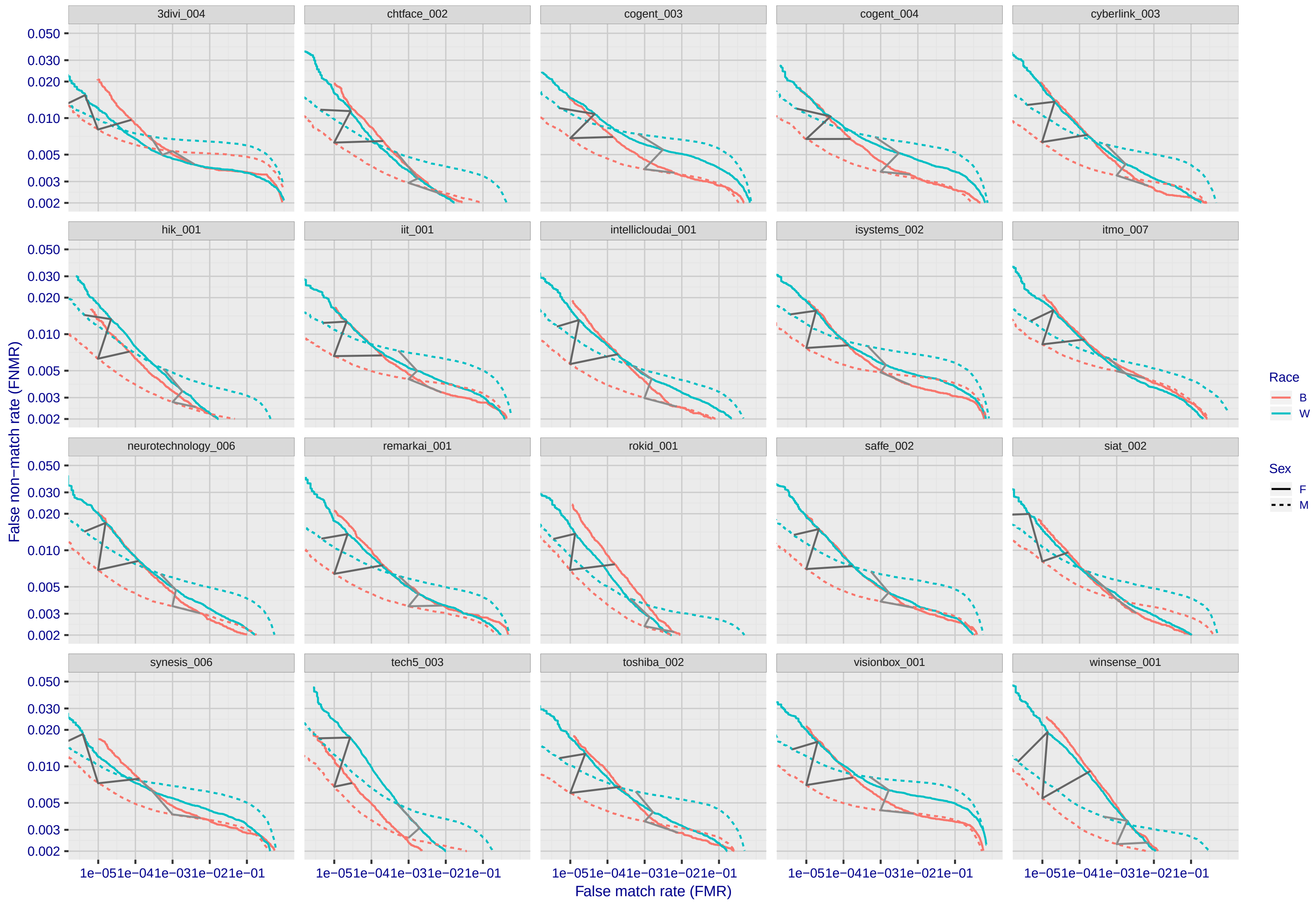


Figure 77: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

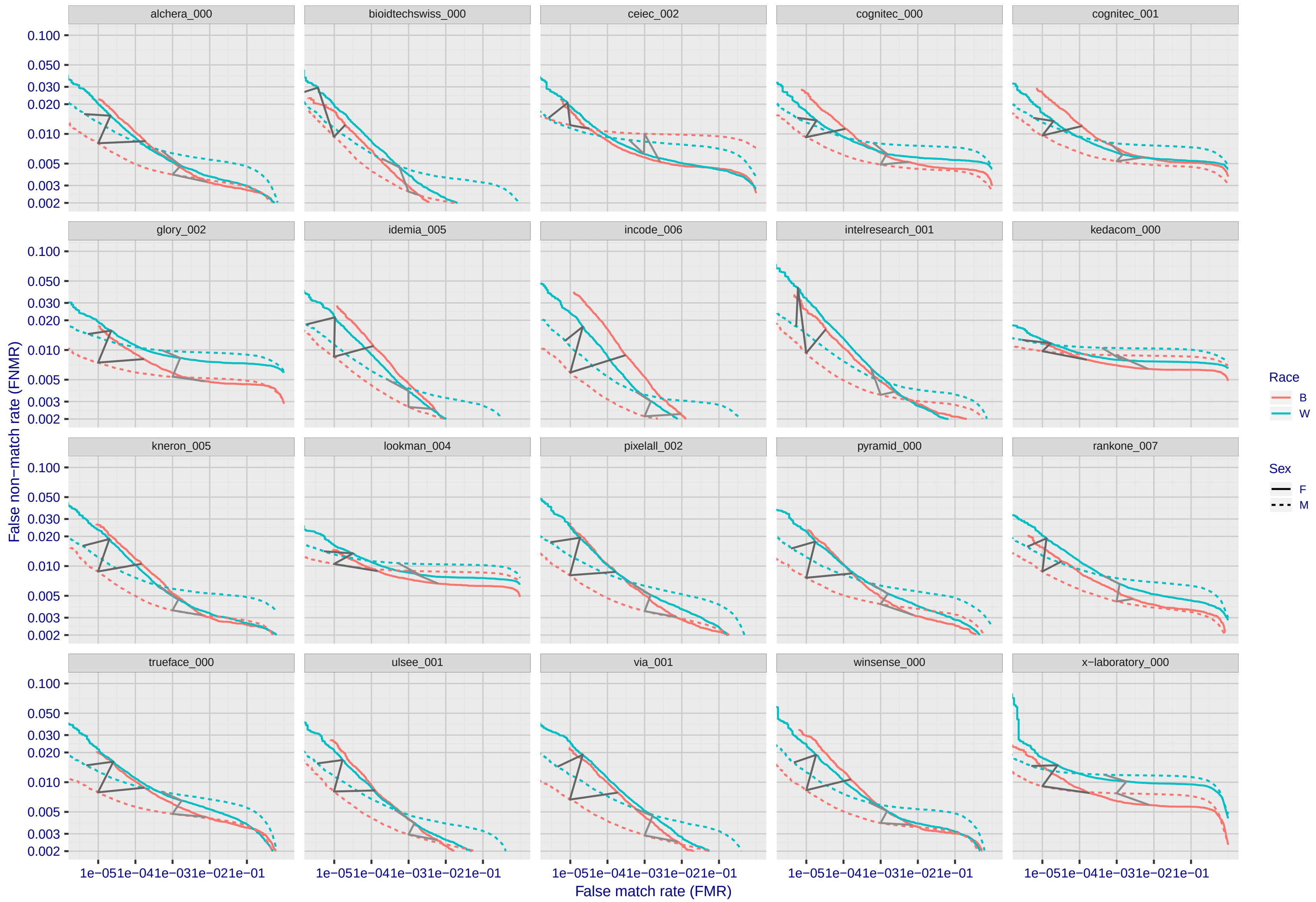


Figure 78: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

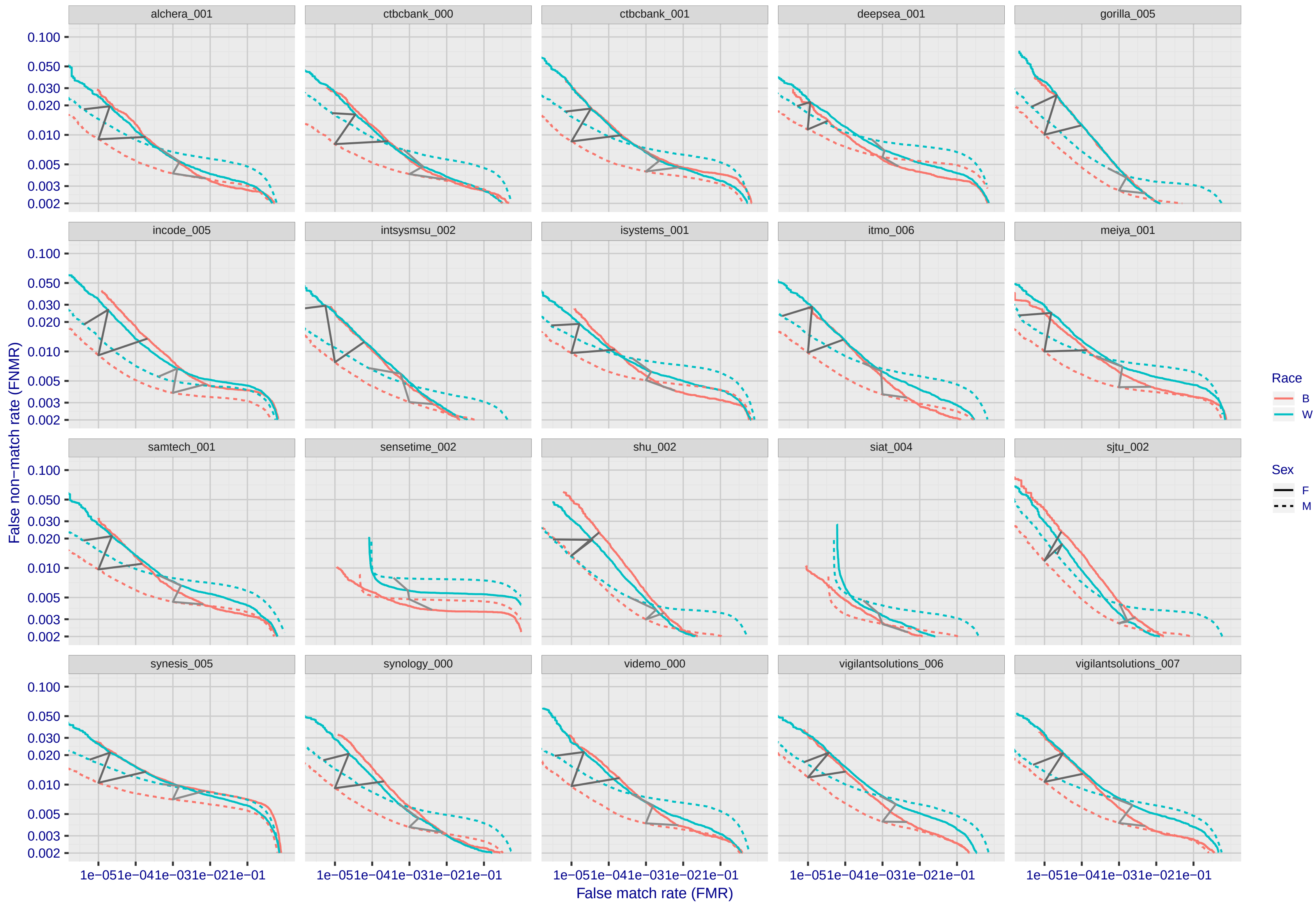


Figure 79: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

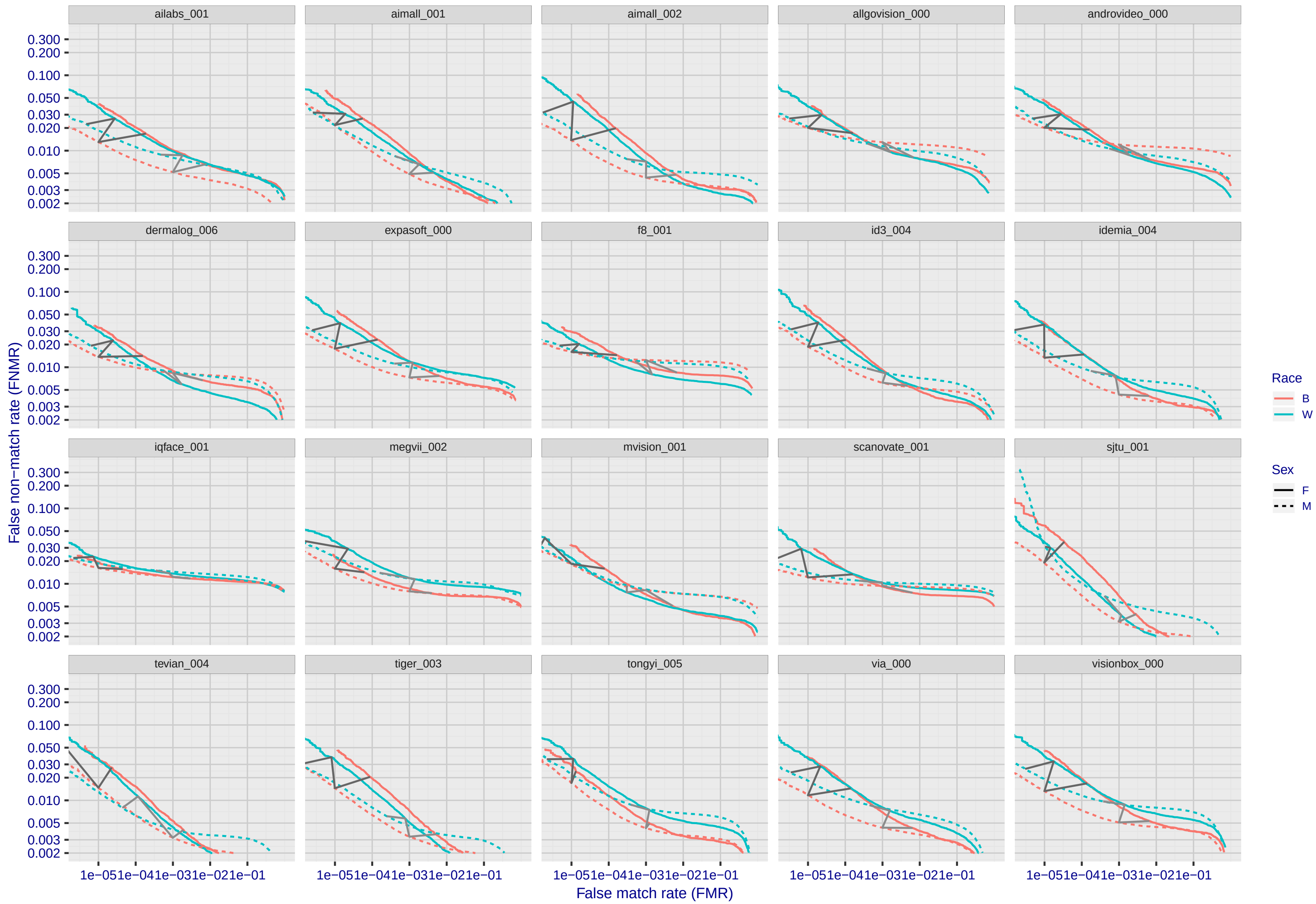


Figure 80: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

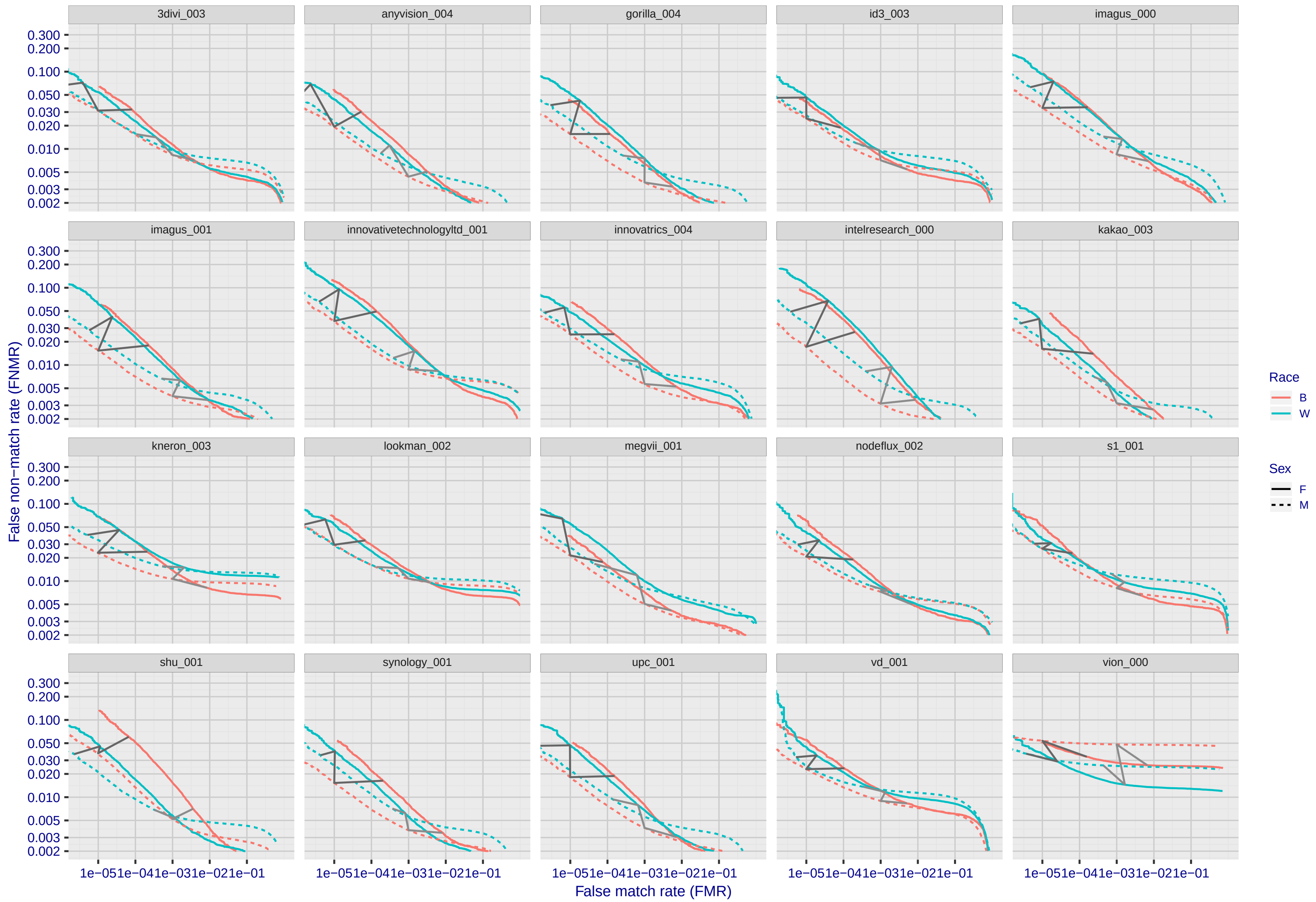


Figure 81: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

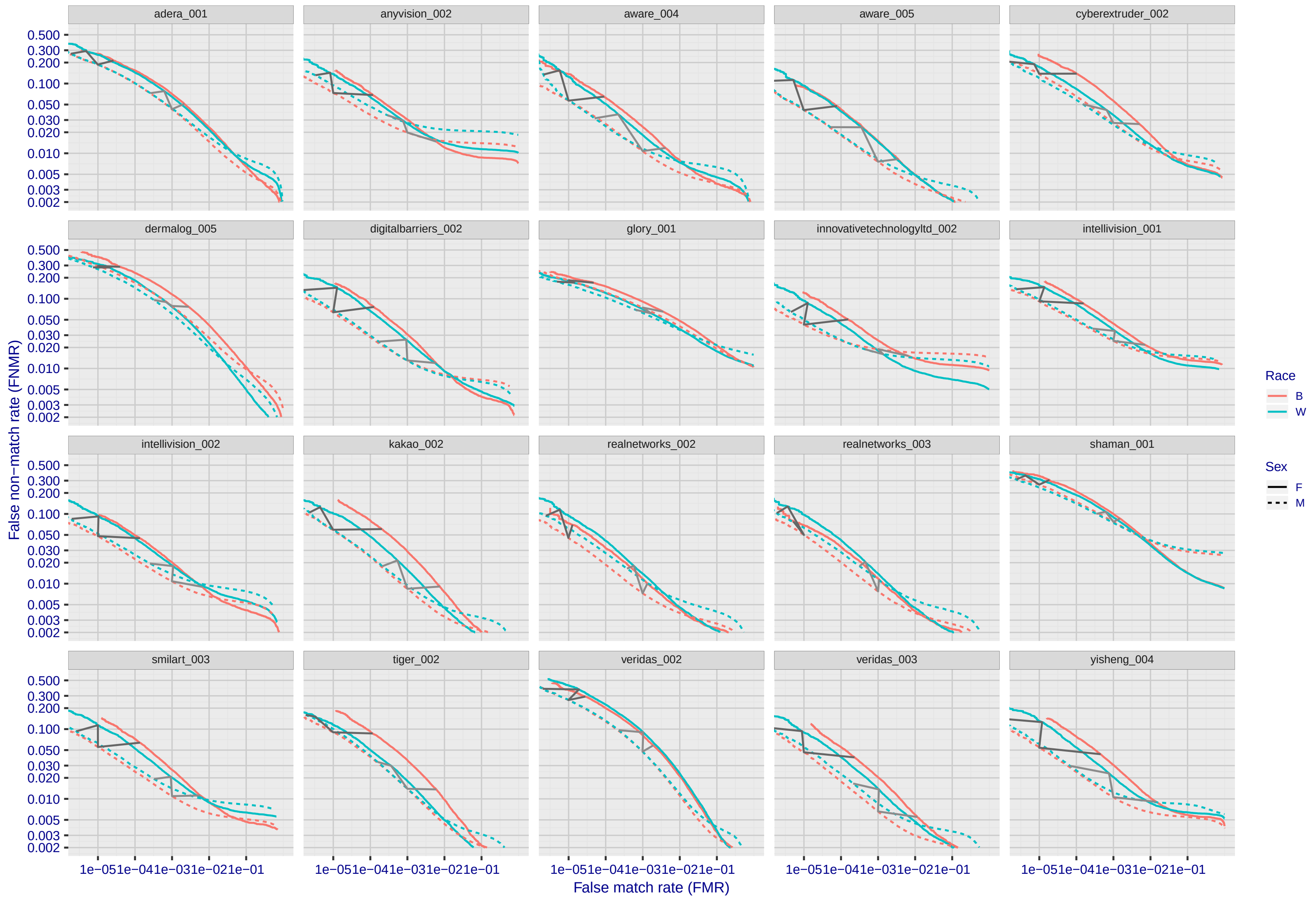


Figure 82: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

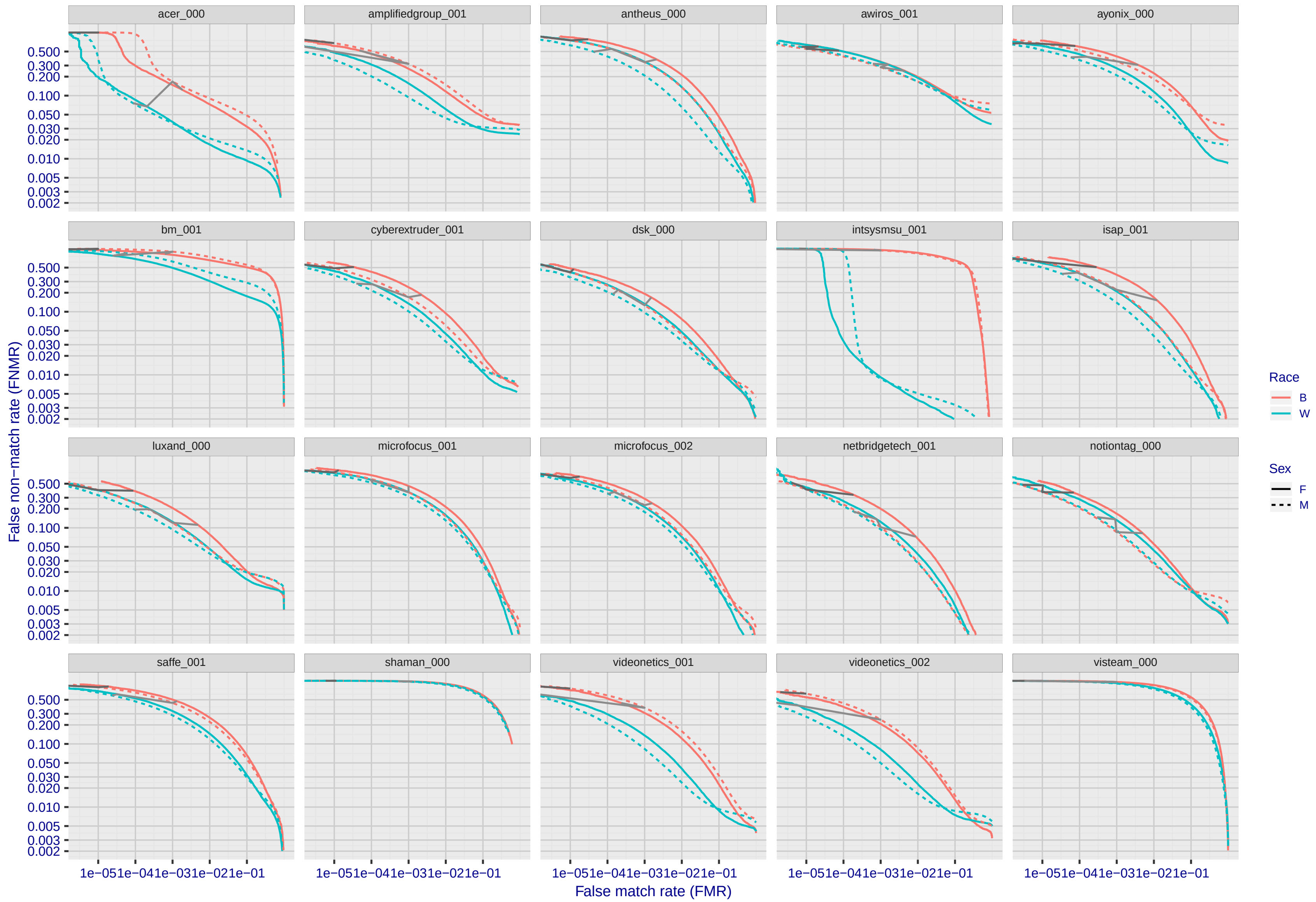


Figure 83: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

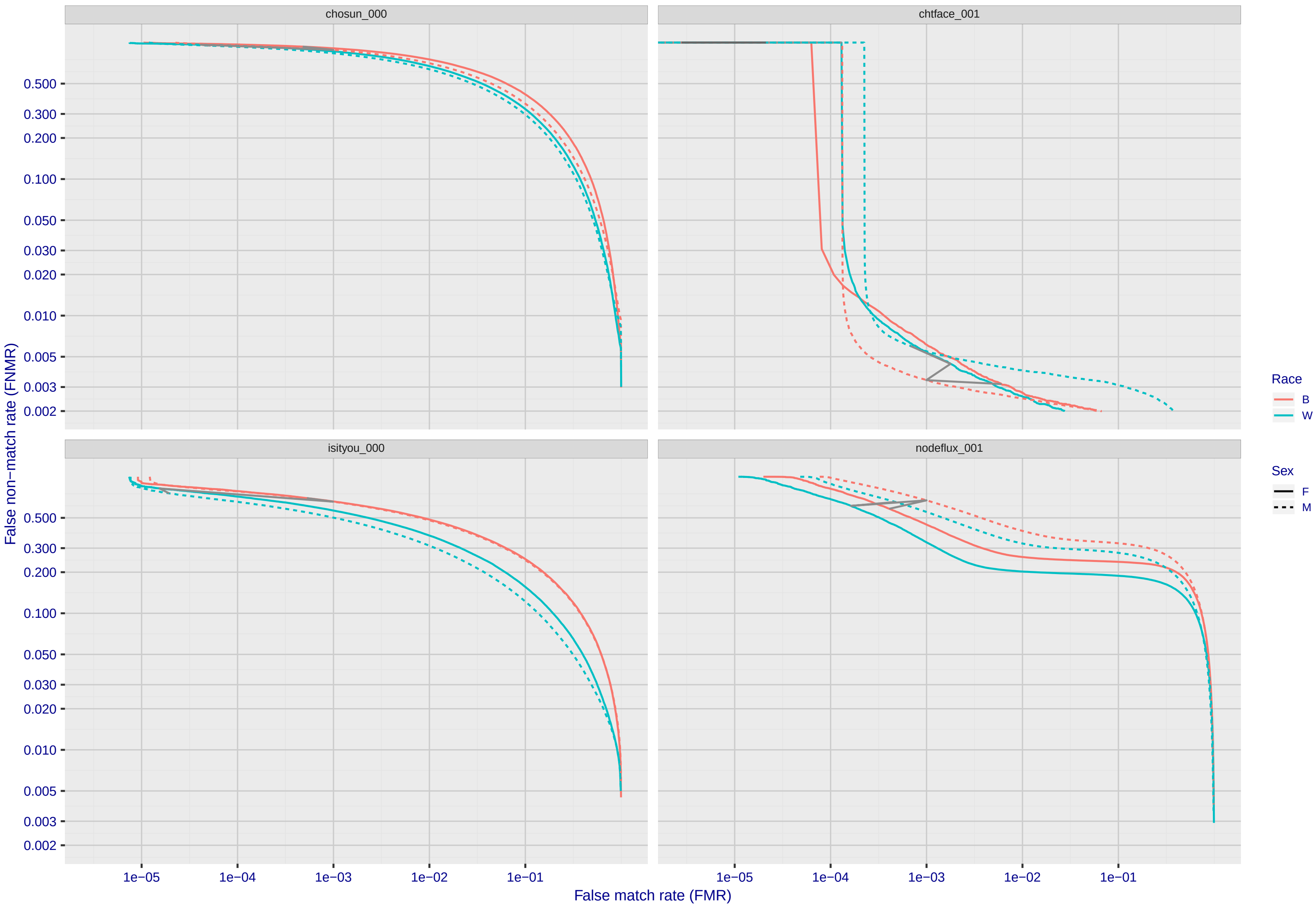


Figure 84: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The Z-shaped grey lines correspond to fixed thresholds, showing both *FNMR and FMR vary at one T value. Note: Many of the plots will naively be read as saying women gives worse error rates than men because the solid traces lie above the dotted ones. However, this is misleading and incomplete: The grey lines show the traces reveal horizontal shifts. Thus for the cogent-003 algorithm FNMR for men is* higher *than for women* at a fixed threshold *but, at the same time, FMR is higher for women - see Figure 113. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

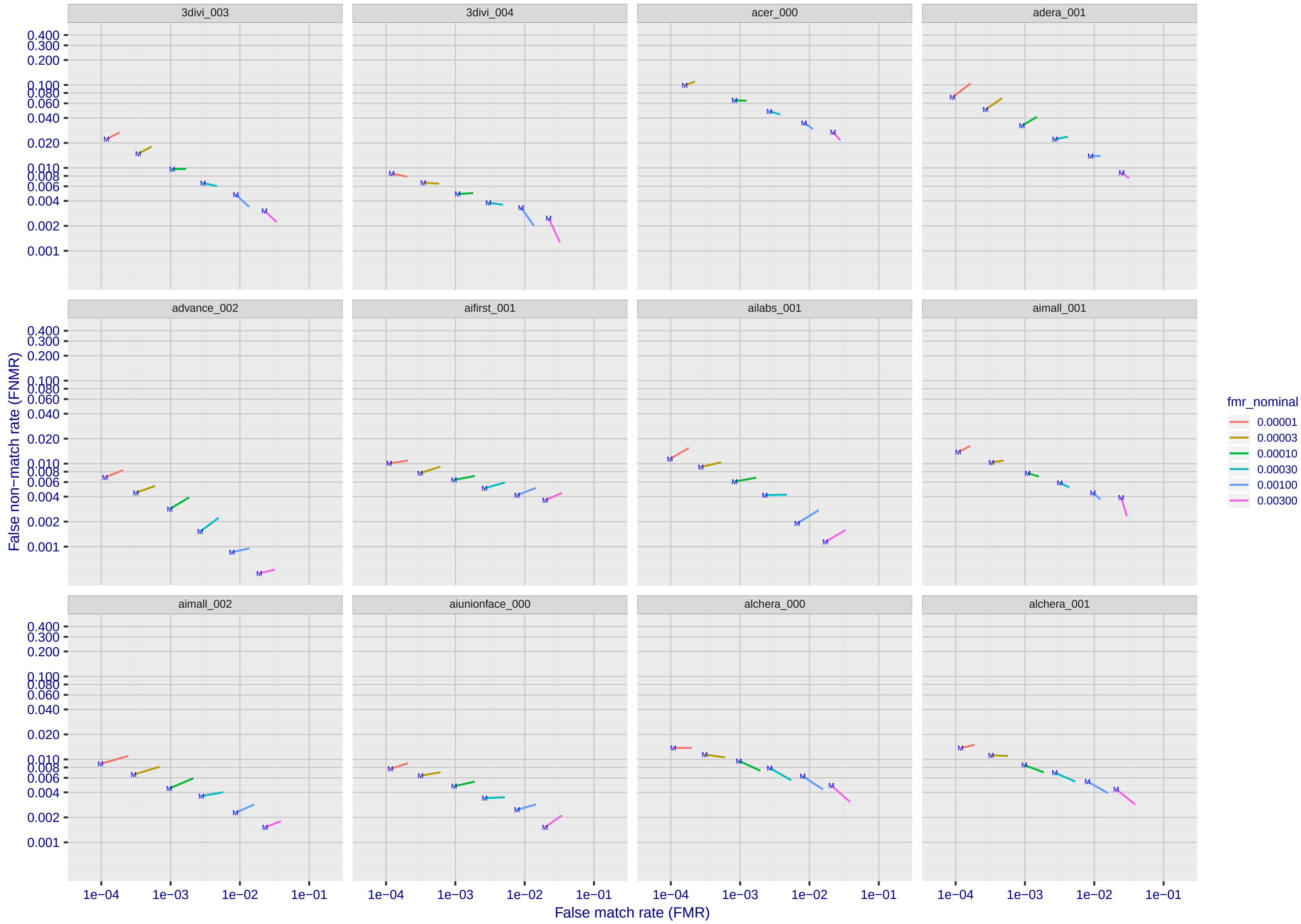


Figure 85: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{MALE}$ and $(FMR,FNMR)_{FEMALE}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

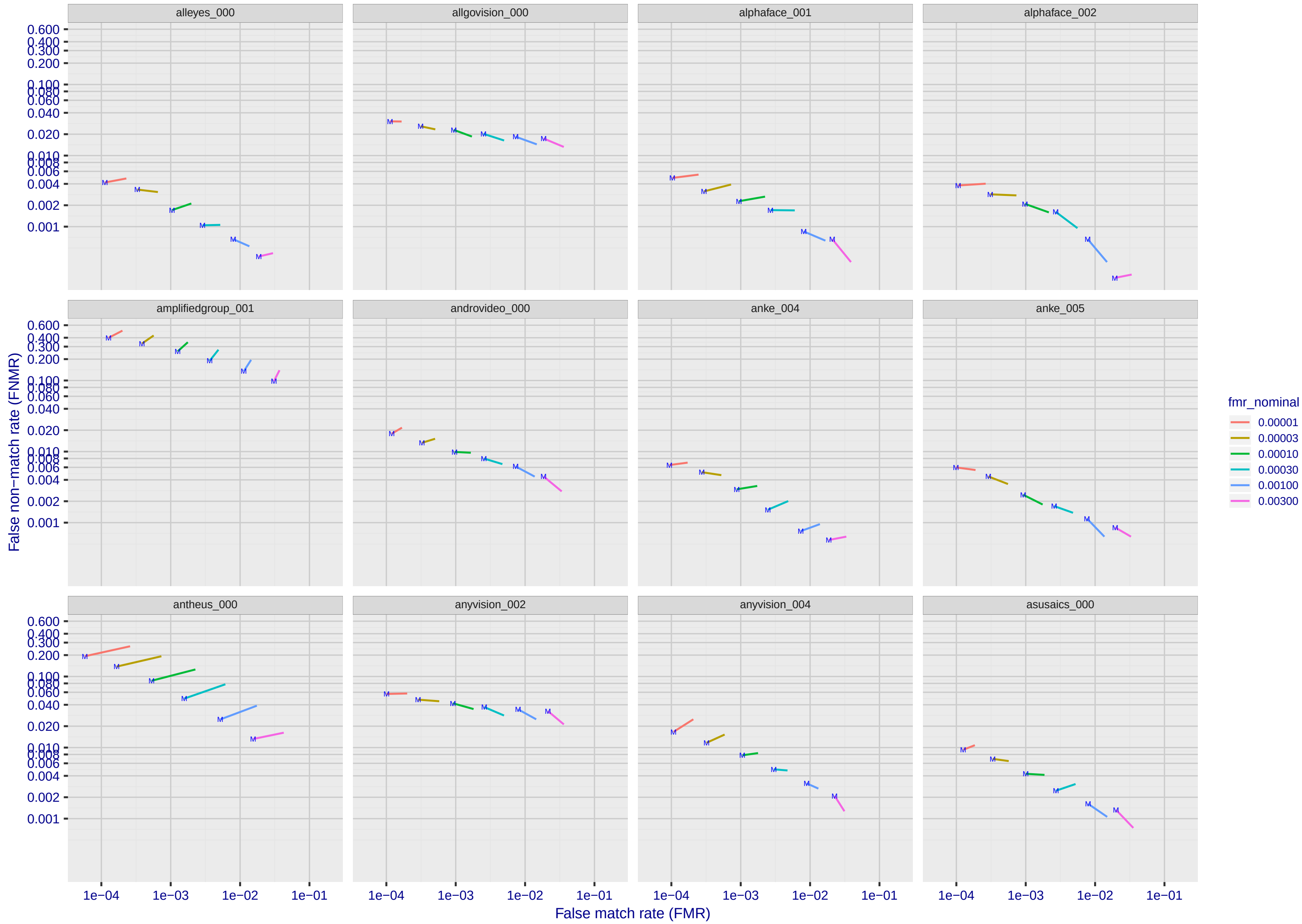


Figure 86: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

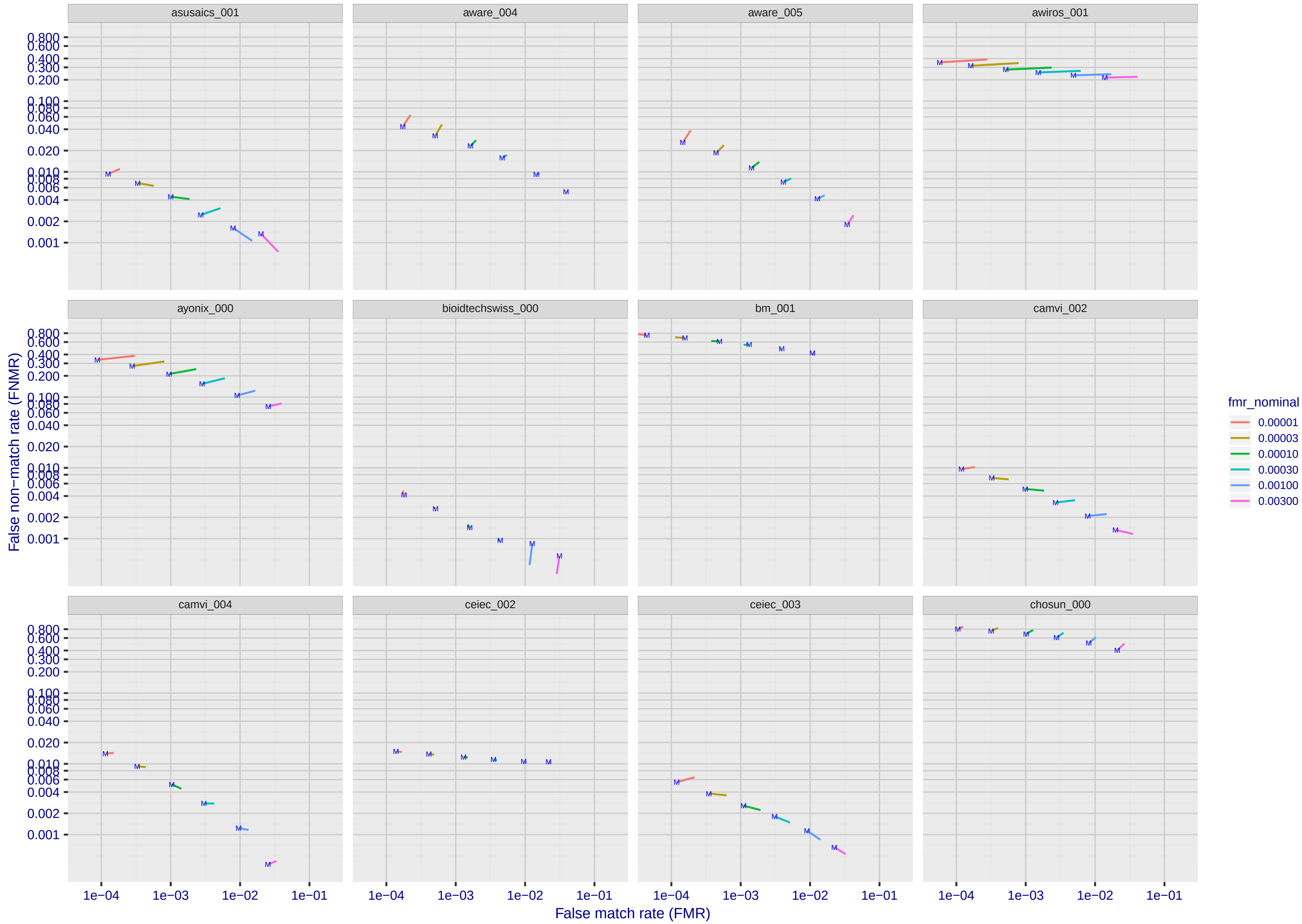


Figure 87: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(\text{FMR,FNMR})_{\text{MALE}}$ and $(\text{FMR,FNMR})_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

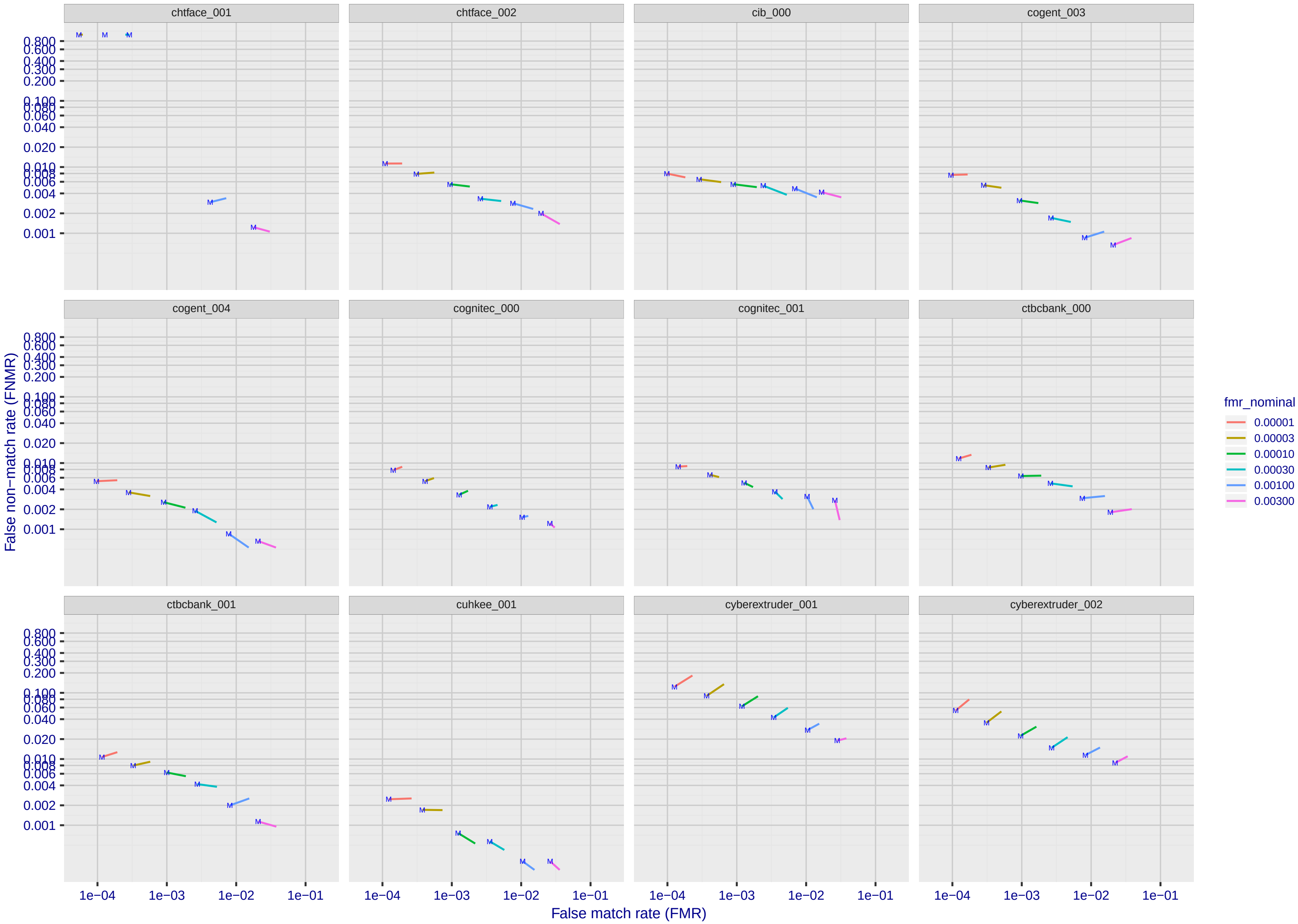


Figure 88: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\text{MALE}}$ and $(FMR,FNMR)_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

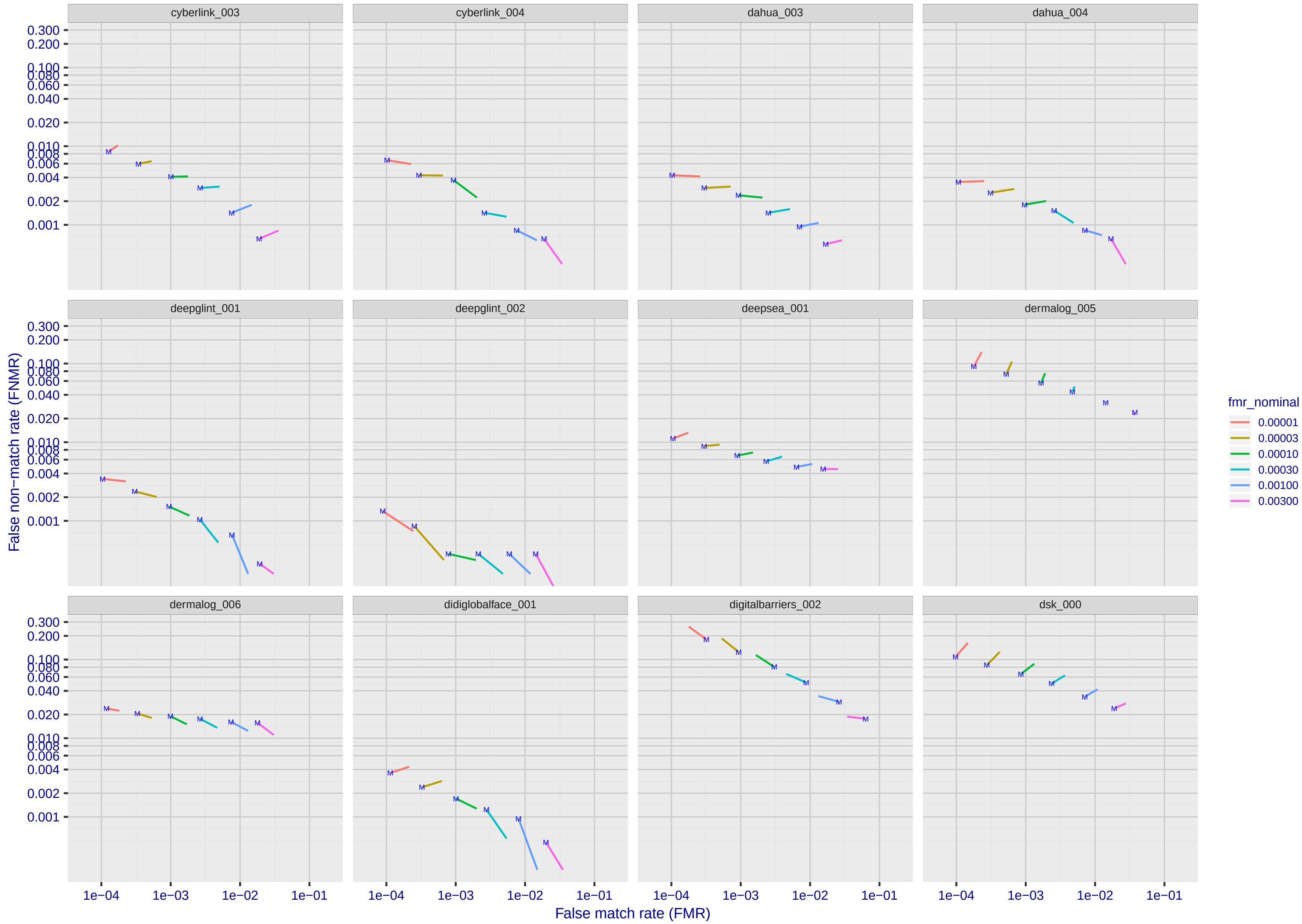


Figure 89: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{MALE}$ and $(FMR,FNMR)_{FEMALE}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

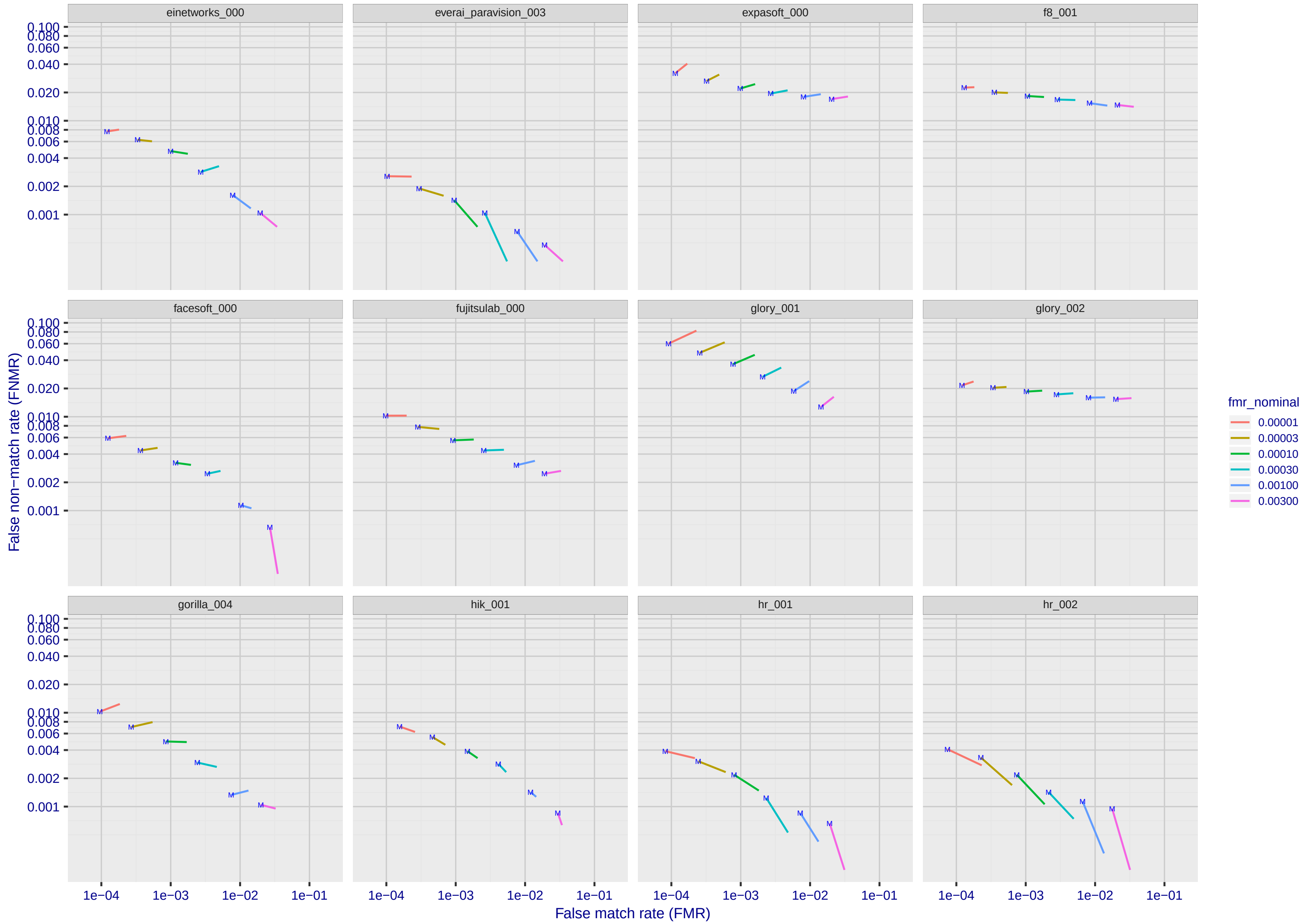


Figure 90: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\text{MALE}}$ and $(FMR,FNMR)_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

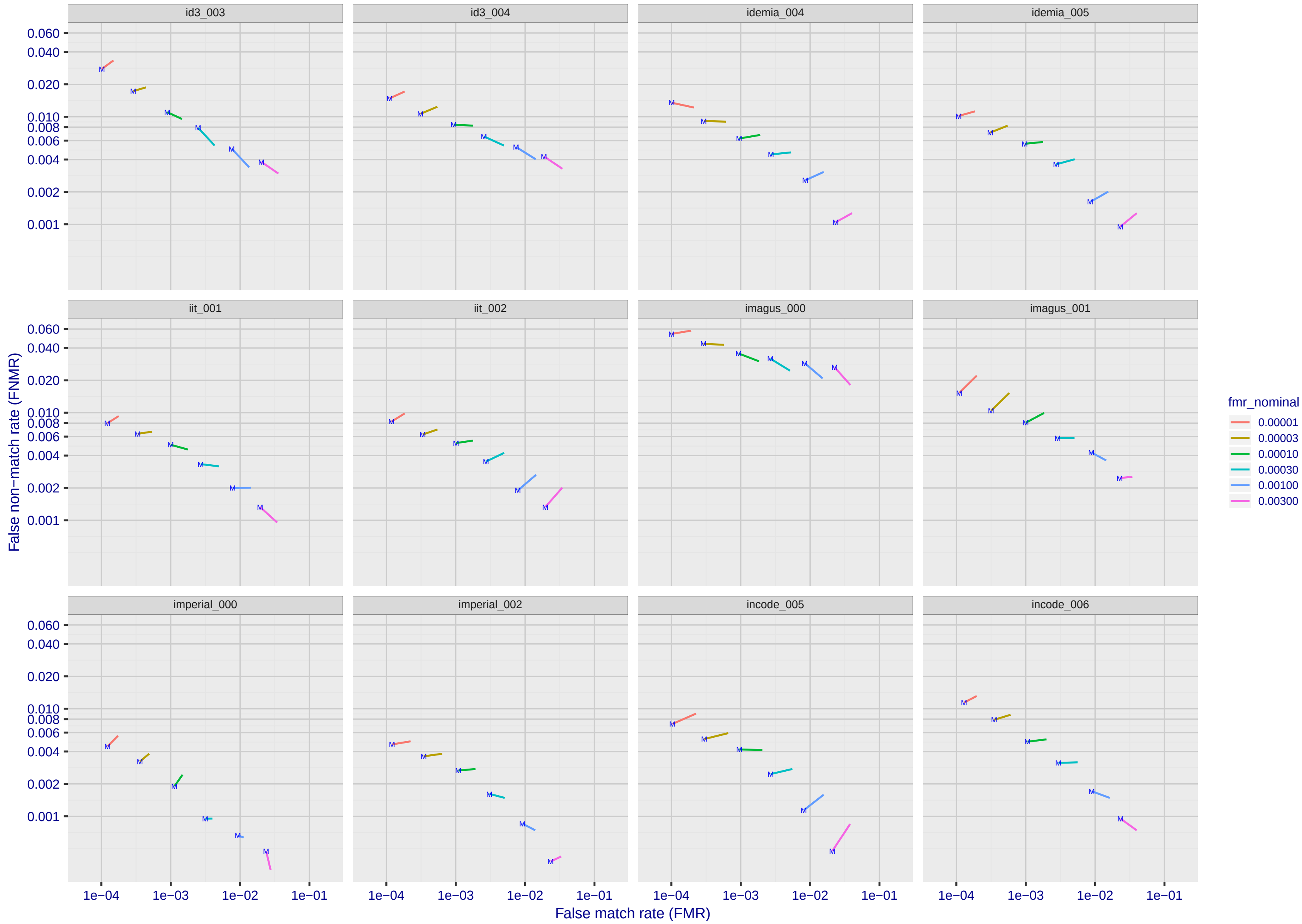


Figure 91: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

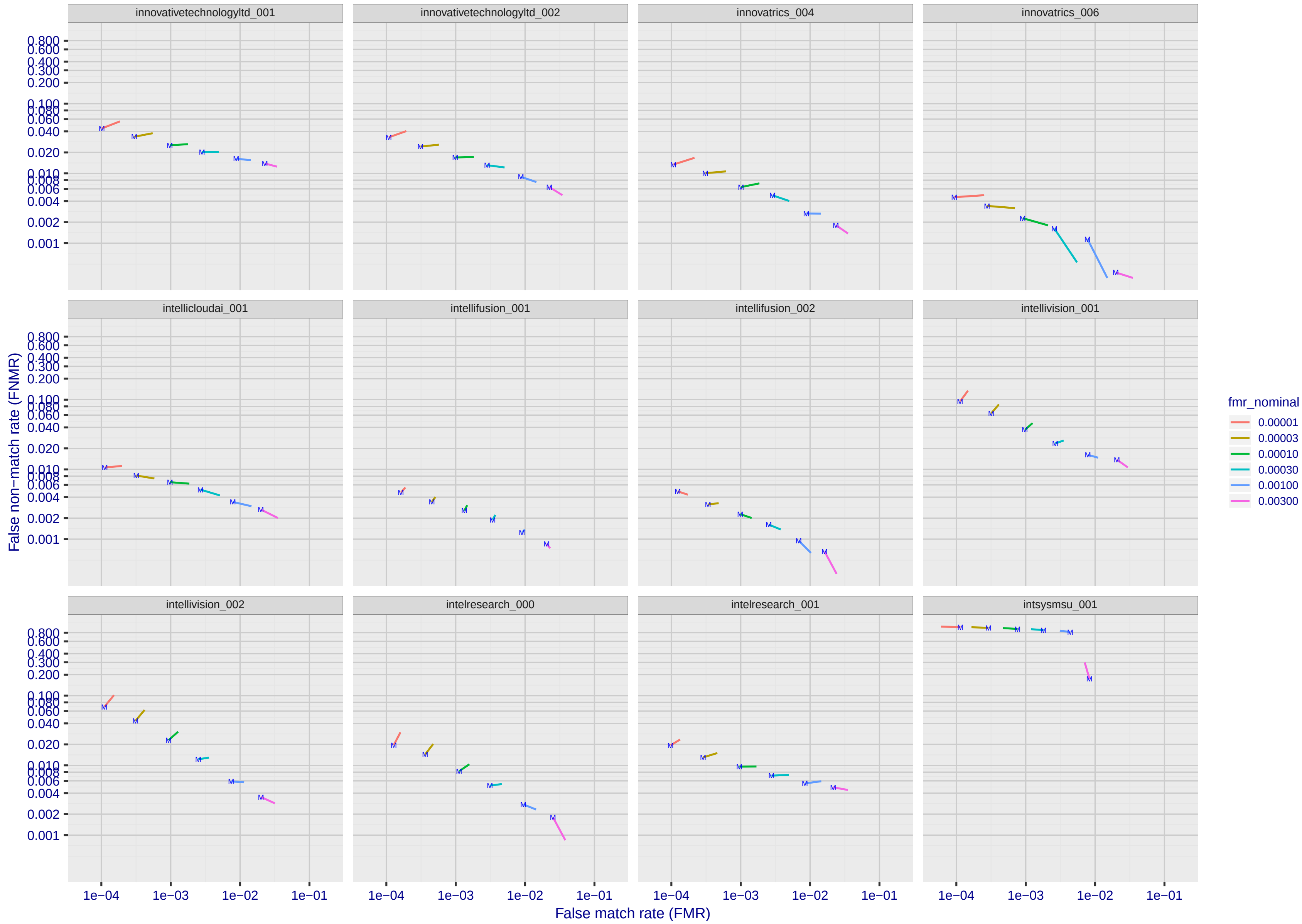


Figure 92: *For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between* $(FMR,FNMR)_{\mathrm{MALE}}$ *and* $(FMR,FNMR)_{\mathrm{FEMALE}}$ *showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).*

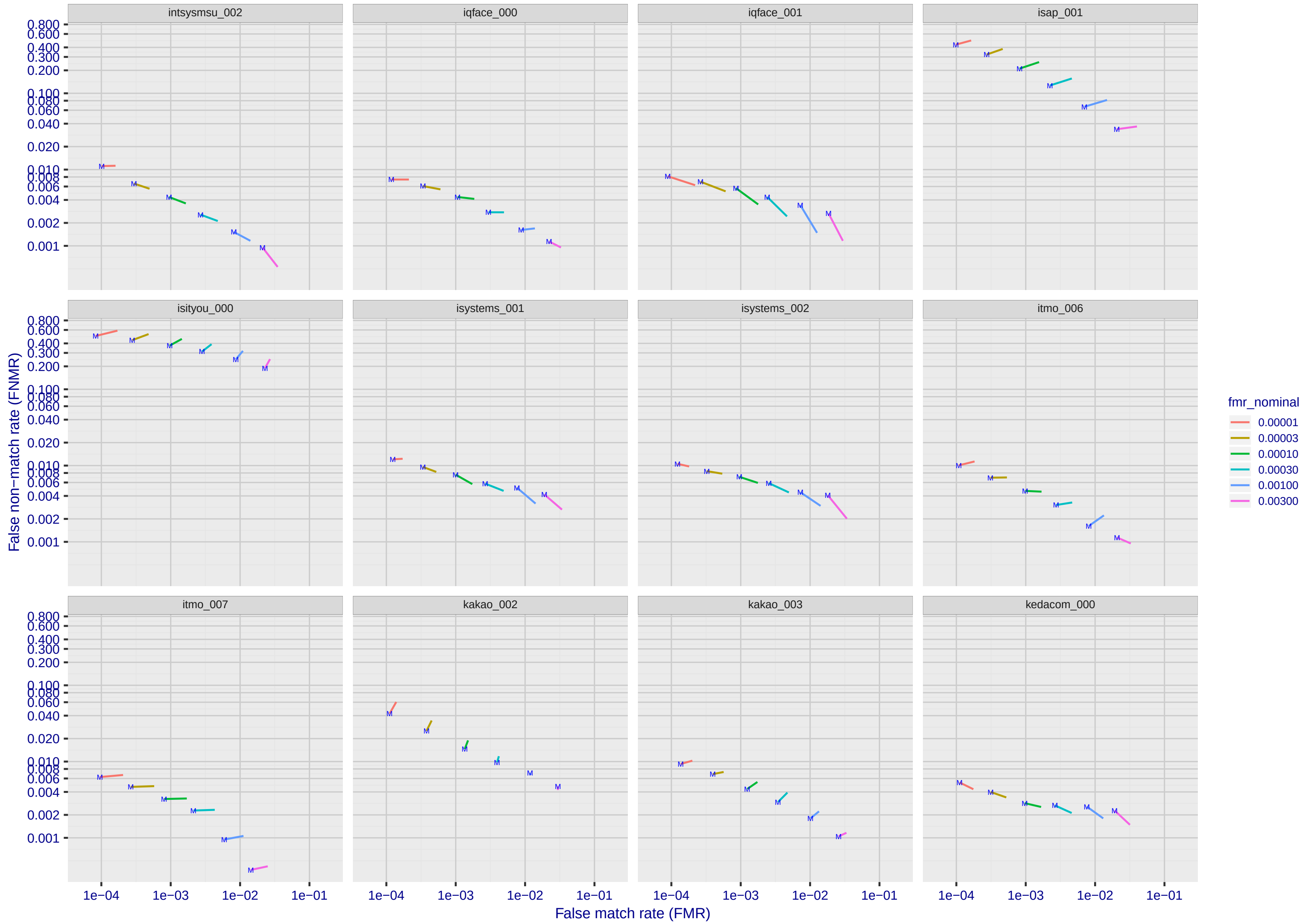


Figure 93: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

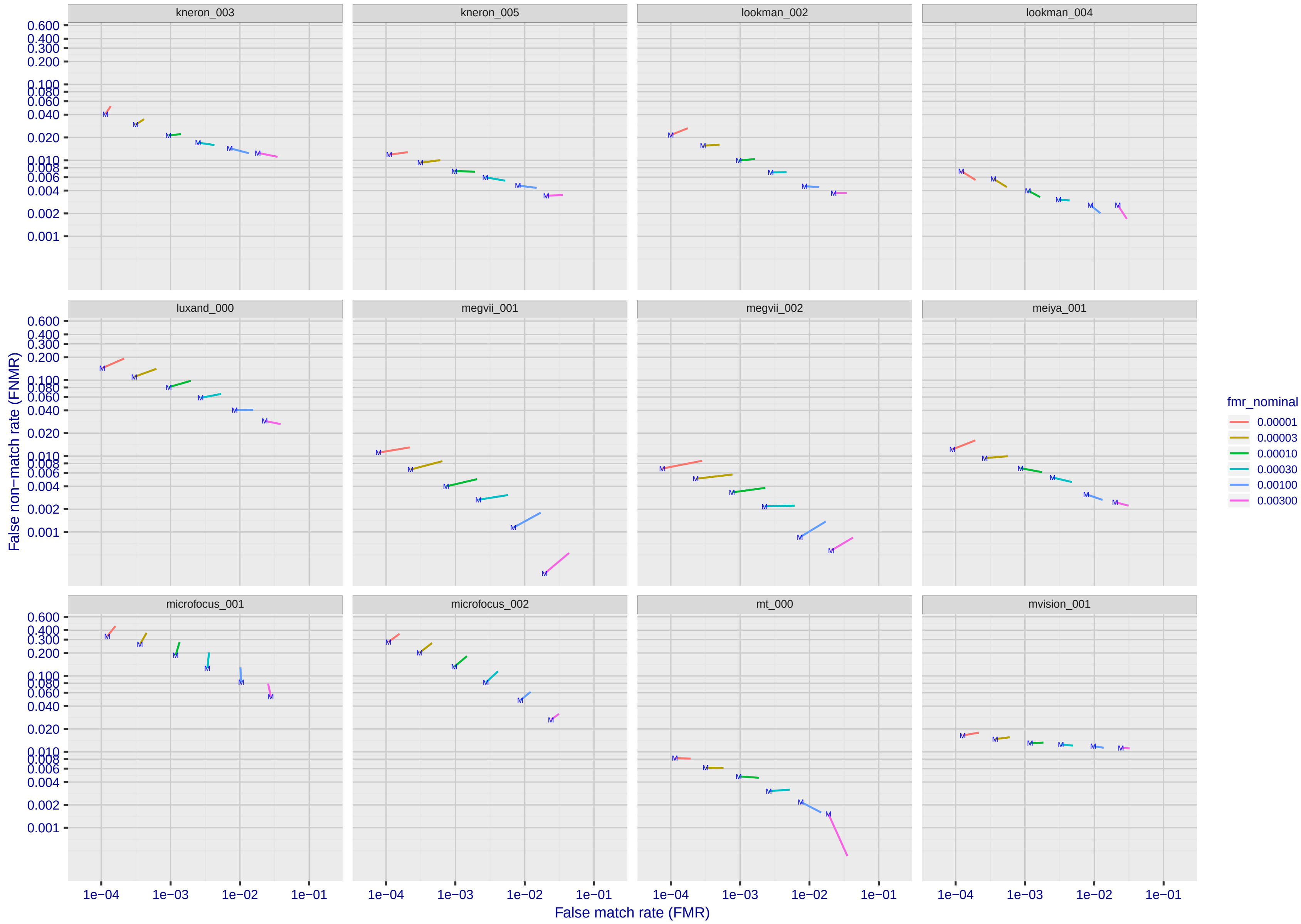


Figure 94: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\mathrm{MALE}}$ and $(FMR,FNMR)_{\mathrm{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

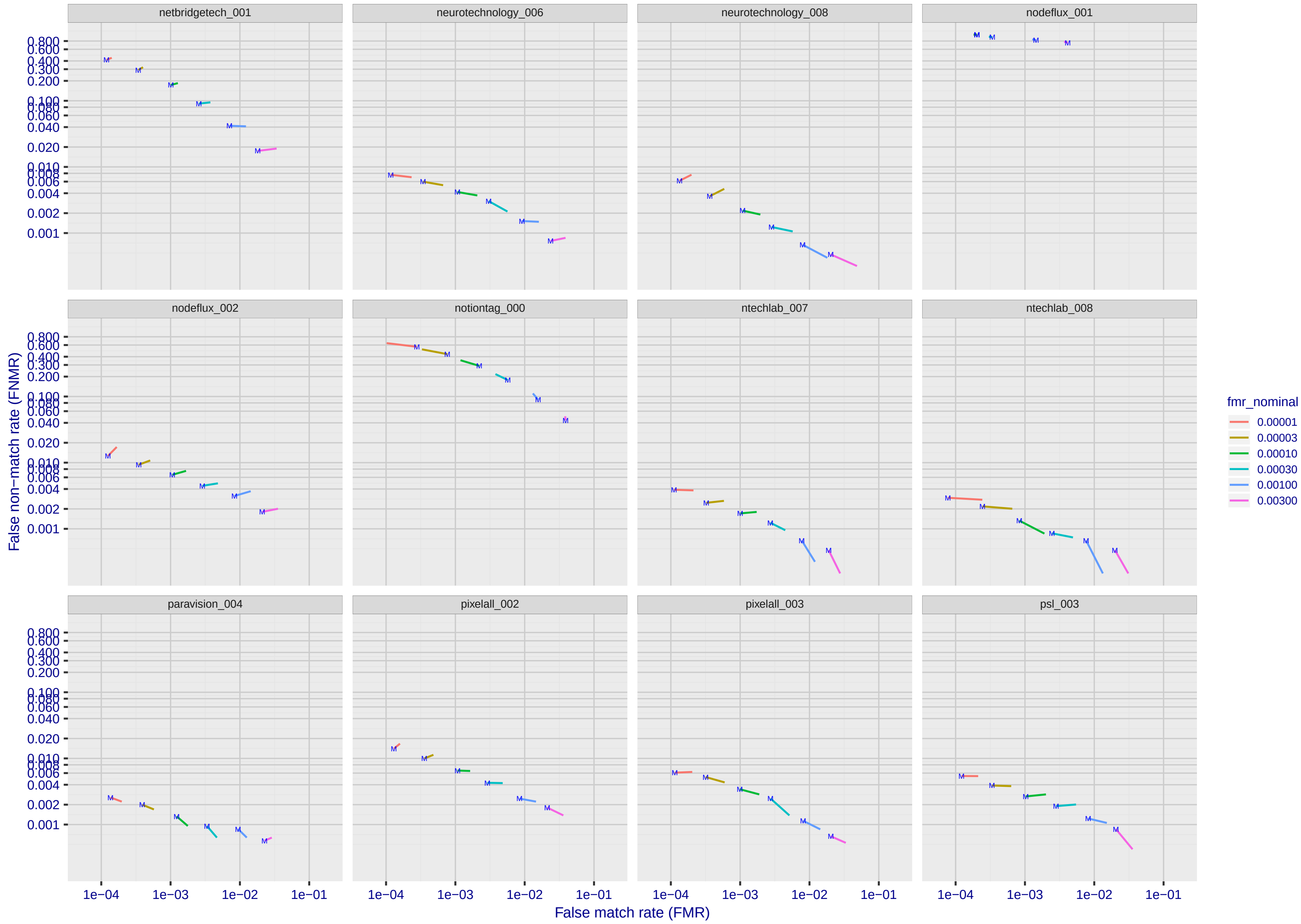


Figure 95: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

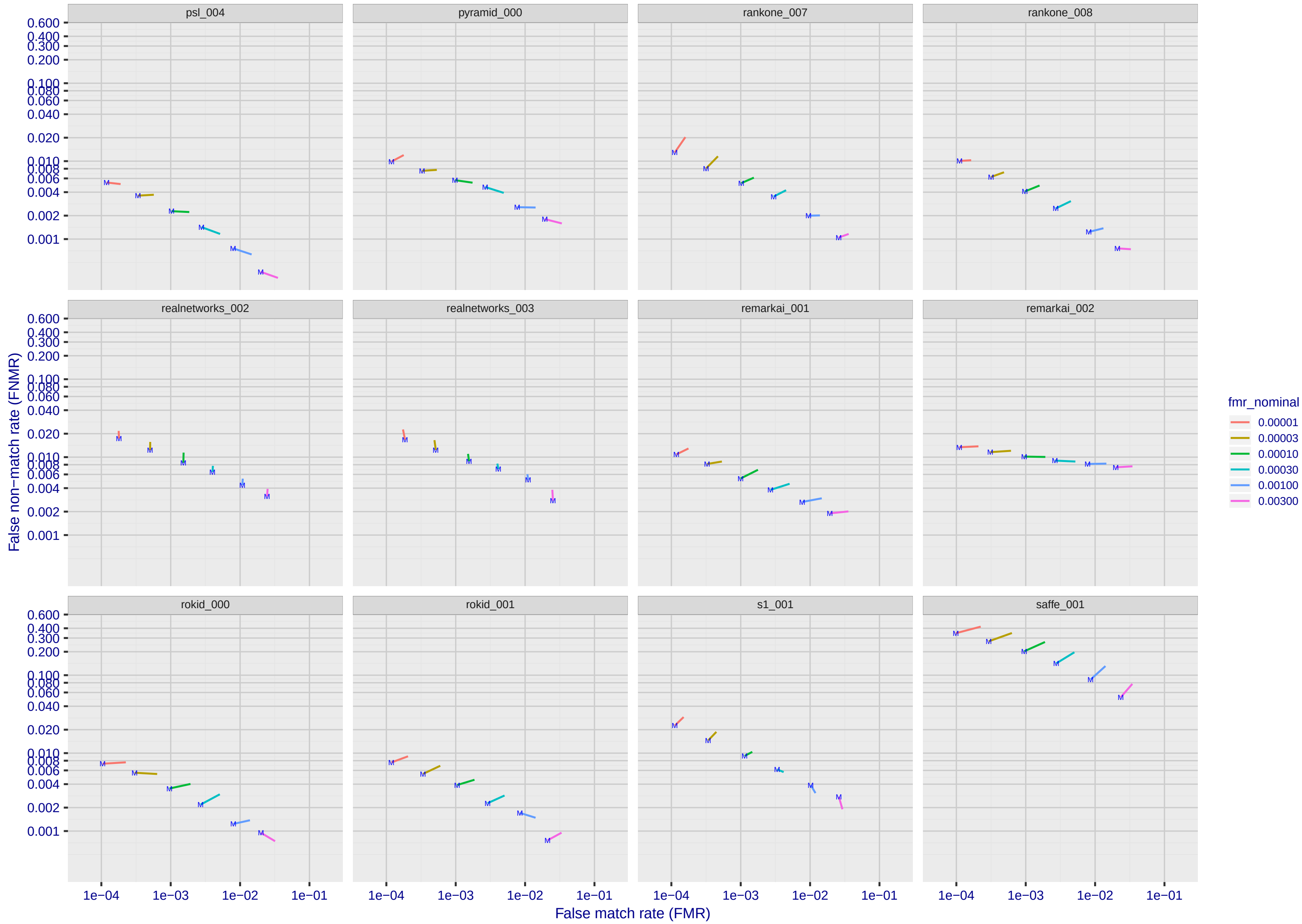


Figure 96: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\mathrm{MALE}}$ and (FMR,FNMR)$_{\mathrm{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

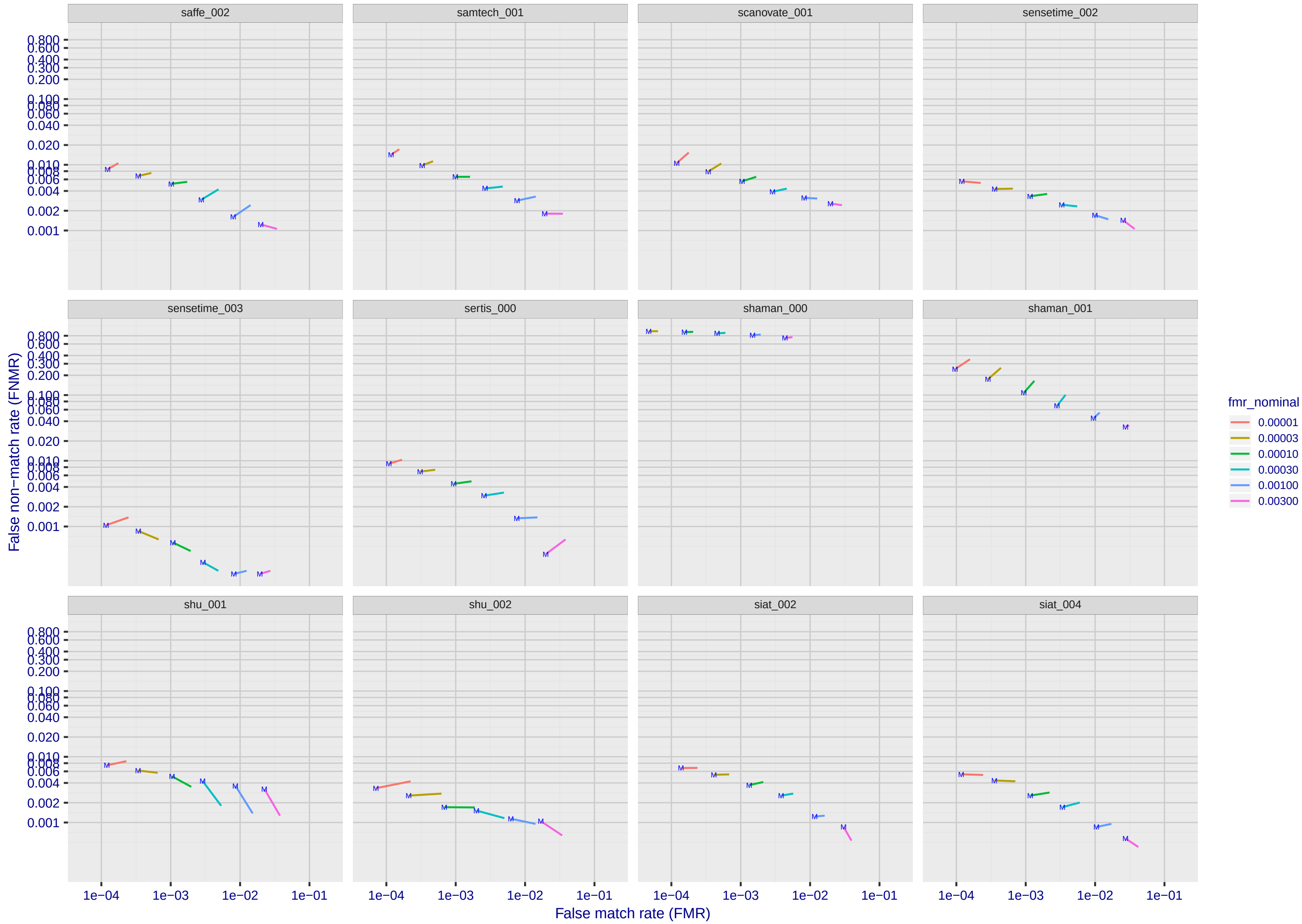


Figure 97: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{MALE}$ and $(FMR,FNMR)_{FEMALE}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

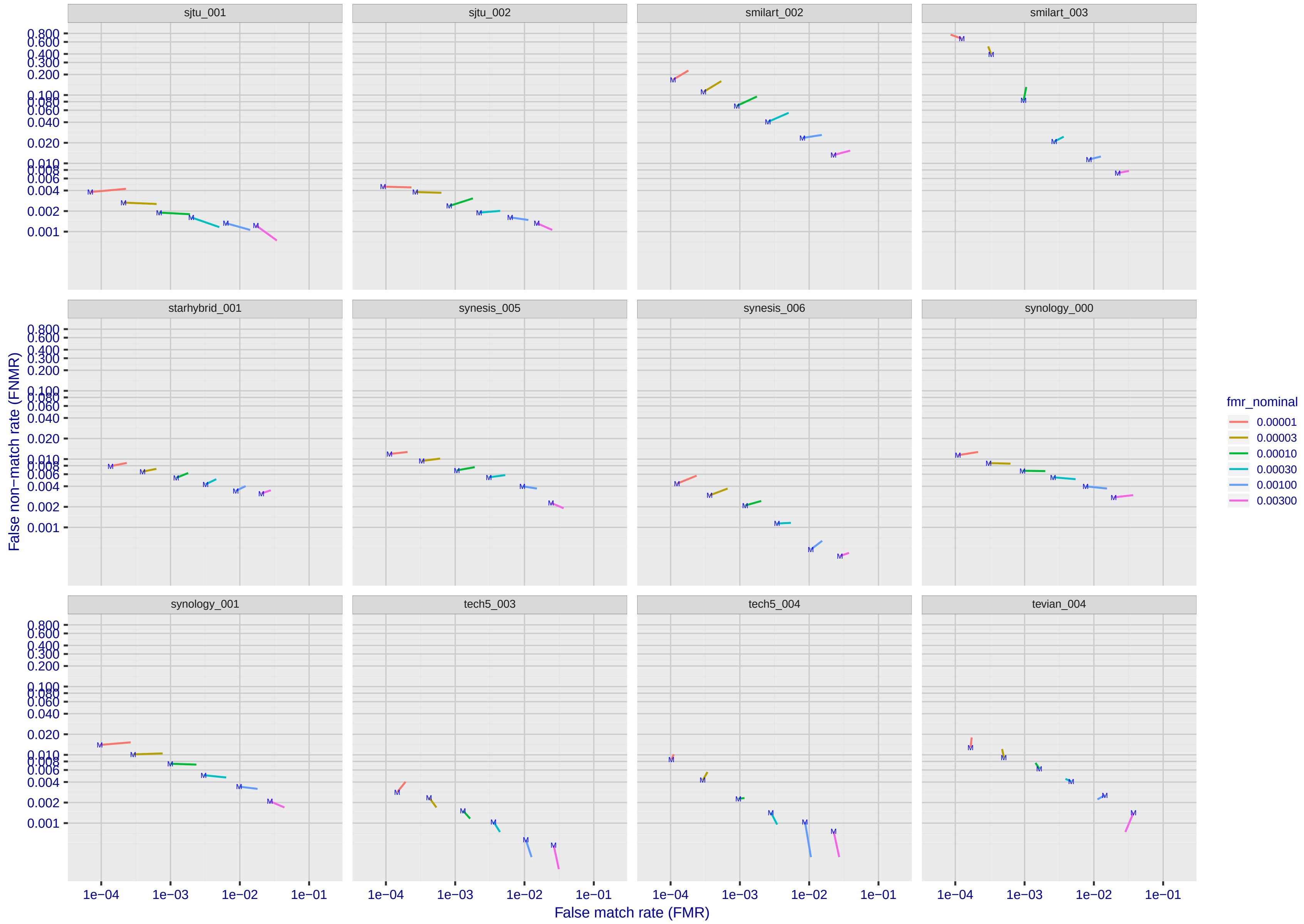


Figure 98: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\text{MALE}}$ *and* $(FMR,FNMR)_{\text{FEMALE}}$ *showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

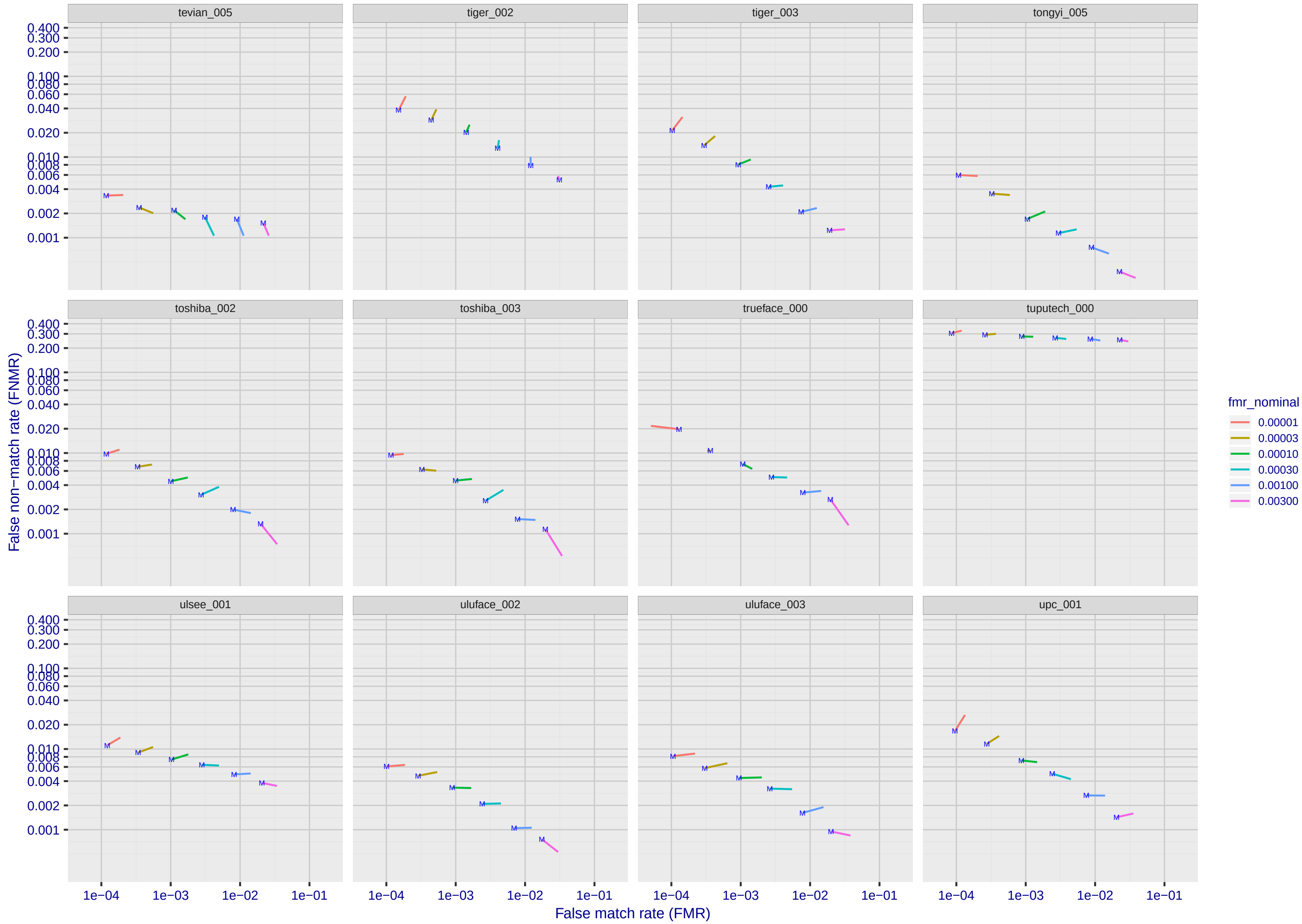


Figure 99: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

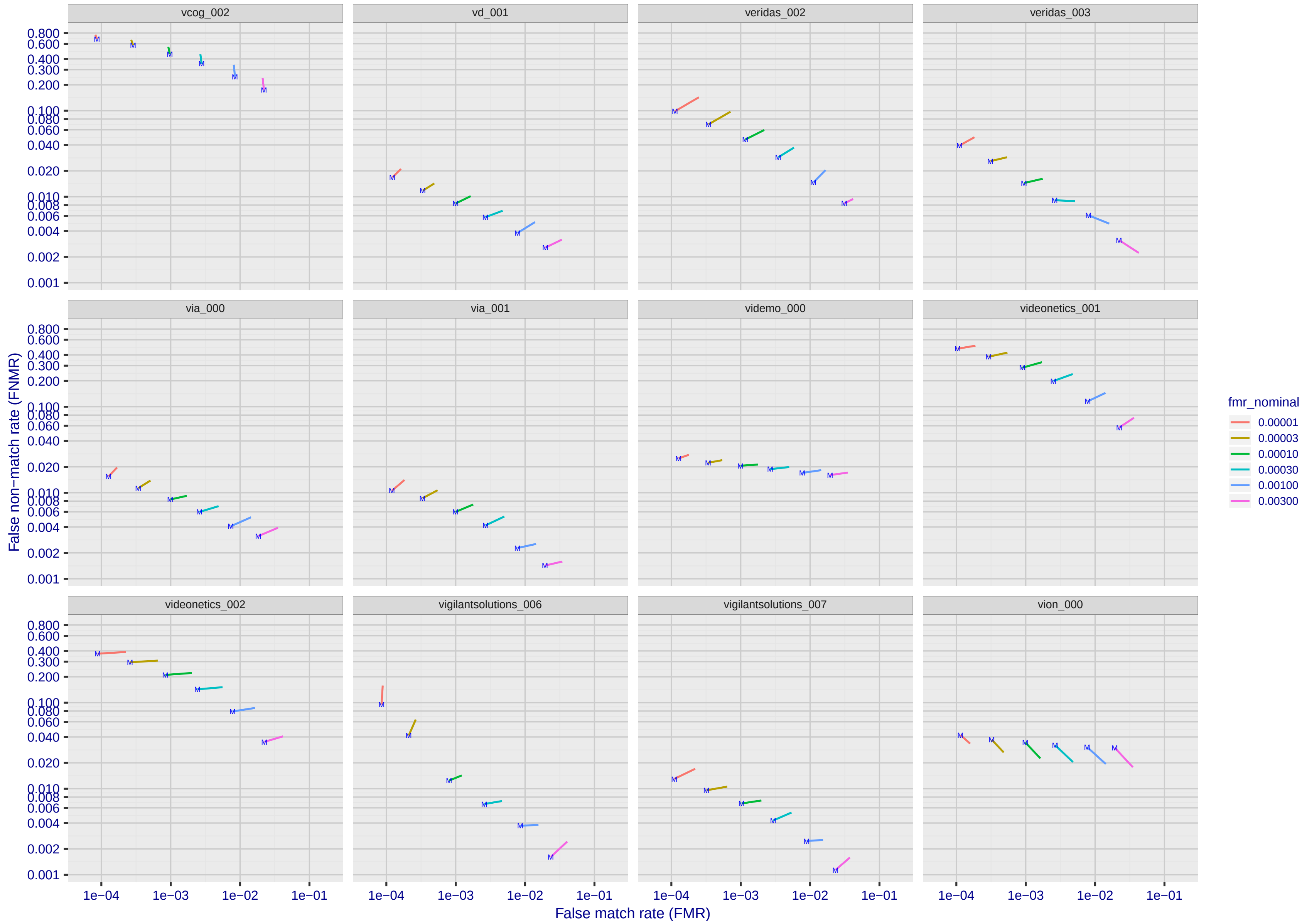


Figure 100: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\mathrm{MALE}}$ and $(FMR,FNMR)_{\mathrm{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

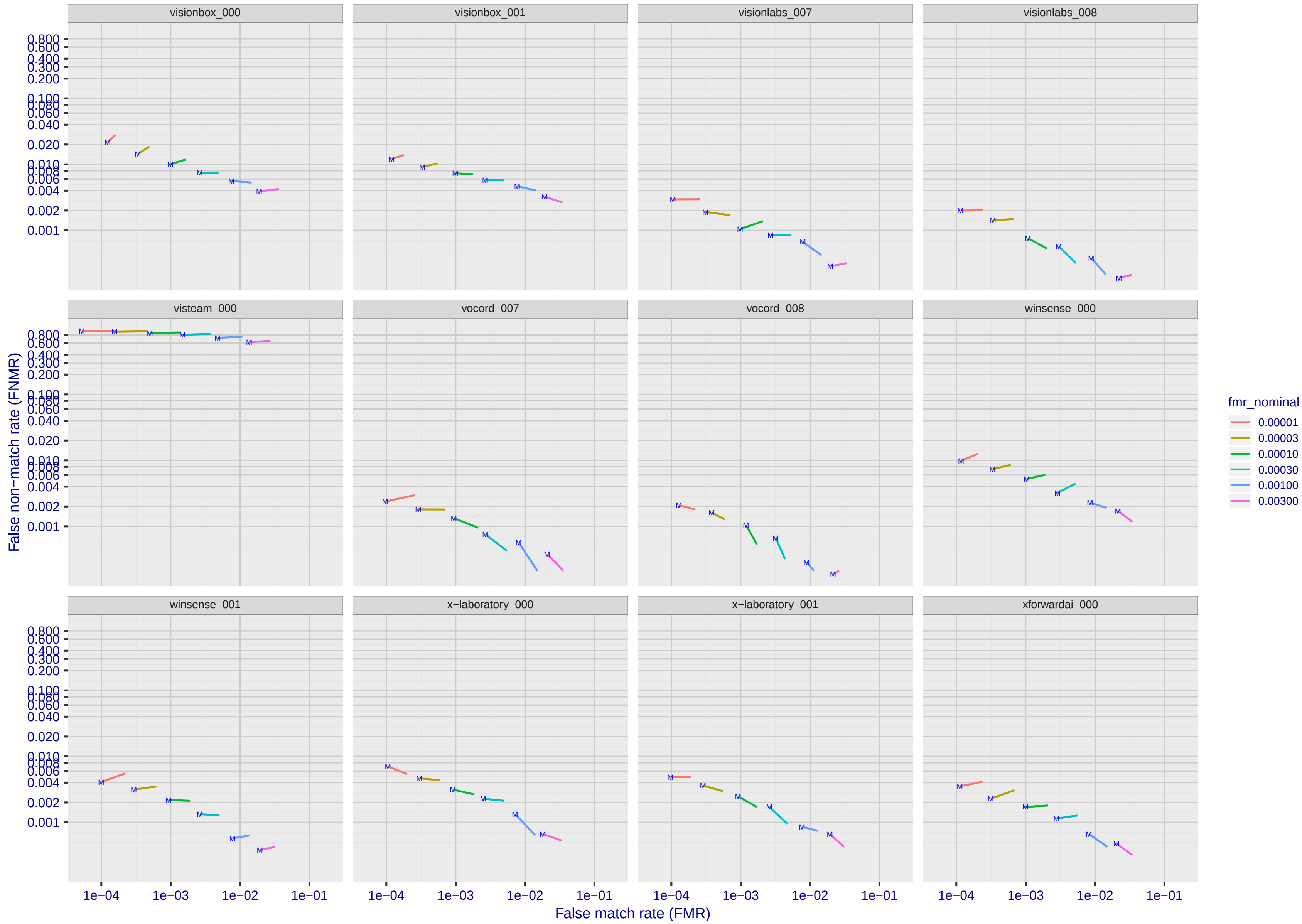


Figure 101: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

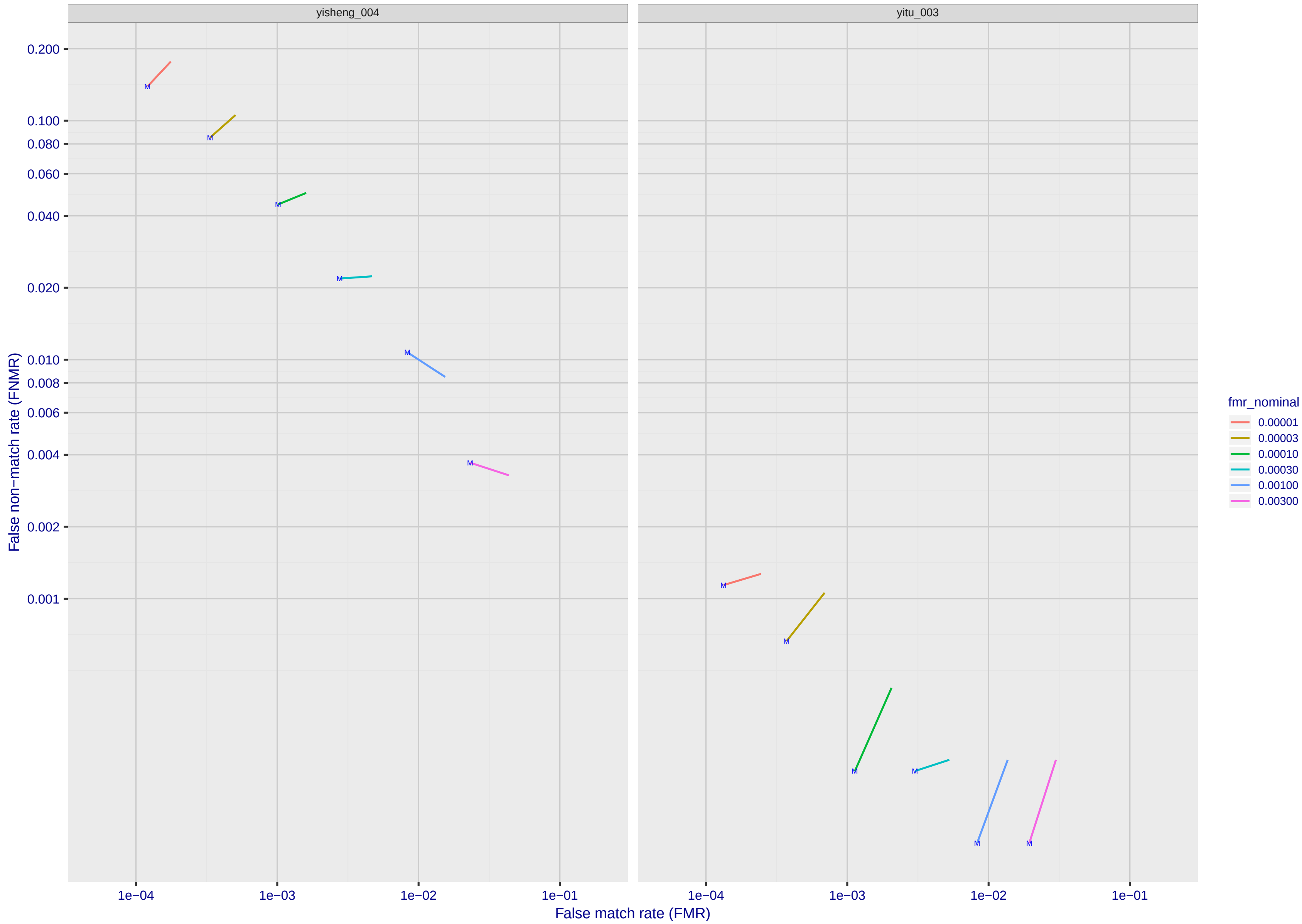


Figure 102: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(FMR,FNMR)_{\mathrm{MALE}}$ and $(FMR,FNMR)_{\mathrm{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 3.6.1), and the same age group (see section 3.6.2).

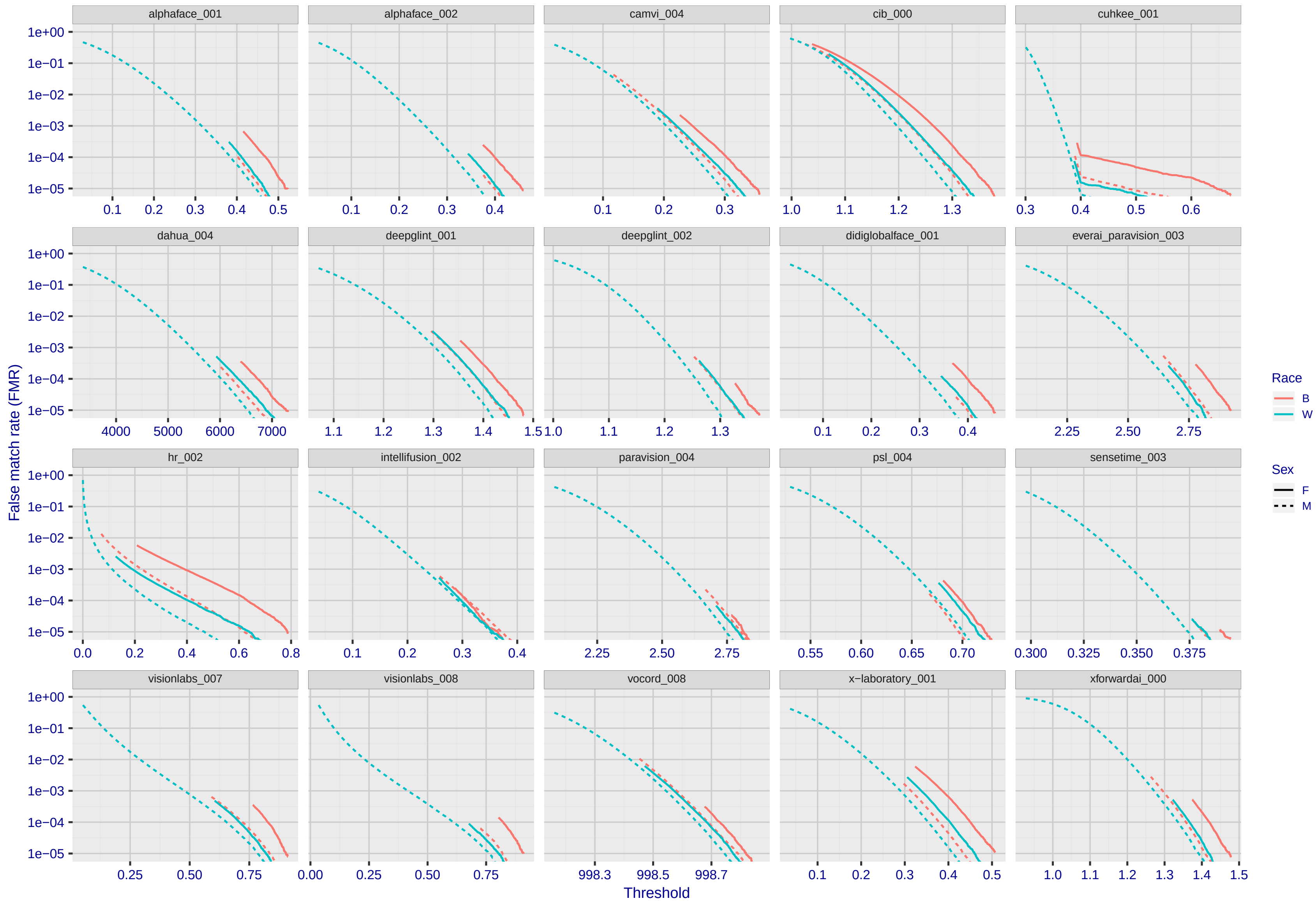


Figure 103: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

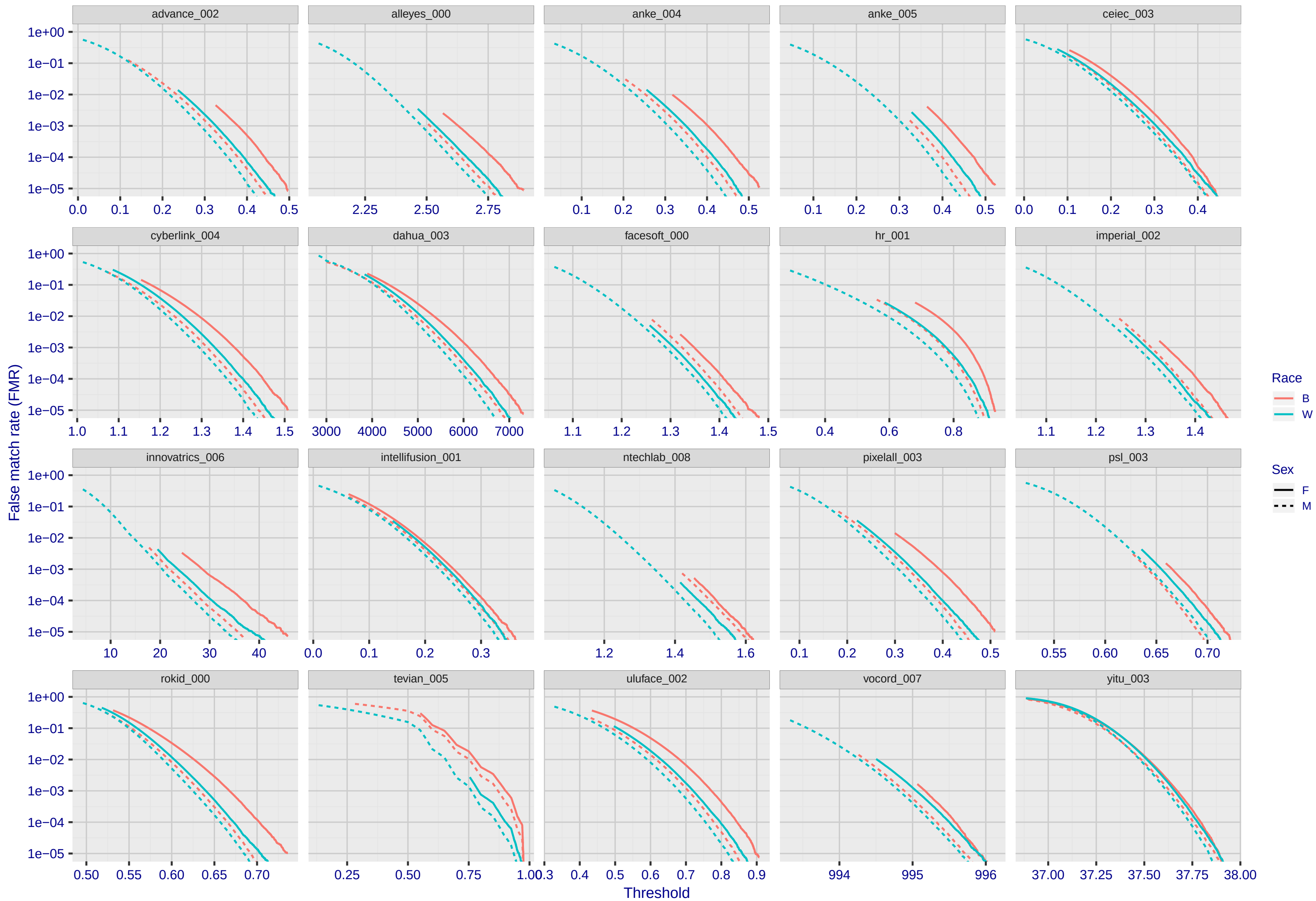


Figure 104: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

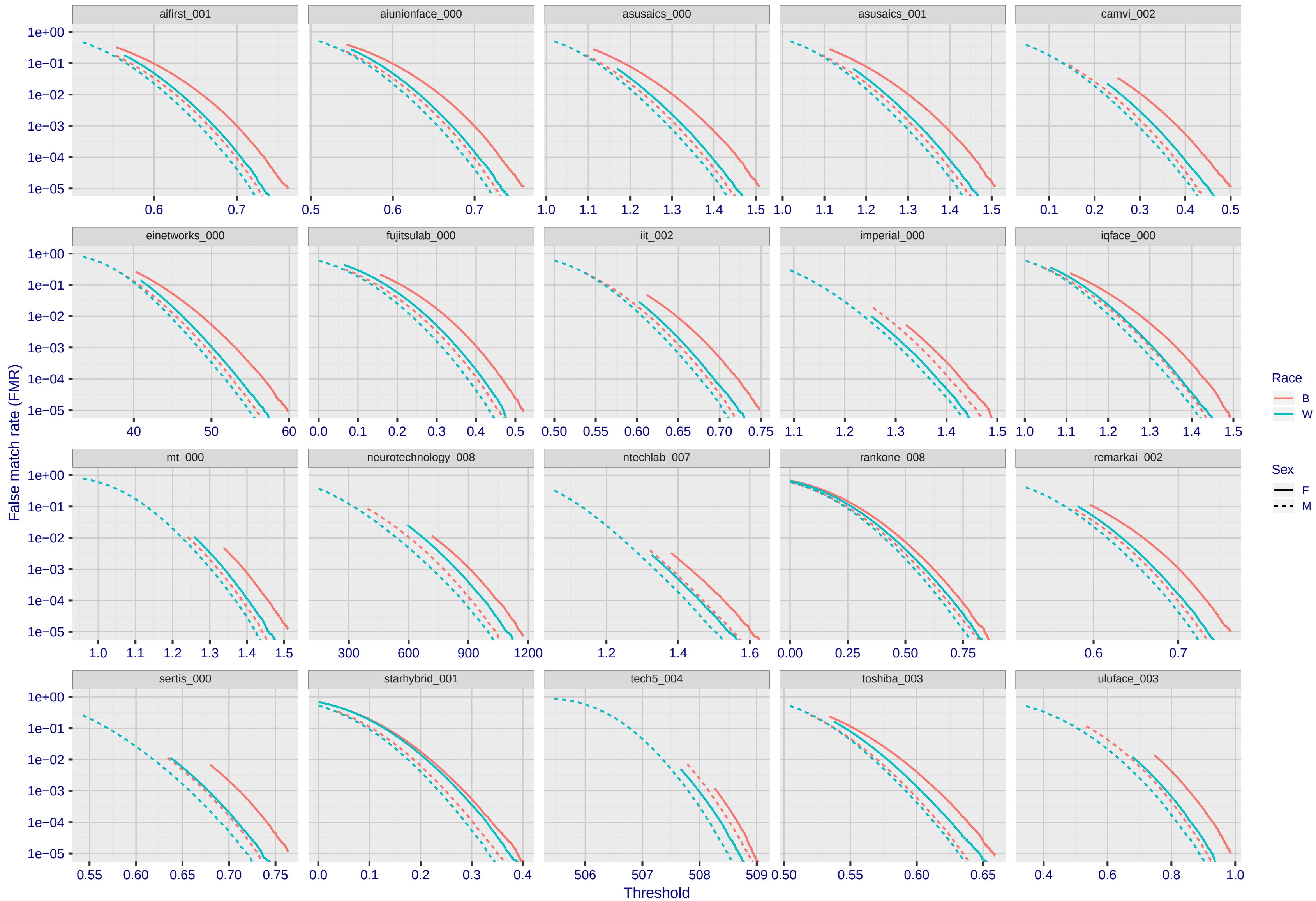


Figure 105: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

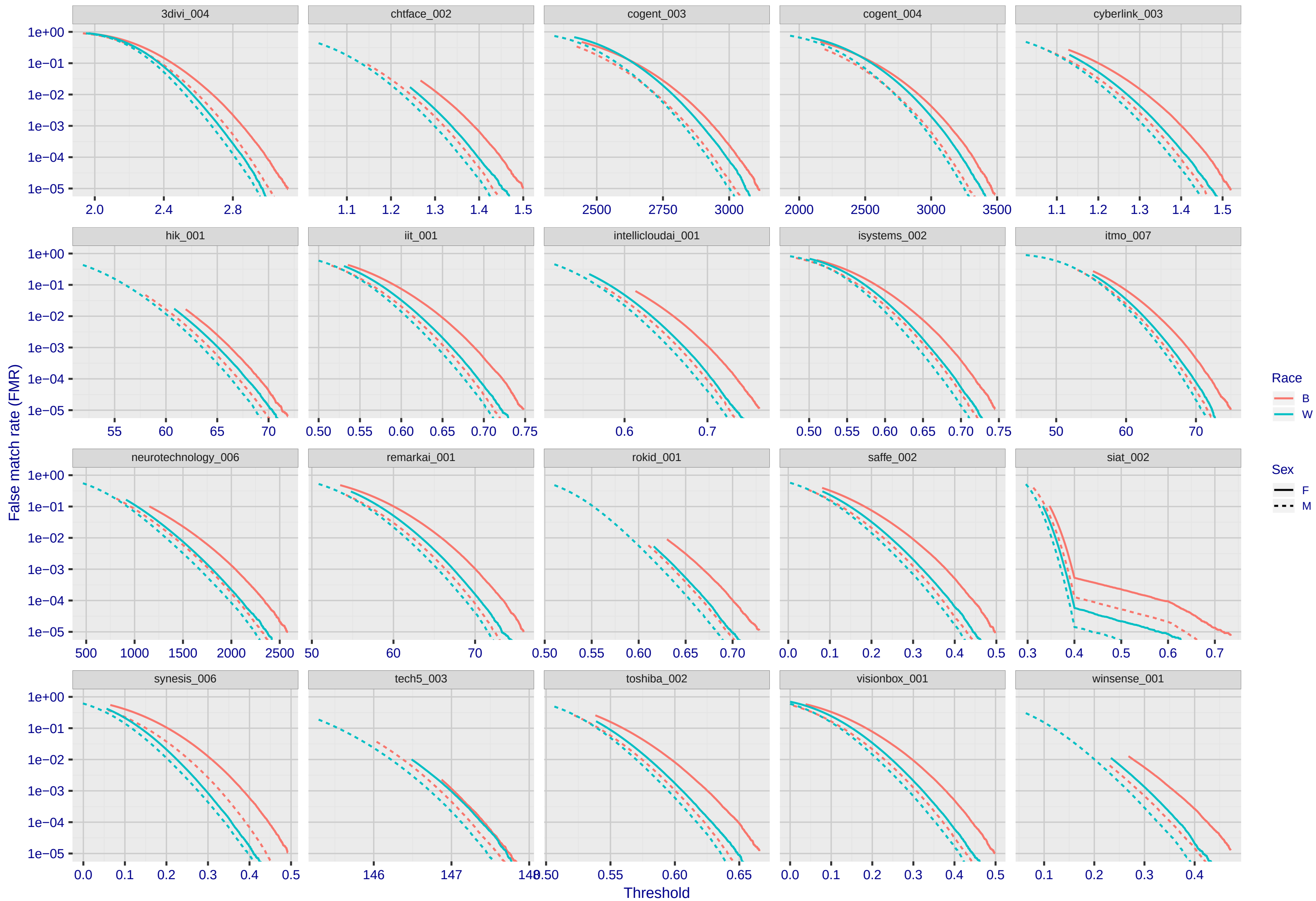


Figure 106: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

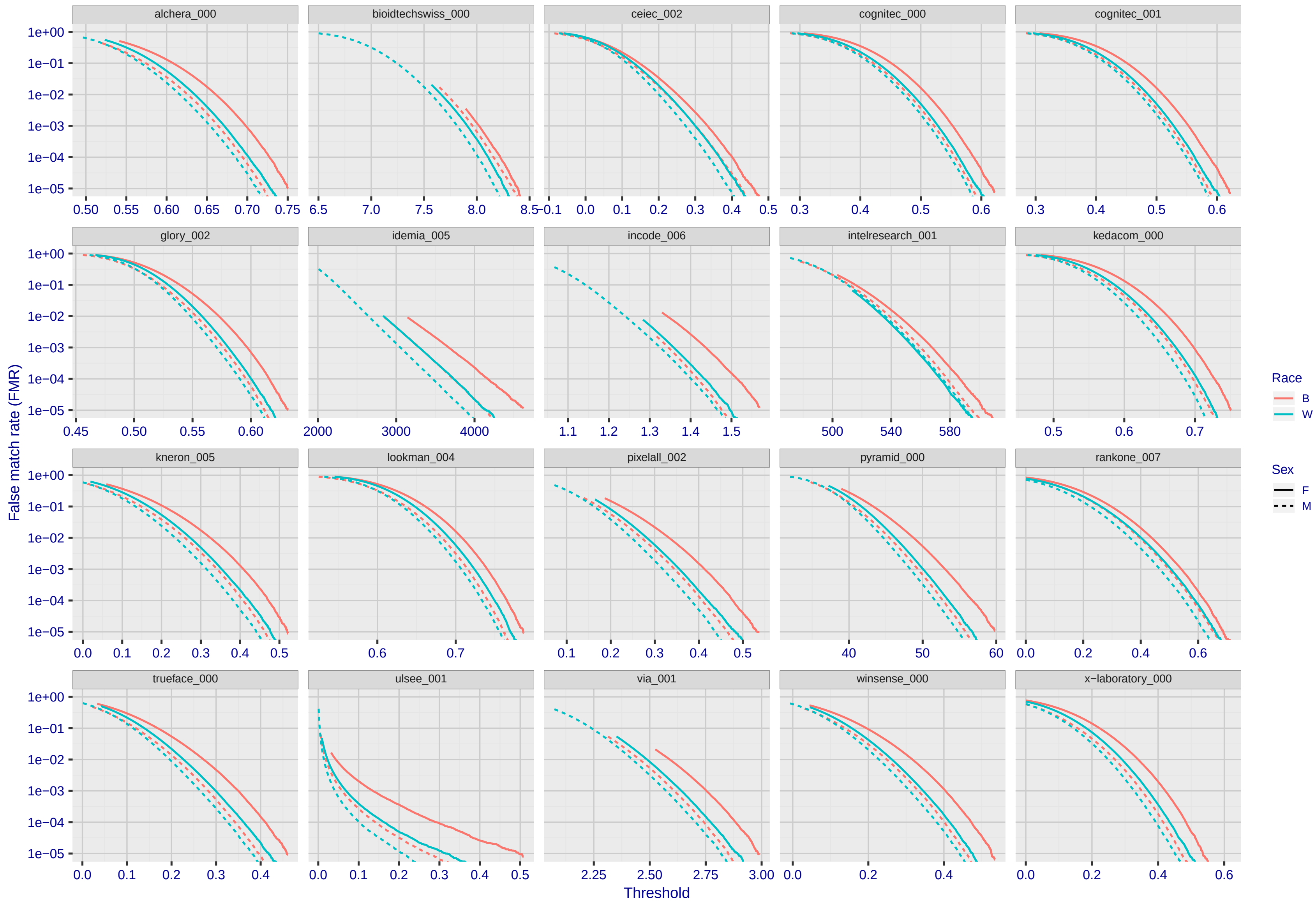


Figure 107: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

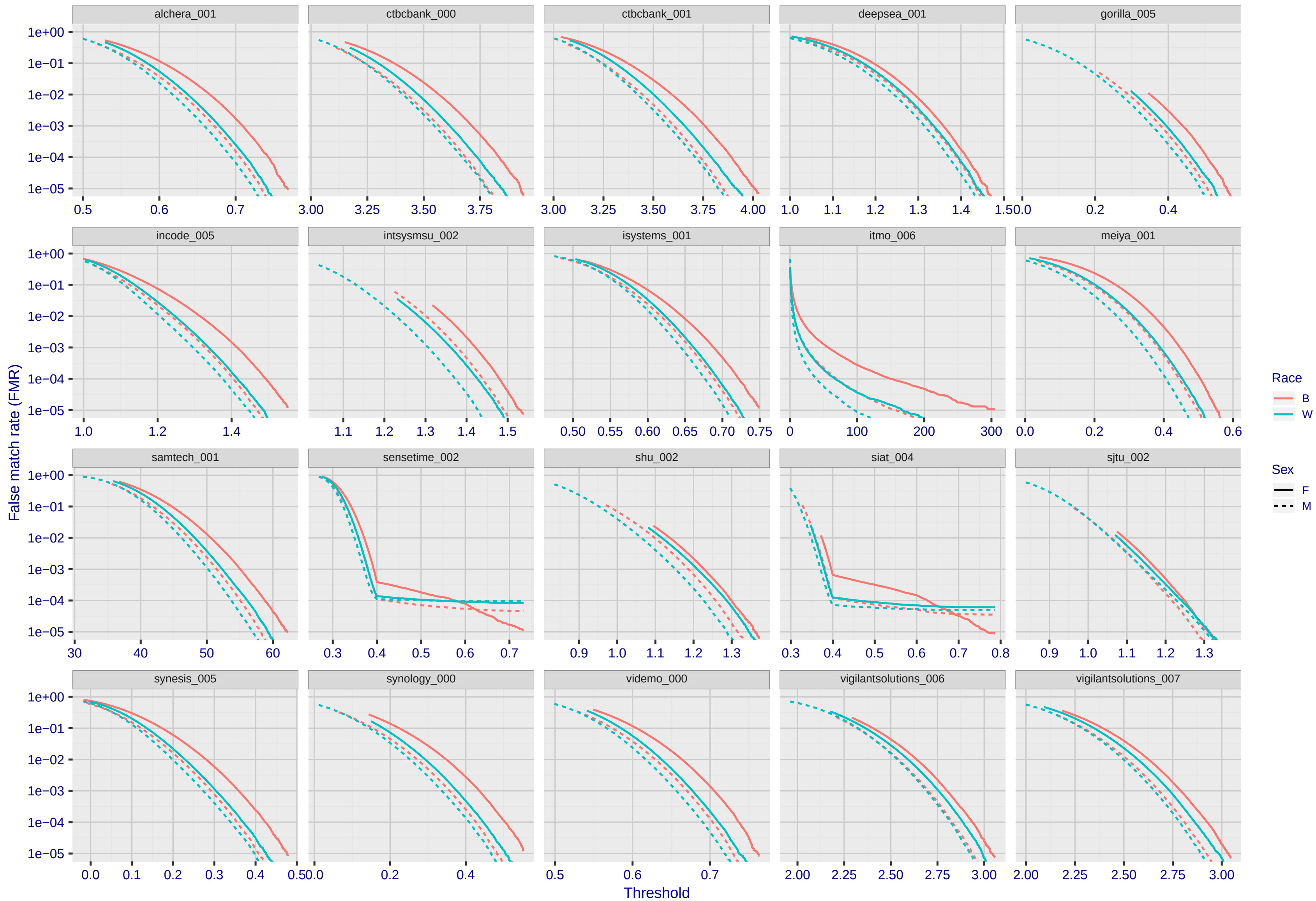


Figure 108: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

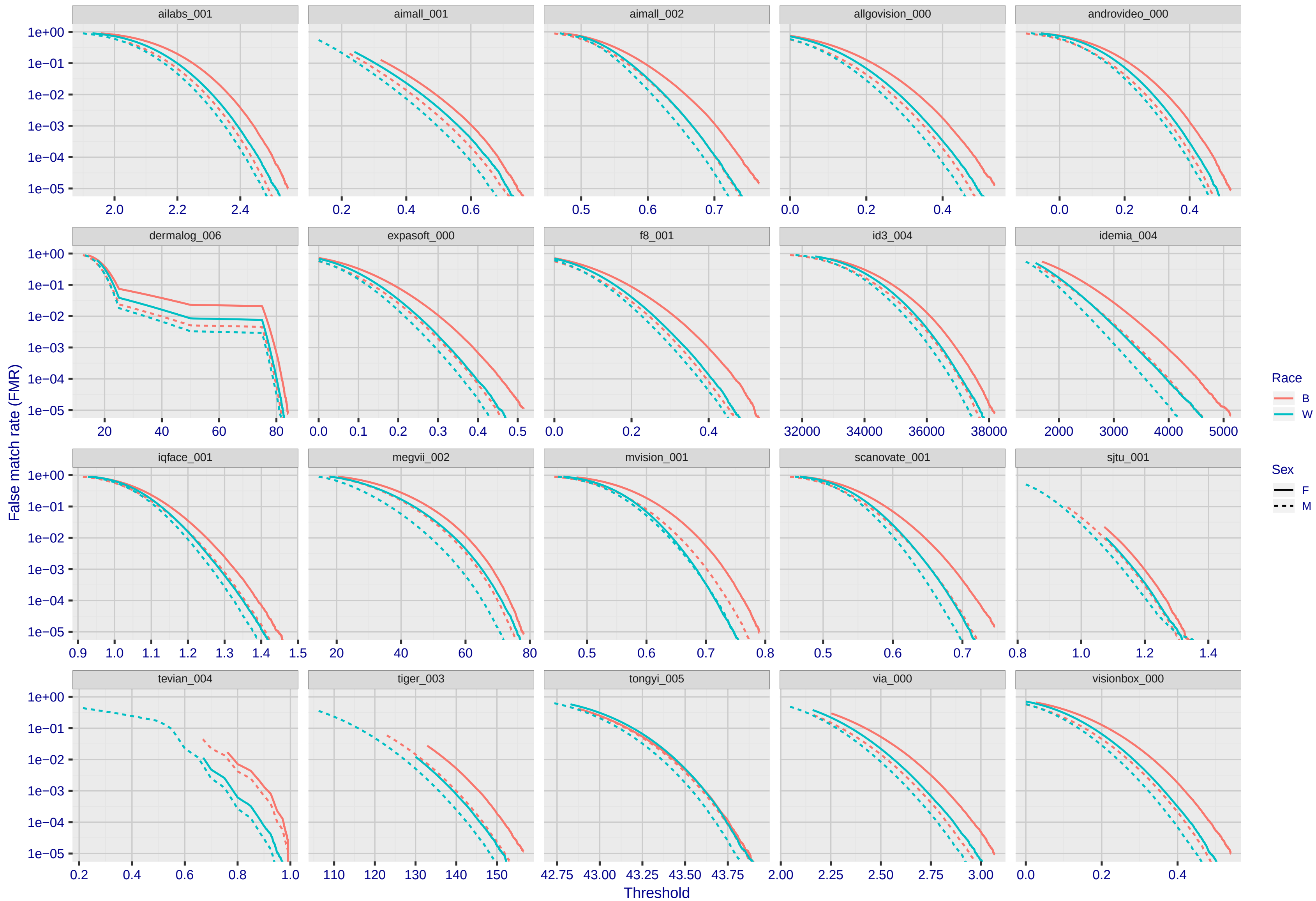


Figure 109: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

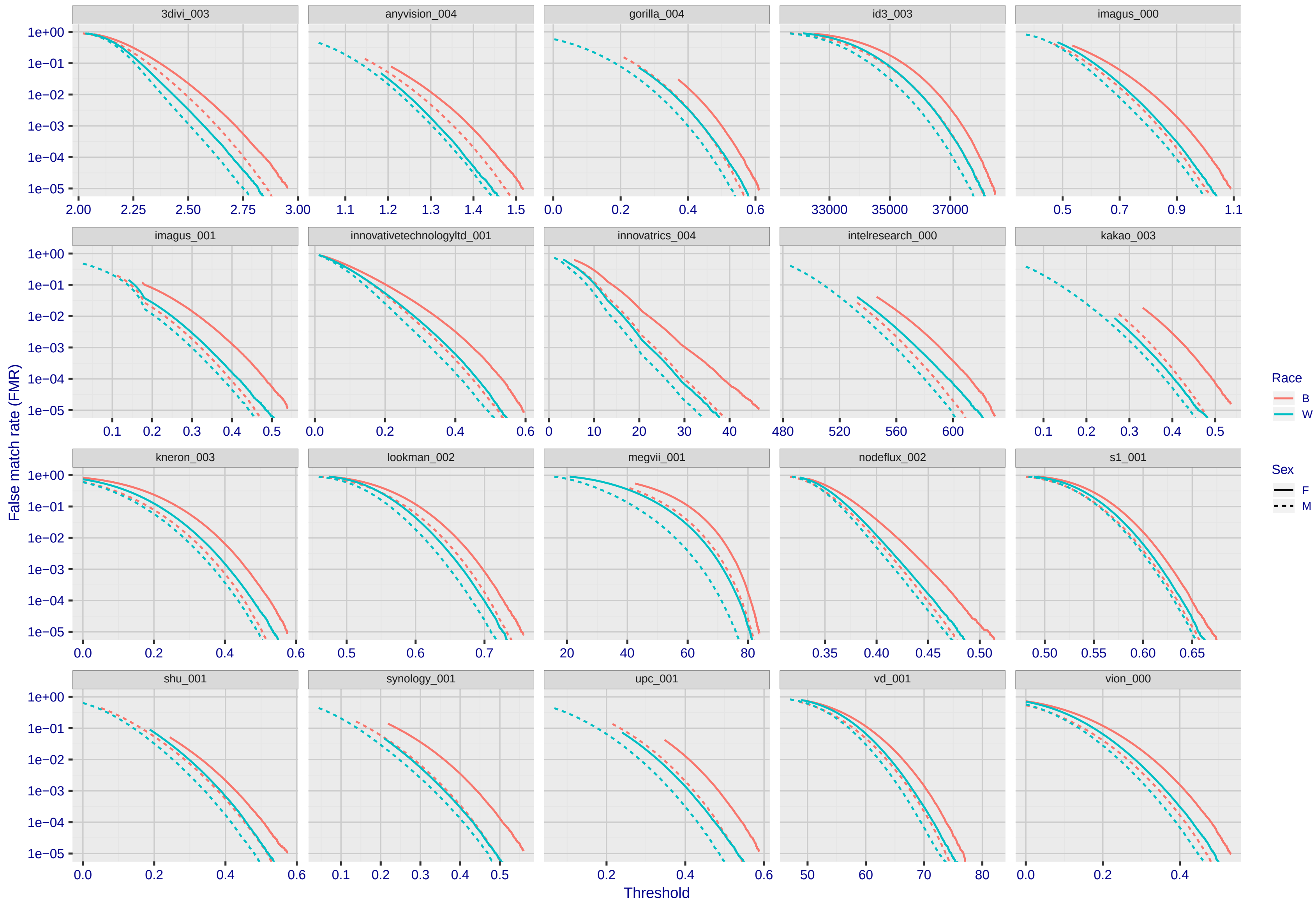


Figure 110: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

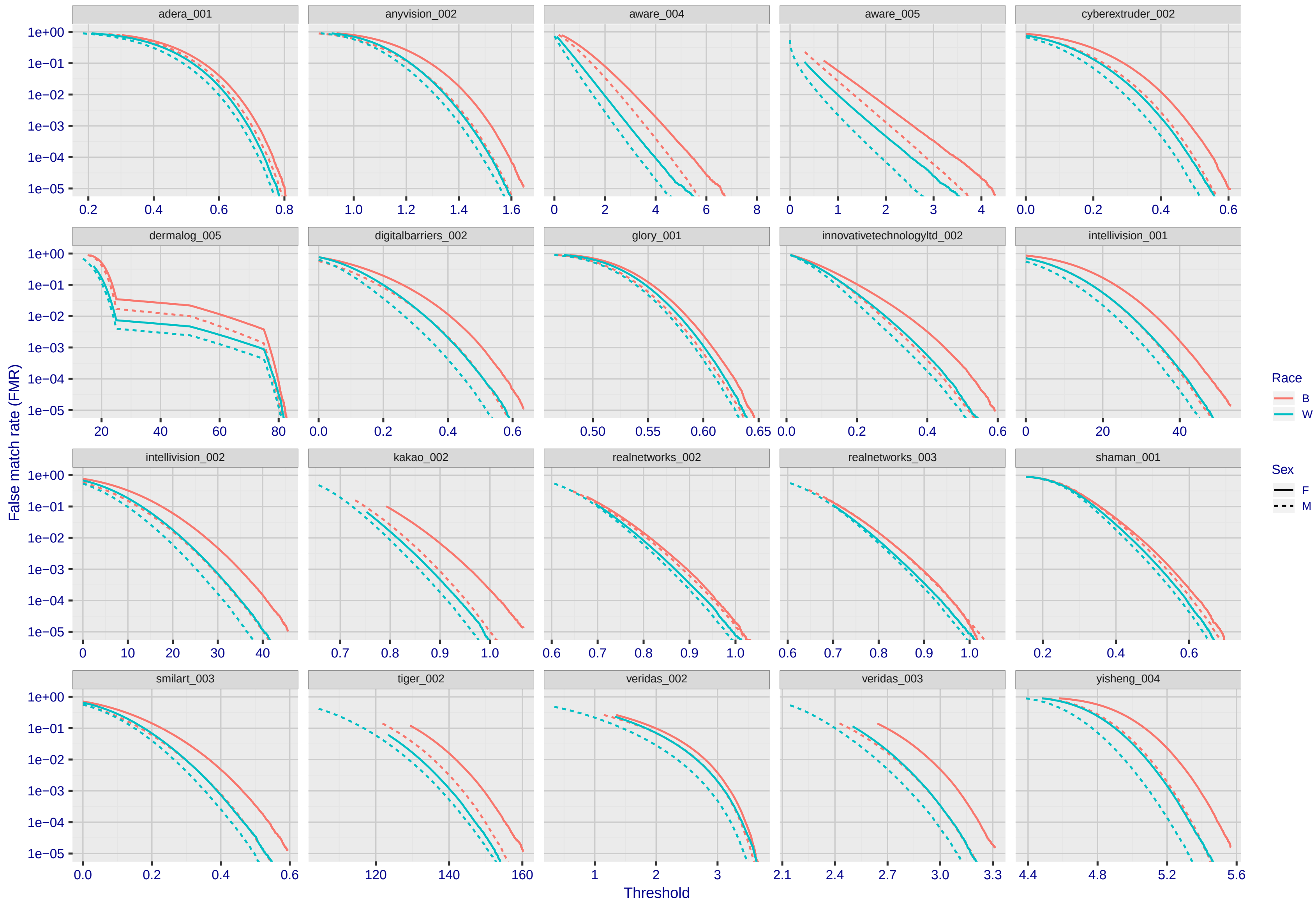


Figure 111: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

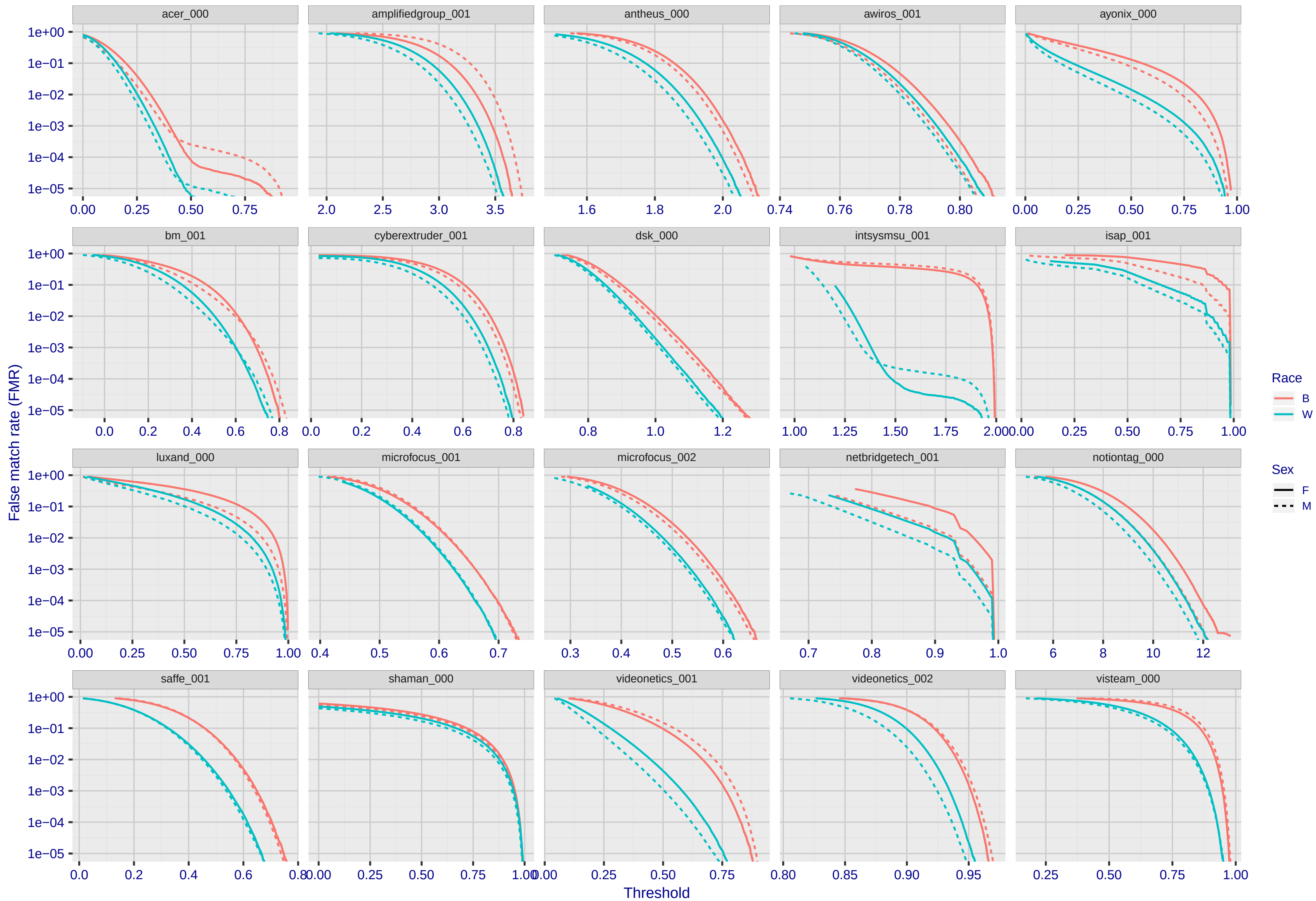


Figure 112: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

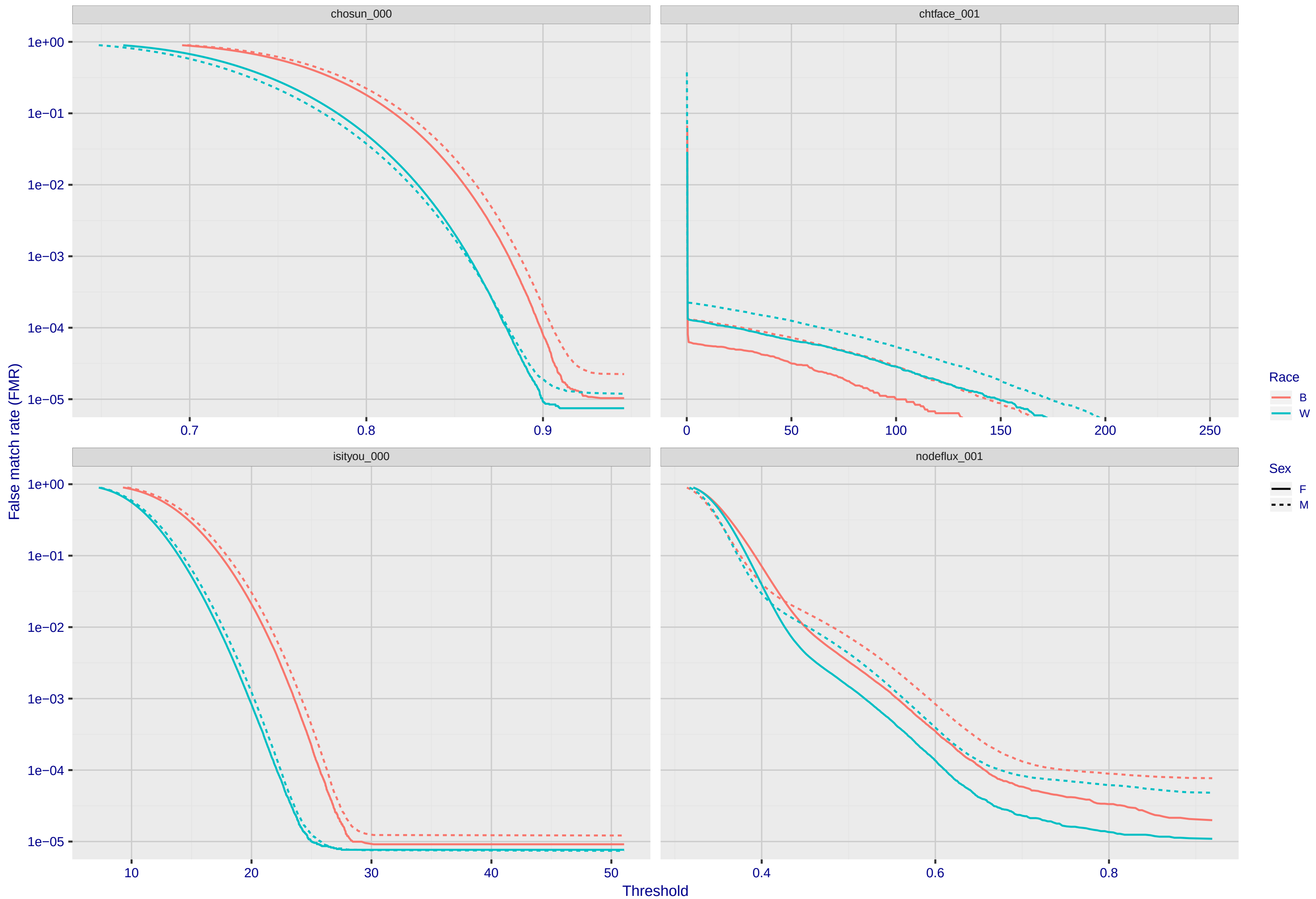


Figure 113: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

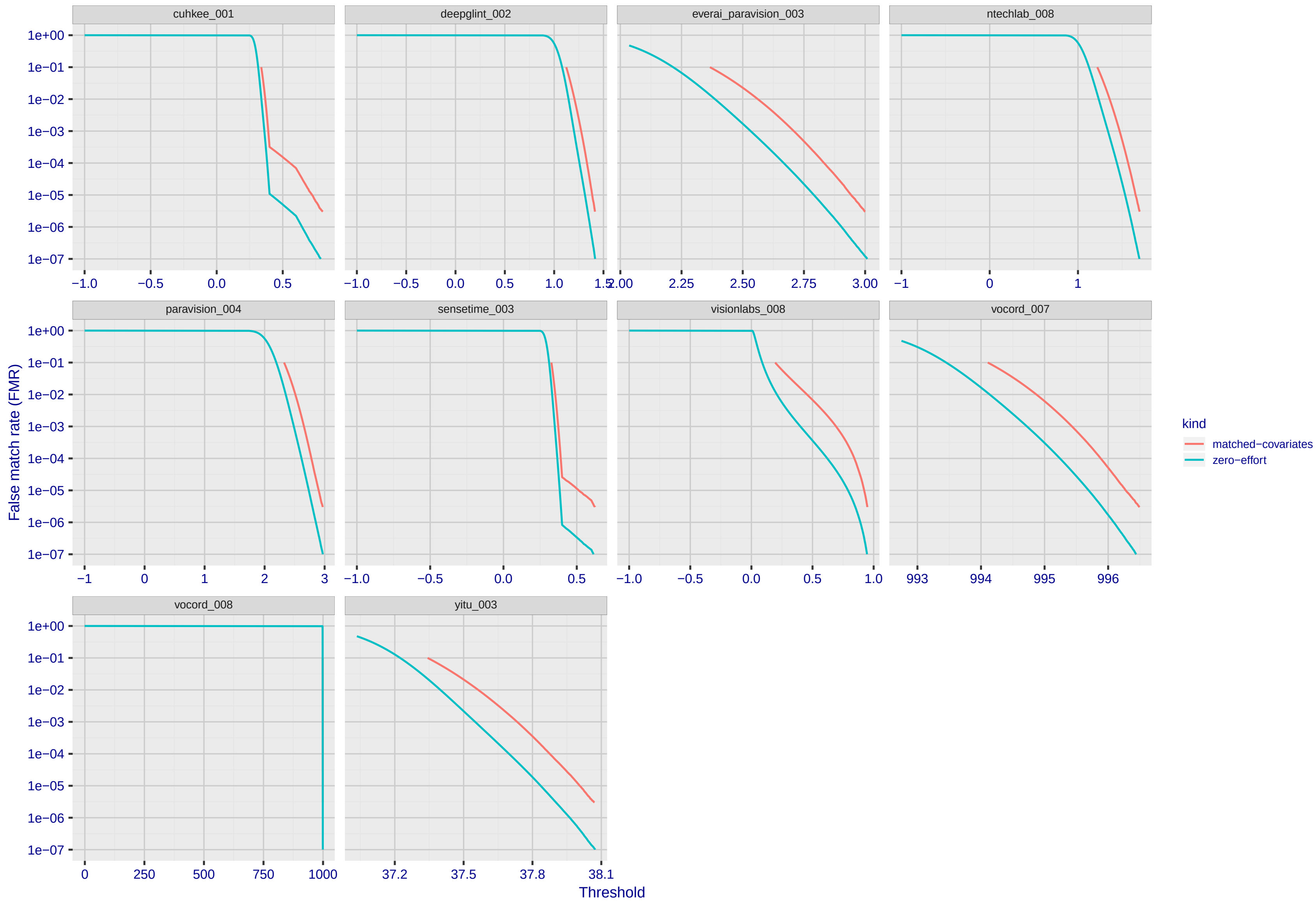


Figure 114: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

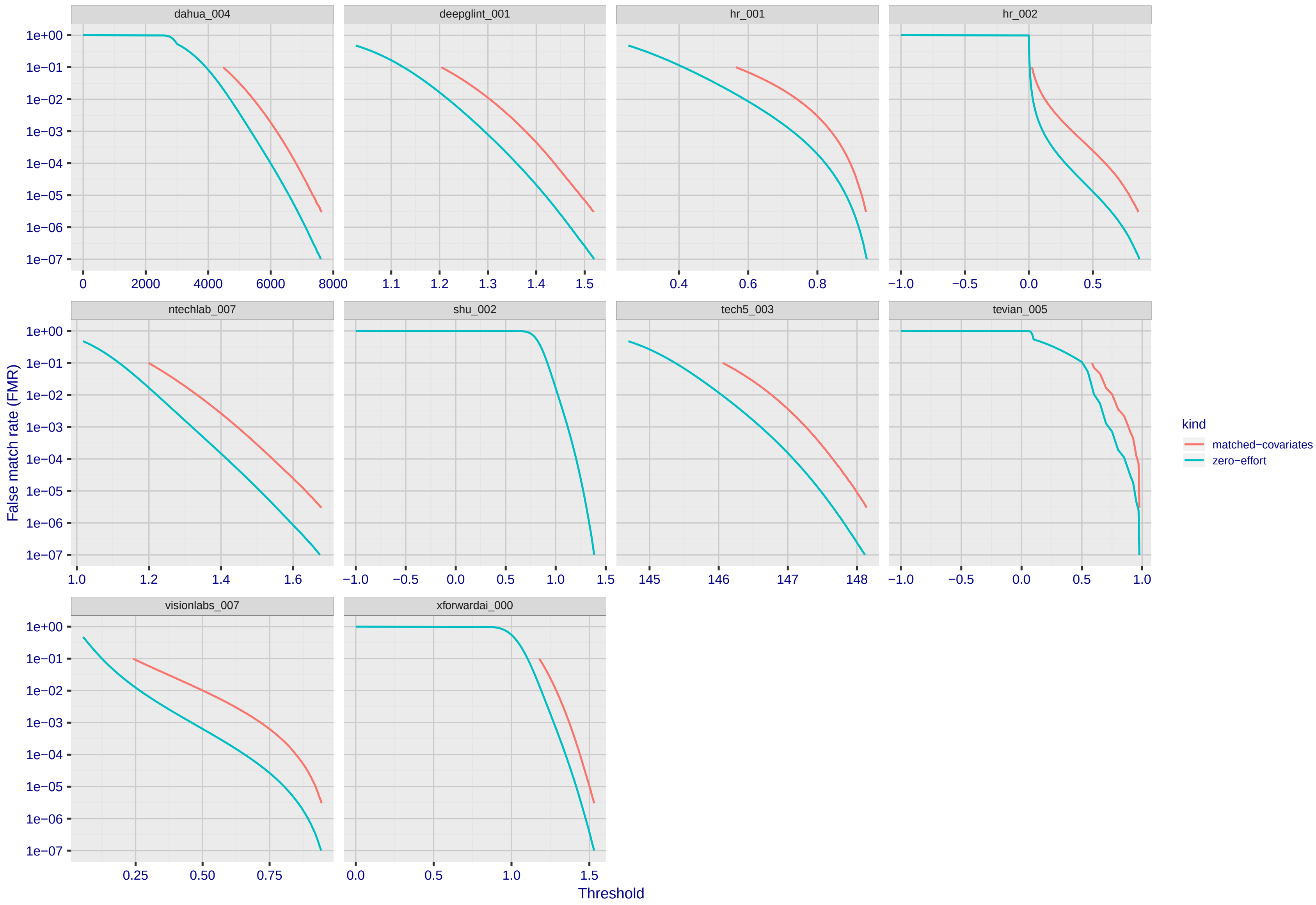


Figure 115: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

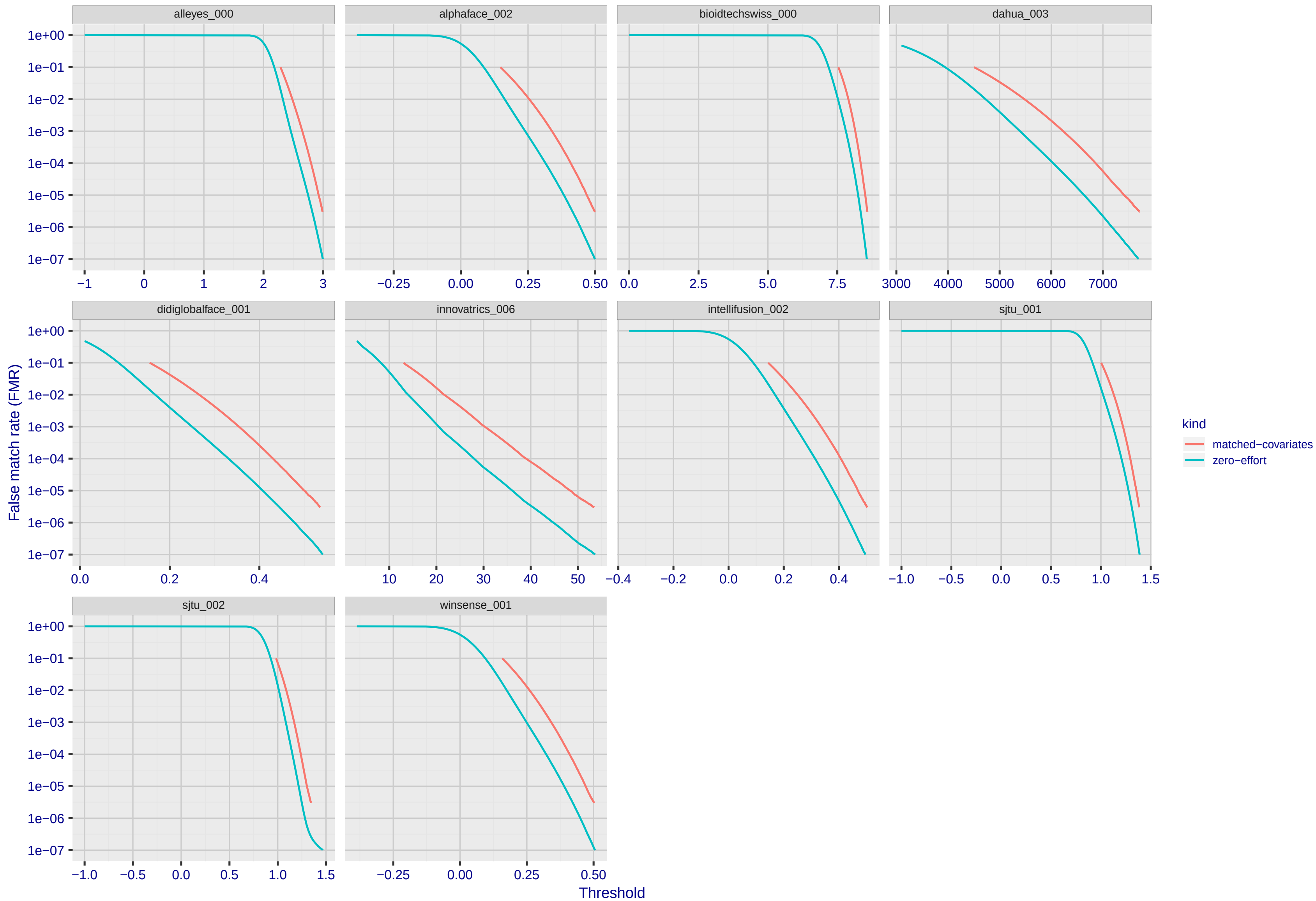


Figure 116: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

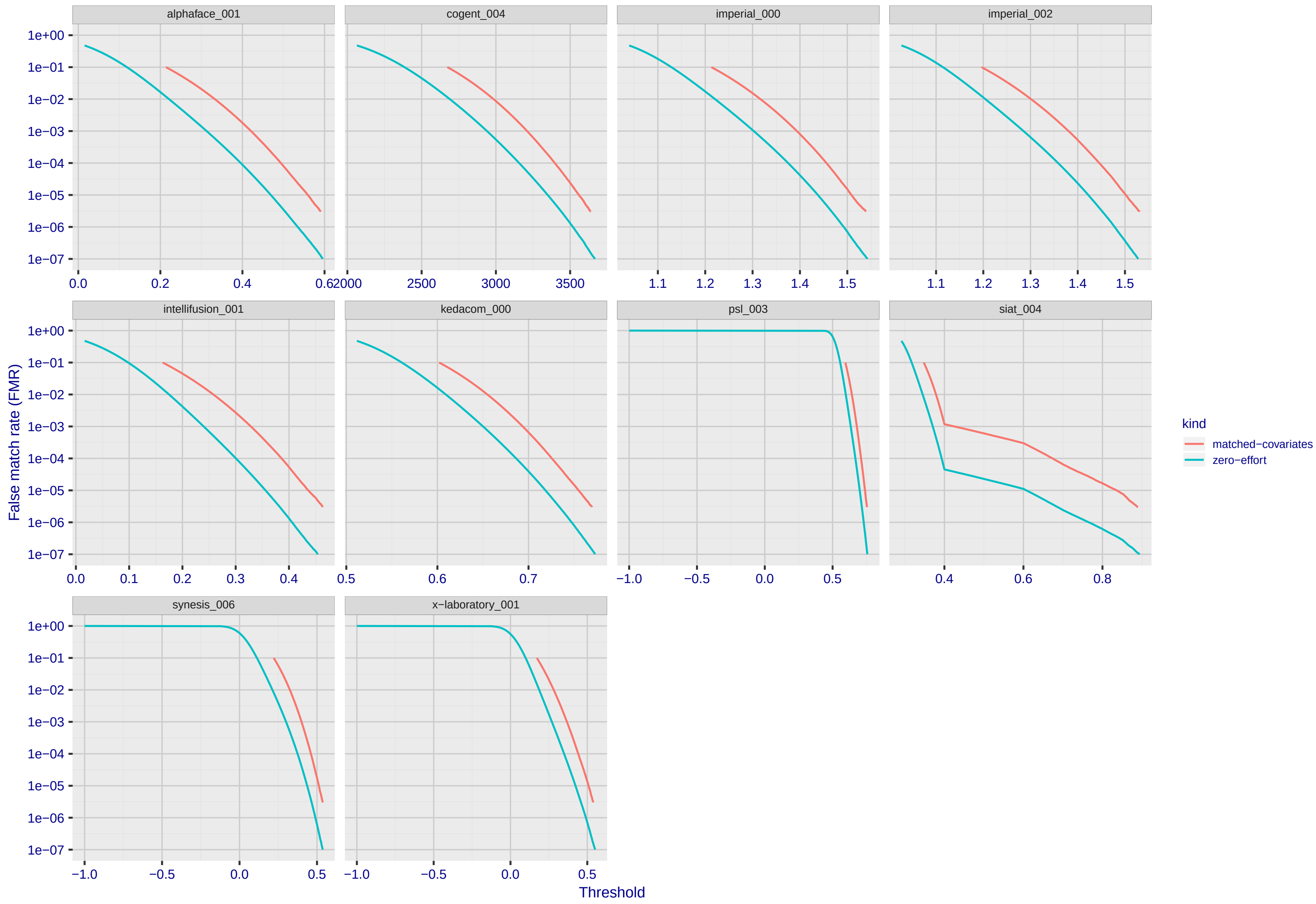


Figure 117: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

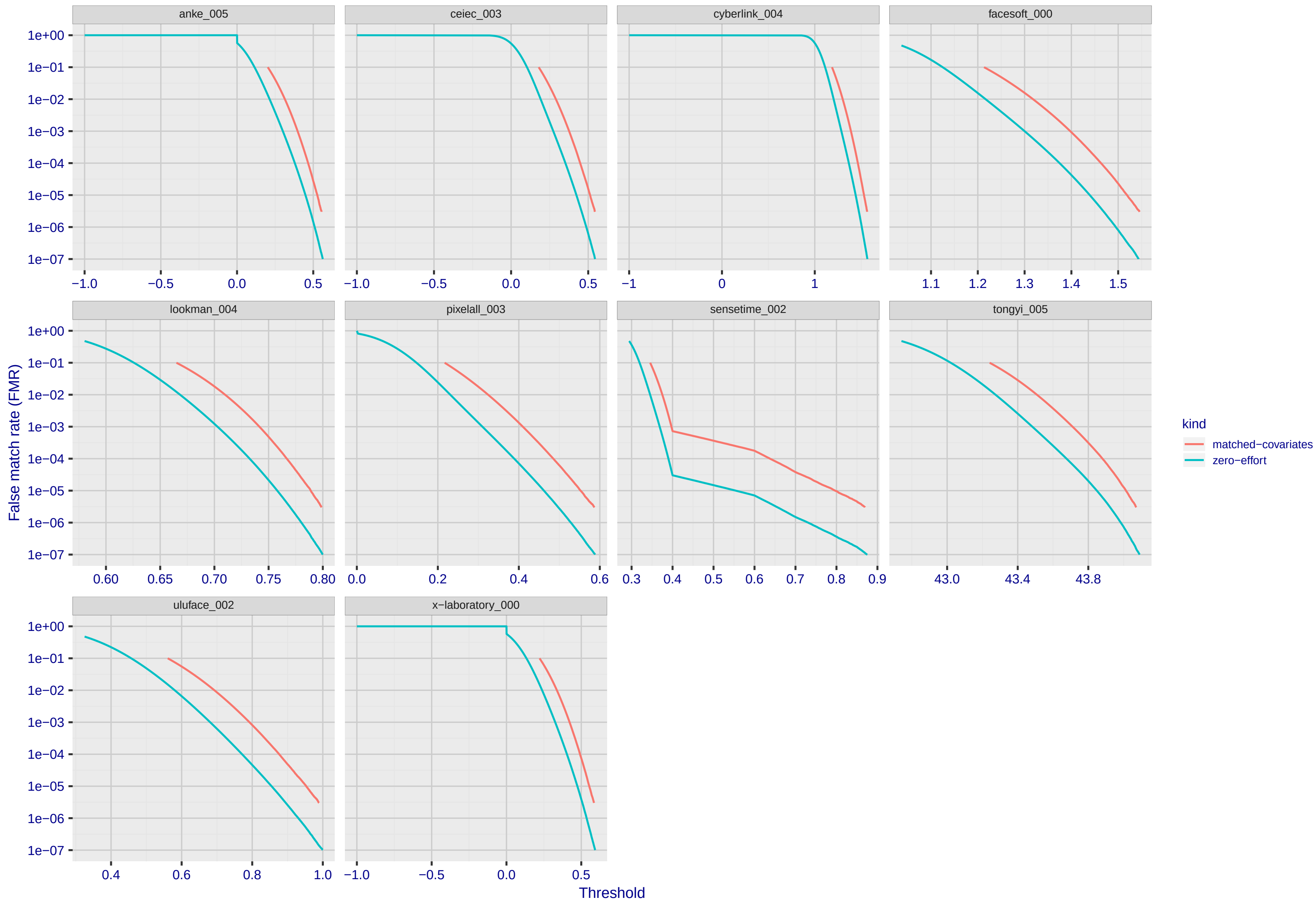


Figure 118: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

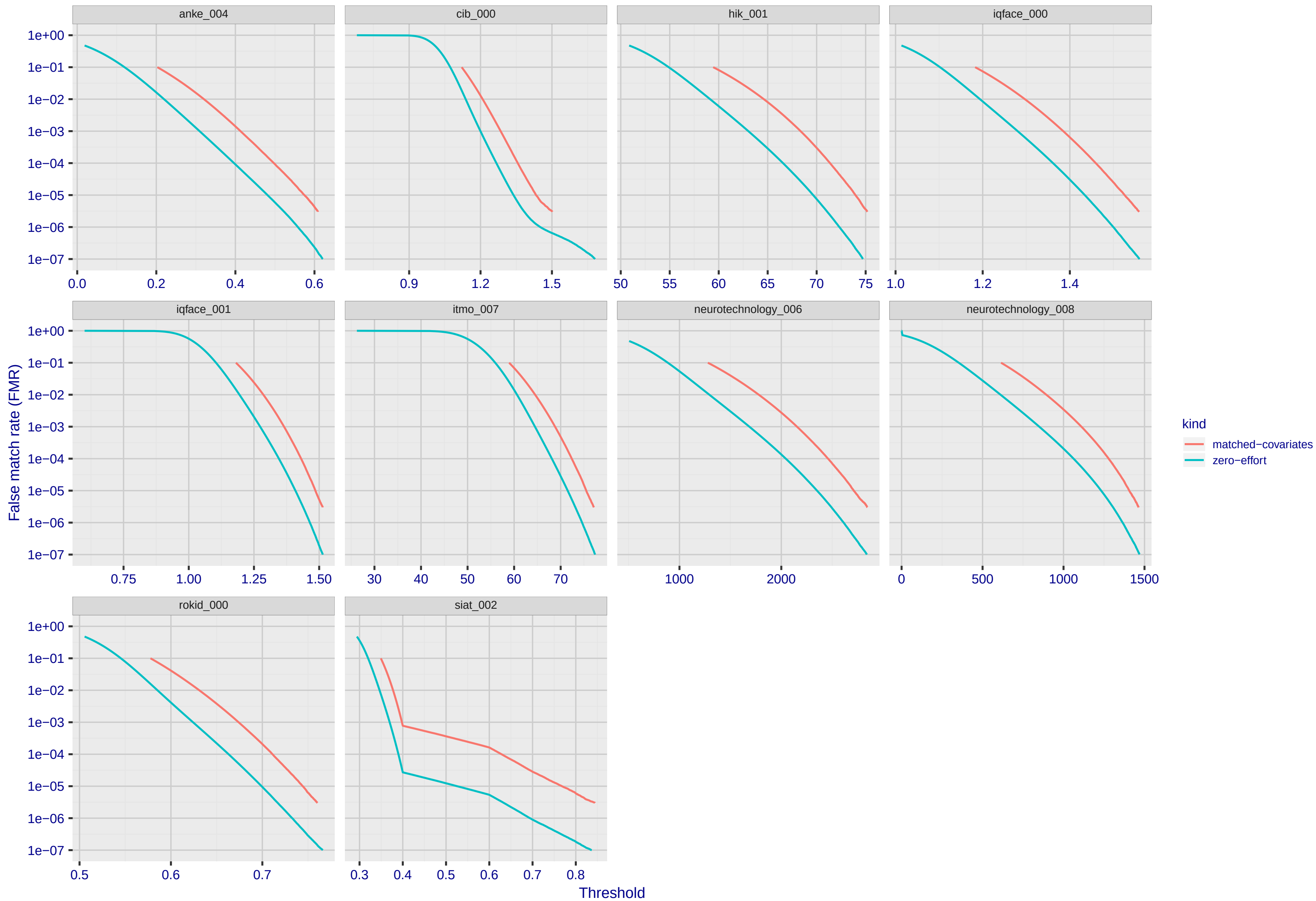


Figure 119: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

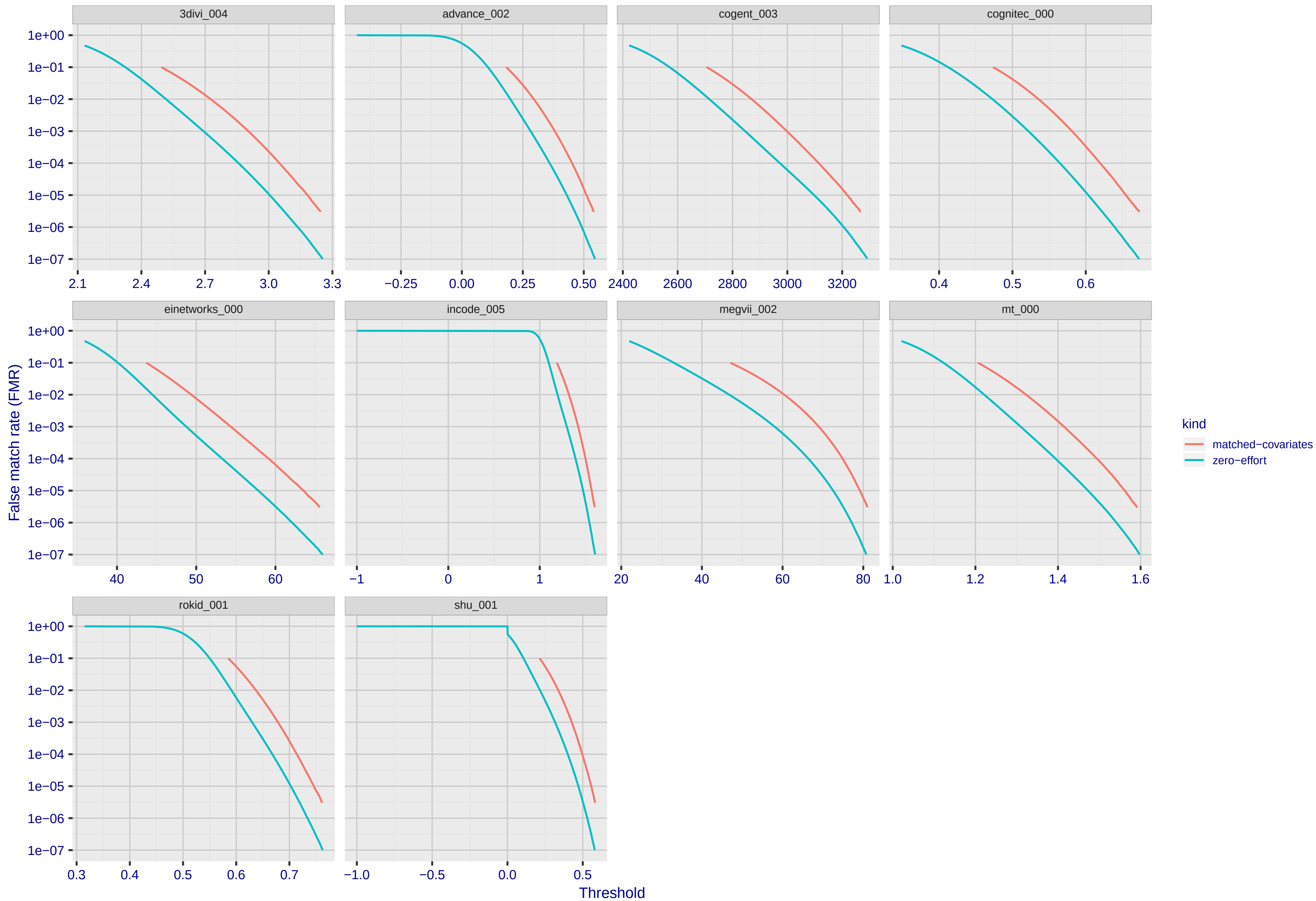


Figure 120: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

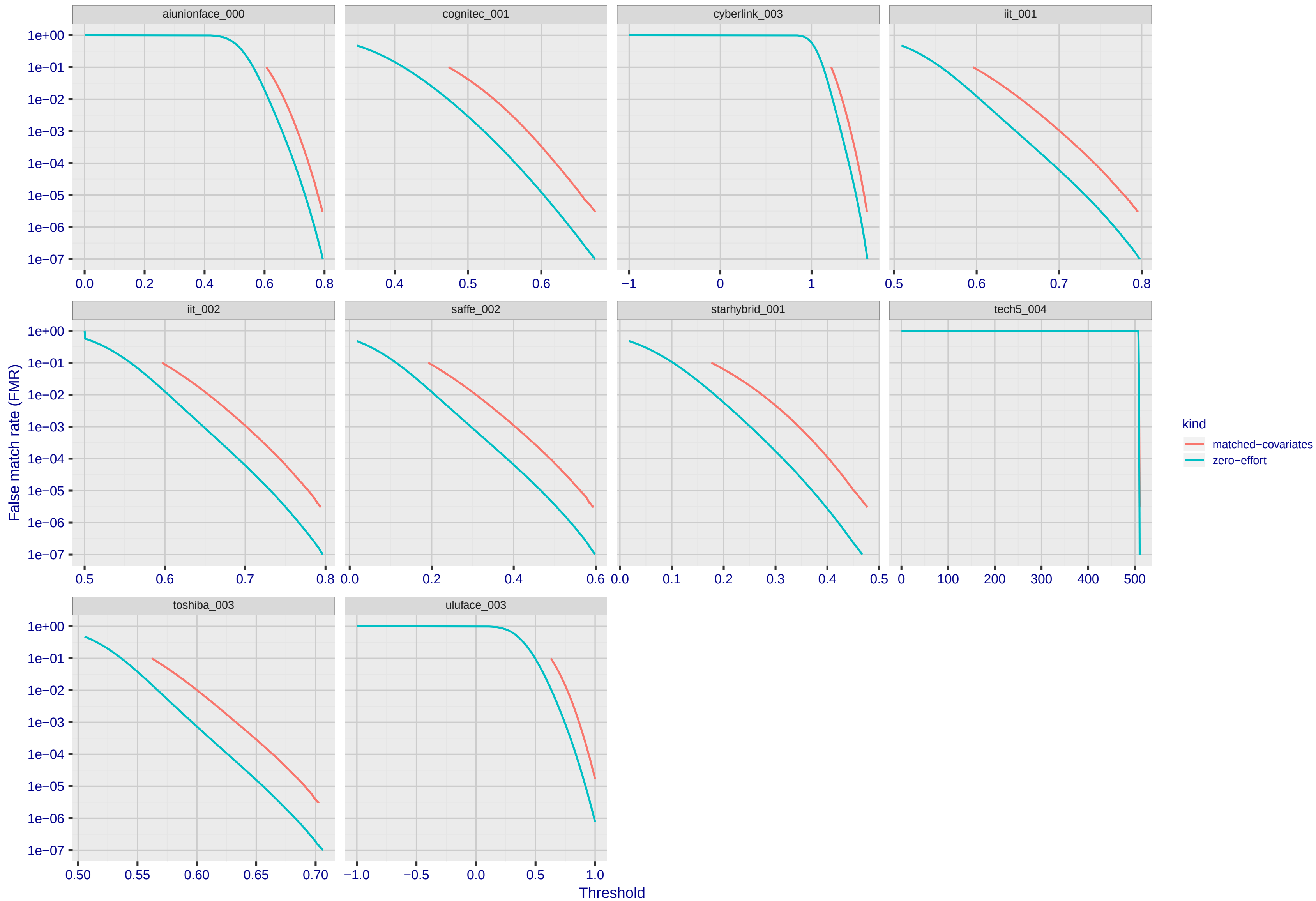


Figure 121: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

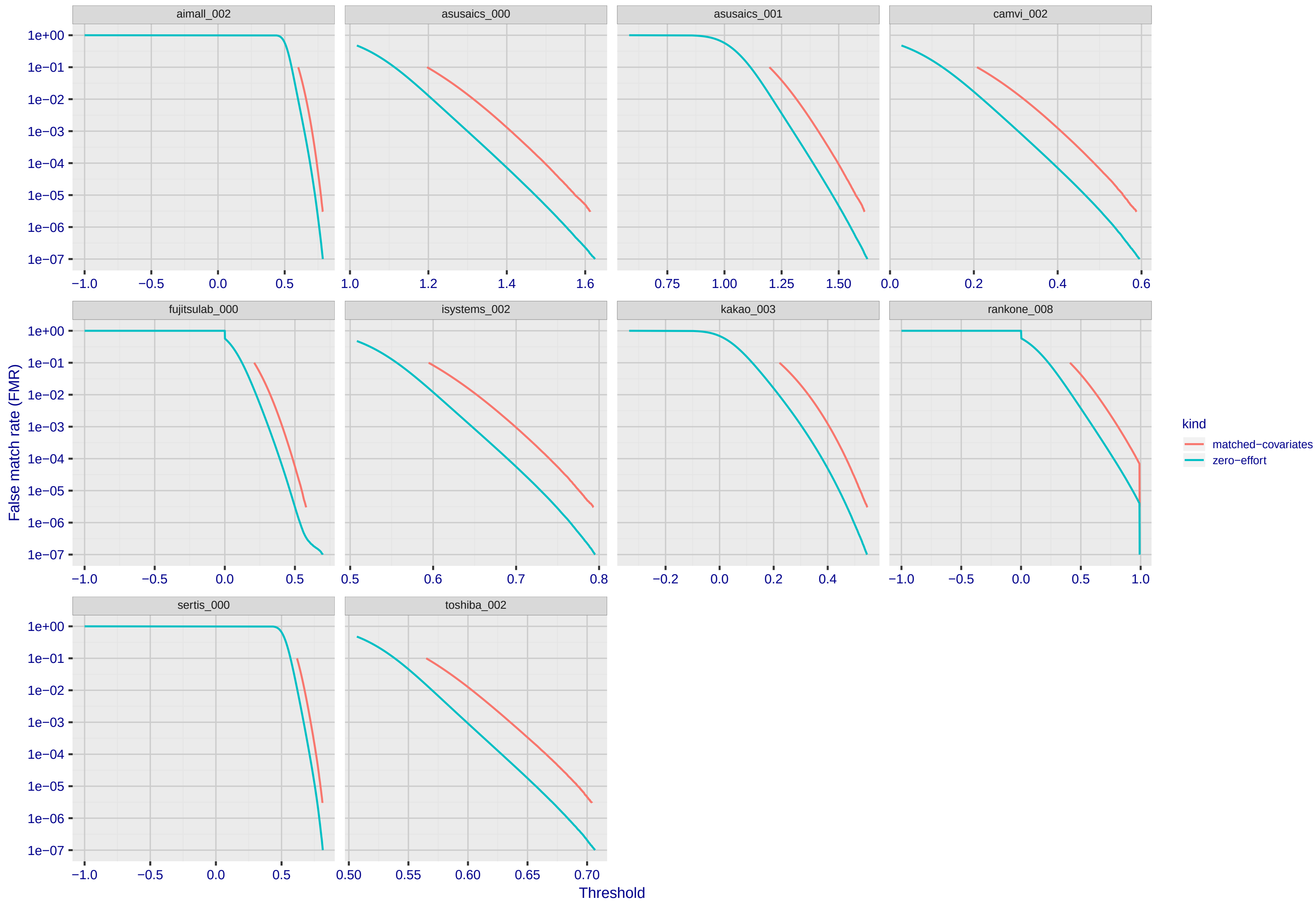


Figure 122: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

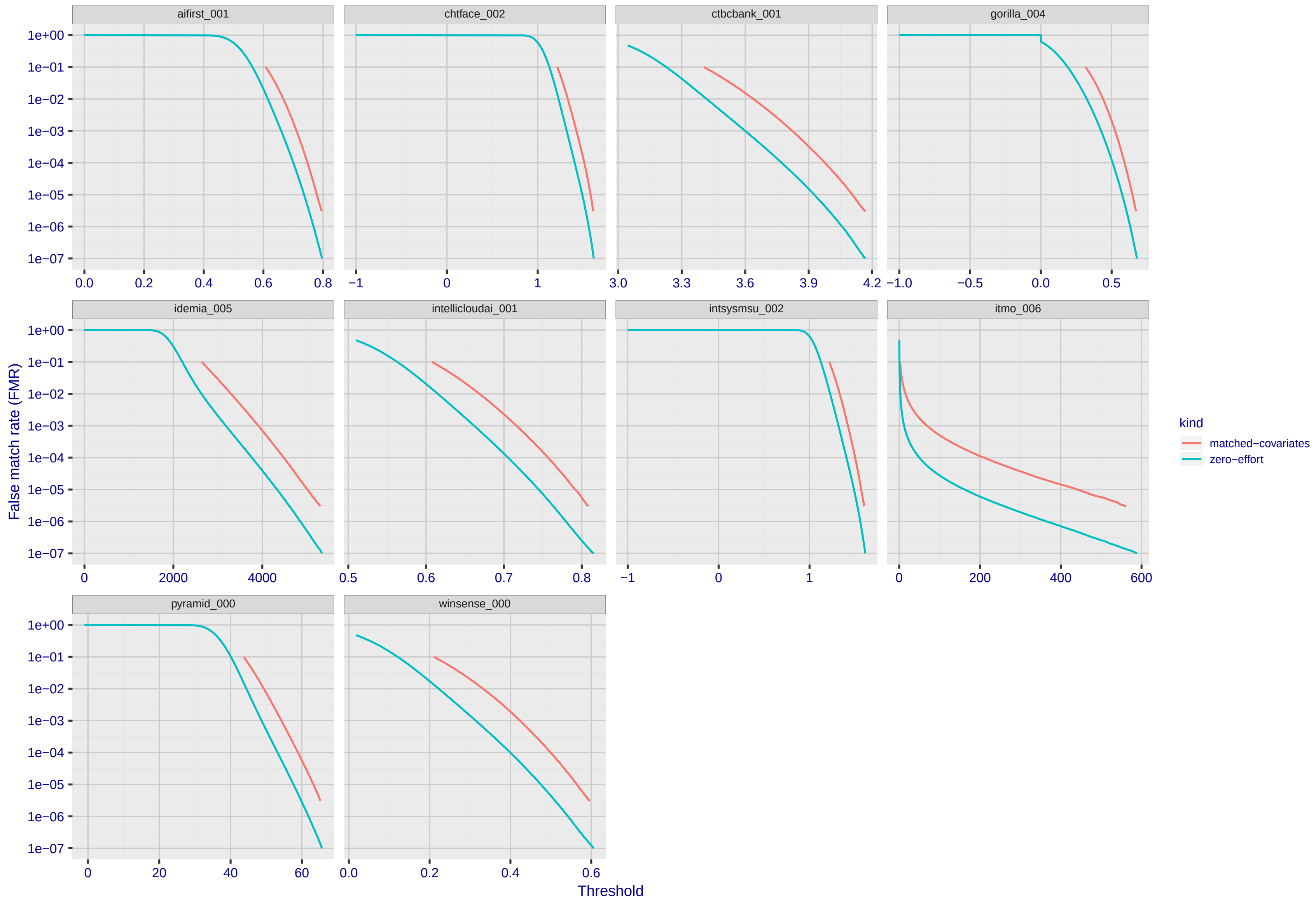


Figure 123: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

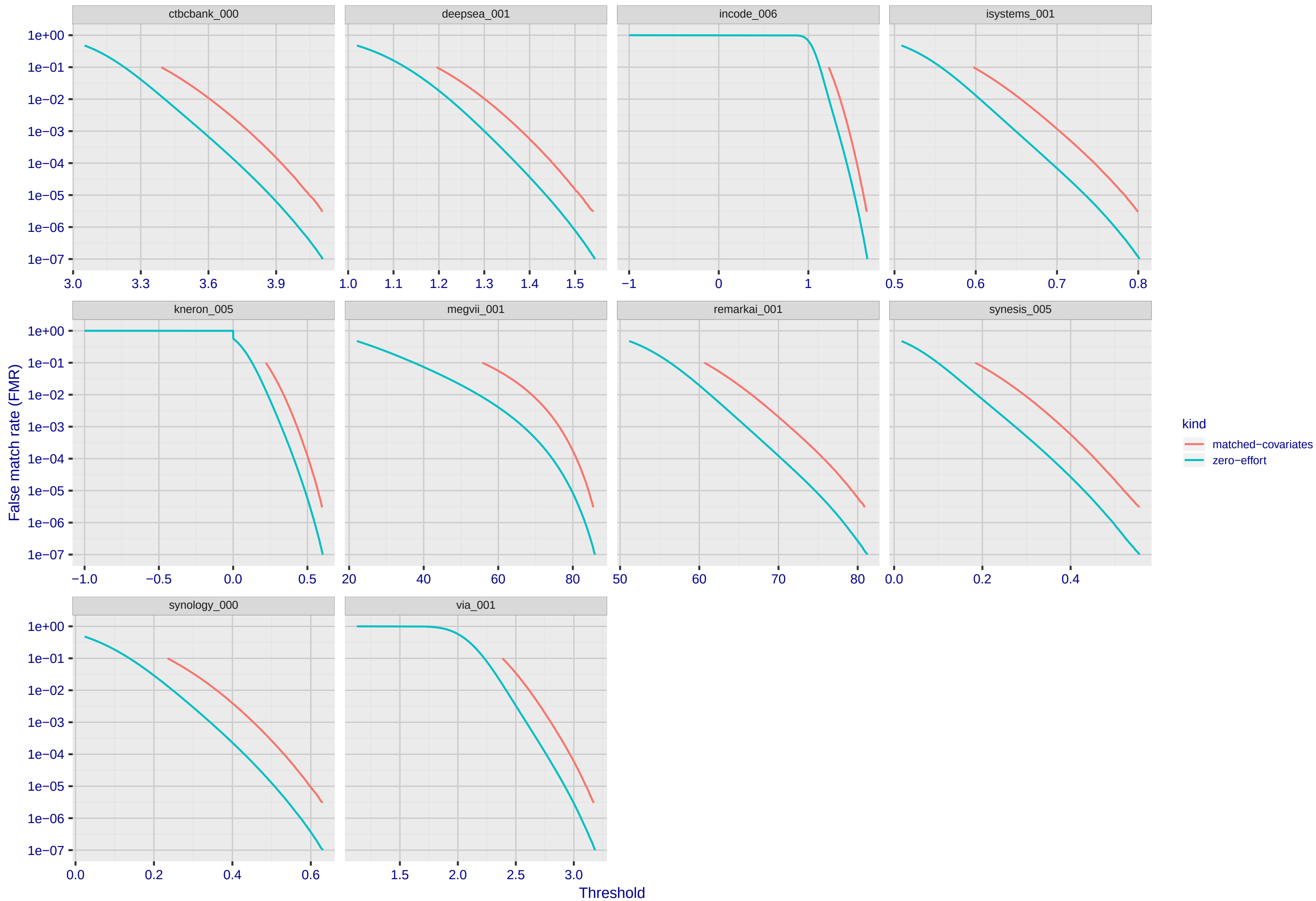


Figure 124: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

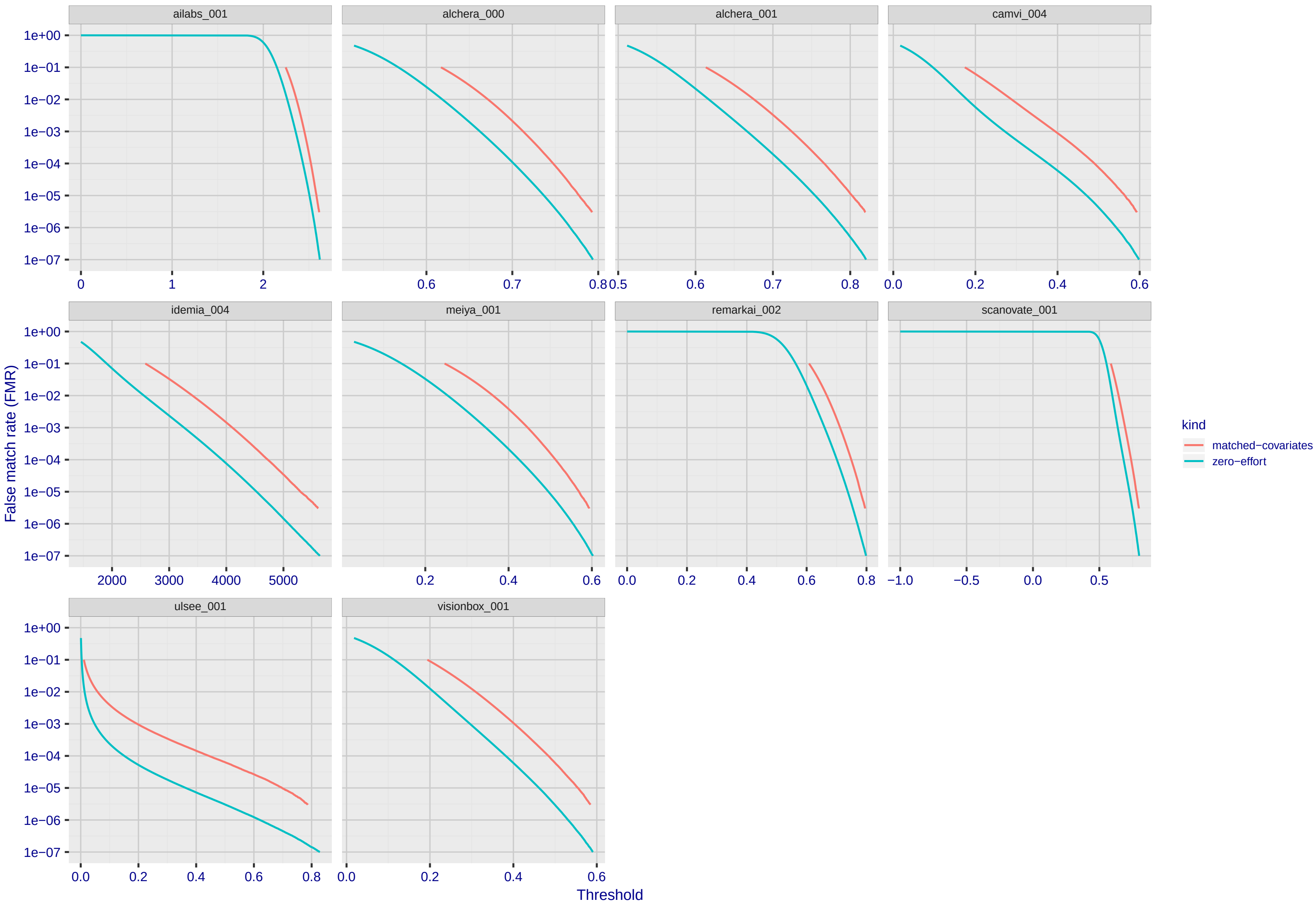


Figure 125: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

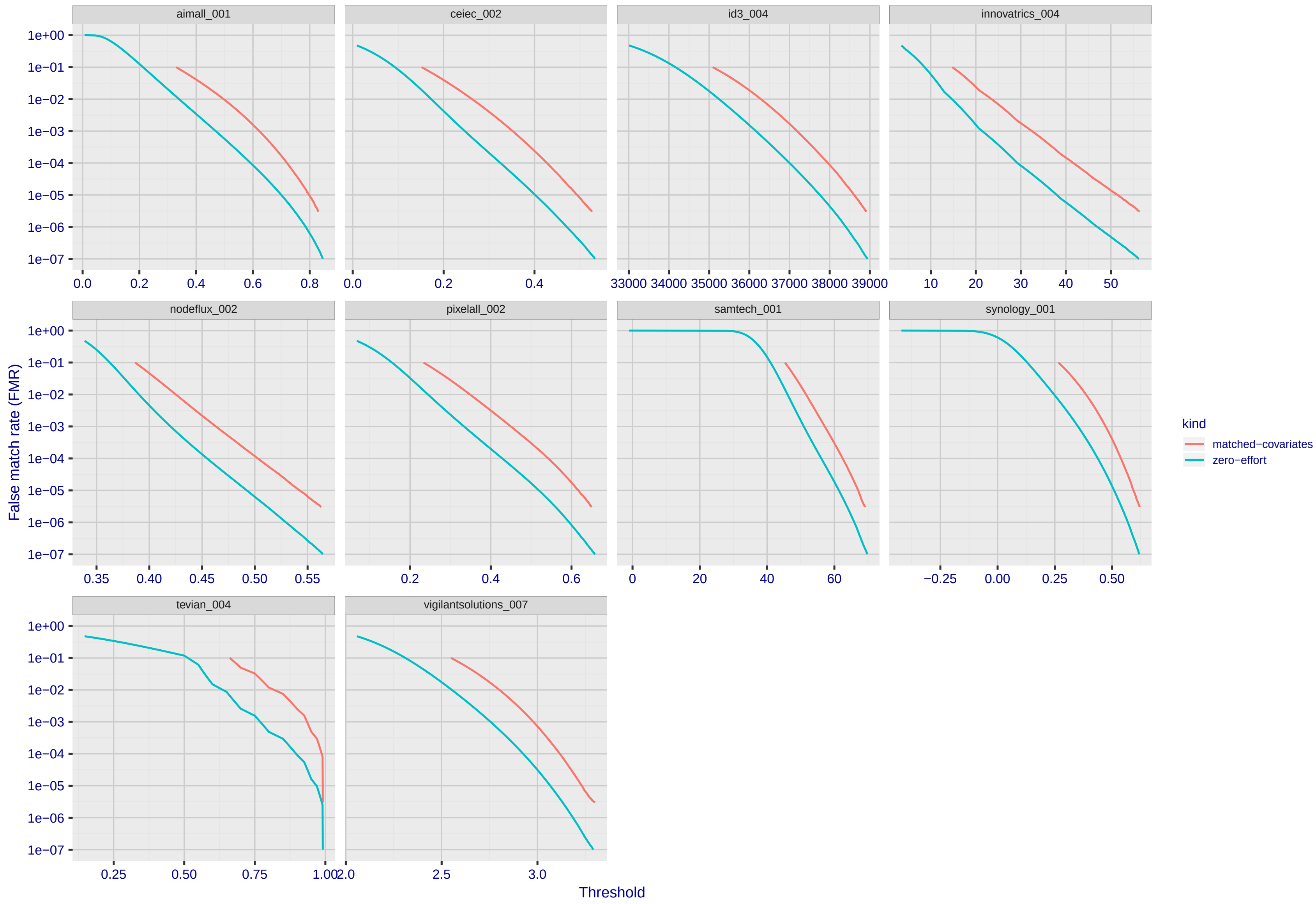


Figure 126: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

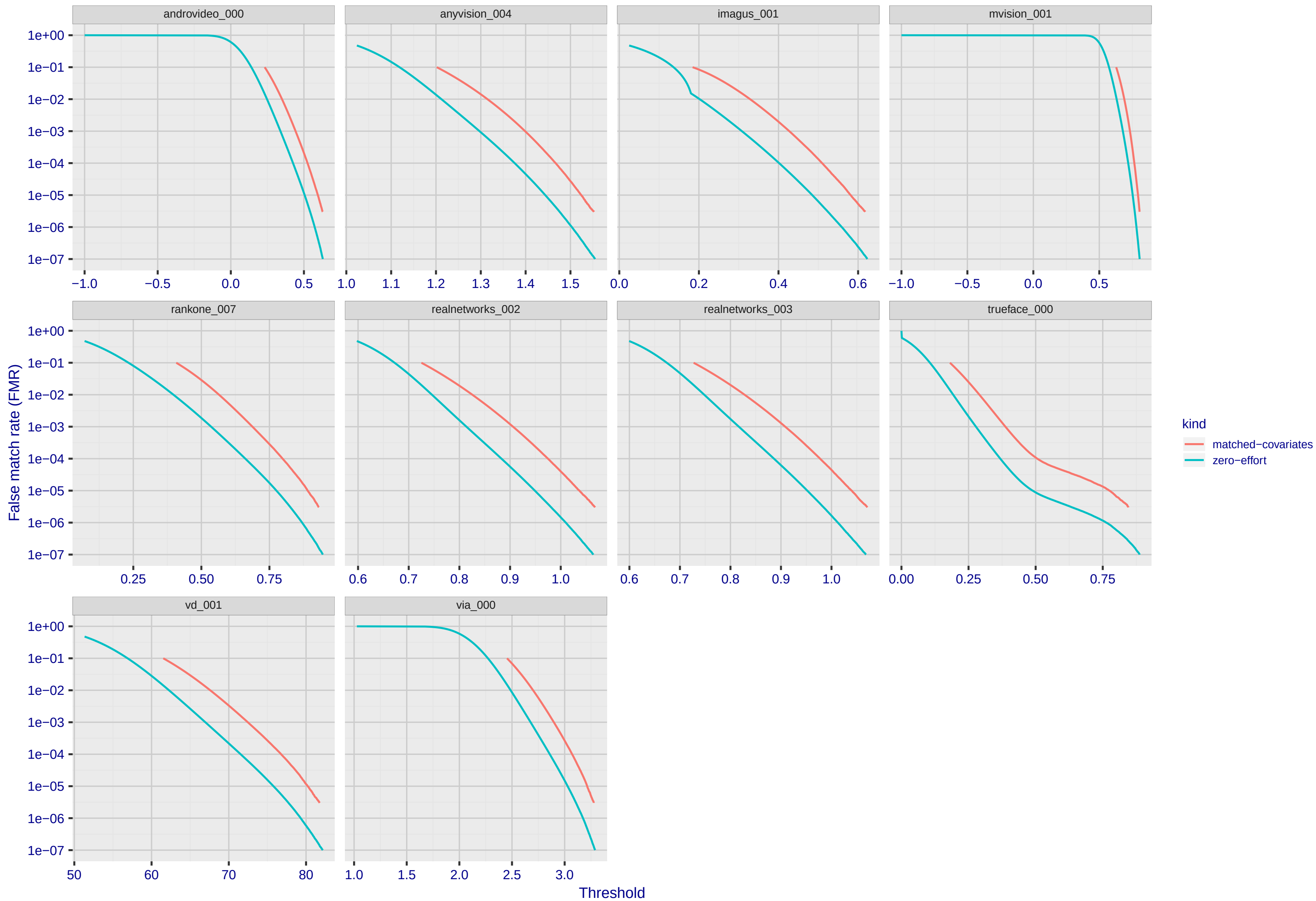


Figure 127: *For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.*

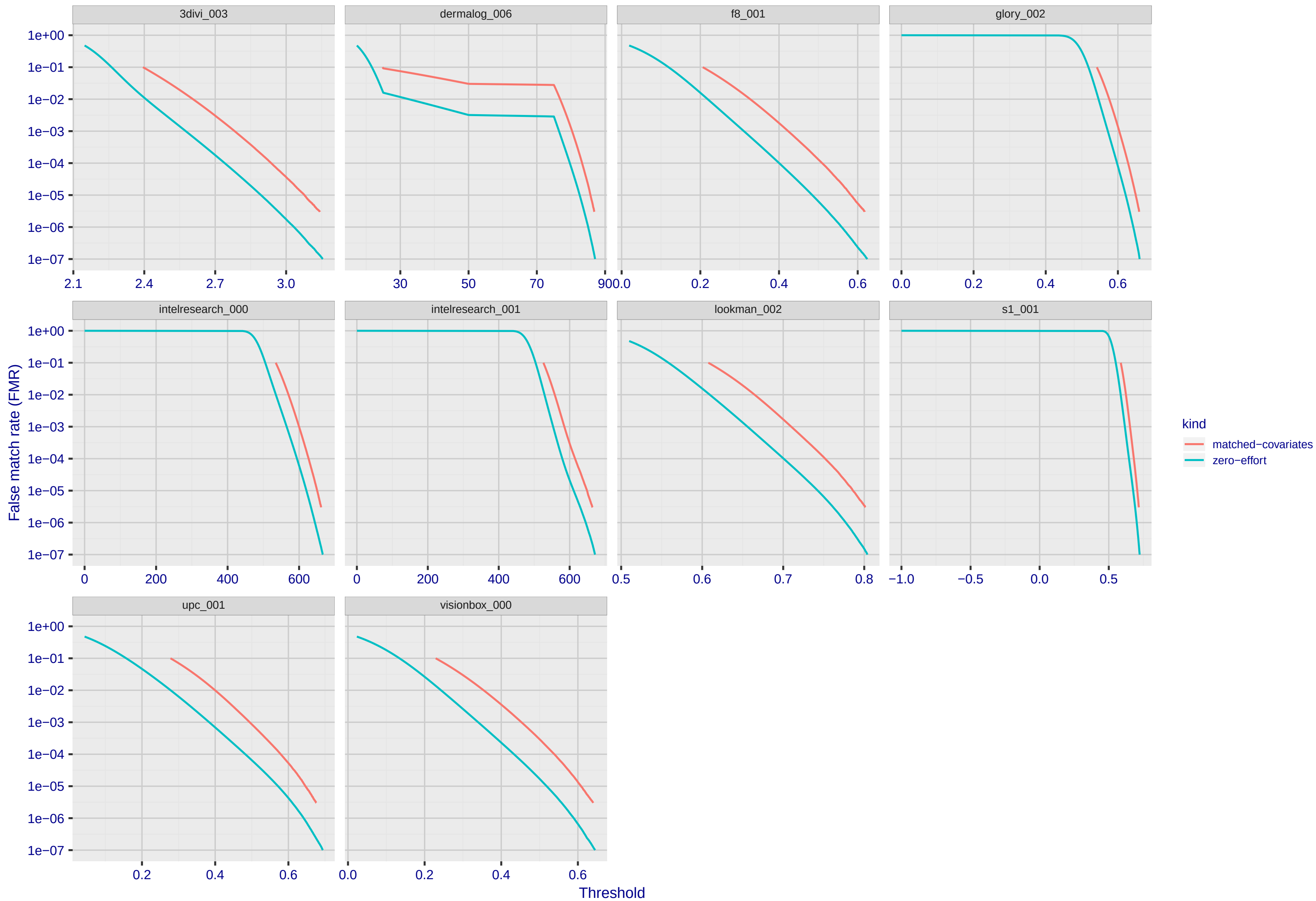


Figure 128: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

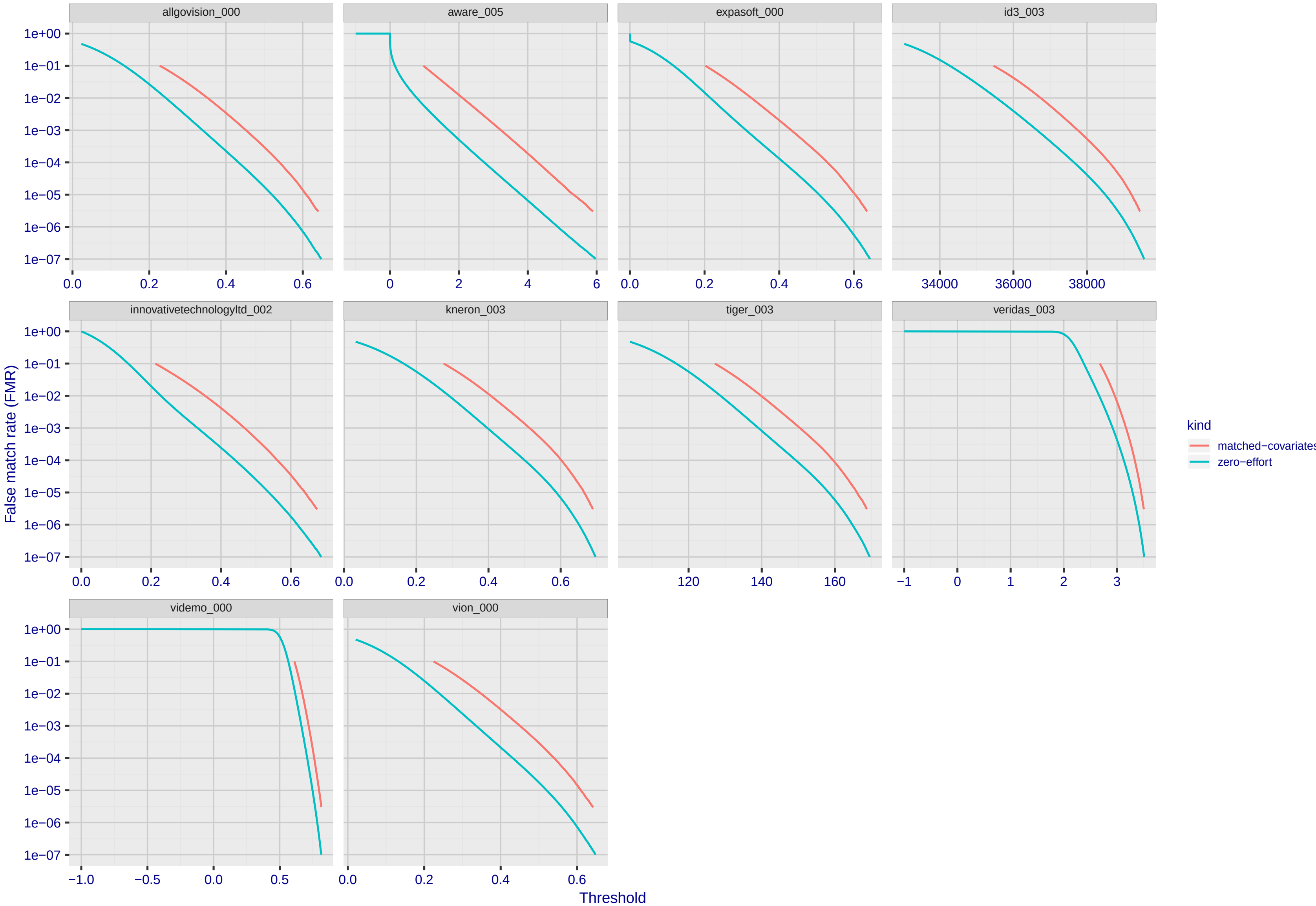


Figure 129: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

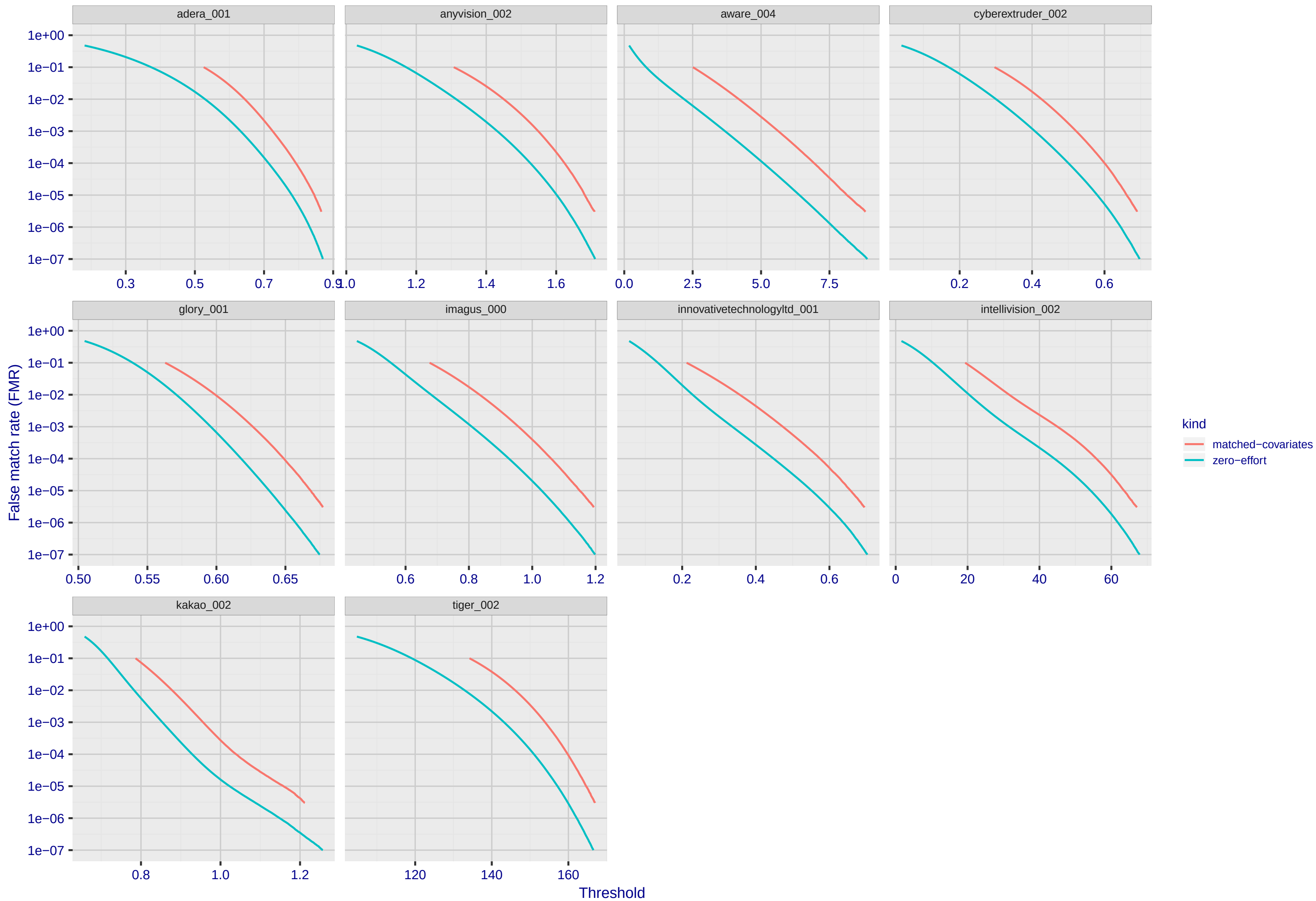


Figure 130: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

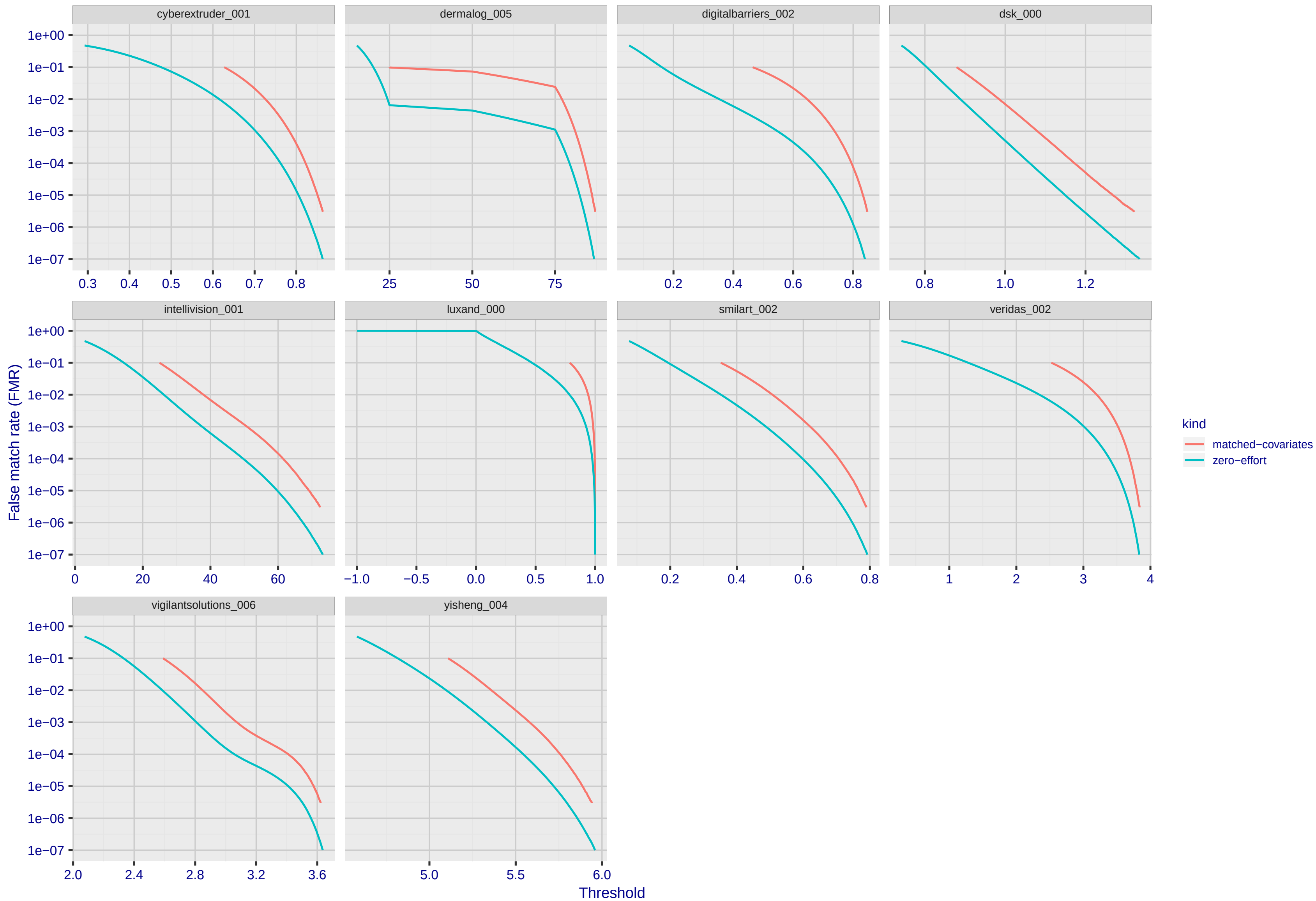


Figure 131: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

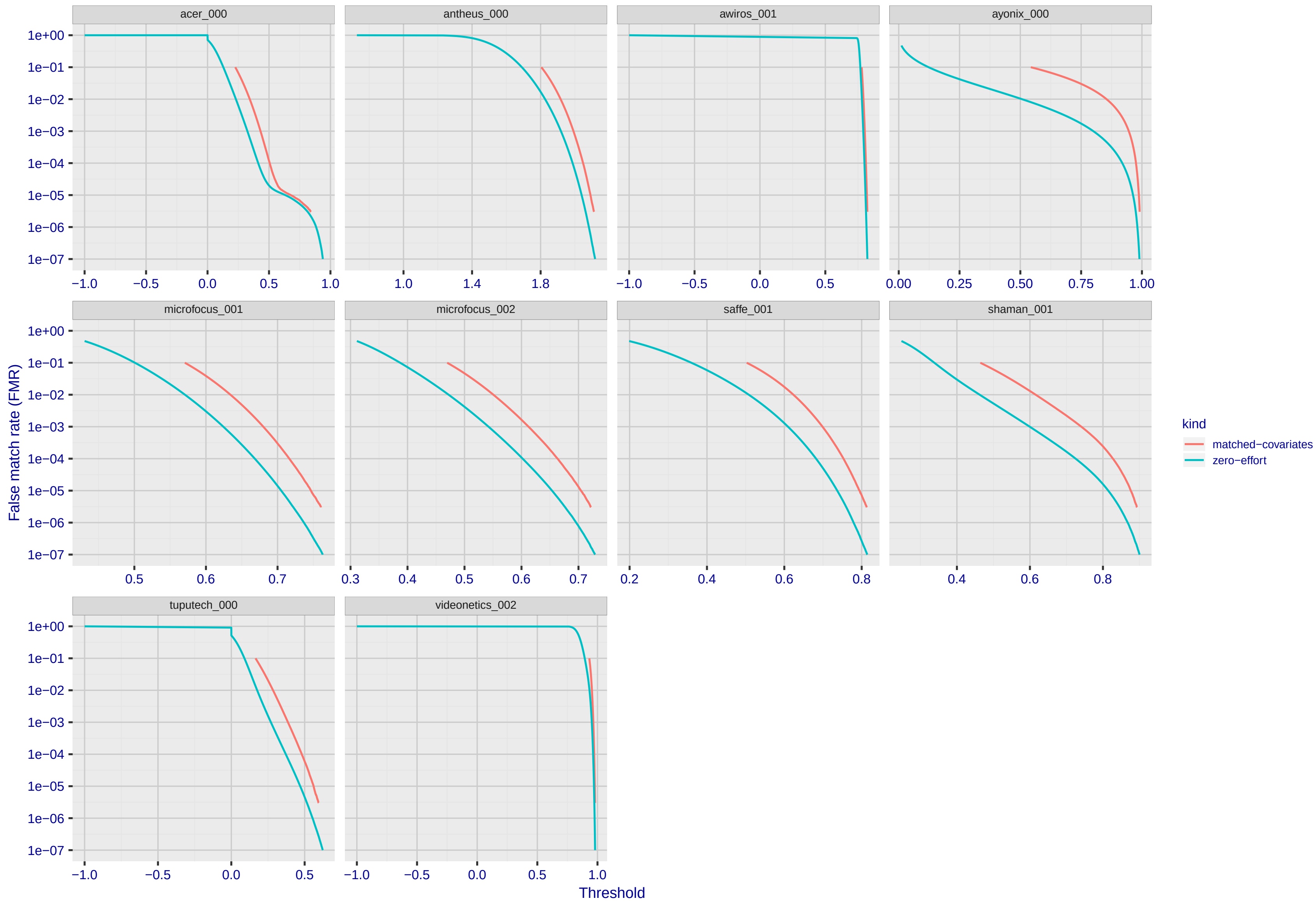


Figure 132: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

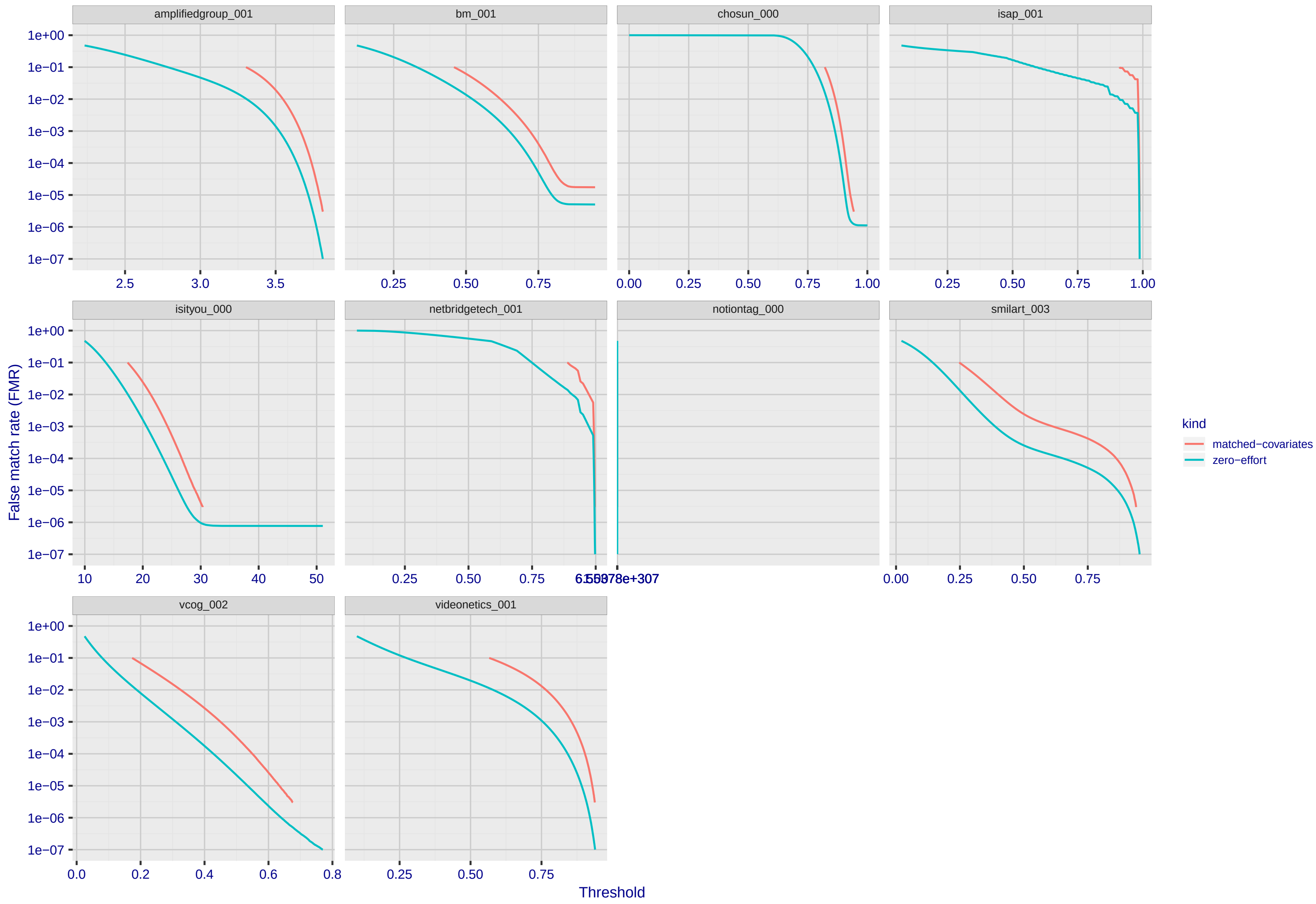


Figure 133: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

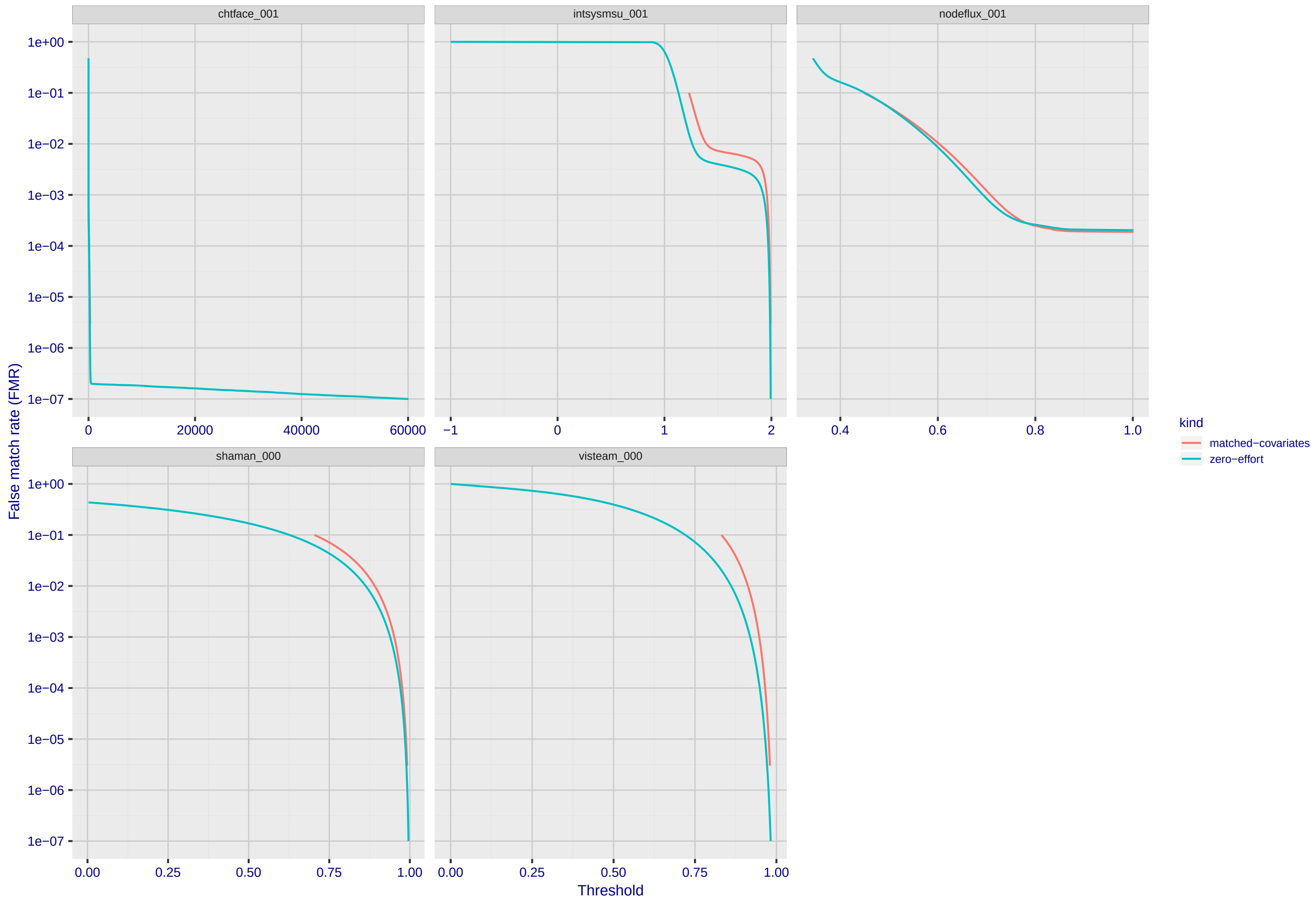


Figure 134: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

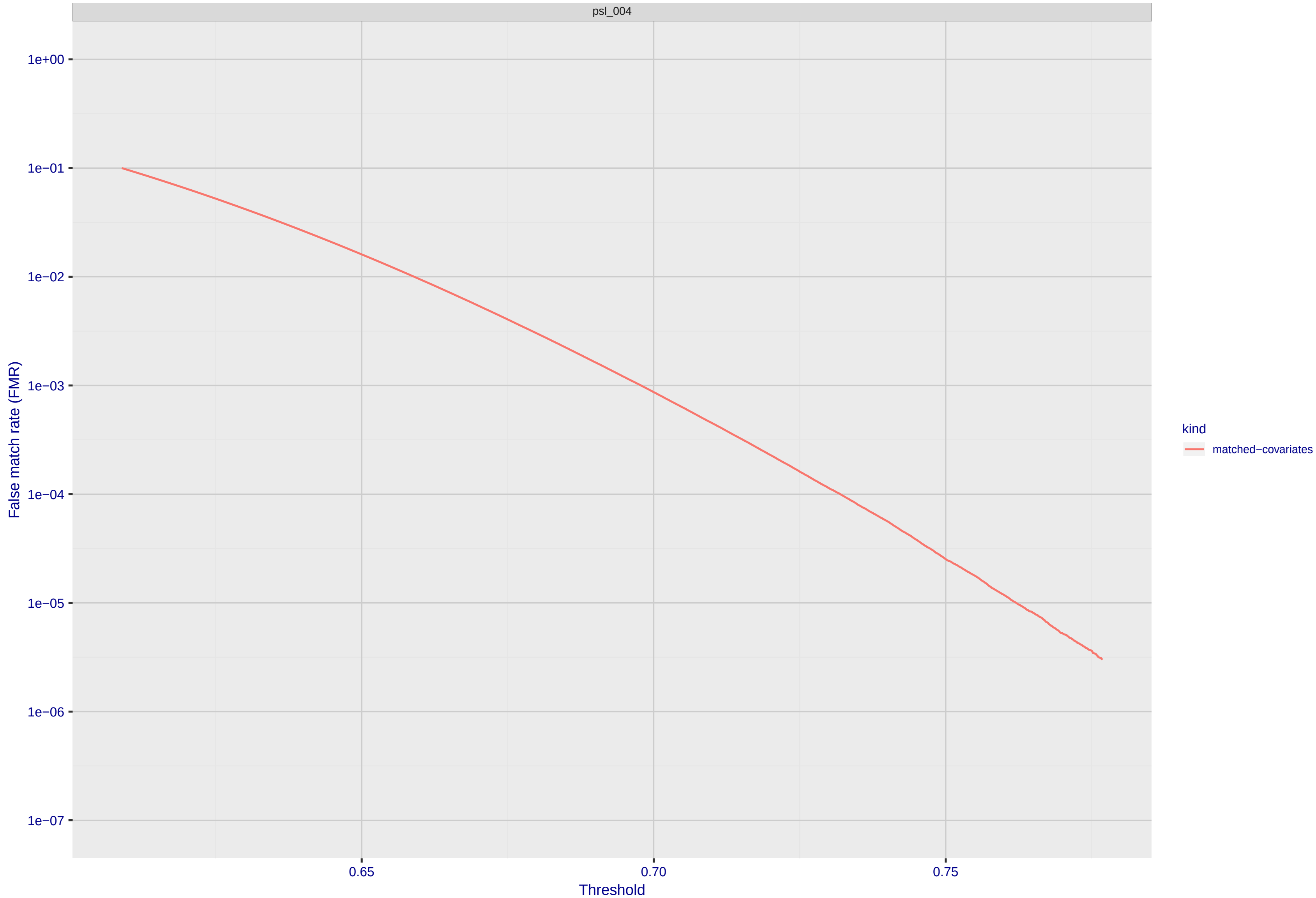


Figure 135: For the visa images, the false match calibration curves show FMR vs. threshold, T. The blue (lower) curves are for zero-effort impostors (i.e. comparing all images against all). The red (upper) curves are for persons of the same-sex, same-age, and same national-origin. This shows that FMR is underestimated (by a factor of 10 or more) by using a zero-effort impostor calculation to calibrate T. As shown later (sec. 3.6), FMR is higher for demographic-matched impostors.

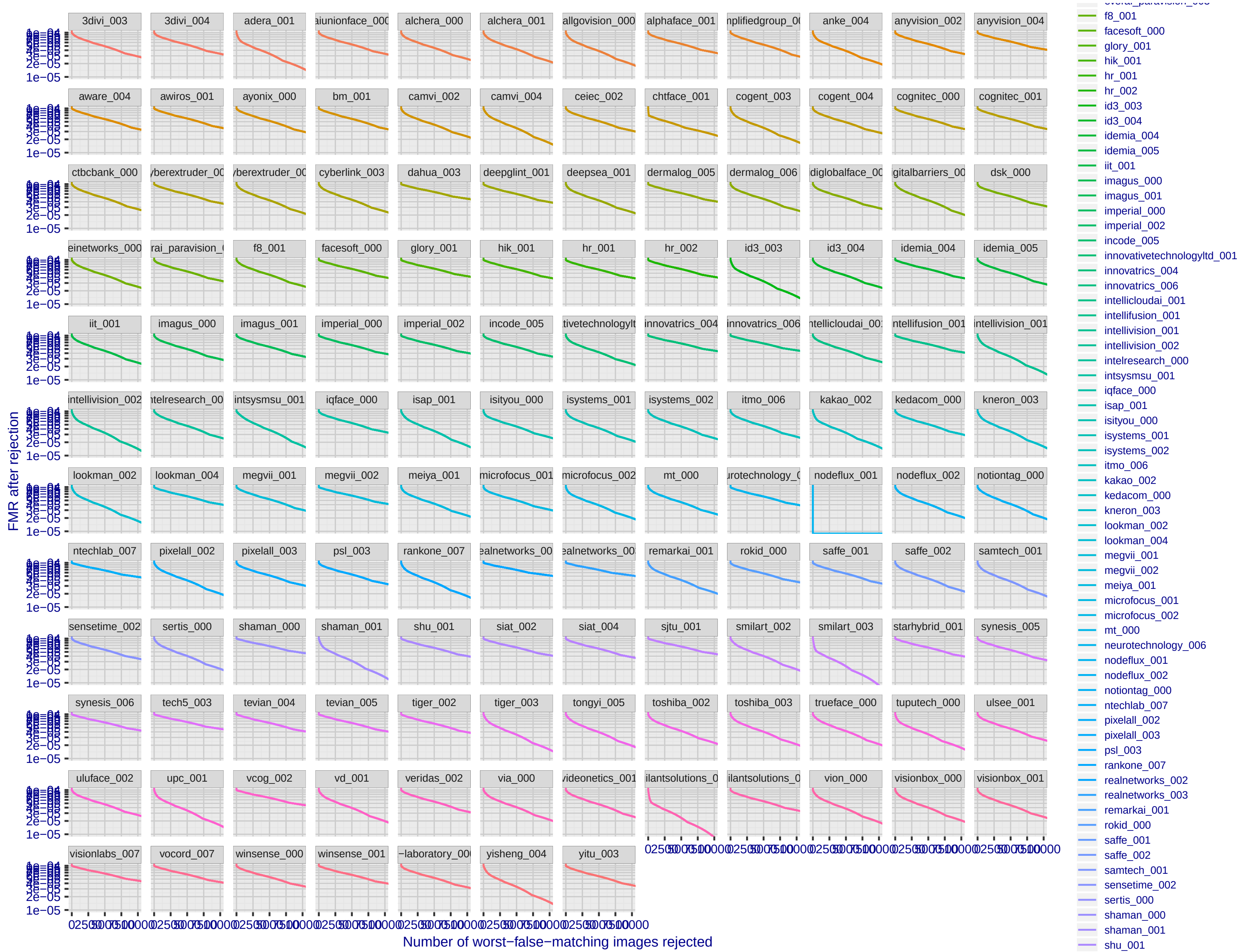


Figure 136: For the visa images, the curves show how false matches are concentrated in certain images. Specifically each line plots FMR(k) with k the number of images rejected in decreasing order of how many false matches that image was involved in. FMR$(0) = 10^{-4}$. In terms of the biometric zoo, the most "wolf-ish" images are rejected first i.e. those enrollment or verification images most often involved in false matches. A flatter response is considered superior. A steeply descending response indicates that certain kinds of images false match against others, e.g. if hypothetically images of men with particular mustaches would falsely match others.

3.5 Genuine distribution stability

3.5.1 Effect of birth place on the genuine distribution

Background: Both skin tone and bone structure vary geographically. Prior studies have reported variations in FNMR and FMR.

Goal: To measure false non-match rate (FNMR) variation with country of birth.

Methods: Thresholds are determined that give FMR = $\{0.001, 0.0001\}$ over the entire impostor set. Then FNMR is measured over 1000 bootstrap replications of the genuine scores. Only those countries with at least 140 individuals are included in the analysis.

Results: Figure 153 shows FNMR by country of birth for the two thresholds.

Figure 137: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

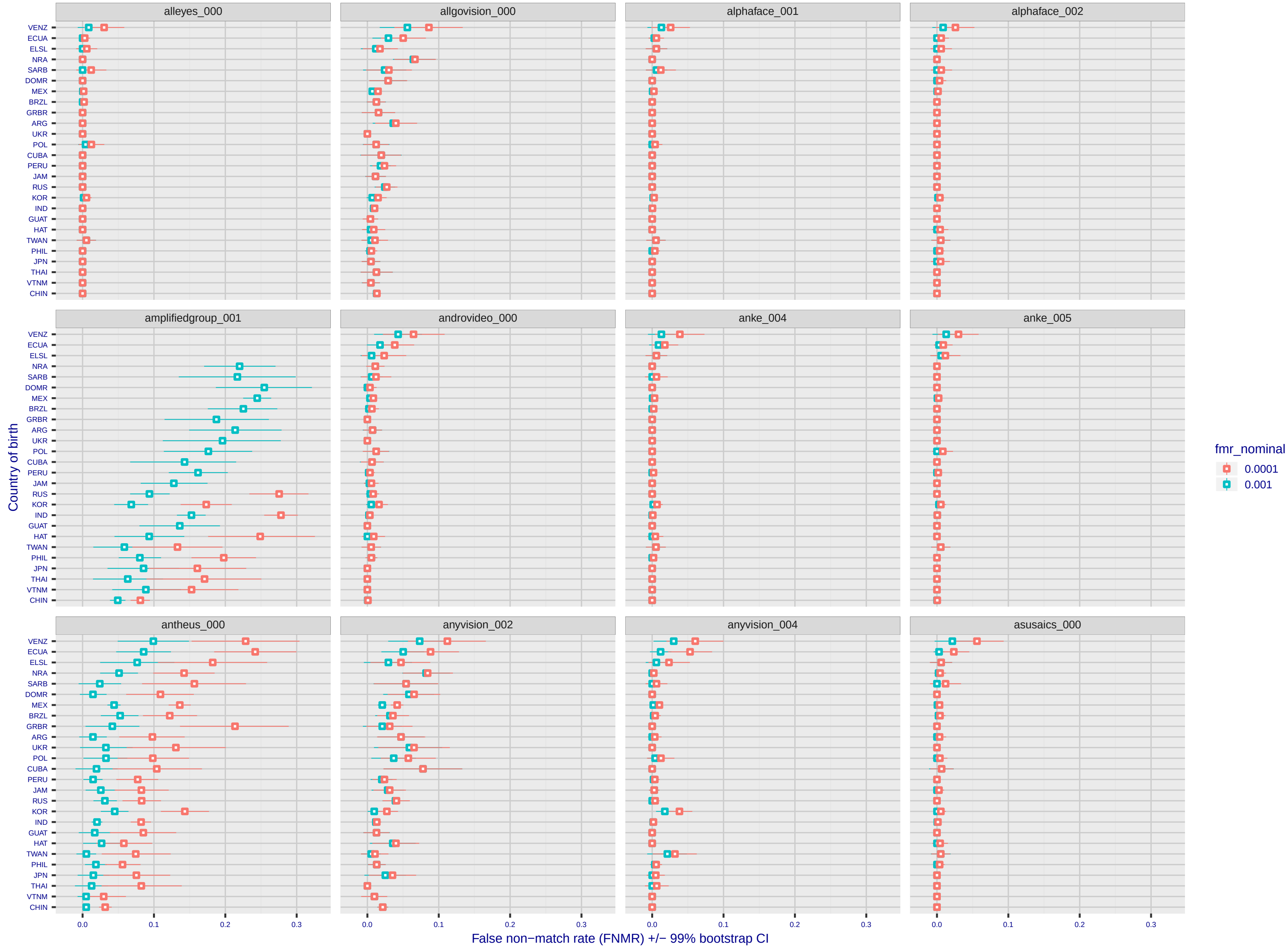


Figure 138: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

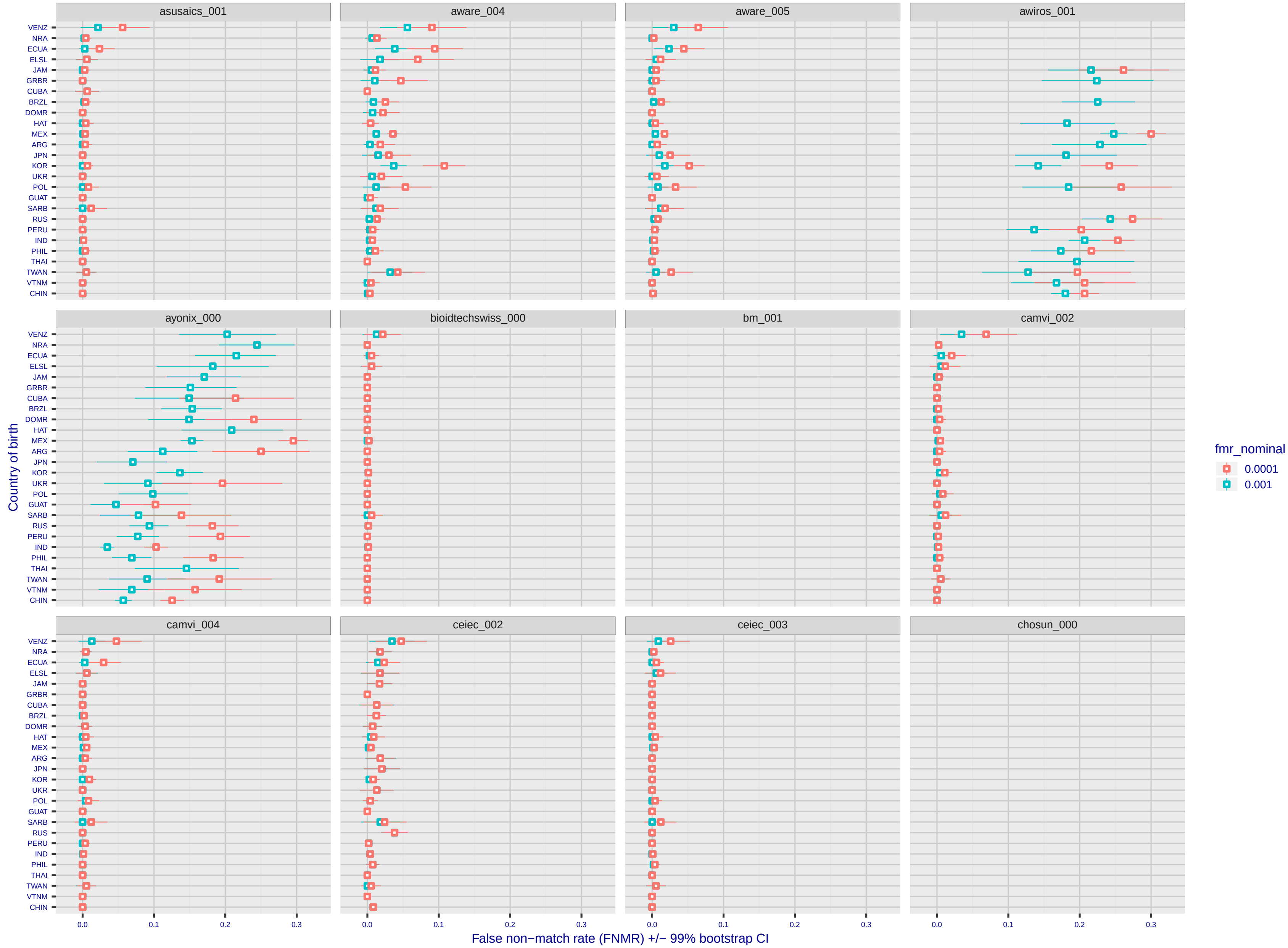


Figure 139: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

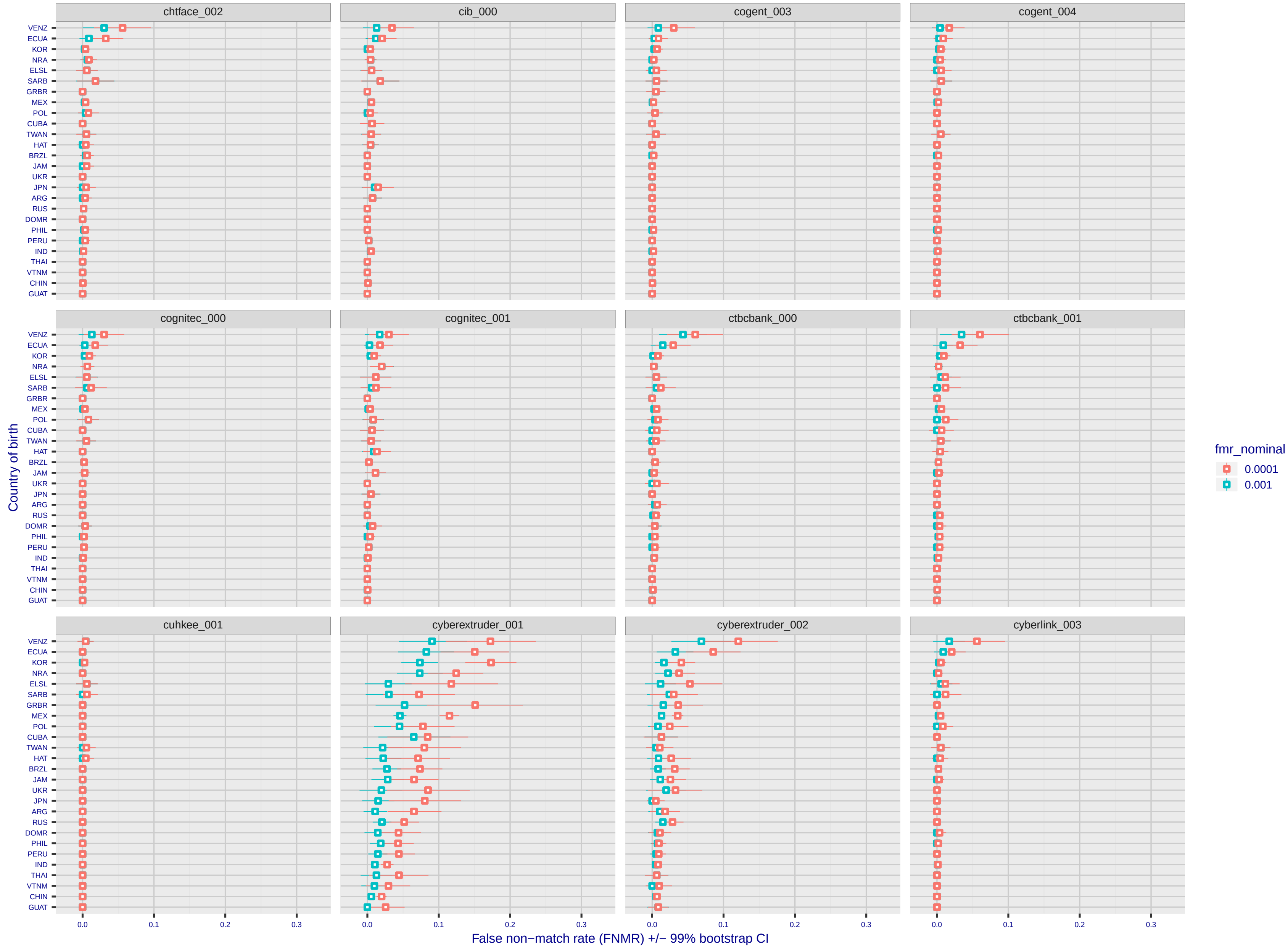


Figure 140: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

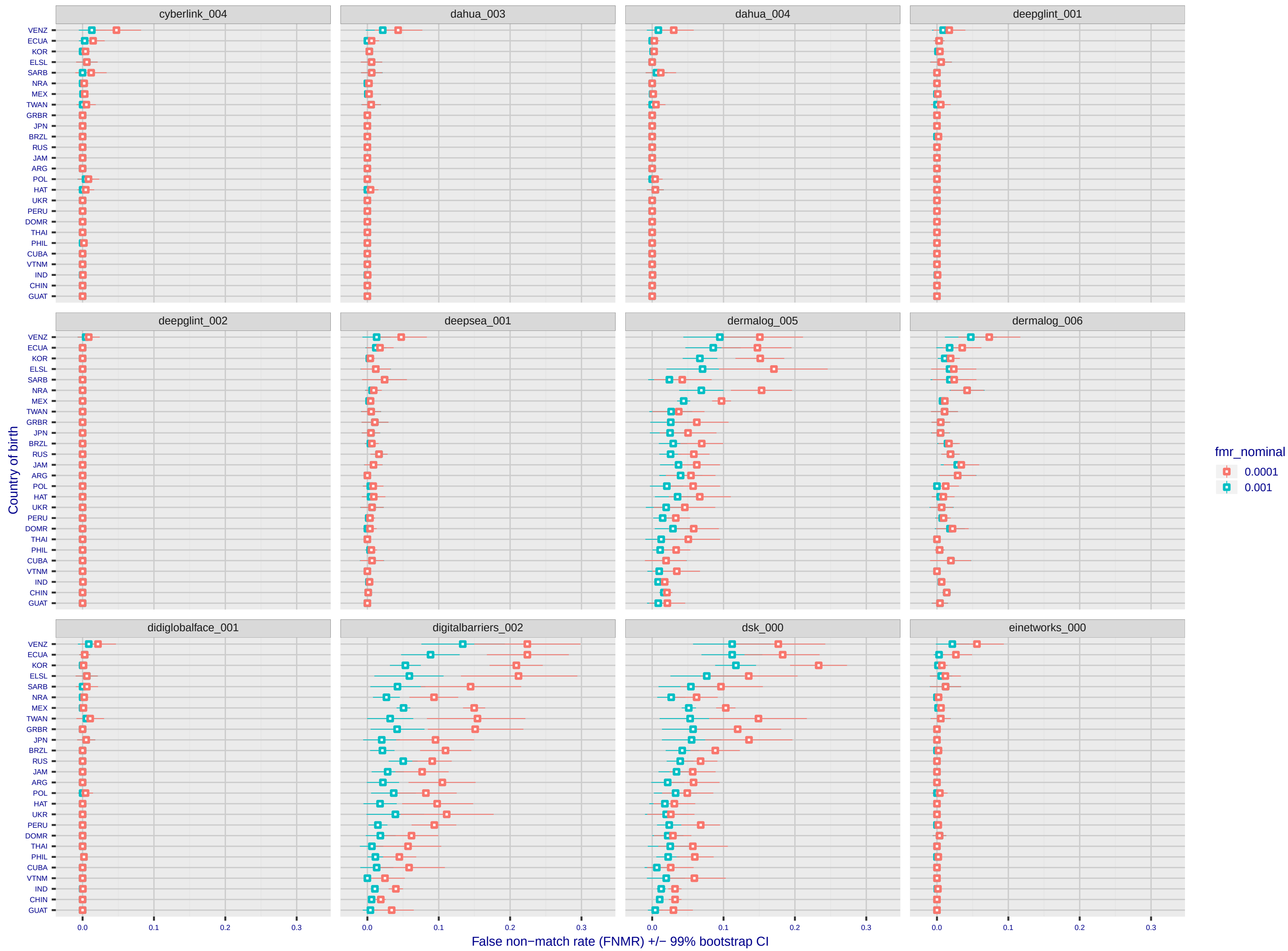


Figure 141: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

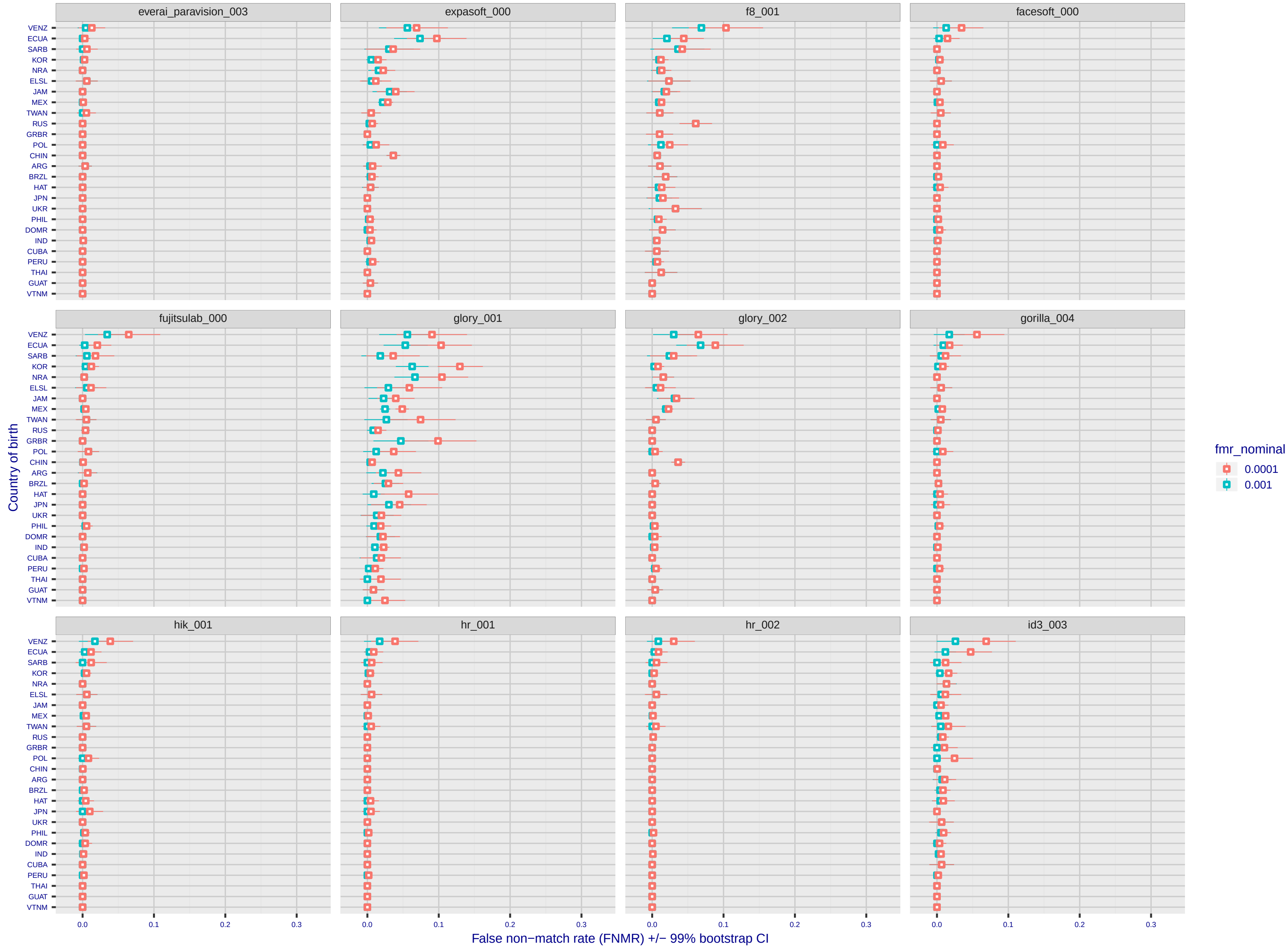


Figure 142: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

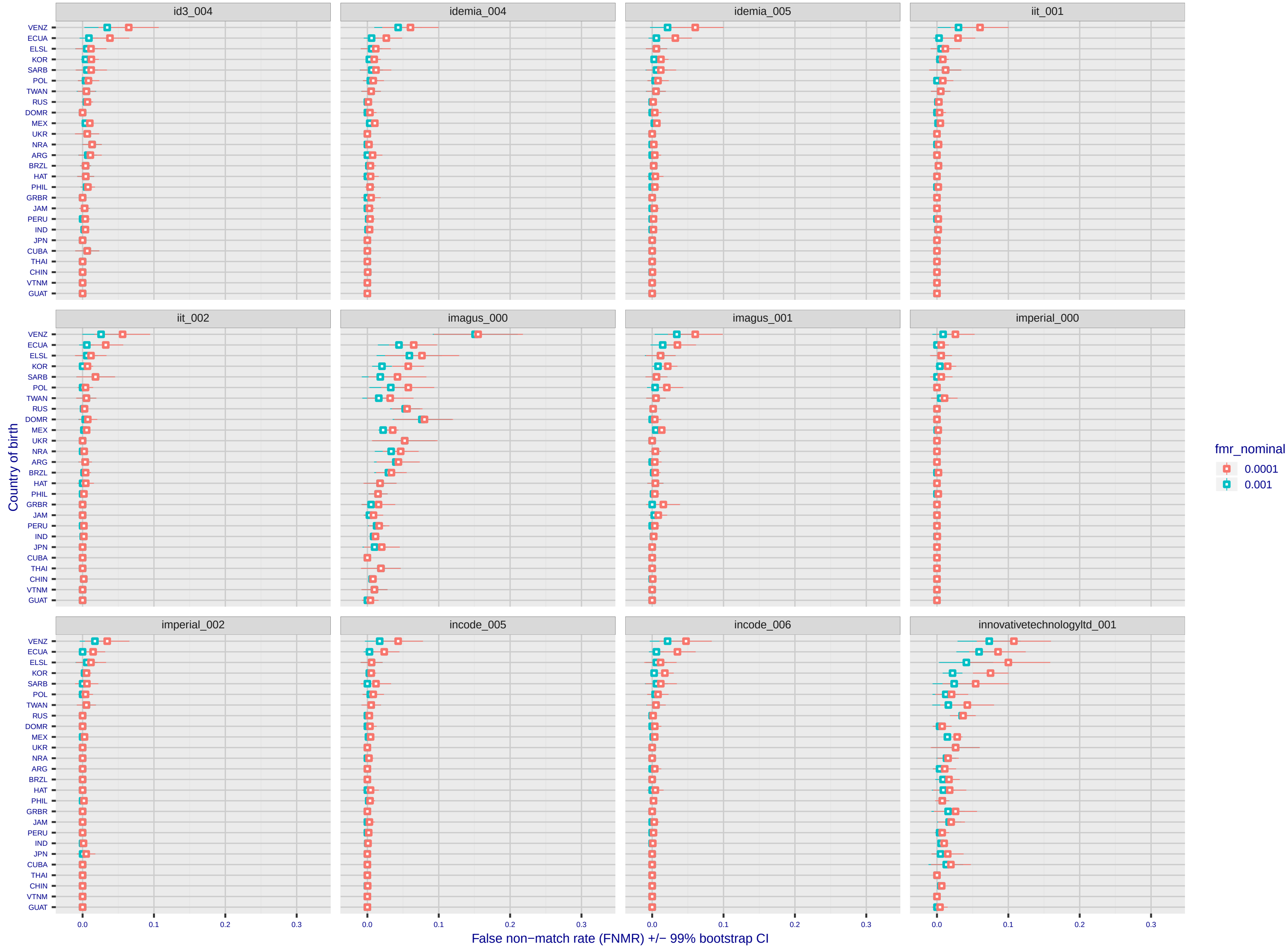


Figure 143: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

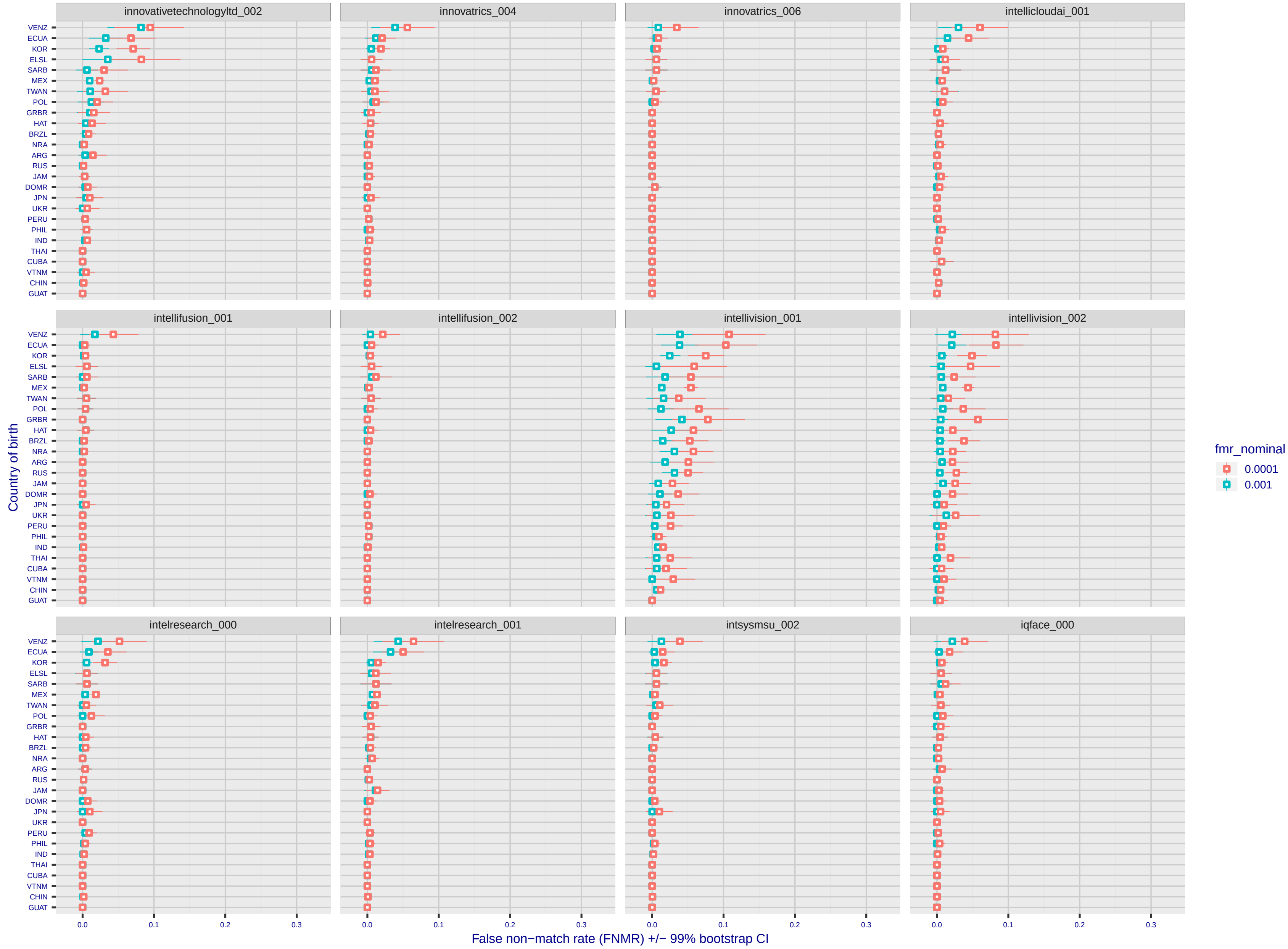


Figure 144: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

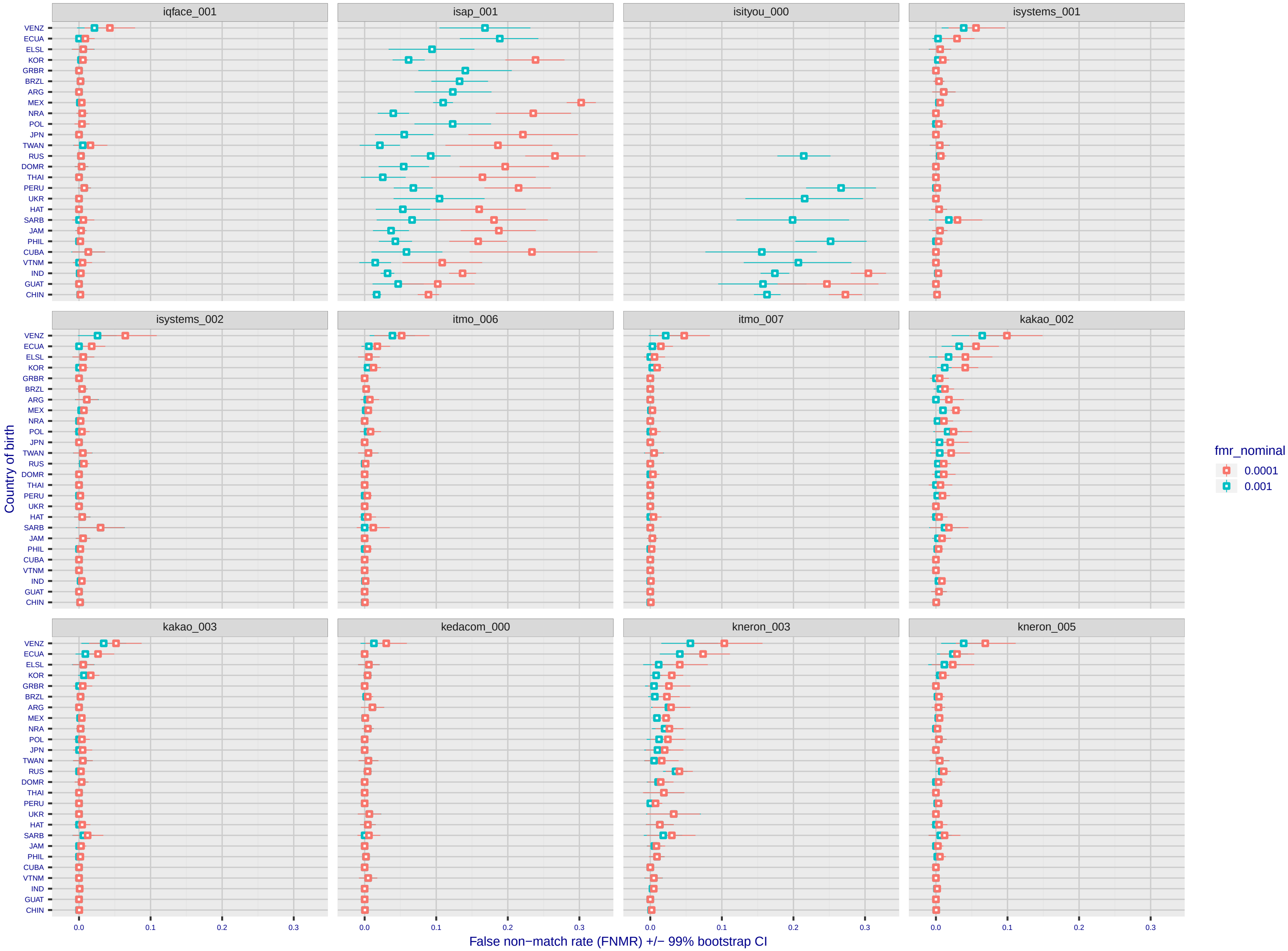


Figure 145: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

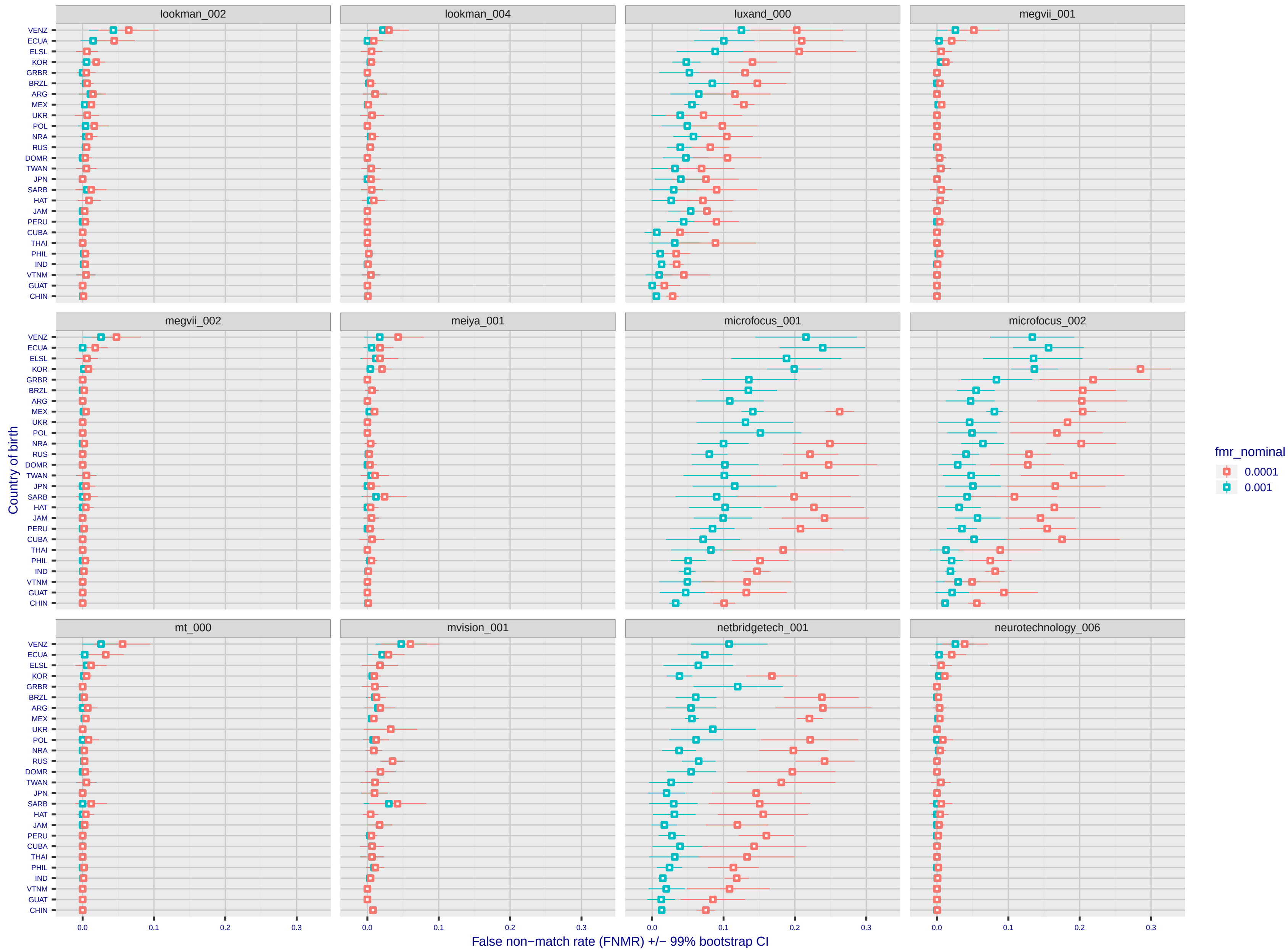


Figure 146: *For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.*

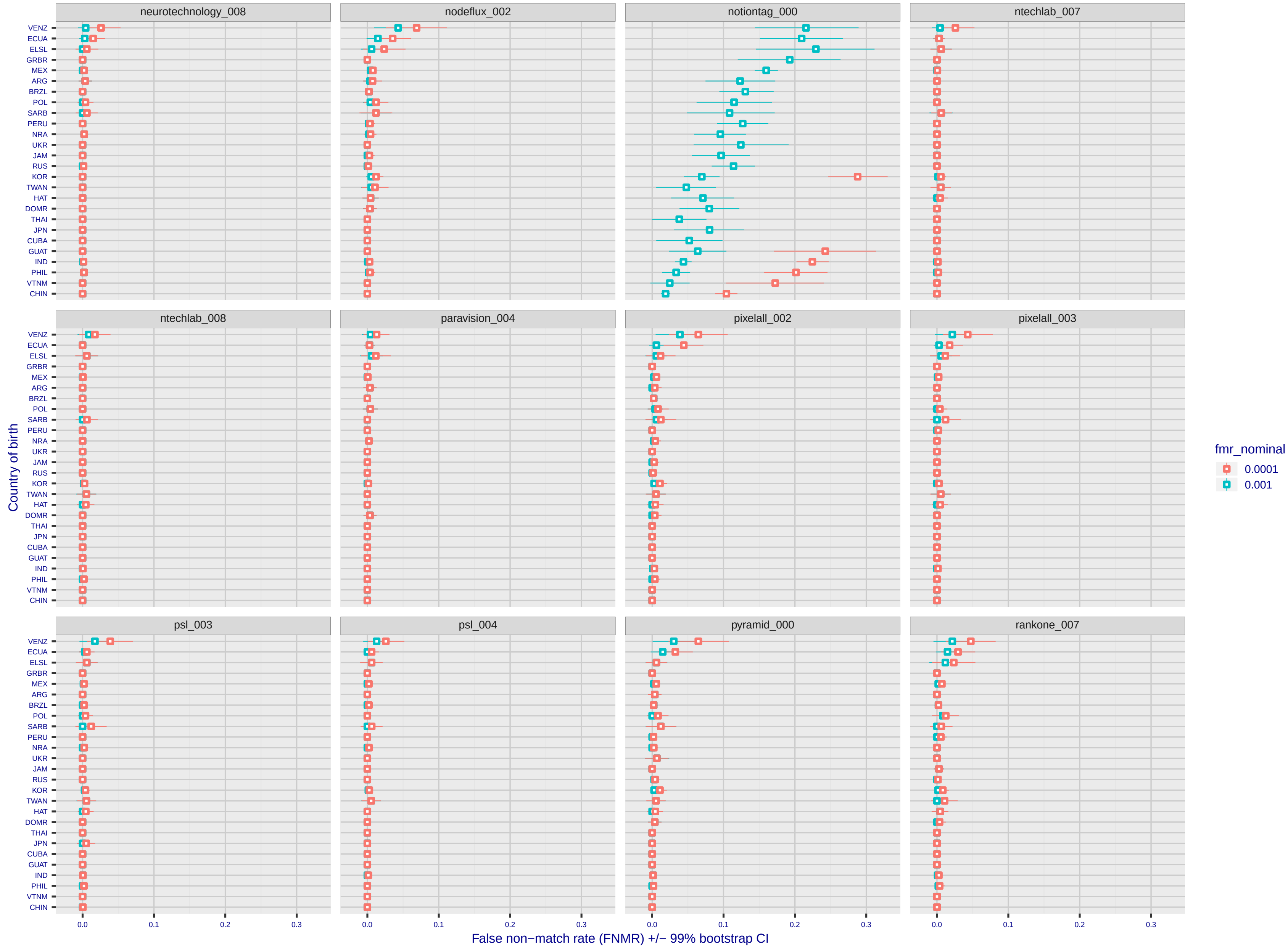


Figure 147: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

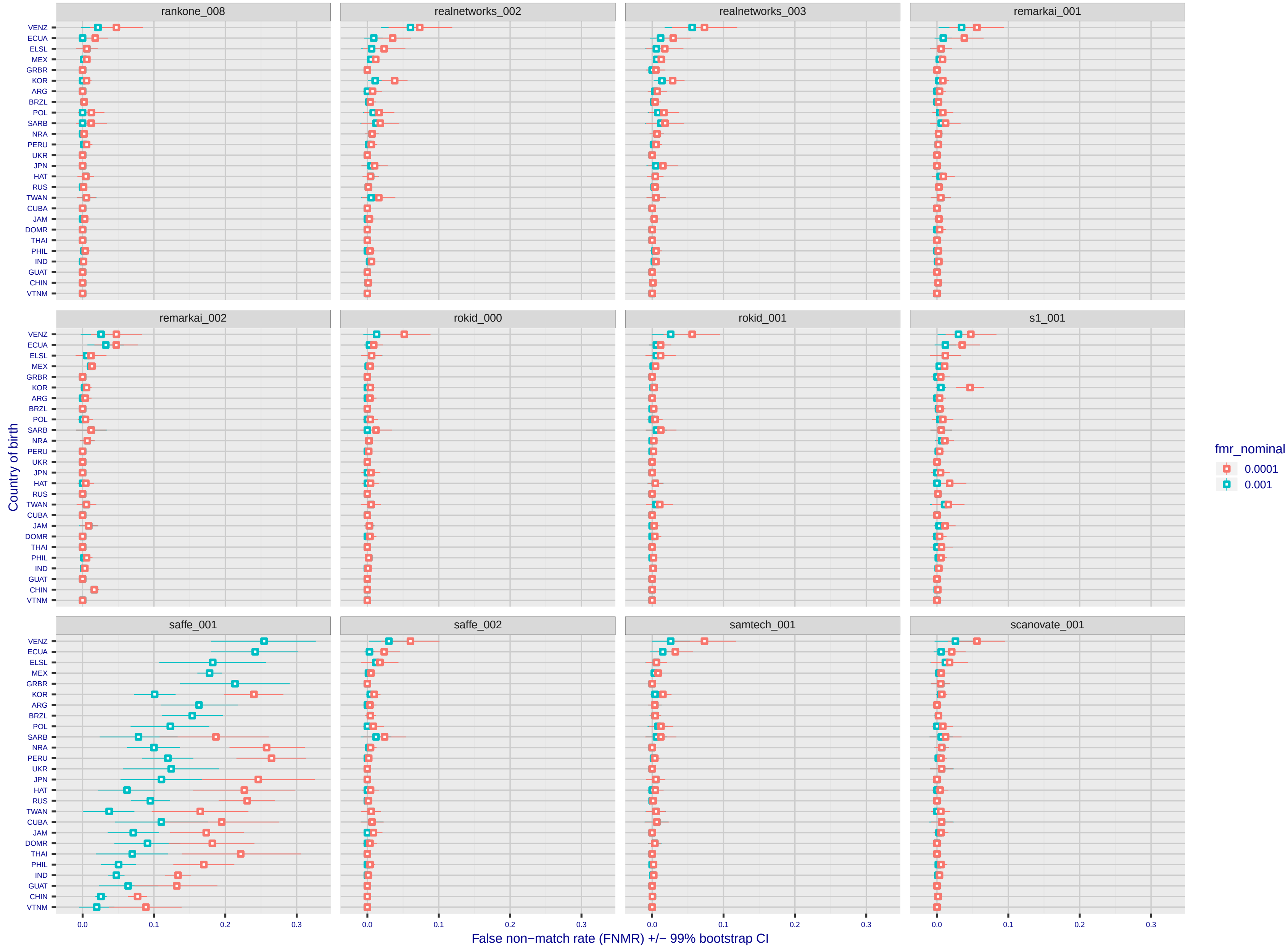


Figure 148: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

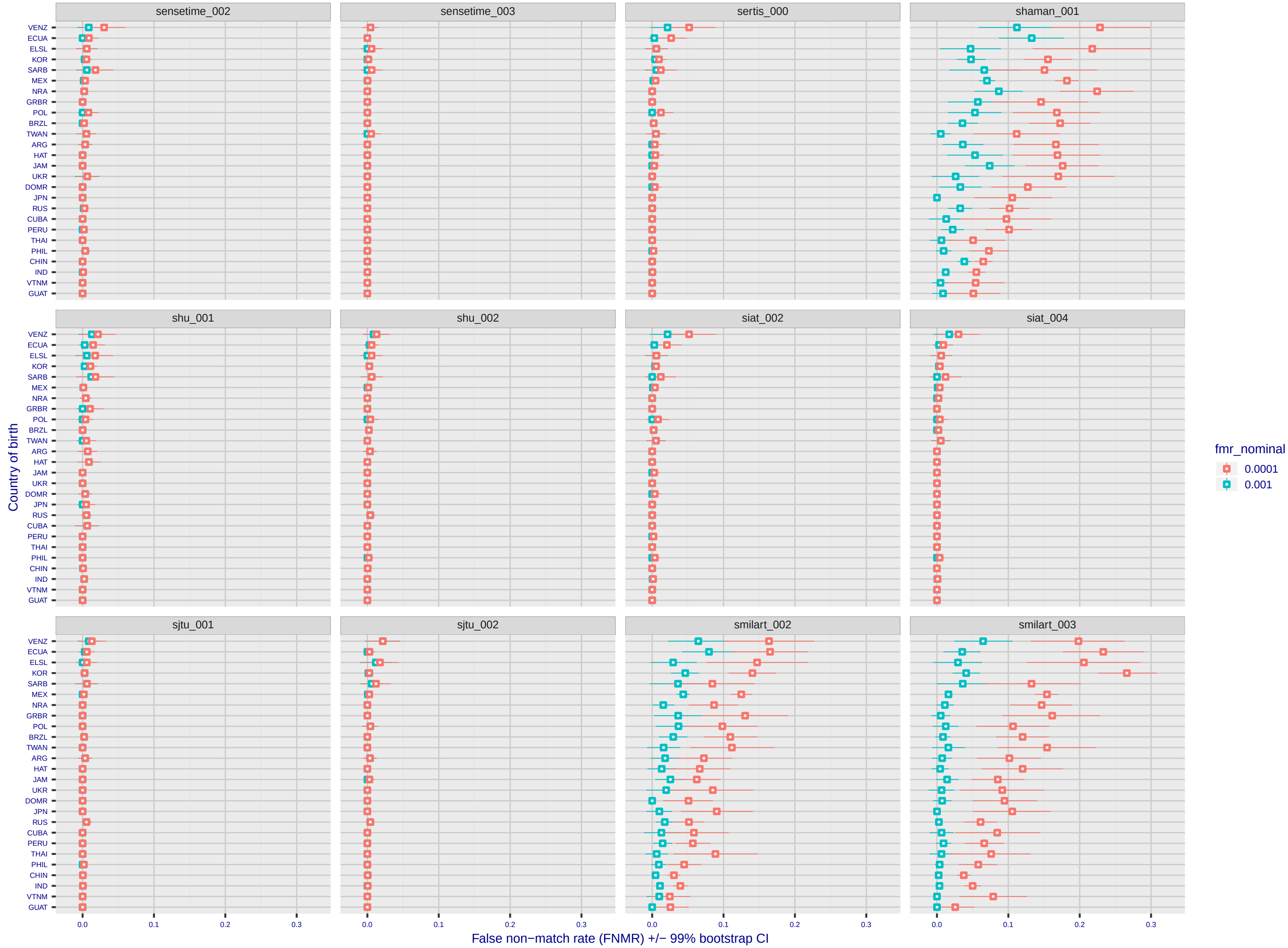


Figure 149: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

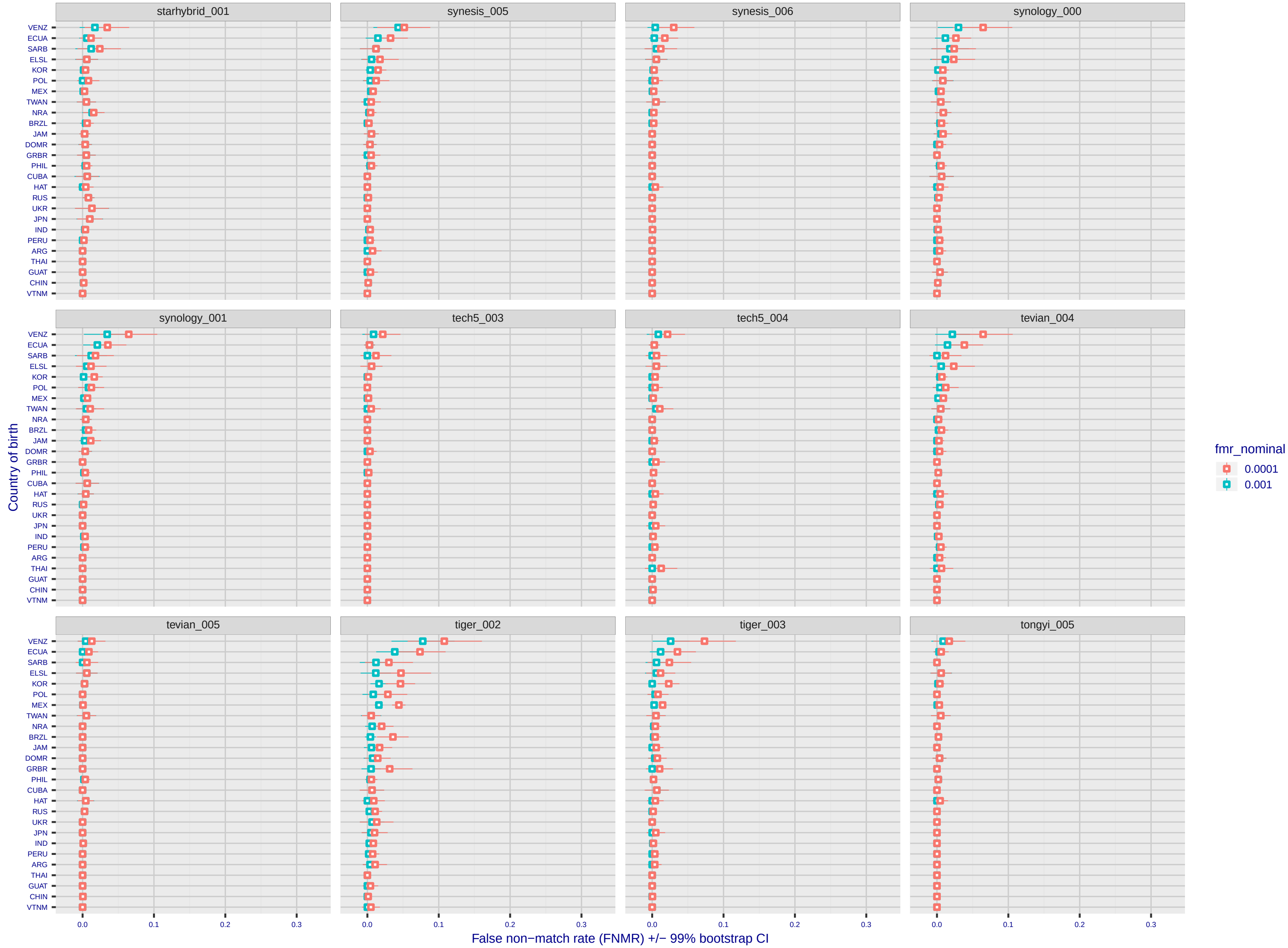


Figure 150: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

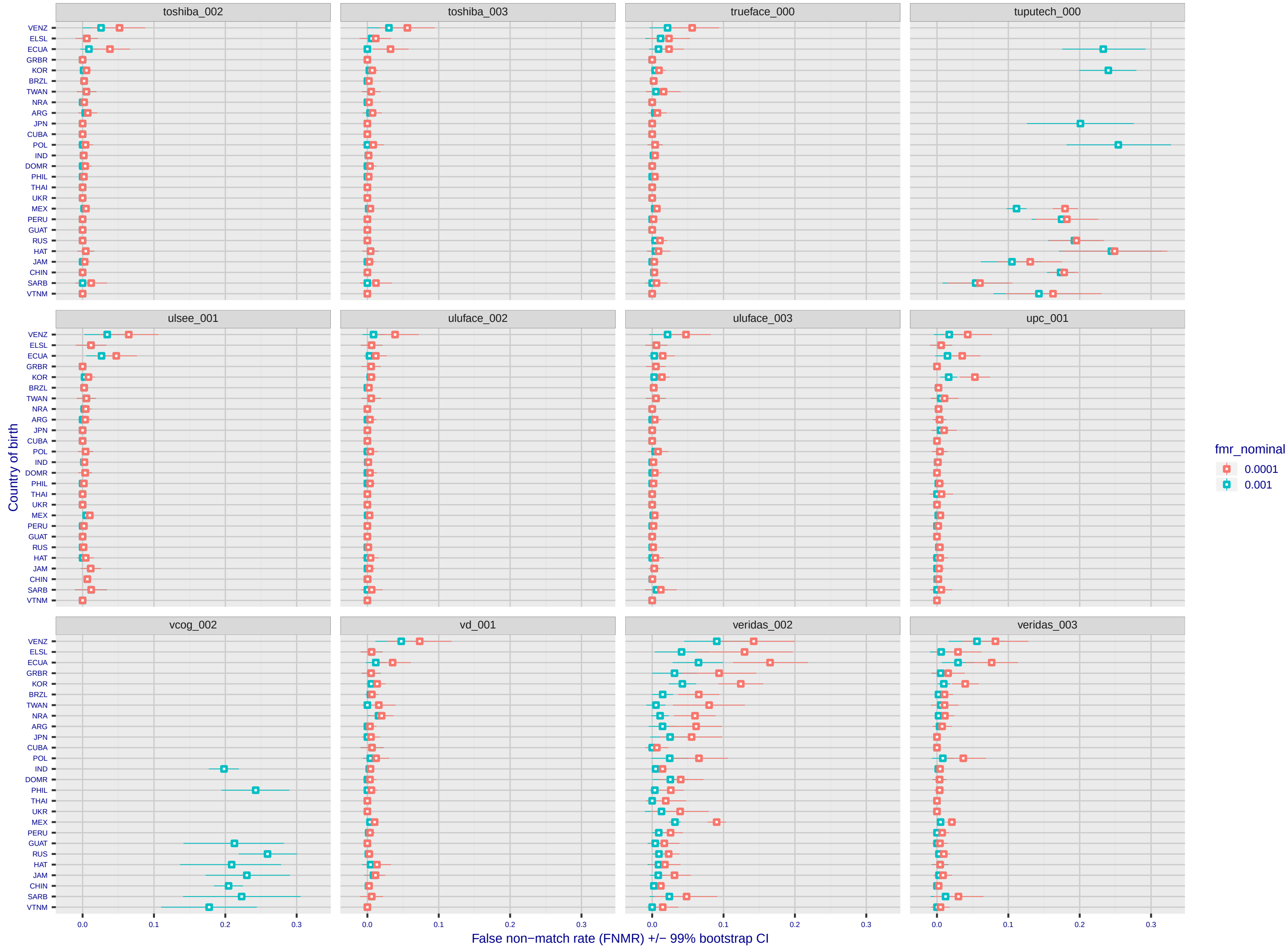


Figure 151: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

Figure 152: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

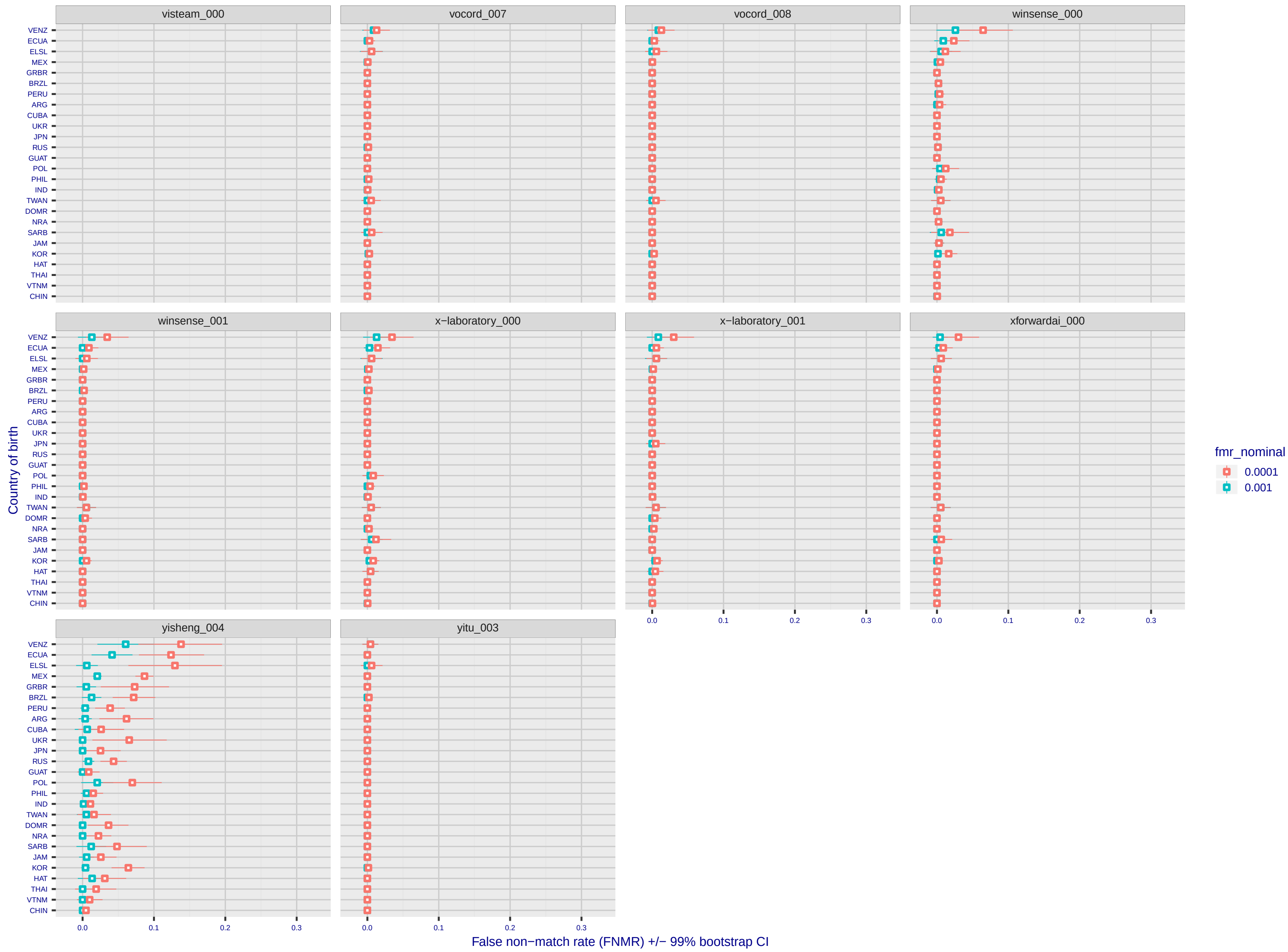


Figure 153: For the visa images, the dots show FNMR by country of birth for two globally set operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.1. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are likely due quality variations, then demographics like age and race. The error rates in some cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

Caveats: The results may not relate to subject-specific properties. Instead they could reflect image-specific quality differences, which could occur due to collection protocol or software processing variations.

3.5.2 Effect of ageing

Background: Faces change appearance throughout life. This change gradually reduces similarity of a new image to an earlier image. Face recognition algorithms give reduced similarity scores and more frequent false rejections.

Goal: To quantify false non-match rates (FNMR) as a function of elapsed time in an adult population.

Methods: Using the mugshot images, a threshold is set to give FMR = 0.00001 over the entire impostor set. Then FNMR is measured over 1000 bootstrap replications of the genuine scores.

Results: For the visa images, Figure 166 shows how false non-match rates for genuine users, as a function of age group.

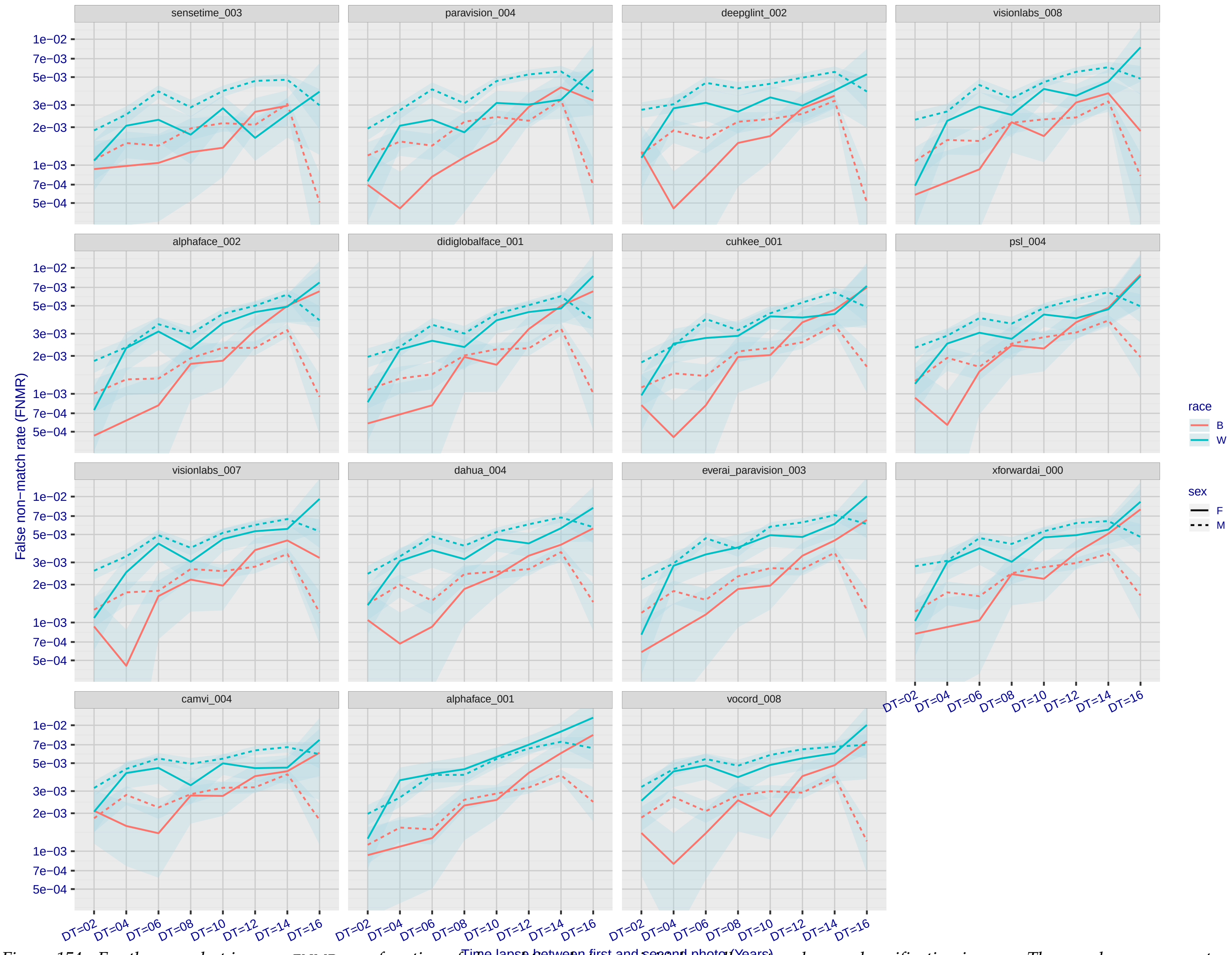


Figure 154: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

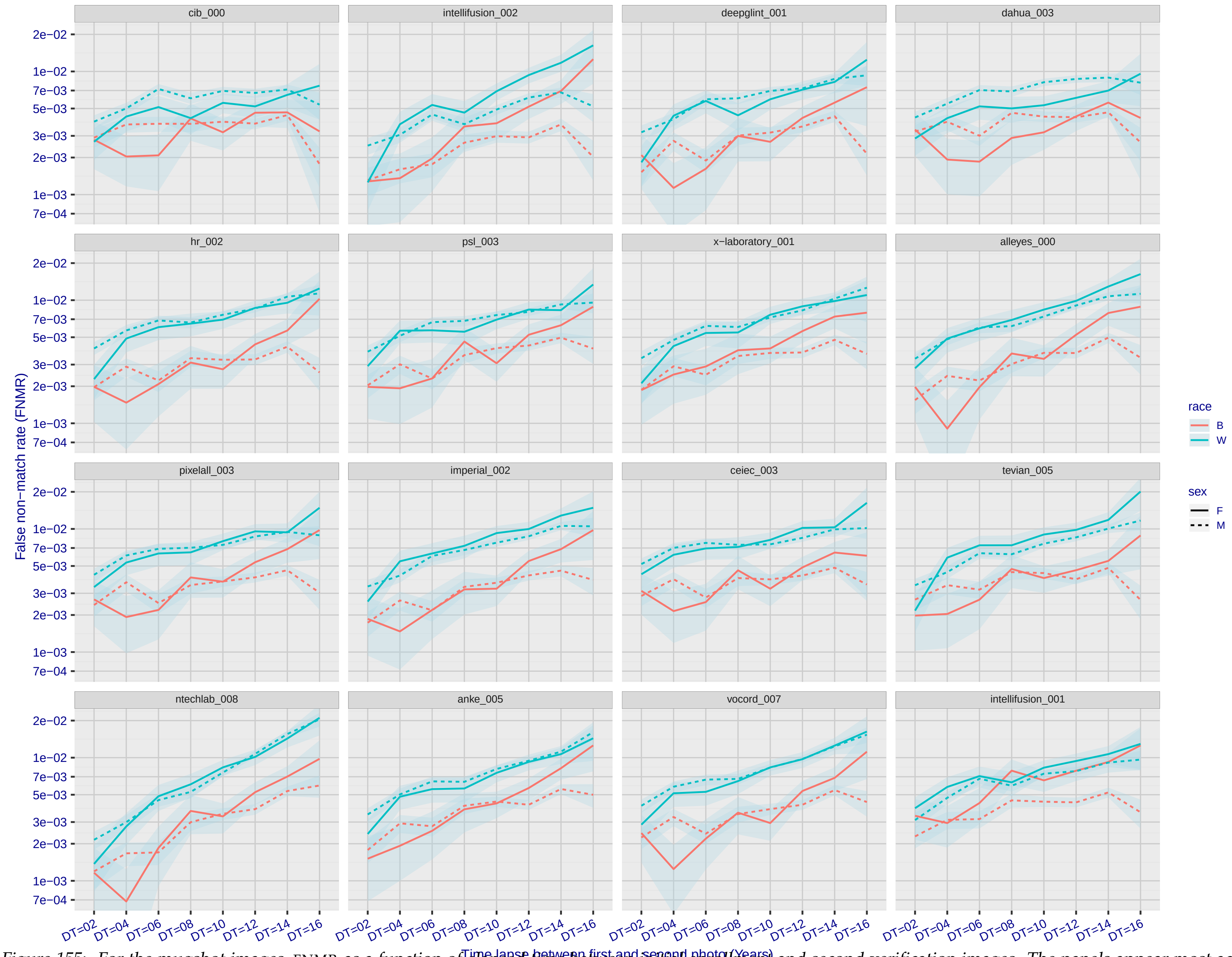


Figure 155: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all* (10^8) *impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR* < 0.001*) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

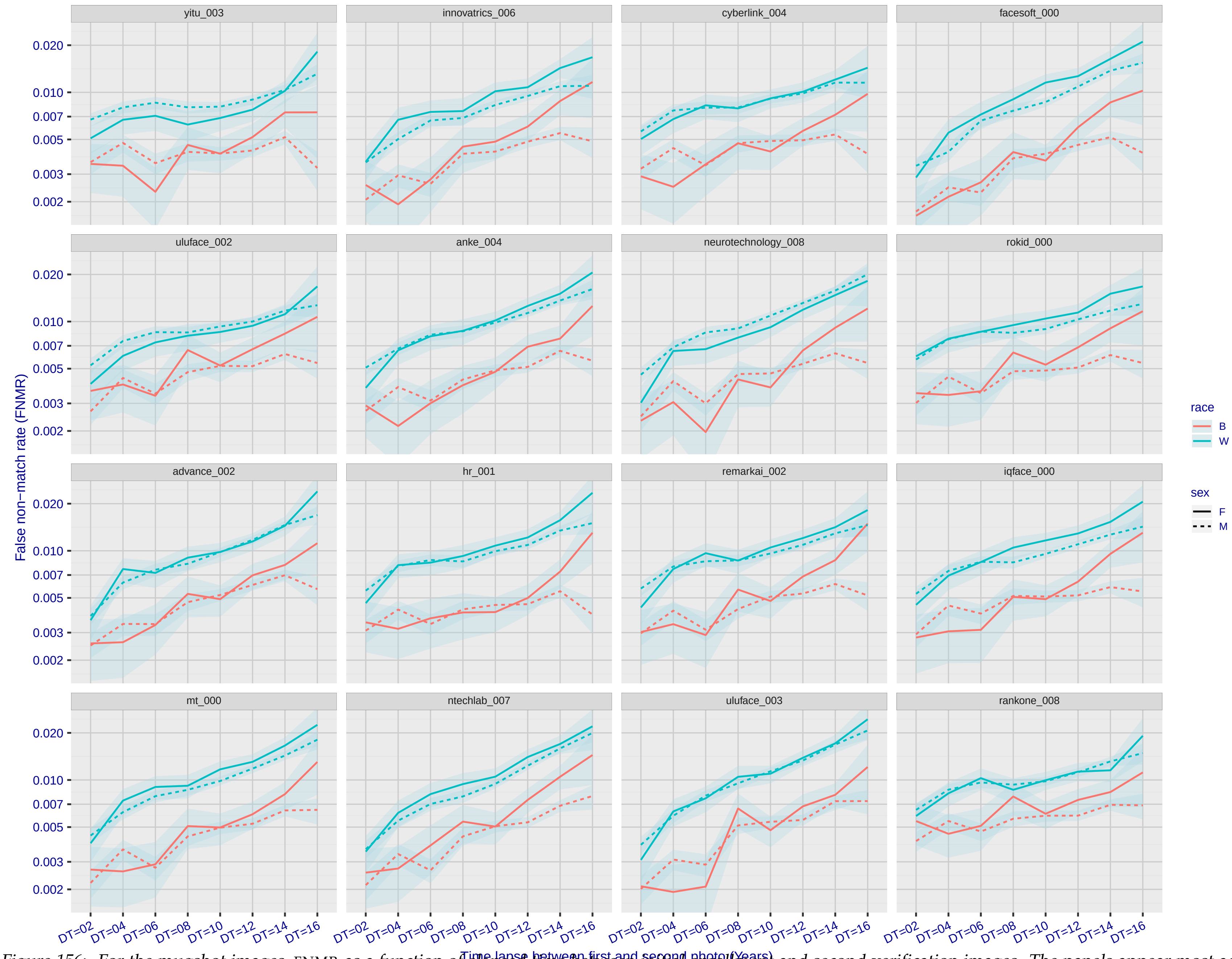


Figure 156: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

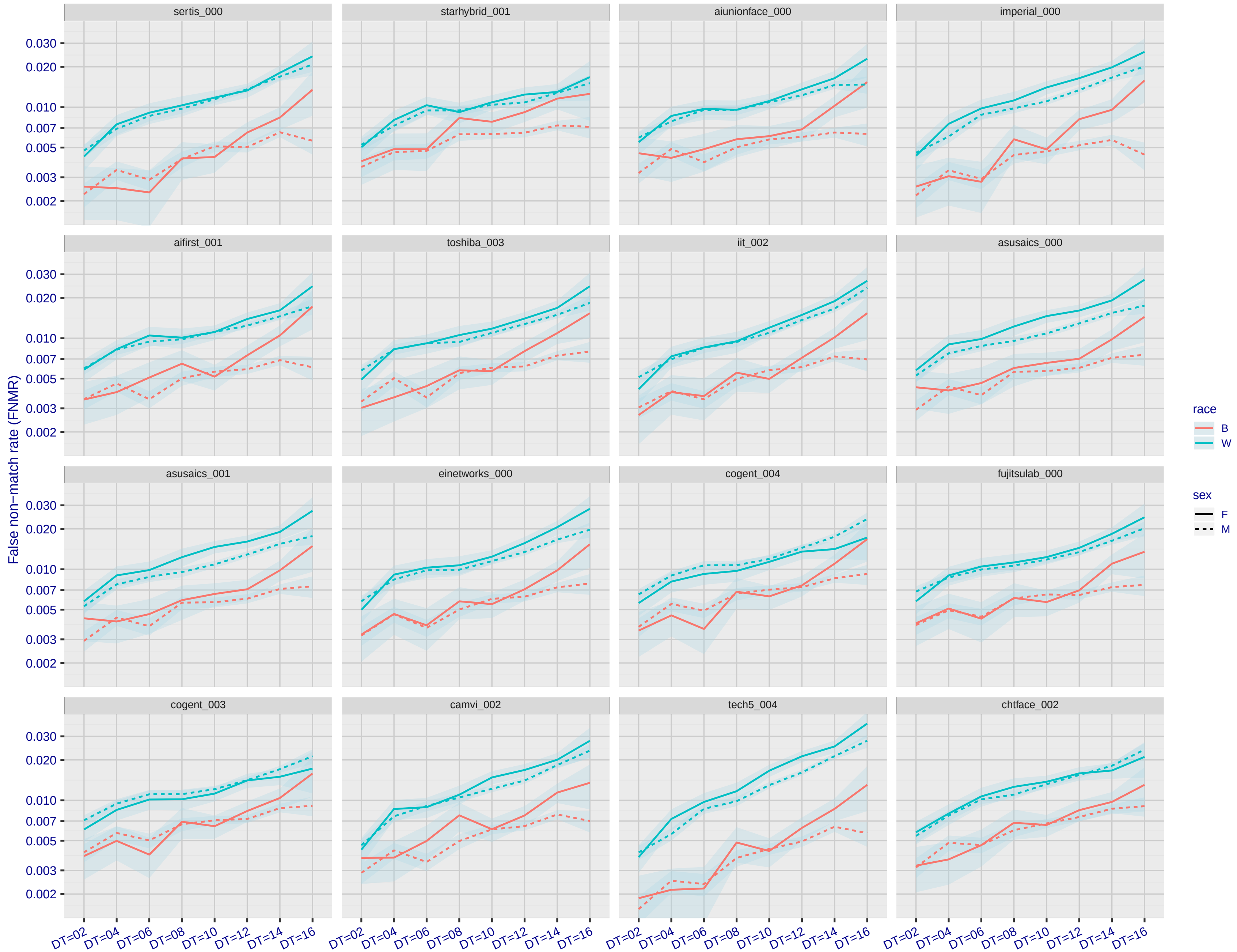


Figure 157: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

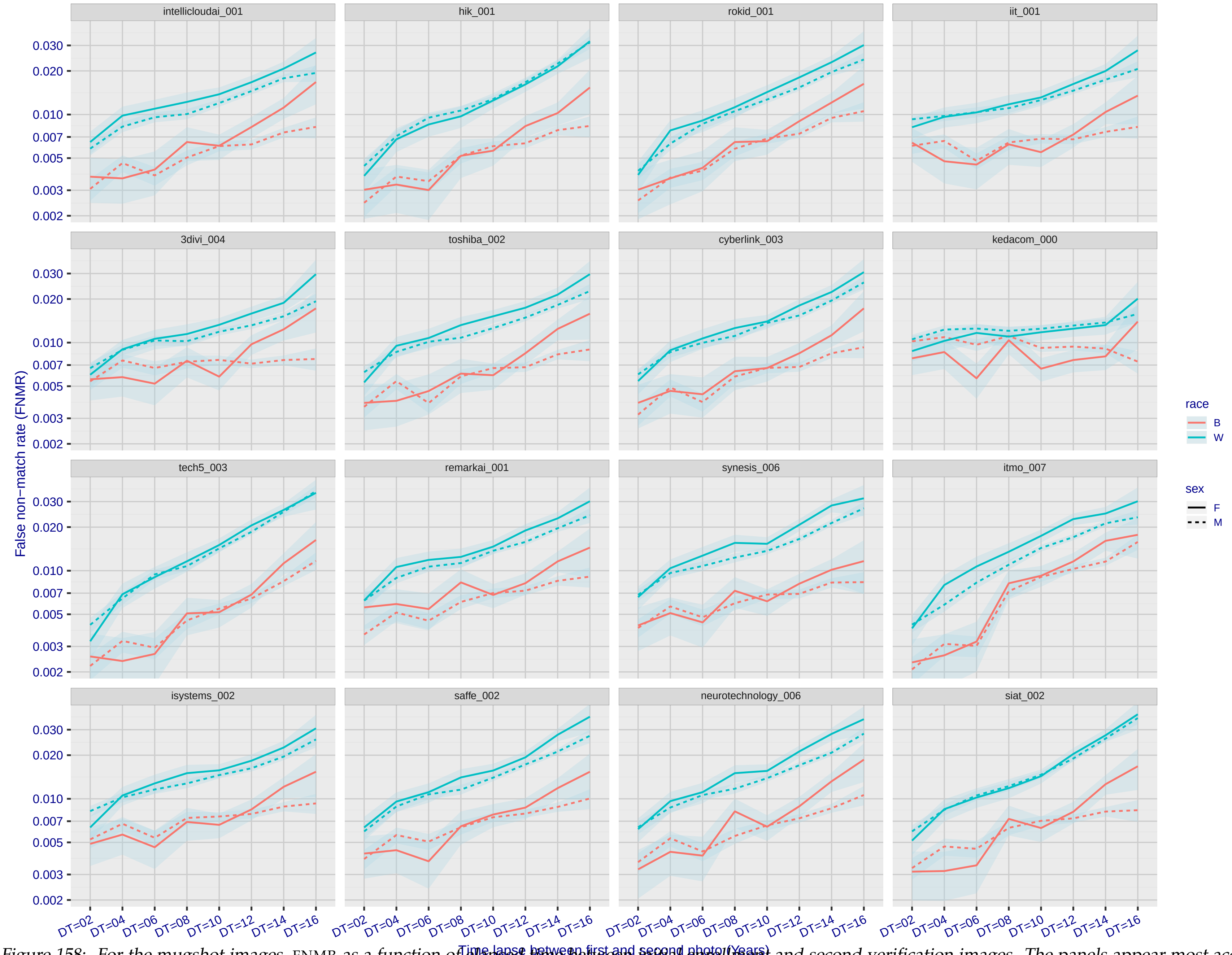


Figure 158: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

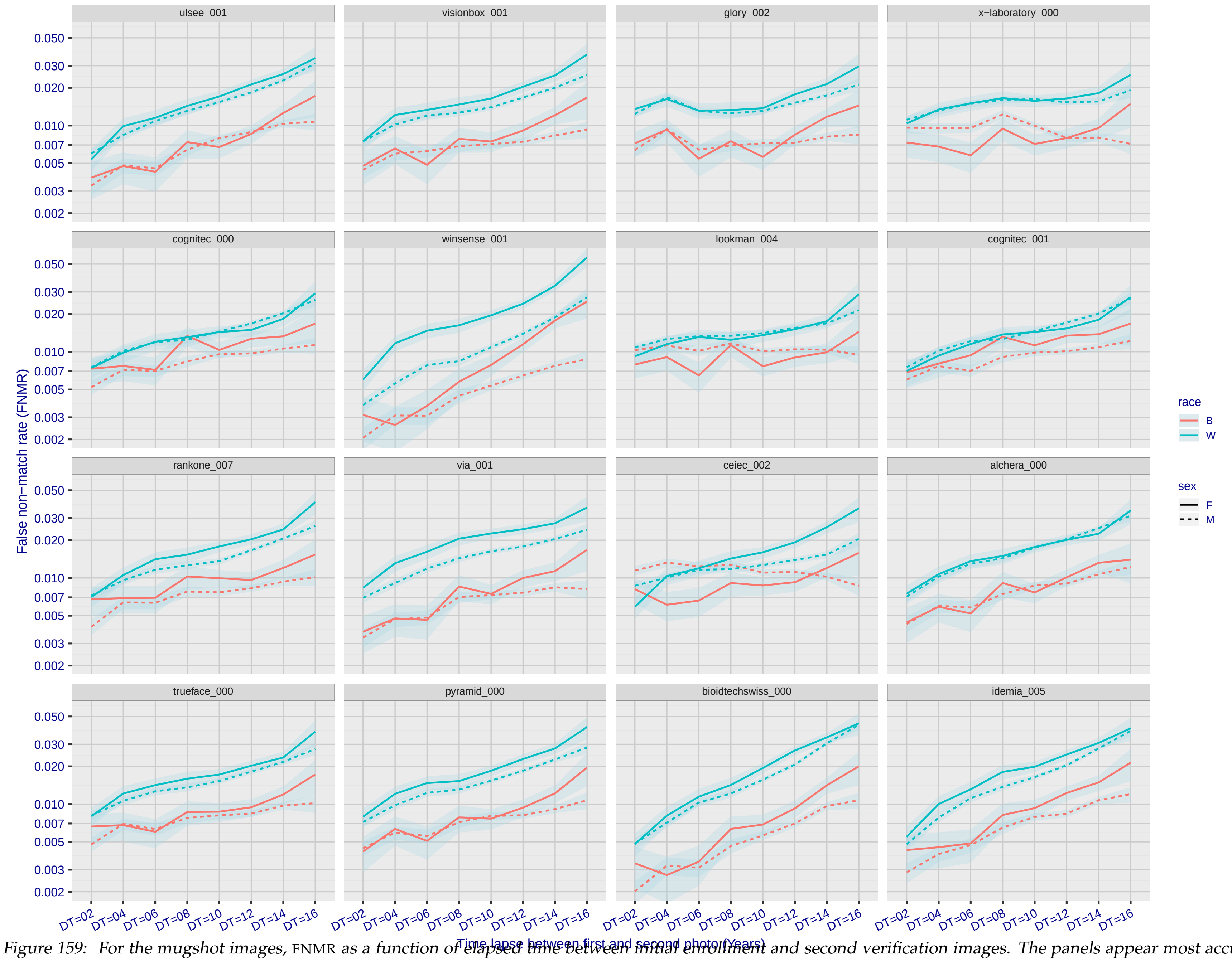


Figure 159: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all* (10^8) *impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR* < 0.001*) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

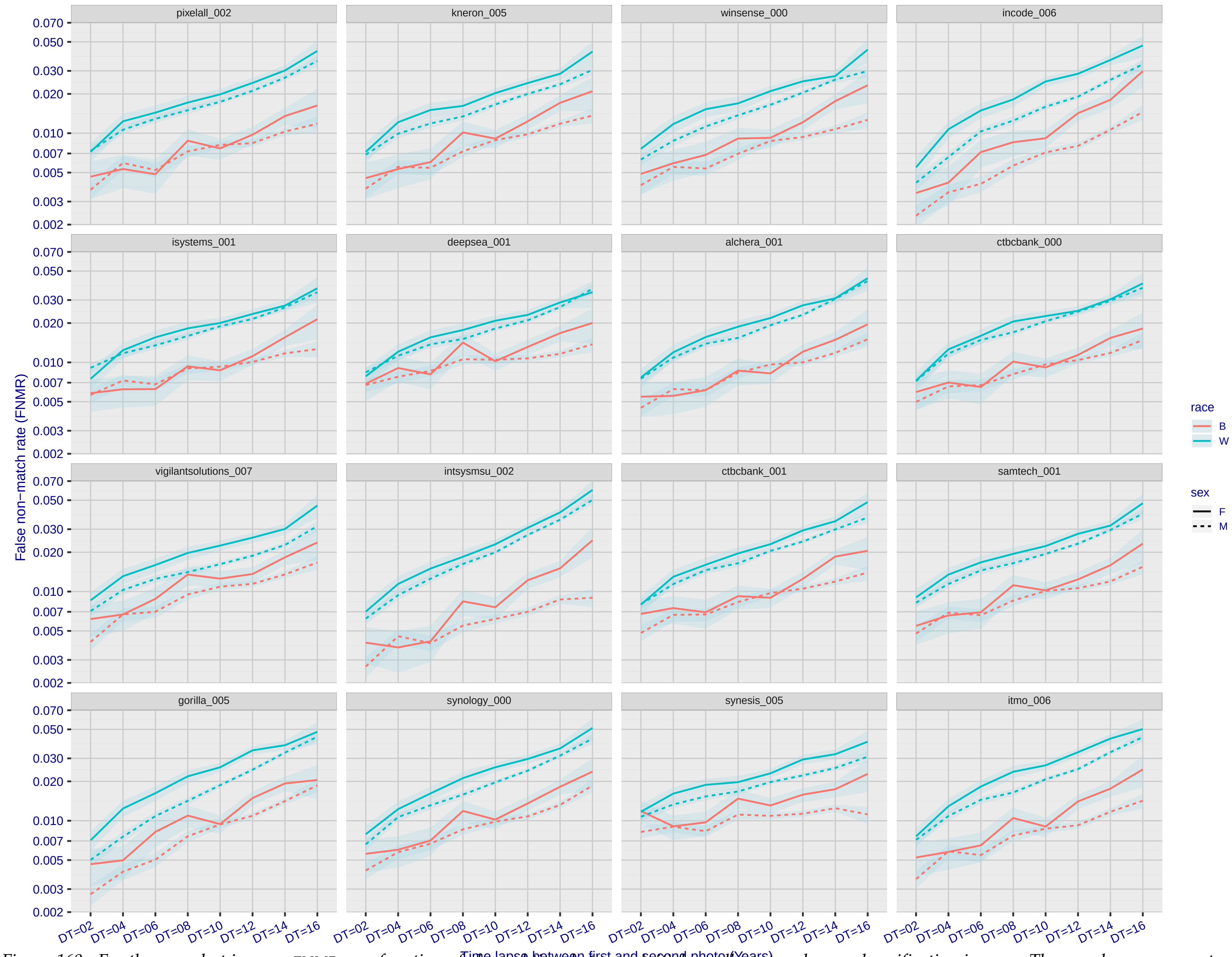


Figure 160: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all* (10^8) *impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR* < 0.001*) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

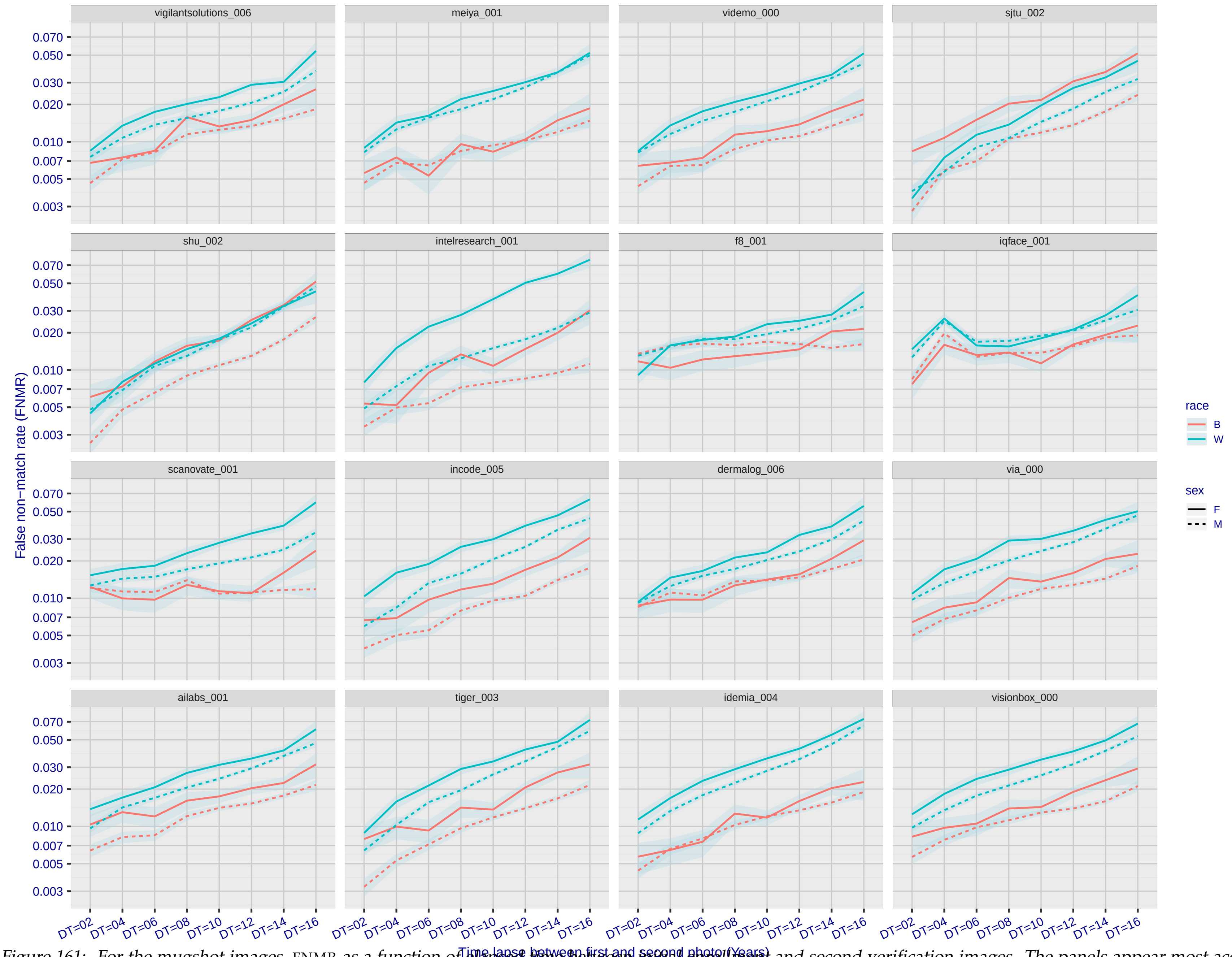


Figure 161: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR $<$ 0.001) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

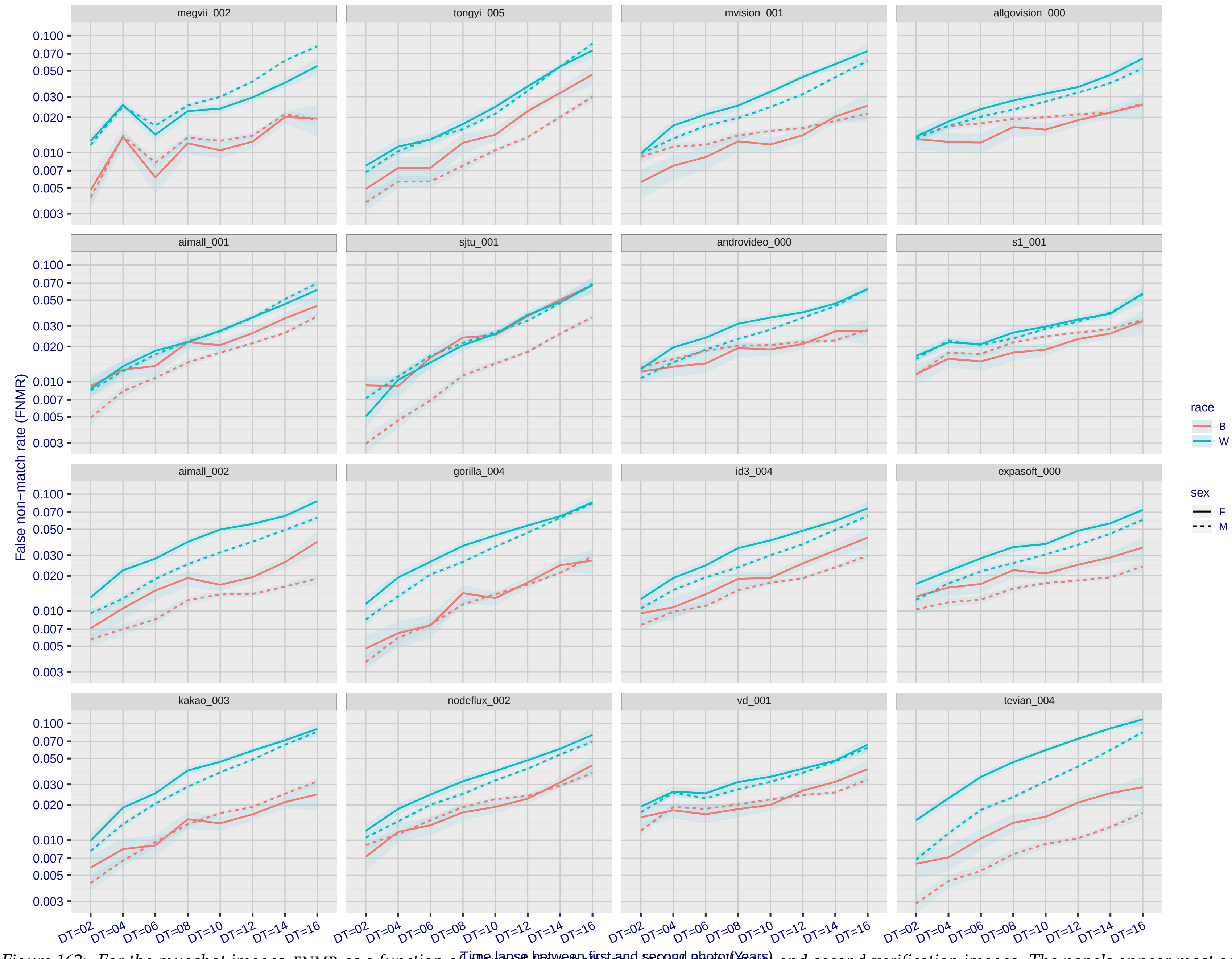


Figure 162: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR $<$ 0.001) so that the uncertainty estimates are high.*

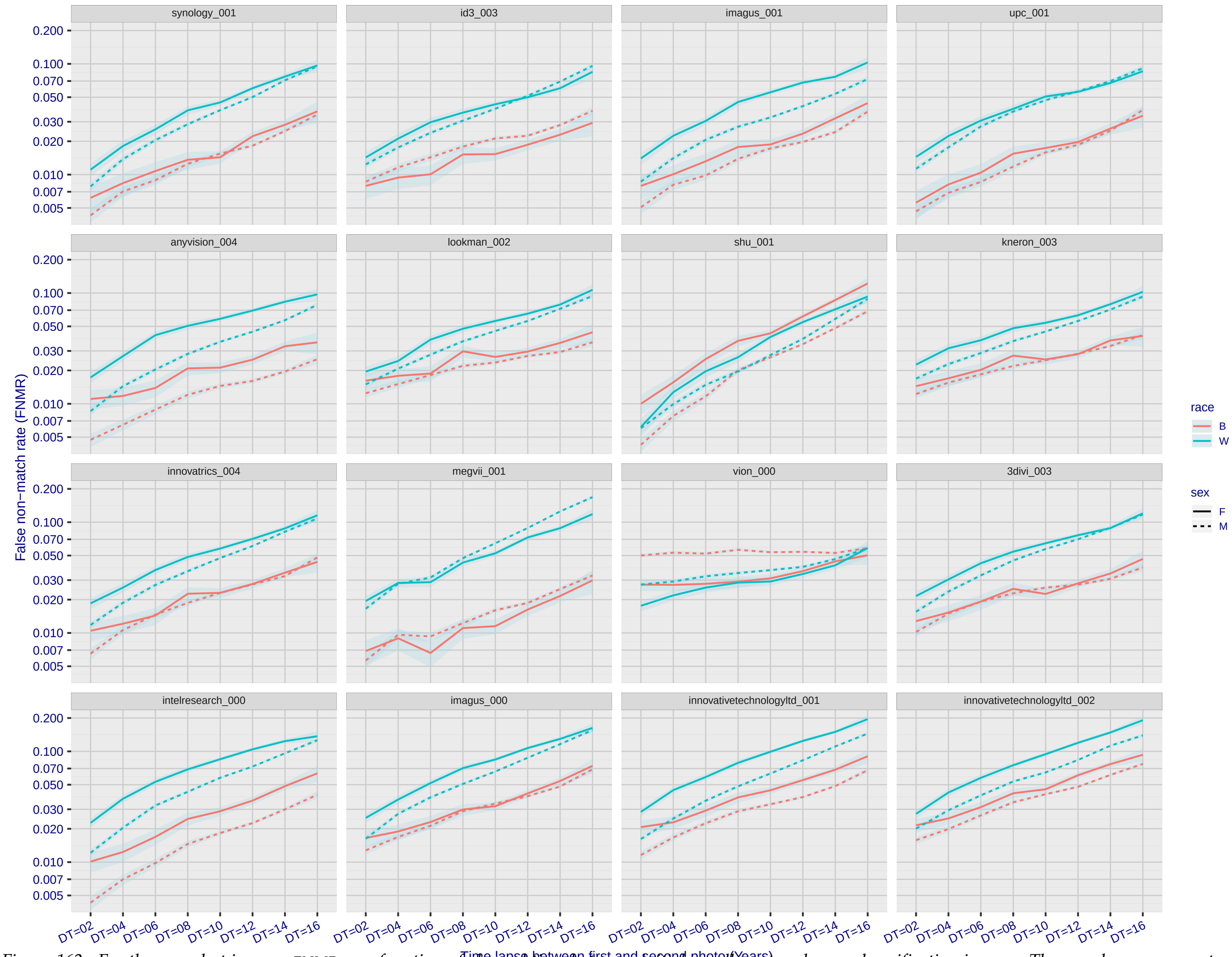


Figure 163: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (*10^8*) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR* < 0.001*) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

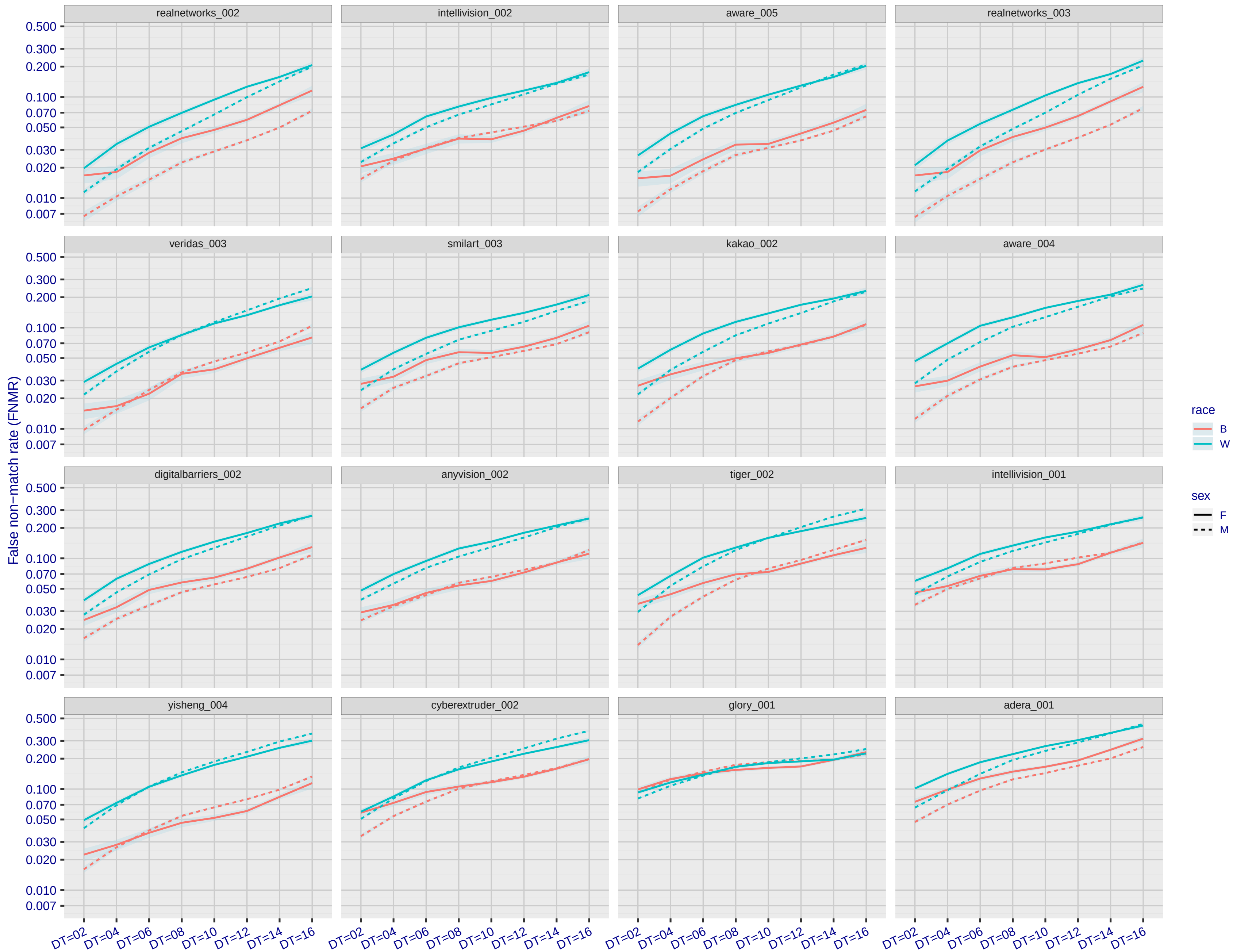


Figure 164: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

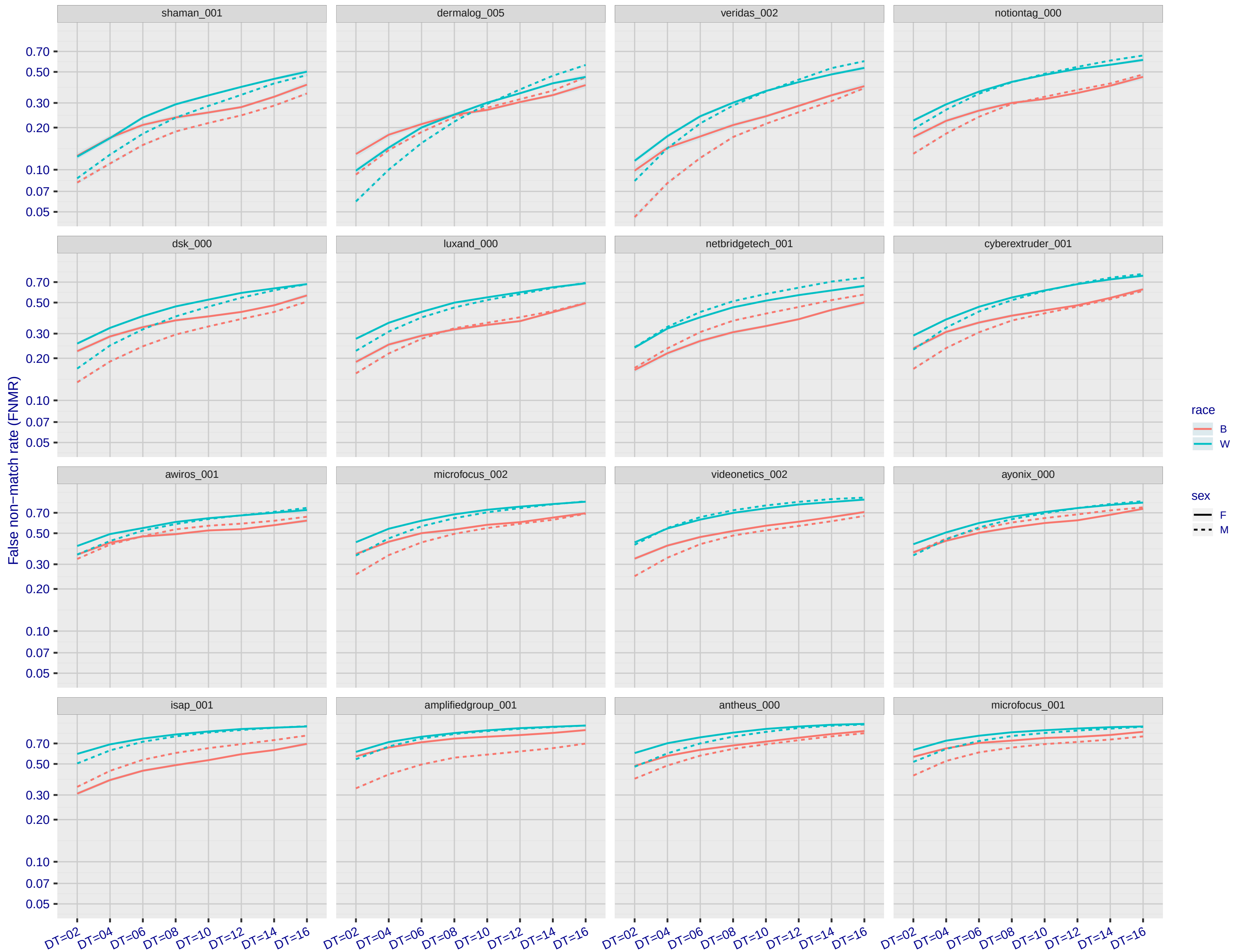


Figure 165: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR $<$ 0.001) so that the uncertainty estimates are high.*

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

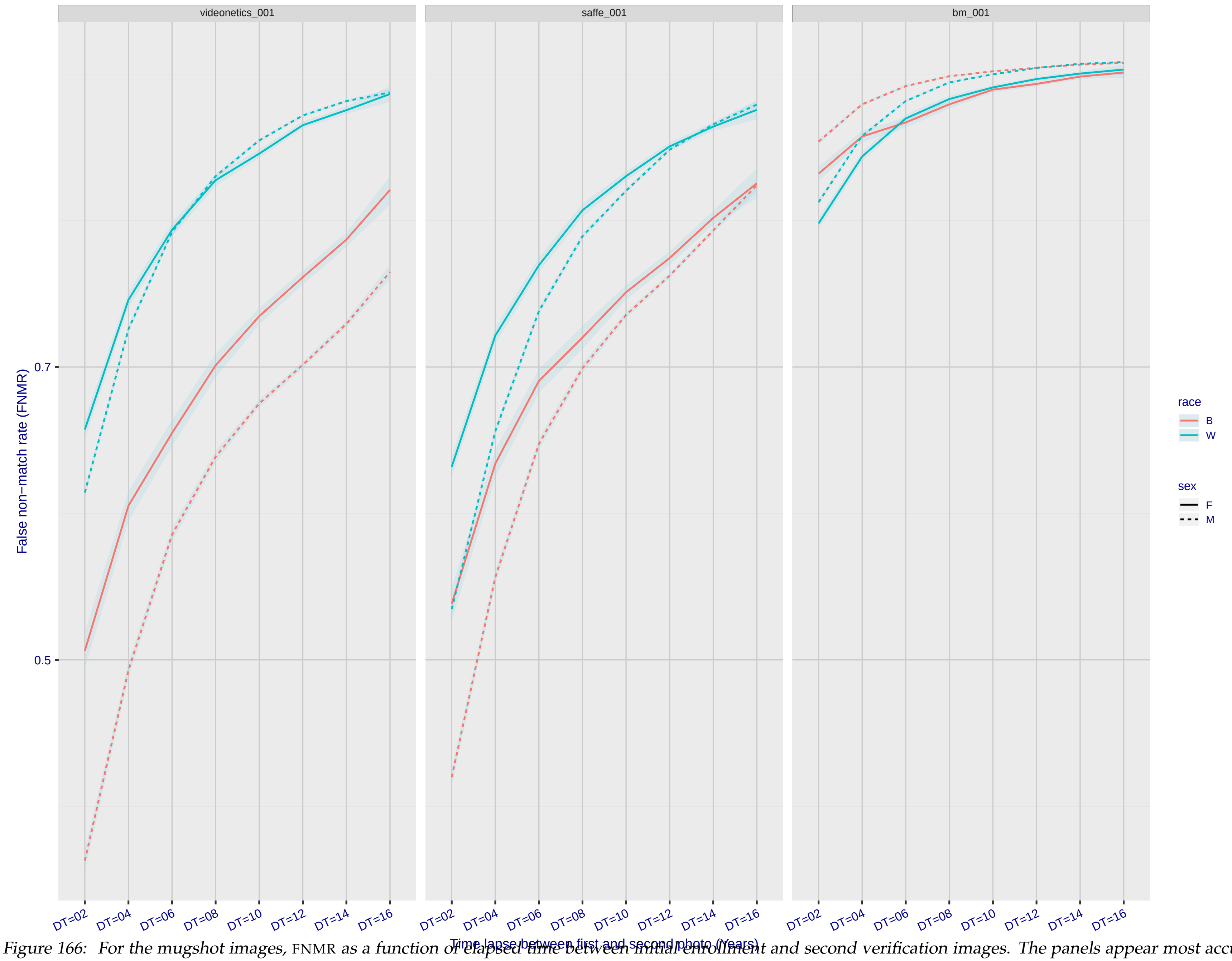


Figure 166: For the mugshot images, FNMR *as a function of elapsed time between initial enrollment and second verification images. The panels appear most accurate first, and vertical scale changes on each page. The four traces correspond to images annotated with codes for black female, black male, white female, white male. The threshold is fixed for each algorithm to give FMR = 0.00001 over all (10^8) impostor comparisons. For short time-lapses, the most accurate algorithms give very few errors (FNMR < 0.001) so that the uncertainty estimates are high.*

3.5.3 Effect of age on genuine subjects

Background: Faces change appearance throughout life. Face recognition algorithms have previously beeen reported to give better accuracy on older individuals (See NIST IR 8009).

Goal: To quantify false non-match rates (FNMR) as a function of age, without an ageing component.

Methods: Using the visa images, which span fewer than five years, thresholds are determined that give FMR = 0.001 and 0.0001 over the entire impostor set. Then FNMR is measured over 1000 bootstrap replications of the genuine scores.

Results: For the visa images, Figure 184 shows how false non-match rates for genuine users, as a function of age group.

The notable aspects are:

- ▷ Younger subjects give considerably higher FNMR. This is likely due to rapid growth and change in facial appearance.
- ▷ FNMR trends down throughout life. The last bin, AGE $>$ 72, contains fewer than 140 mated pairs, and may be affected by small sample size.

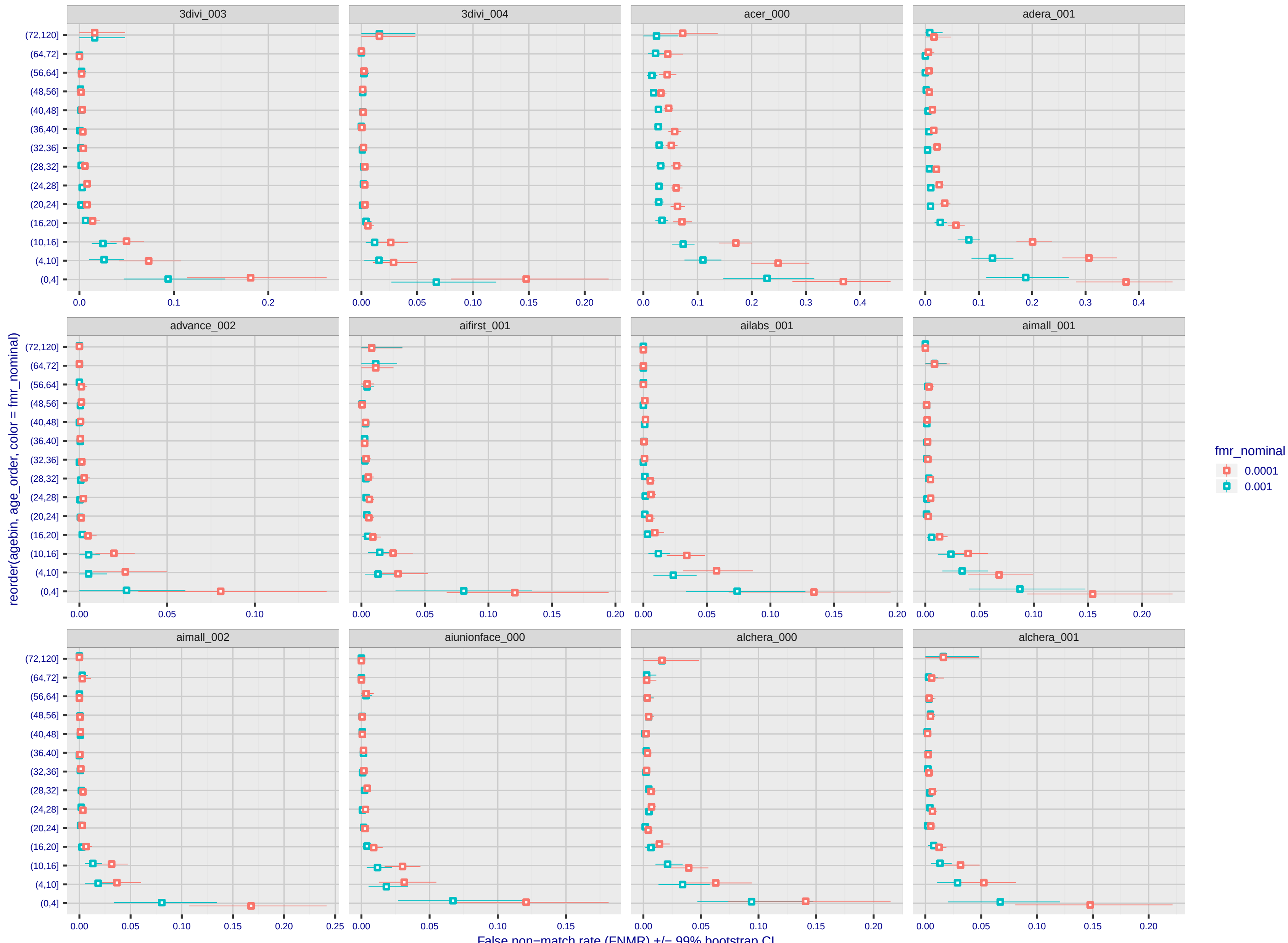


Figure 167: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

Figure 168: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

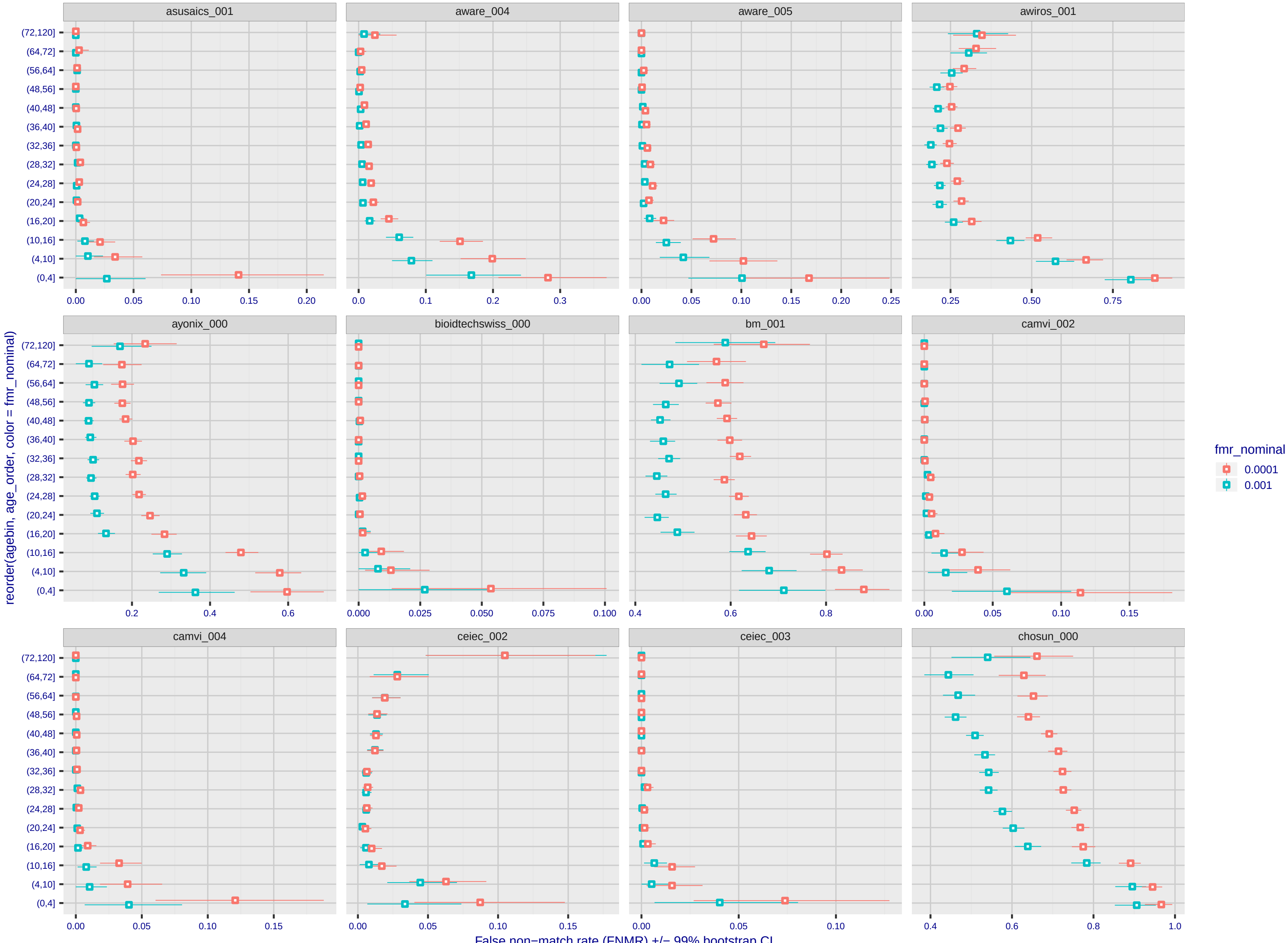


Figure 169: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

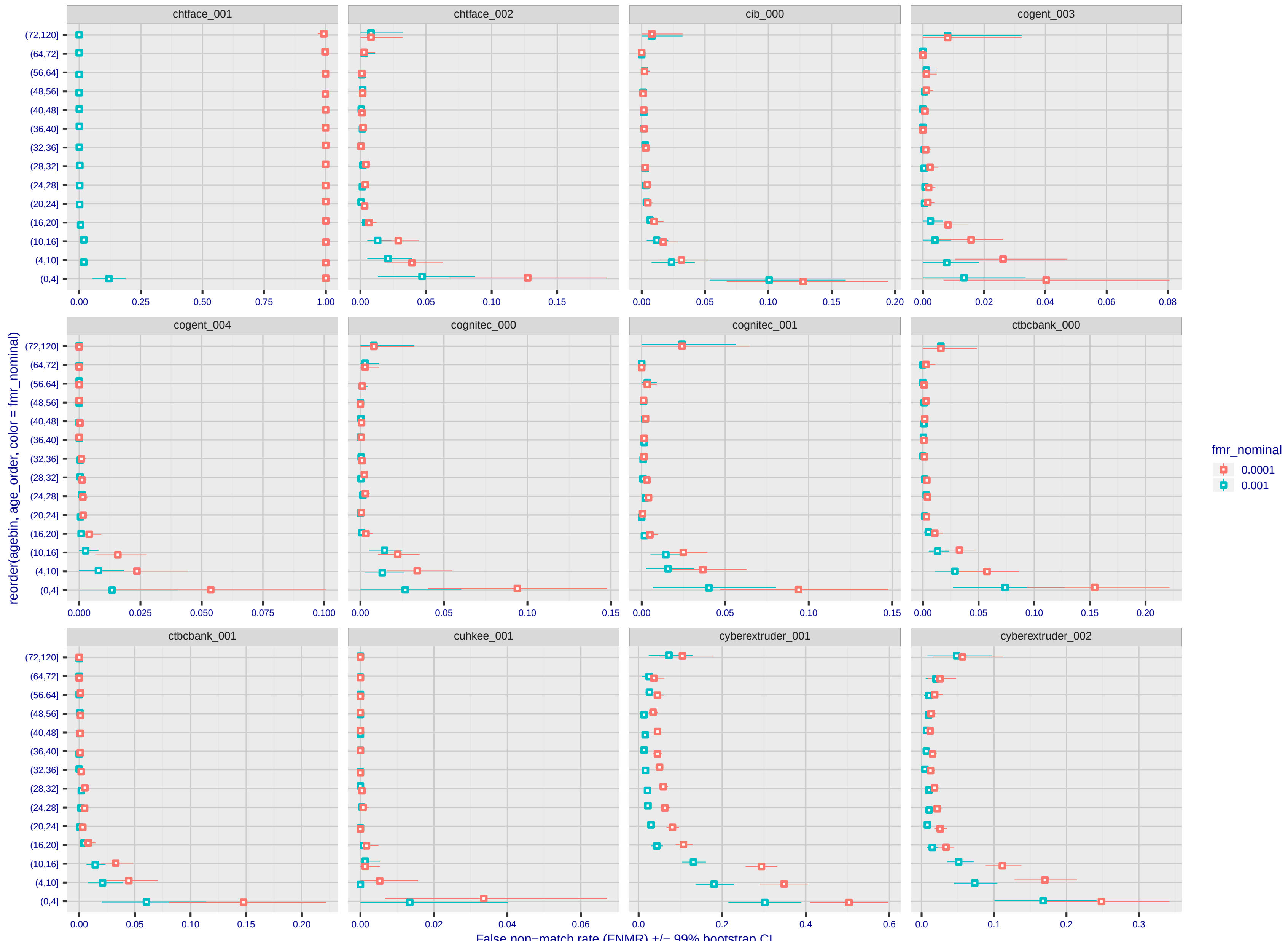


Figure 170: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

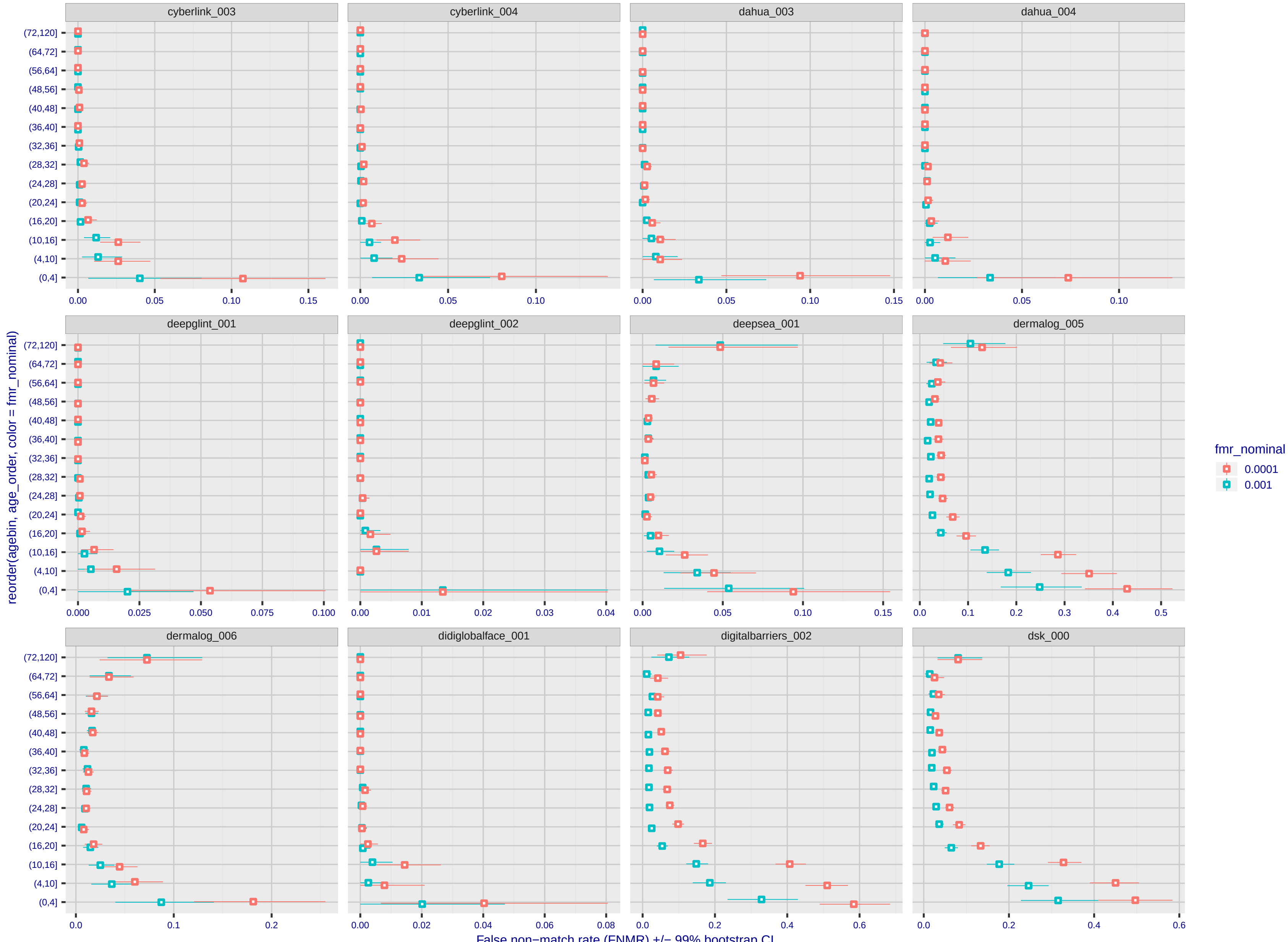


Figure 171: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

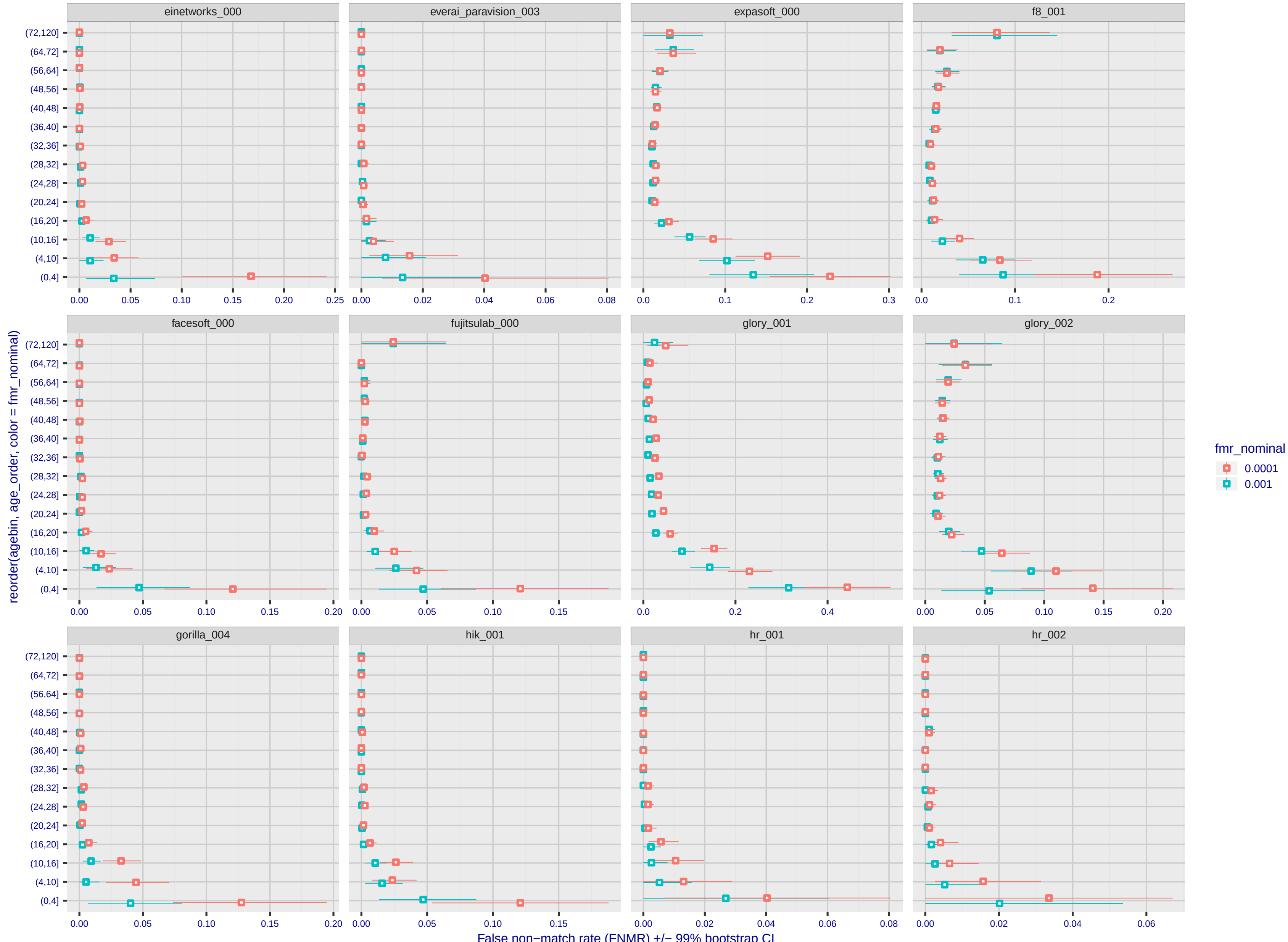


Figure 172: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

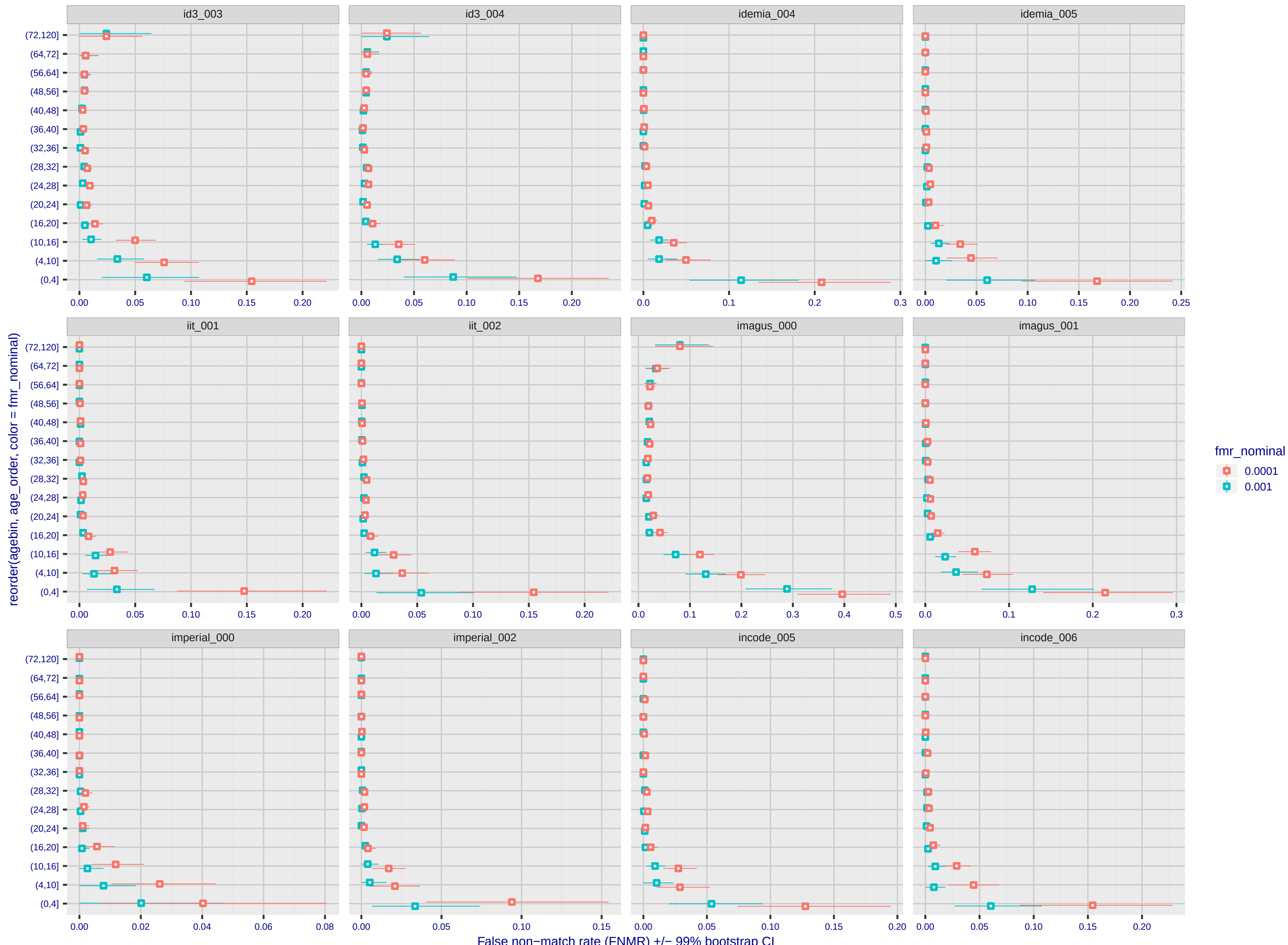


Figure 173: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

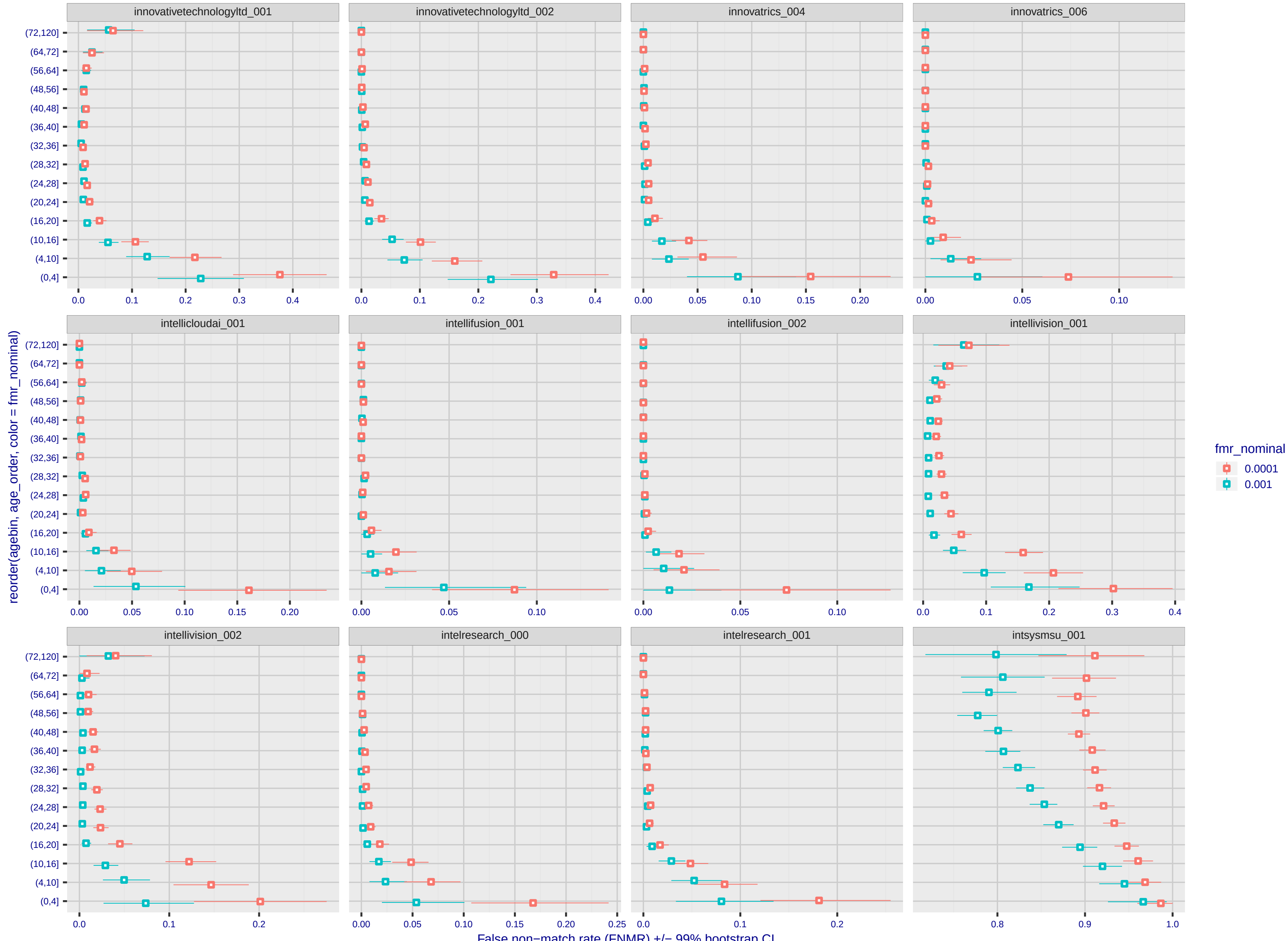


Figure 174: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

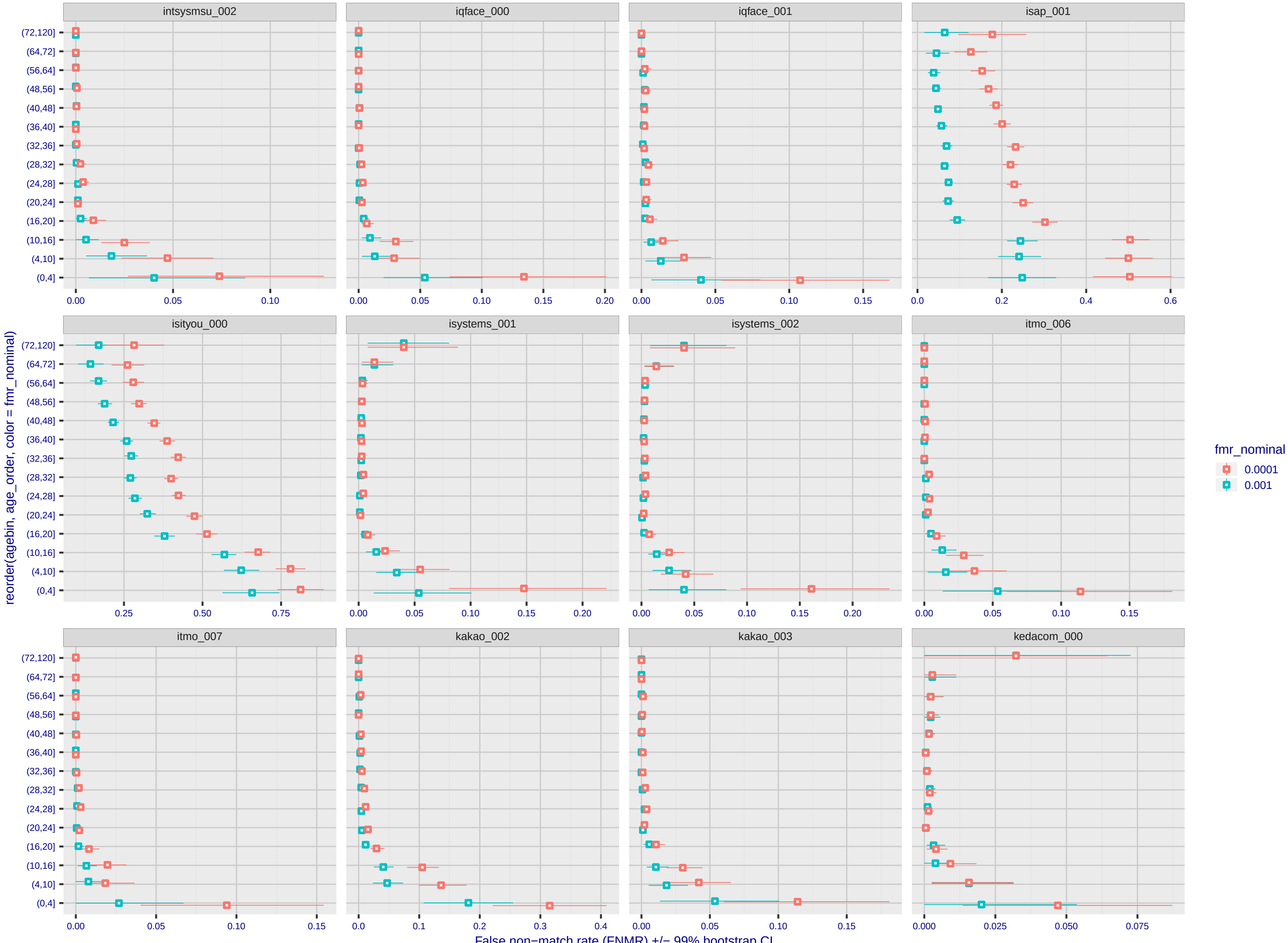


Figure 175: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

Figure 176: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

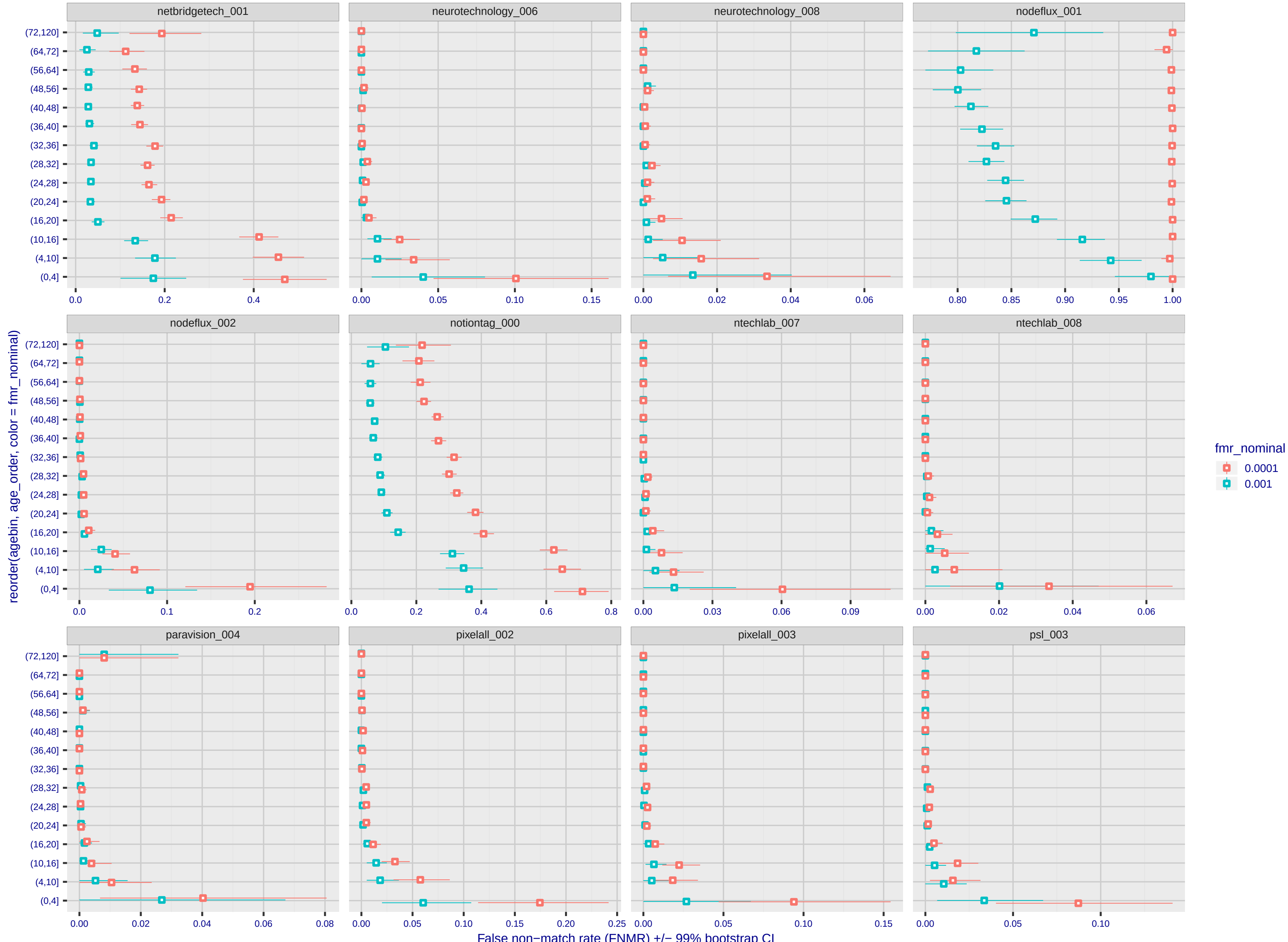


Figure 177: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

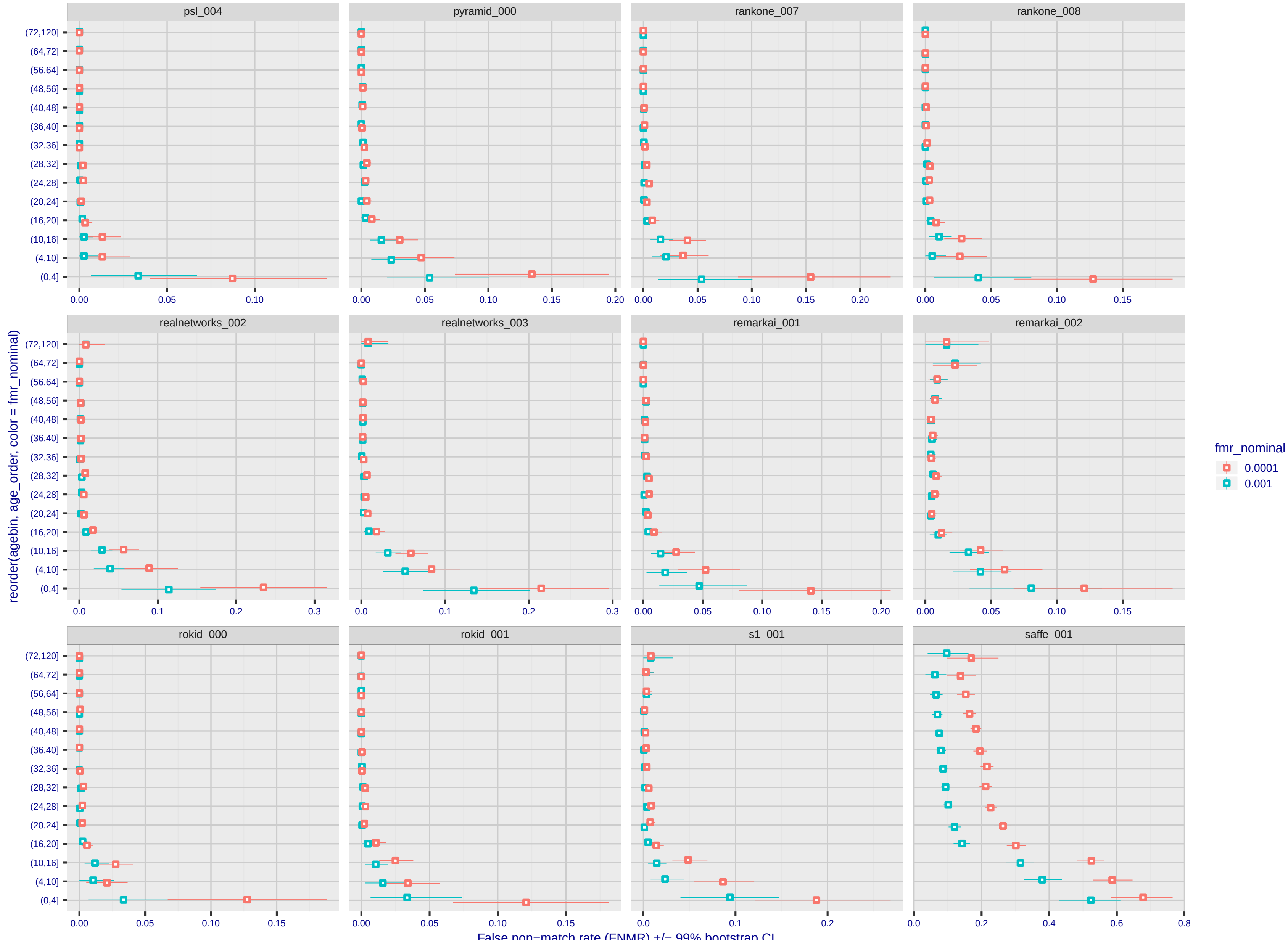


Figure 178: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

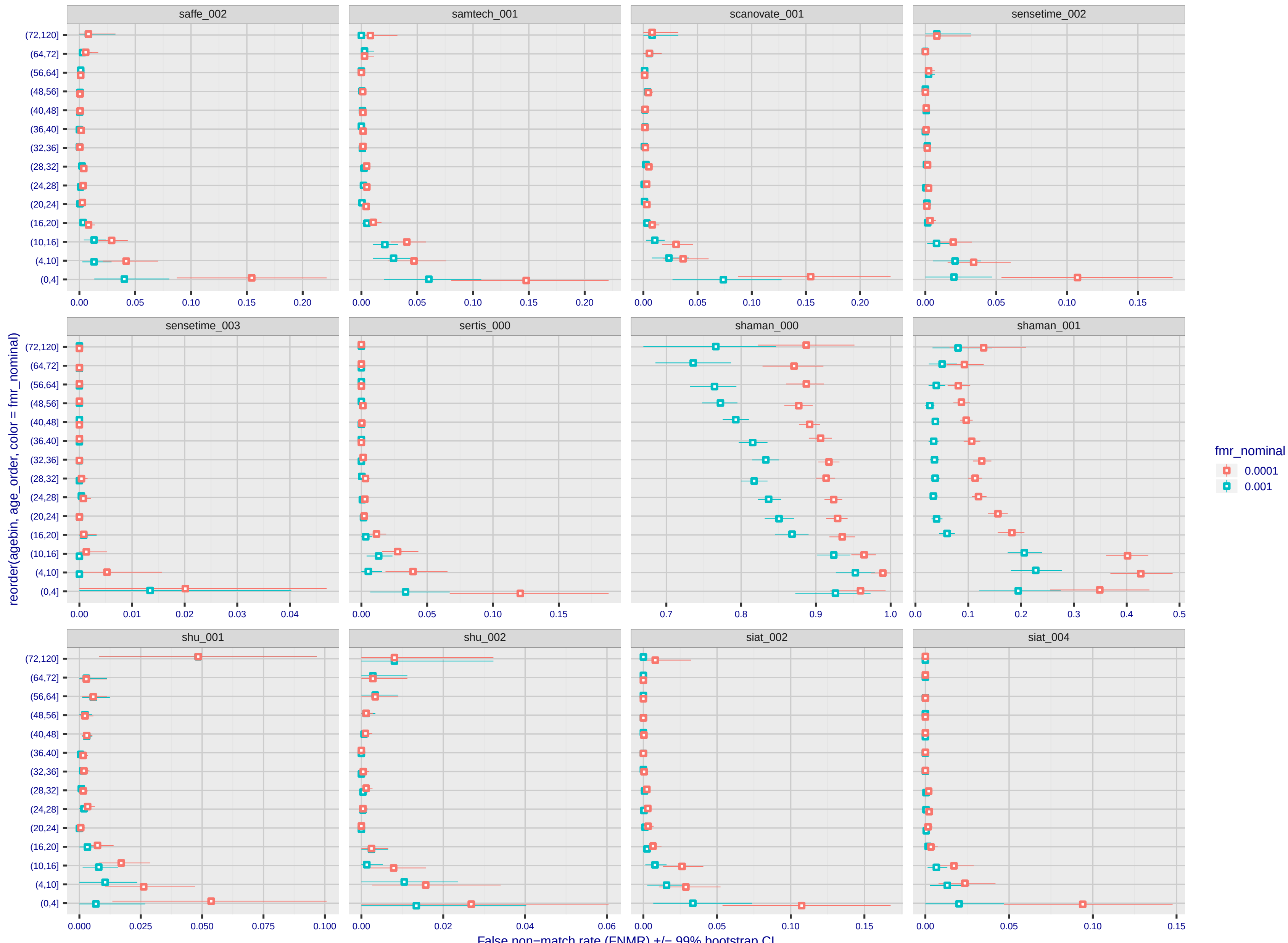


Figure 179: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

Figure 180: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

Figure 181: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

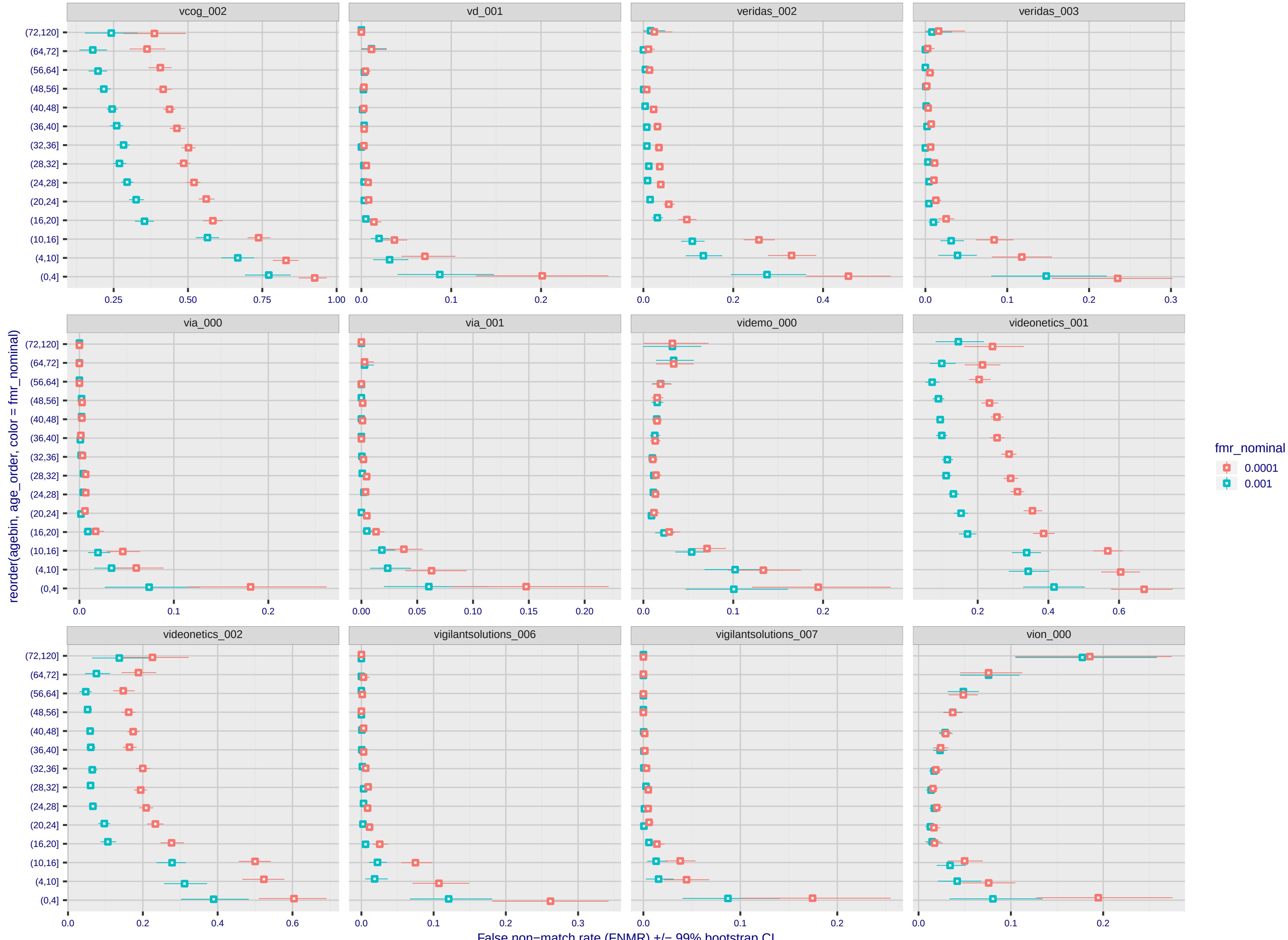


Figure 182: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

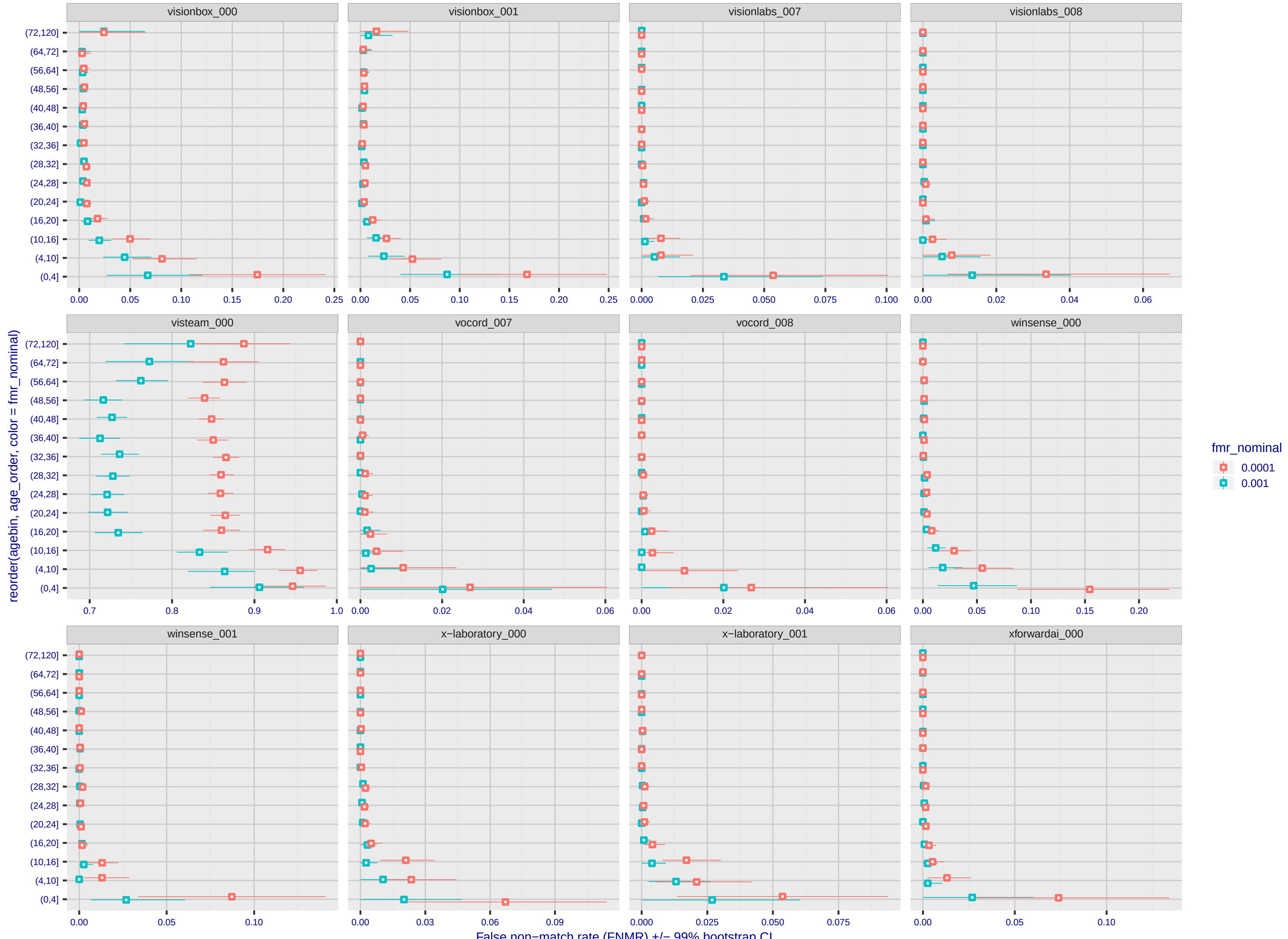


Figure 183: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all on the order of 10^{10} impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.

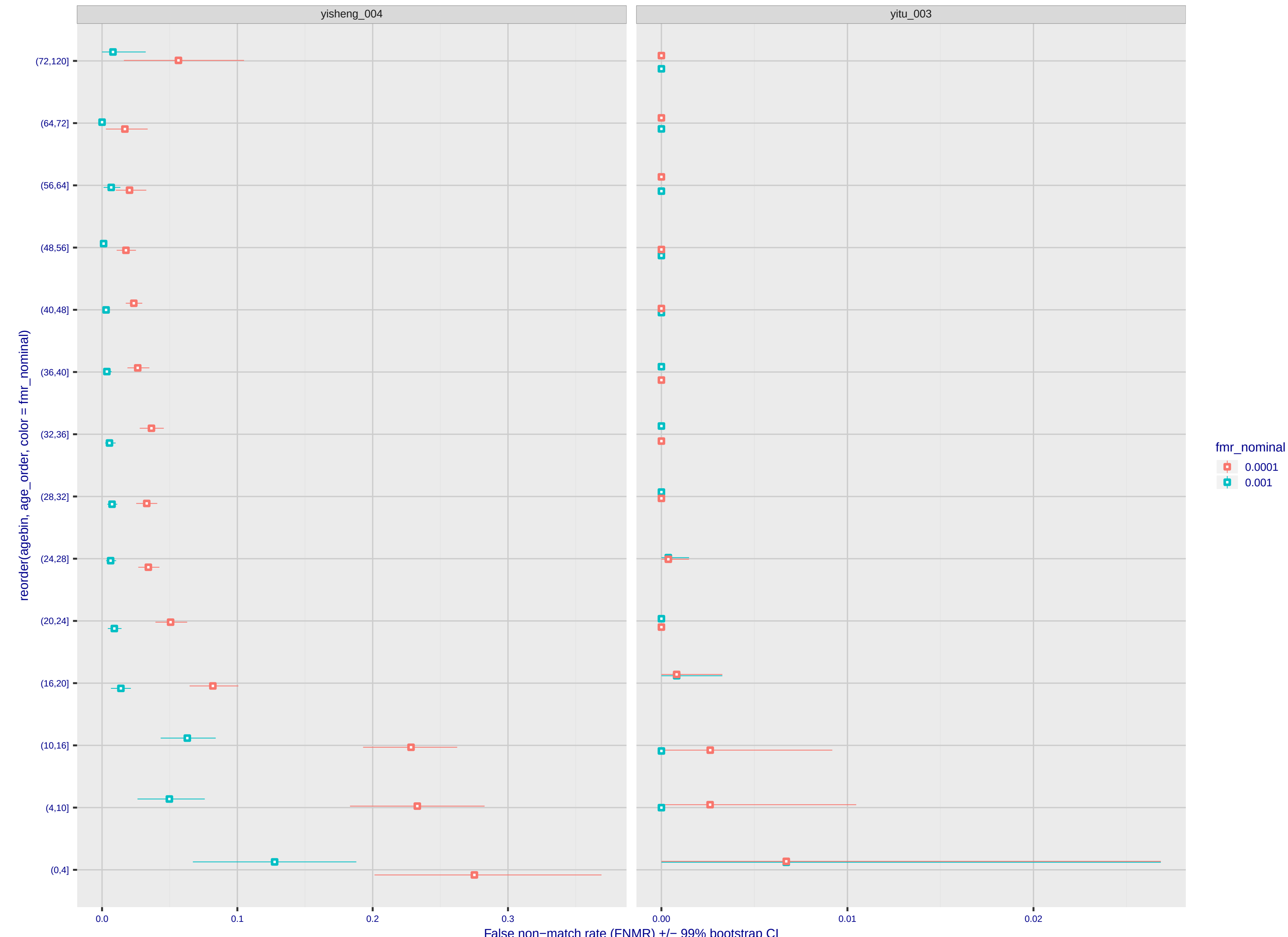


Figure 184: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ *computed over all on the order of* 10^{10} *impostor scores. The FMR in each bin will vary also - see subsequent impostor heatmaps in sec. 3.6.2. Given a pair of face images taken at different times, we assign the comparison to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands, however the first and last bins are computed over 149 and 124 respectively. The error rates in some (adult) cases are zero, and in others the DET is flat so the error rates at the two thresholds are identical. The lines span 1% and 99% of boostrap replicated FNMR estimates.*

Caveats: None.

3.6 Impostor distribution stability

3.6.1 Effect of birth place on the impostor distribution

Background: Facial appearance varies geographically, both in terms of skin tone, cranio-facial structure and size. This section addresses whether false match rates vary intra- and inter-regionally.

Goals:

- ▷ To show the effect of birth region of the impostor and enrollee on false match rates.
- ▷ To determine whether some algorithms give better impostor distribution stability.

Methods:

- ▷ For the visa images, NIST defined 10 regions: Sub-Saharan Africa, South Asia, Polynesia, North Africa, Middle East, Europe, East Asia, Central and South America, Central Asia, and the Caribbean.
- ▷ For the visa images, NIST mapped each country of birth to a region. There is some arbitrariness to this. For example, Egypt could reasonably be assigned to the Middle East instead of North Africa. An alternative methodology could, for example, assign the Philippines to *both* Polynesia and East Asia.
- ▷ FMR is computed for cases where all face images of impostors born in region r_2 are compared with enrolled face images of persons born in region r_1.

$$\mathrm{FMR}(r_1, r_2, T) = \frac{\sum_{i=1}^{N_{r_1,r_2}} H(s_i - T)}{N_{r_1,r_2}} \tag{5}$$

  where the same threshold, T, is used in all cells, and H is the unit step function. The threshold is set to give FMR(T) = 0.001 over the entire set of visa image impostor comparisons.

- ▷ This analysis is then repeated by country-pair, but only for those country pairs where both have at least 1000 images available. The countries[1] appear in the axes of graphs that follow.
- ▷ The mean number of impostor scores in any cross-region bin is 33 million. The smallest number of impostor scores in any bin is 135000, for Central Asia - North Africa. While these counts are large enough to support reasonable significance, the number of individual faces is much smaller, on the order of $N^{0.5}$.
- ▷ The numbers of impostor scores in any cross-country bin is shown in Figure 187.

Results: Subsequent figures show heatmaps that use color to represent the base-10 logarithm of the false match rate. Red colors indicate high (bad) false match rates. Dark colors indicate benign false match rates. There are two series of graphs corresponding to aggregated geographical regions, and to countries. The notable observations are:

- ▷ The on-diagonal elements correspond to within-region impostors. FMR is generally above the nominal value of FMR = 0.001. Particularly there is usually higher FMR in, Sub-Saharan Africa, South Asia, and the Caribbean. Europe and Central Asia, on the other hand, usually give FMR closer to the nominal value.
- ▷ The off-diagonal elements correspond to across-region impostors. The highest FMR is produced between the Caribbean and Sub-Saharan Africa.
- ▷ Algorithms vary.

[1]These are Argentina, Australia, Brazil, Chile, China, Costa Rica, Cuba, Czech Republic, Dominican Republic, Ecuador, Egypt, El Salvador, Germany, Ghana, Great Britain, Greece, Guatamala, Haiti, Hong Kong, Honduras, Indonesia, India, Israel, Jamaica, Japan, Kenya, Korea, Lebanon, Mexico, Malaysia, Nepal, Nigeria, Peru, Philippines, Pakistan, Poland, Romania, Russia, South Africa, Saudi Arabia, Thailand, Trinidad, Turkey, Taiwan, Ukraine, Venezuela, and Vietnam.

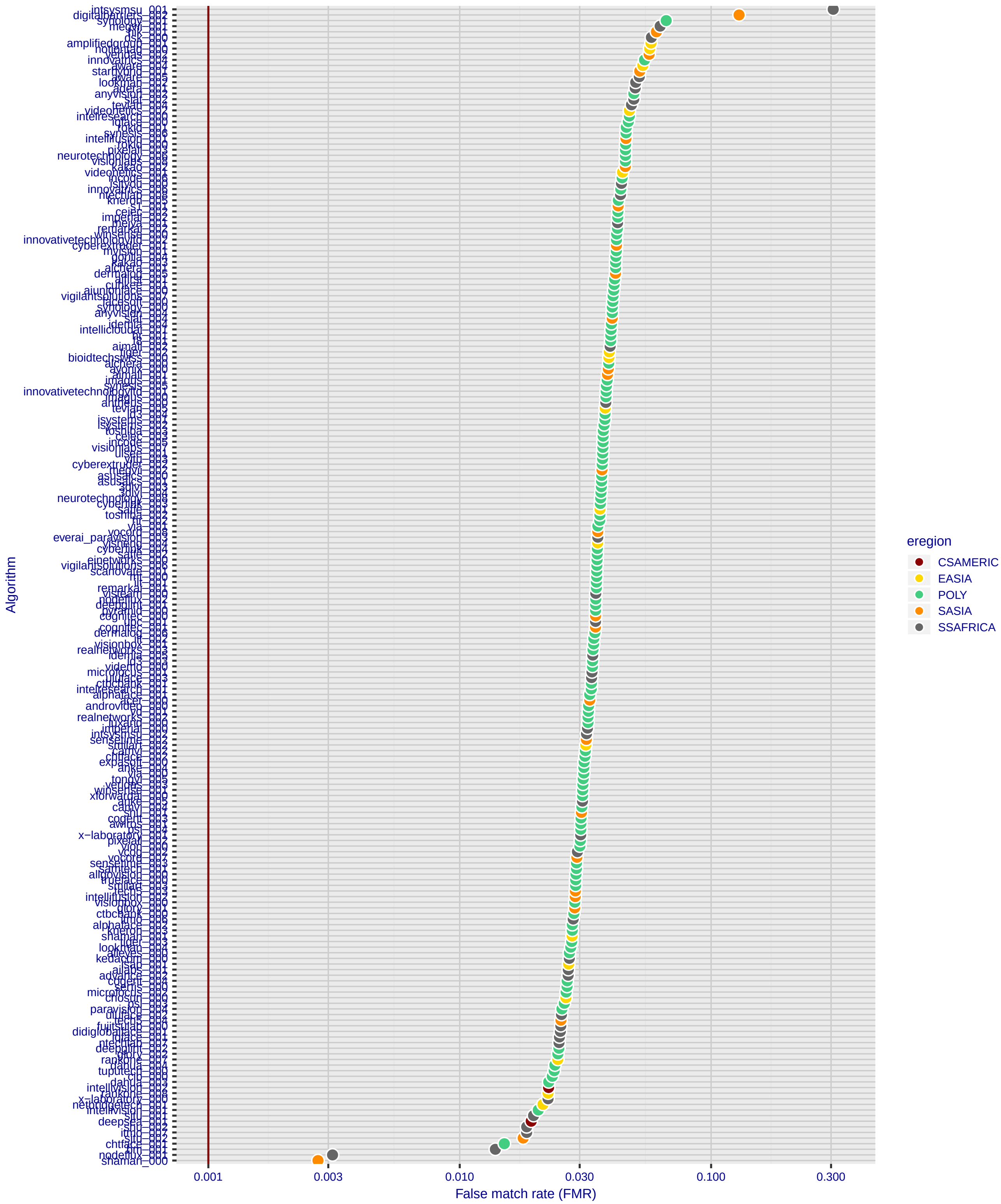


Figure 185: For the visa images, the dots show FMR for impostor comparisons of individuals of the same sex and same age group for the region of the world that gives the worst (highest) FMR when the threshold is set to give FMR = 0.001 (red vertical line) over all on the order of 10^{10} impostor scores i.e. zero-effort. The shift of the dots to right shows massive increases in FMR when impostors have the same sex, age, and region of birth. The color code indicates which region gives the worst case FMR. If the observed variation is due to the prevalence of one kind of images in the training imagery, then algorithms developed on one kind of data might be expected to give higher FMR on other kinds.

▷ We computed the same quantities for a global FMR = 0.0001. The effects are similar.

Caveats:

▷ The effects of variable impostor rates on one-to-many identification systems may well differ from what's implied by these one-to-one verification results. Two reasons for this are a) the enrollment galleries are usually imbalanced across countries of birth, age and sex; b) one-to-many identification algorithms often implement techniques aimed at stabilizing the impostor distribution. Further research is necessary.

▷ In principle, the effects seen in this subsection could be due to differences in the image capture process. We consider this unlikely since the effects are maintained across geography - e.g. Caribbean vs. Africa, or Japan vs. China.

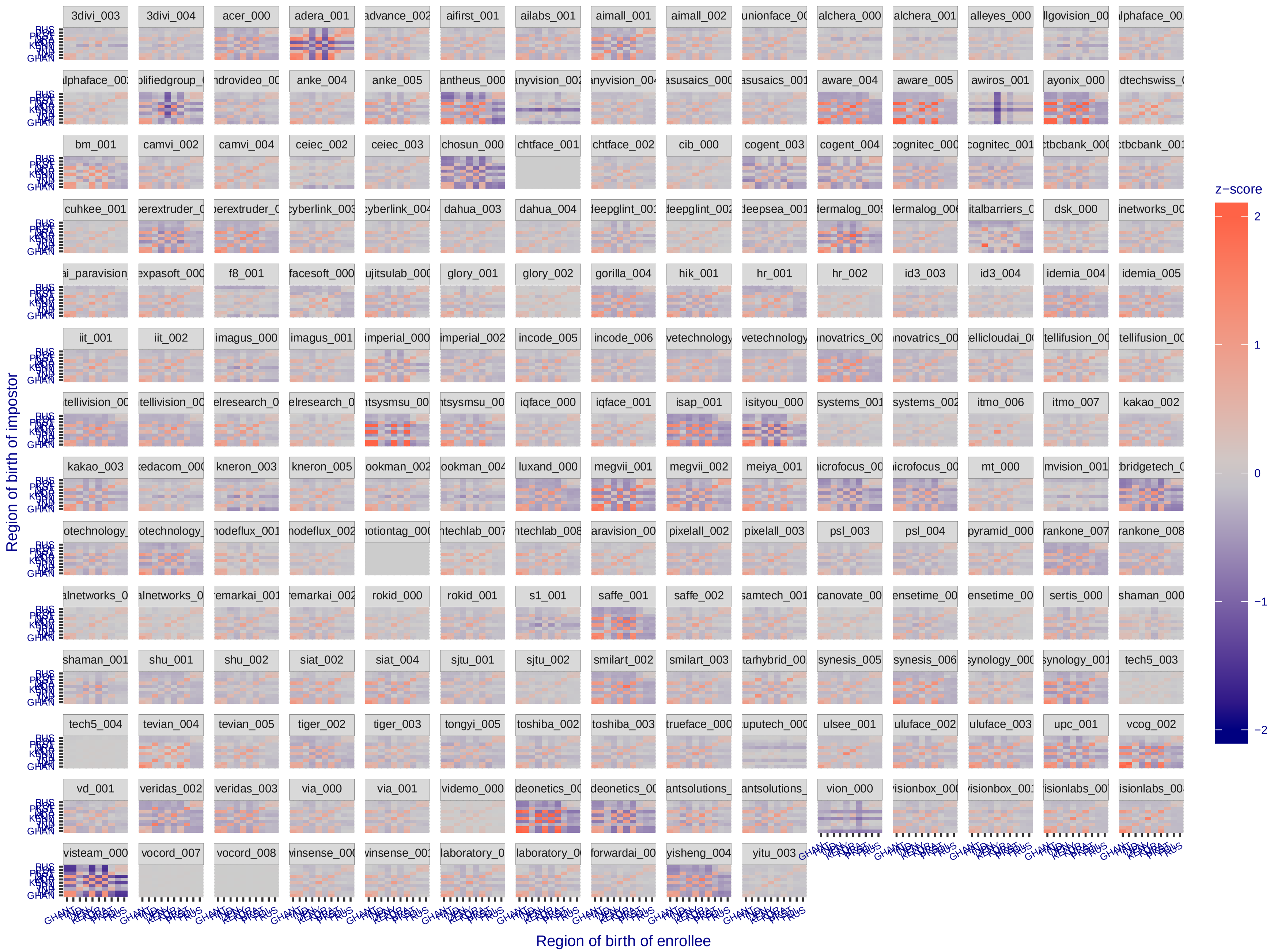


Figure 186: *For visa images, the heatmap shows how the mean of the impostor distribution for the country pair (a,b) is shifted relative to the mean of the global impostor distribution, expressed as a number of standard deviations of the global impostor distribution. This statistic is designed to show shifts in the entire impostor distribution, not just tail effects that manifest as the anomalously high (or low) false match rates that appear in the subsequent figures. The countries are chosen to show that skin tone alone does not explain impostor distribution shifts. The reduced shift in Asian populations with the Yitu and TongYiTrans algorithms, is accompanied by positive shifts in the European populations. This reversal relative to most other algorithms, may derive from use of nationally weighted training sets. The figure is computed from same-sex and same-age impostor pairs.*

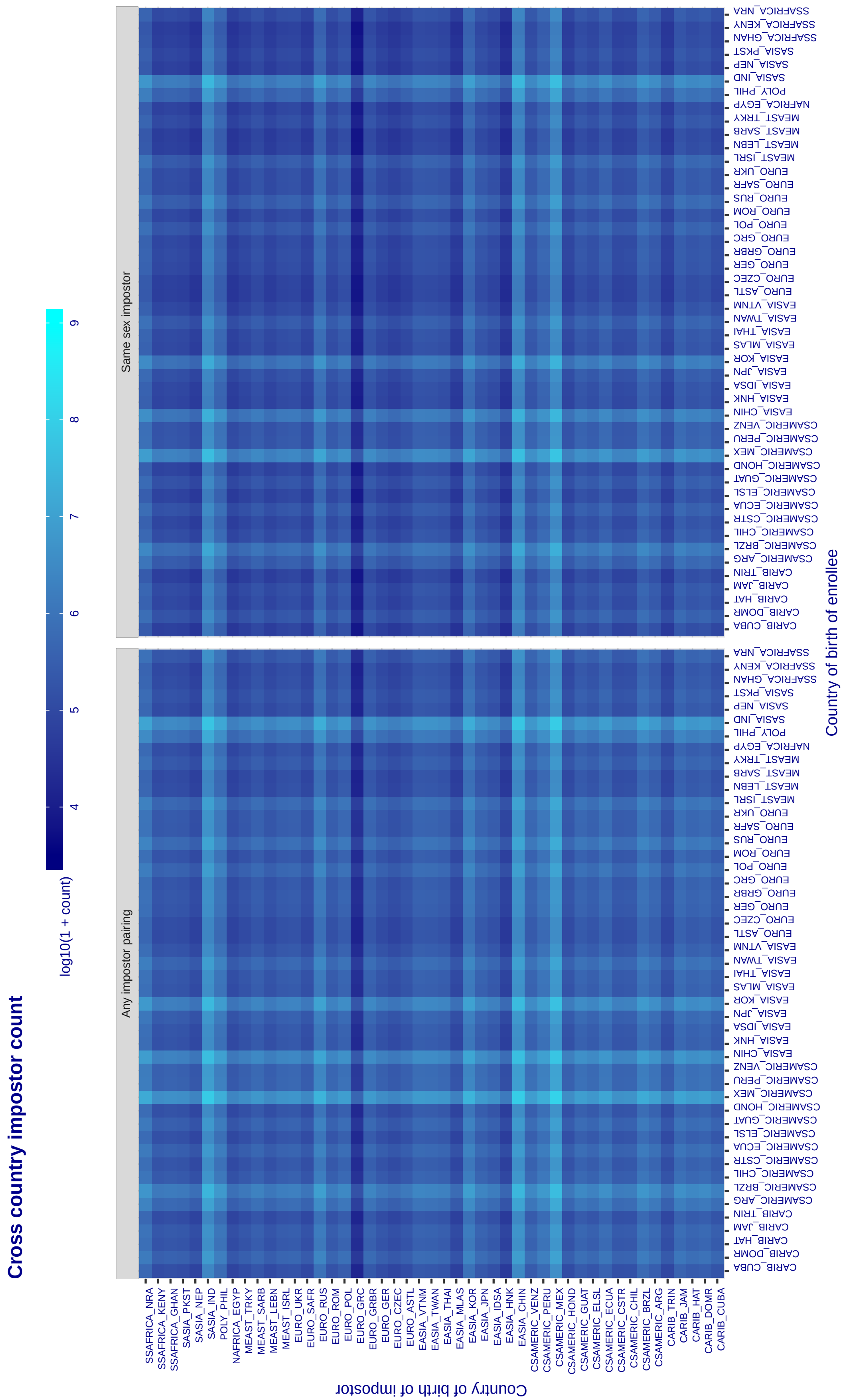


Figure 187: For visa images, the heatmap shows the count of impostor comparisons of faces from different individuals who were born in the given country pair.

3.6.2 Effect of age on impostors

Background: This section shows the effect of age on the impostor distribution. The ideal behaviour is that the age of the enrollee and the impostor would not affect impostor scores. This would support FMR stability over sub-populations.

Goals:

- ▷ To show the effect of relative ages of the impostor and enrollee on false match rates.
- ▷ To determine whether some algorithms have better impostor distribution stability.

Methods:

- ▷ Define 14 age group bins, spanning 0 to over 100 years old.
- ▷ Compute FMR over all impostor comparisons for which the subjects in the enrollee and impostor images have ages in two bins.
- ▷ Compute FMR over all impostor comparisons for which the subjects are additionally of the same sex, and born in the same geographic region.

Results:

The notable aspects are:

- ▷ Diagonal dominance: Impostors are more likely to be matched against their same age group.
- ▷ Same sex and same region impostors are more successful. On the diagonal, an impostor is more likely to succeed by posing as someone of the same sex. If $\Delta \log_{10} \text{FMR} = 0.2$, then same-sex same-region FMR exceeds the all-pairs FMR by factor of $10^{0.2} = 1.6$.
- ▷ Young children impostors give elevated FMR against young children. Older adult impostor give elevated FMR against older adults. These effects are quite large, for example if $\Delta \log_{10} \text{FMR} = 1.0$ larger than a 32 year old, then these groups have higher FMR by a factor of $10^1 = 10$. This would imply an FMR above 0.01 for a nominal (global) FMR = 0.001.
- ▷ Algorithms vary.
- ▷ We computed the same quantities for a global FMR = 0.0001. The effects are similar.

Note the calculations in this section include impostors paired across all countries of birth.

Accuracy Terms + Definitions

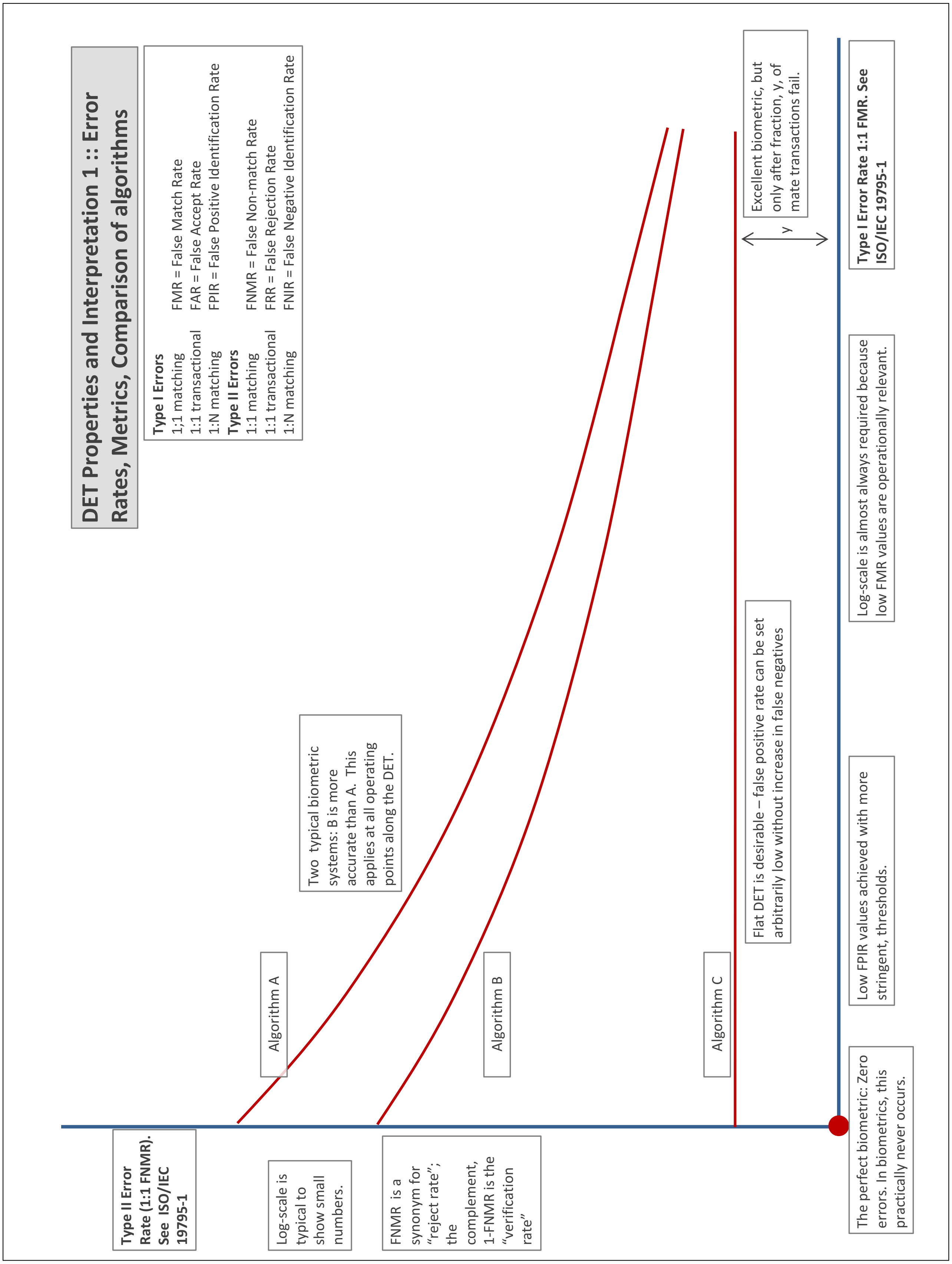
In biometrics, Type II errors occur when two samples of one person do not match – this is called a **false negative**. Correspondingly, Type I errors occur when samples from two persons do match – this is called a **false positive**. Matches are declared by a biometric system when the native comparison score from the recognition algorithm meets some **threshold**. Comparison scores can be either **similarity scores**, in which case higher values indicate that the samples are more likely to come from the same person, or **dissimilarity scores**, in which case higher values indicate different people. Similarity scores are traditionally computed by **fingerprint** and **face** recognition algorithms, while dissimilarities are used in **iris recognition**. In some cases, the dissimilarity score is a distance; this applies only when **metric** properties are obeyed. In any case, scores can be either **mate** scores, coming from a comparison of one person's samples, or **nonmate** scores, coming from comparison of different persons' samples. The words **genuine** or **authentic** are synonyms for mate, and the word **impostor** is used a synonym for nonmate. The words mate and nonmate are traditionally used in identification applications (such as law enforcement search, or background checks) while genuine and impostor are used in verification applications (such as access control).

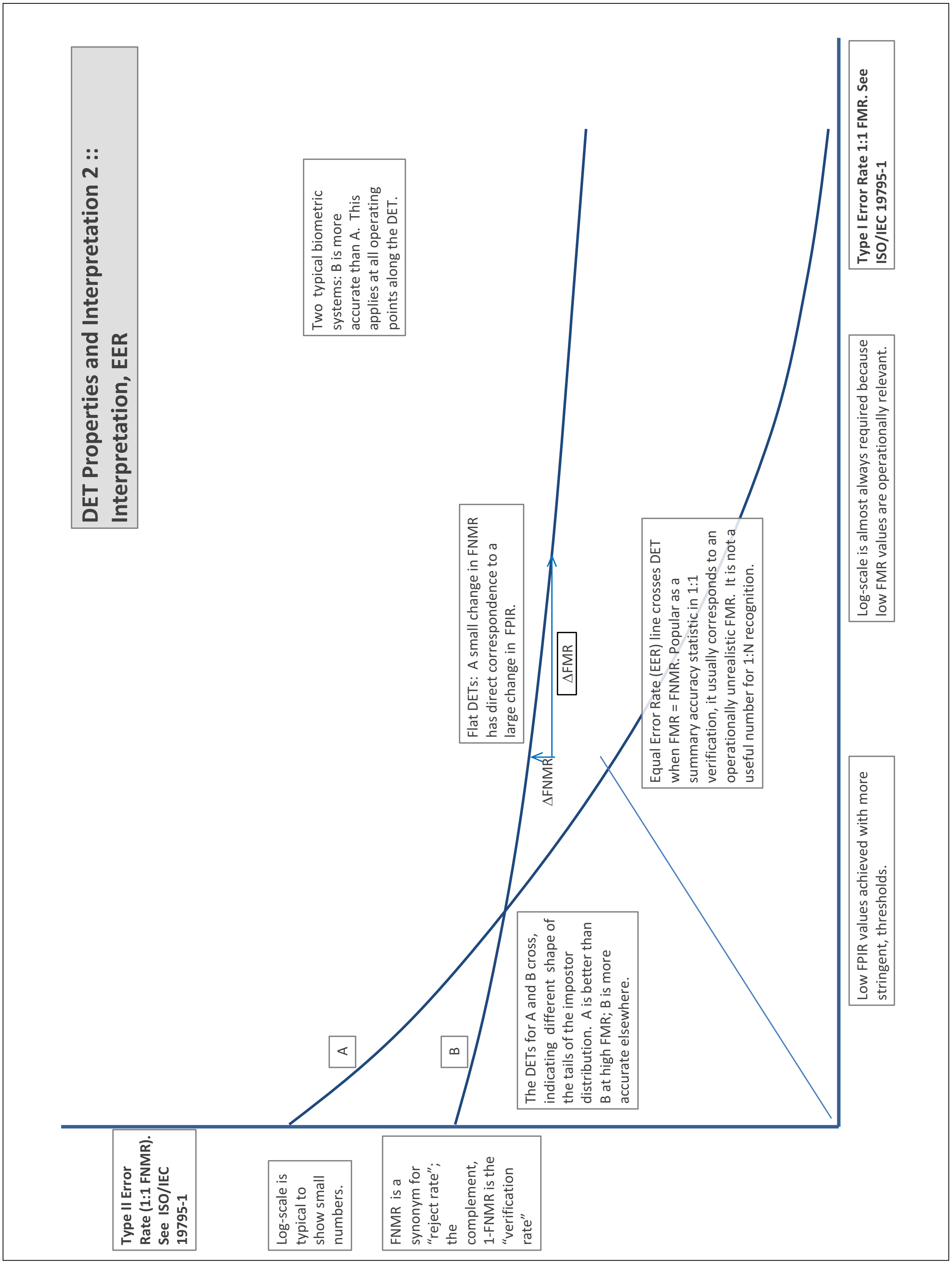
A **error tradeoff** characteristic represents the tradeoff between Type II and Type I classification errors. For verification this plots false non-match rate (FNMR) vs. false match rate (FMR) parametrically with T.

The error tradeoff plots are often called **detection error tradeoff (DET)** characteristics or **receiver operating characteristic** (ROC). These serve the same function but differ, for example, in plotting the complement of an error rate (e.g. TMR = 1 – FNMR) and in transforming the axes most commonly using logarithms, to show multiple decades of FMR. More rarely, the function might be the inverse Gaussian function.

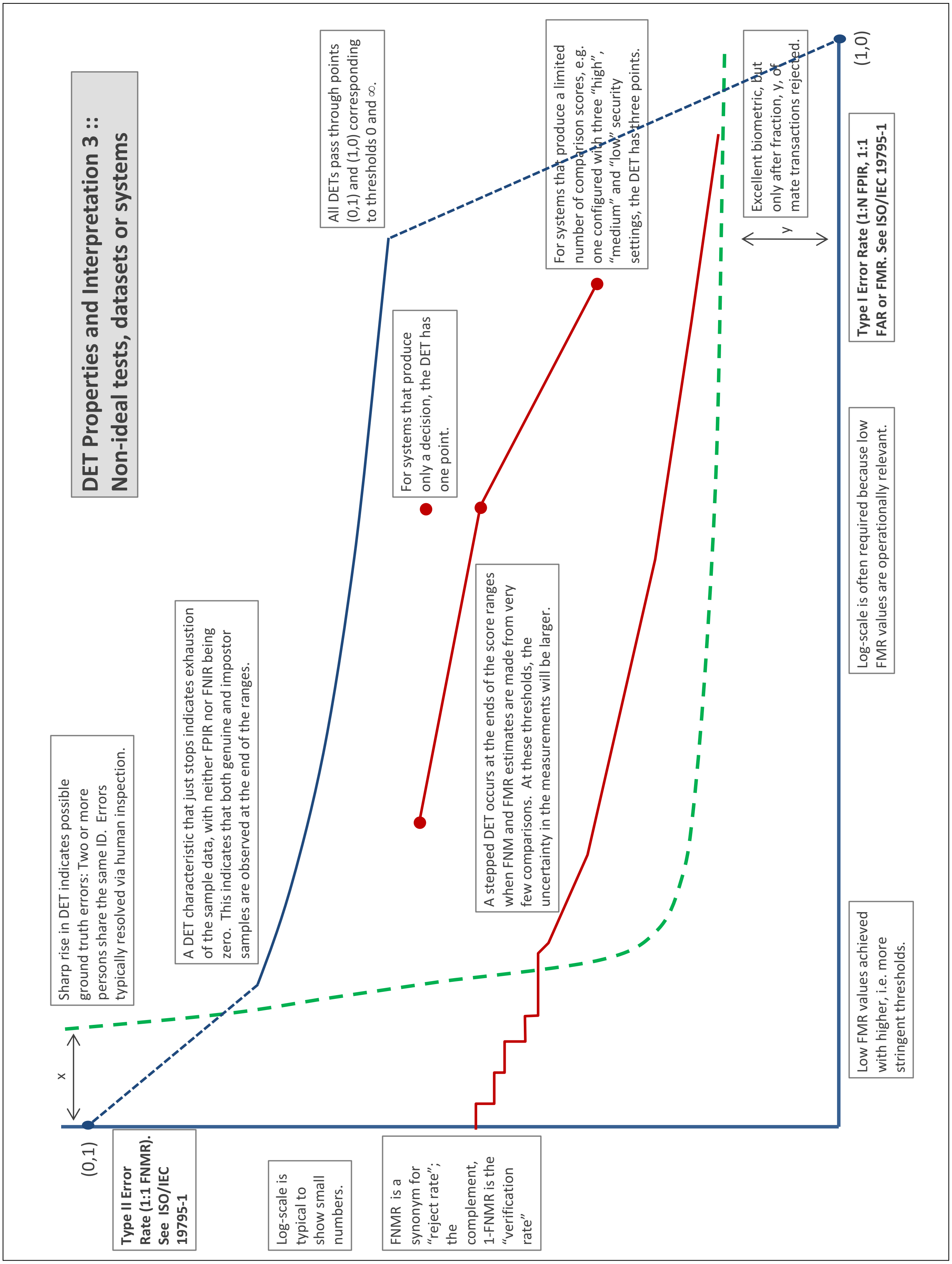
More detail and generality is provided in formal biometrics testing standards, see the various parts of ISO/IEC 19795 Biometrics Testing and Reporting. More terms, including and beyond those to do with accuracy, see ISO/IEC 2382-37 Information technology -- Vocabulary -- Part 37: Harmonized biometric vocabulary

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

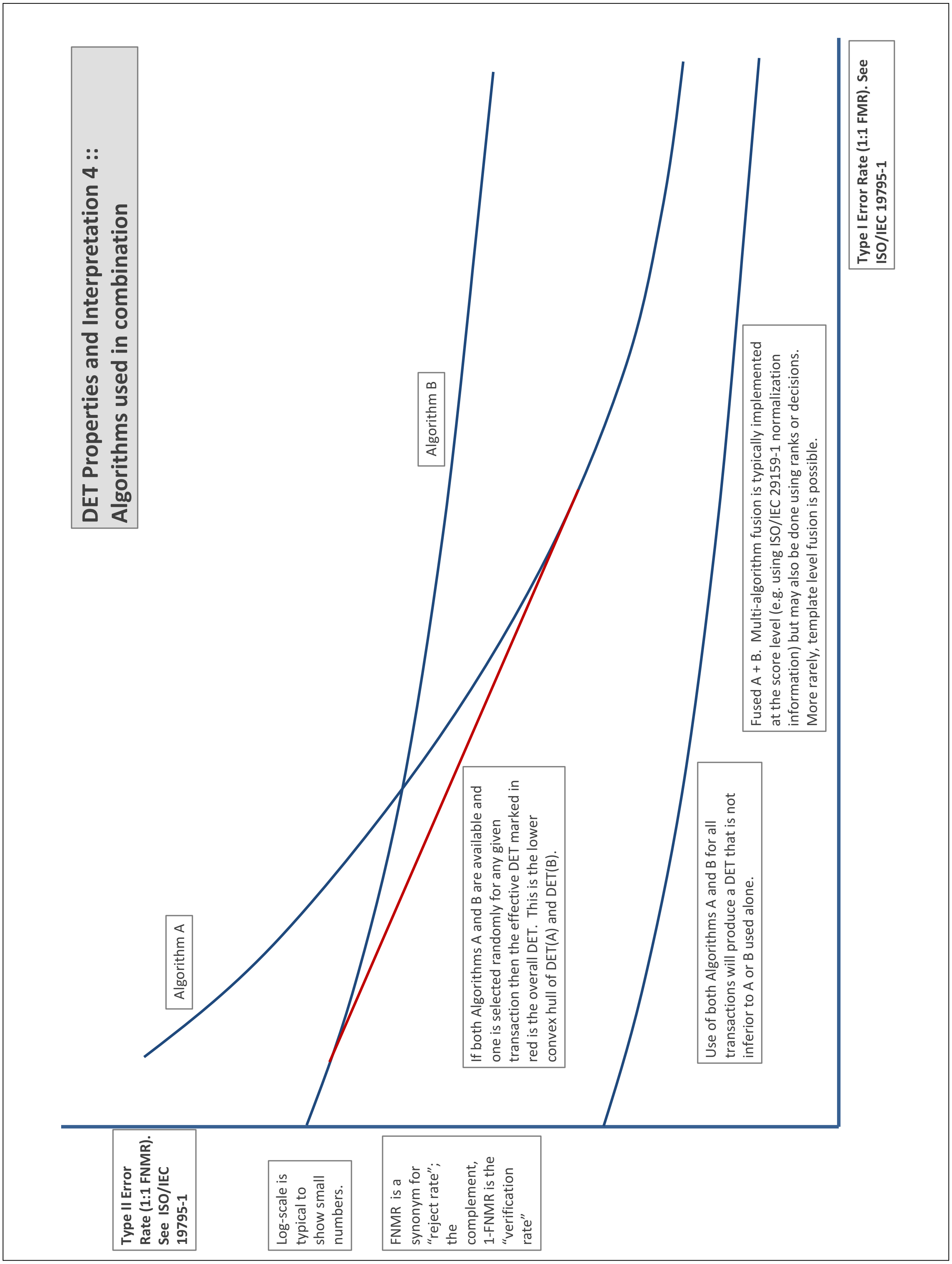

DET Properties and Interpretation 1 :: Error Rates, Metrics, Comparison of algorithms
Type I Errors
1;1 matching FMR = False Match Rate
1:1 transactional FAR = False Accept Rate
1:N matching FPIR = False Positive Identification Rate
Type II Errors
1:1 matching FNMR = False Non-match Rate
1:1 transactional FRR = False Rejection Rate
1:N matching FNIR = False Negative Identification Rate
Two typical biometric systems: B is more accurate than A. This applies at all operating points along the DET.
Algorithm A
Algorithm B
Algorithm C
Flat DET is desirable – false positive rate can be set arbitrarily low without increase in false negatives
Excellent biometric, but only after fraction, y, of mate transactions fail.
y
Type I Error Rate 1:1 FMR. See ISO/IEC 19795-1
Log-scale is almost always required because low FMR values are operationally relevant.
Low FPIR values achieved with more stringent, thresholds.
The perfect biometric: Zero errors. In biometrics, this practically never occurs.
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate", the complement, 1-FNMR is the "verification rate"

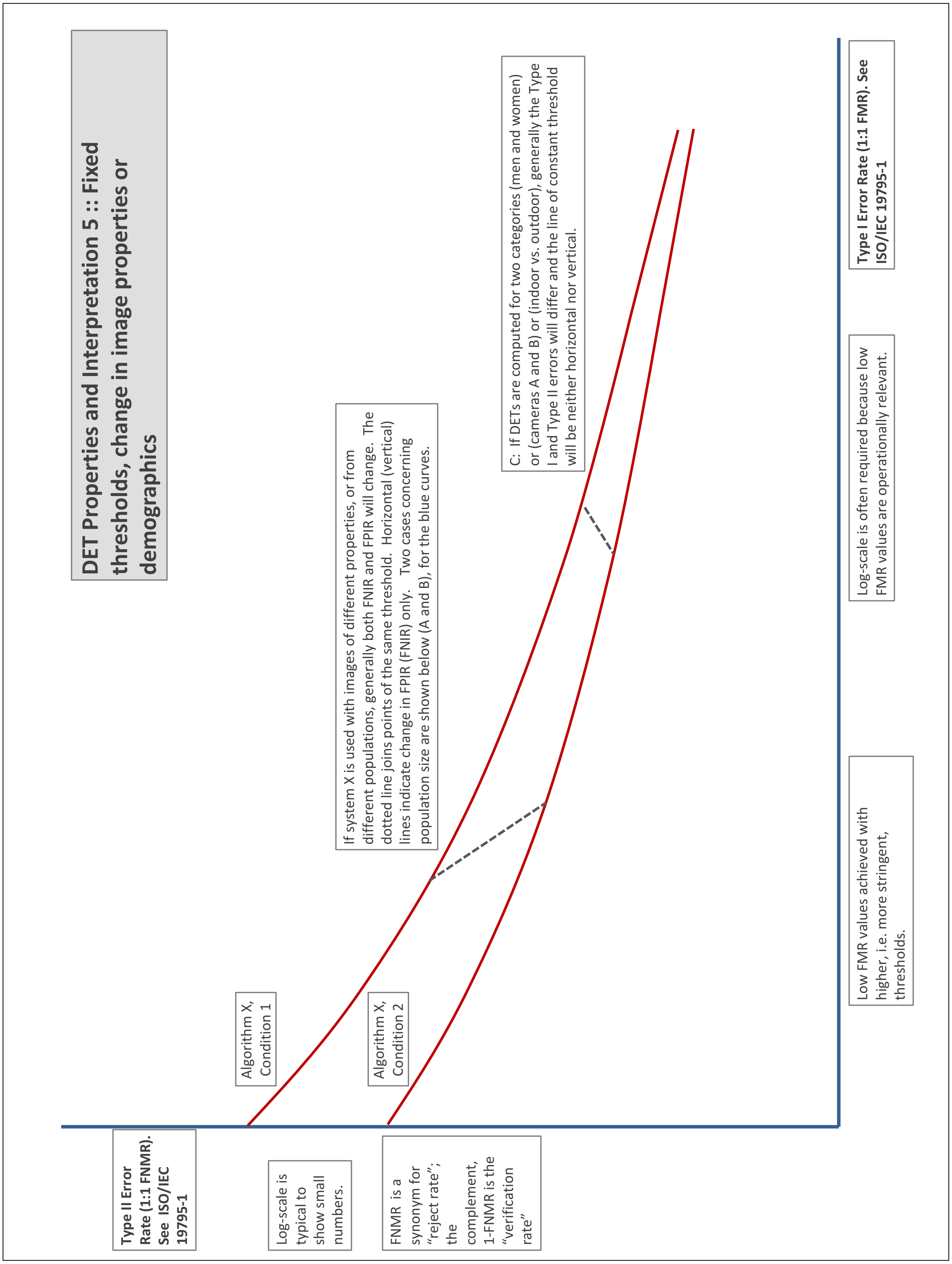


FNMR(T) “False non-match rate”
FMR(T) “False match rate”



FNMR(T) “False non-match rate”
FMR(T) “False match rate”


DET Properties and Interpretation 4 :: Algorithms used in combination
Algorithm A
Algorithm B
If both Algorithms A and B are available and one is selected randomly for any given transaction then the effective DET marked in red is the overall DET. This is the lower convex hull of DET(A) and DET(B).
Use of both Algorithms A and B for all transactions will produce a DET that is not inferior to A or B used alone.
Fused A + B. Multi-algorithm fusion is typically implemented at the score level (e.g. using ISO/IEC 29159-1 normalization information) but may also be done using ranks or decisions. More rarely, template level fusion is possible.
Type I Error Rate (1:1 FMR). See ISO/IEC 19795-1
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate", the complement, 1-FNMR is the "verification rate"


DET Properties and Interpretation 5 :: Fixed thresholds, change in image properties or demographics
If system X is used with images of different properties, or from different populations, generally both FNIR and FPIR will change. The dotted line joins points of the same threshold. Horizontal (vertical) lines indicate change in FPIR (FNIR) only. Two cases concerning population size are shown below (A and B), for the blue curves.
C: If DETs are computed for two categories (men and women) or (cameras A and B) or (indoor vs. outdoor), generally the Type I and Type II errors will differ and the line of constant threshold will be neither horizontal nor vertical.
Algorithm X, Condition 1
Algorithm X, Condition 2
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for “reject rate”; the complement, 1-FNMR is the “verification rate”
Low FMR values achieved with higher, i.e. more stringent, thresholds.
Log-scale is often required because low FMR values are operationally relevant.
Type I Error Rate (1:1 FMR). See ISO/IEC 19795-1

References

[1] P. Jonathon Phillips, Amy N. Yates, Ying Hu, Carina A. Hahn, Eilidh Noyes, Kelsey Jackson, Jacqueline G. Cavazos, Géraldine Jeckeln, Rajeev Ranjan, Swami Sankaranarayanan, Jun-Cheng Chen, Carlos D. Castillo, Rama Chellappa, David White, and Alice J. O'Toole. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, 115(24):6171–6176, 2018.