PARAVISION

Addressing Critical Inclusion Questions for Face Recognition Technology Buyers

Analysis of NIST FRTE 1:1 Demographic Variations

PARAVISION

CONTENTS

Advocating for Inclusion in Face Recognition Technology

This executive document discusses the risks of face recognition systems with poor demographic performance, highlights how companies can advocate for inclusion in biometric systems, explores the latest results from NIST FRTE 1:1 Demographic Variations and showcases Paravision's leadership specifically within the FMR Max results.

01 Background

In the field of face recognition technology, the importance of inclusive systems cannot be overstressed. For financial institutions and other sensitive sectors, choosing the model with the best demographic performance is crucial. Failure to do so can lead to severe legal, financial, and reputational risk.

With the EU AI Act and other regulations on the horizon, **institutions that neglect to address demographic shortcomings in their systems risk non-compliance, hefty fines, and public backlash due to discriminatory practices exposed in the media**. Incorrect matches can result in identity theft, wrongful denial of services, or even unlawful detentions, posing significant financial losses and reputational damage.

02 Role of NIST in Inclusion

The National Institute of Standards and Technology (NIST) plays a pivotal role in combating demographic variations and advocating for inclusion in face recognition technology. Through its rigorous Face Recognition Technology Evaluation (FRTE) 1:1 Verification tests, NIST provides a standardized framework for assessing the performance of face recognition systems across diverse demographic groups.

These evaluations help identify and address disparities in algorithmic accuracy, driving the industry towards greater ethical responsibility and ensuring advancements in face recognition are fair and inclusive. This commitment fosters public trust and promotes the development of technologies that protect the rights of all individuals.

03 Understanding Key Metrics

The NIST FRTE 1:1 Verification includes critical metrics such as:
- *False Non-Match Rate (FNMR) measures* the rate at which true matches are incorrectly not recognized.
- *False Match Rate (FMR)* measures the rate at which an individual is incorrectly matched with another identity.
- *Maximum False Match Rate (FMR Max)* measures the highest rate at which the system incorrectly matches two non-identical individuals from within a single demographic cohort, indicating a system's worst-case performance. A low FMR Max value suggests that the technology is less likely to exhibit demographic differences, as error rates are low for all demographic groups.

PARAVISION'S PERFORMANCE:

#1	**5X**	**24X**	**535**
lowest FMR Max compared to all vendors with FMR Max of 0.00086	lower FMR Max compared to the next best vendor within the top-25	lower FMR Max compared to the worst vendor within the top-25	entries submitted to NIST FRTE 1:1 Demographic Variations in total

NIST FRTE 1:1 Demographic Variations, as of June 2024

Organizational Risks of Demographic Variations in Face Recognition Systems

ETHICAL RISK

Face recognition systems with demographic variations can perpetuate discrimination and inequality, undermining an organization's commitment to ethical practices. The failure to address these biases can lead to systemic injustices, impacting marginalized communities and contradicting corporate social responsibility efforts.

EXPERIENCE

Demographic variations in face recognition systems can negatively impact customer experience, leading to dissatisfaction and loss of loyalty. Customers who experience wrongful denial of services or privacy violations are likely to seek alternatives, reducing customer retention and harming long-term business relationships.

FINANCIAL RISK

Inaccurate face recognition technology can result in significant financial losses. Incorrect matches may lead to fraud, unauthorized access, or wrongful denial of services, causing costly disruptions. Additionally, fines from regulatory bodies due to non-compliance with anti-bias regulations can further strain financial resources.

LEGAL RISK

Organizations using poorly performing face recognition systems face increased legal risks. Discriminatory practices can lead to lawsuits, regulatory enforcement, and hefty penalties under laws like the EU AI Act. Legal challenges not only deplete financial resources but also divert focus from core business activities, impacting overall productivity.

REPUTATION RISK

The use of non-inclusive face recognition technology can severely damage an organization's reputation. Public exposure of discriminatory practices can lead to a loss of trust among customers, investors, and the general public. Rebuilding a tarnished reputation requires significant time, effort, and resources, often with no guarantee of full recovery.

Key Questions to Ask Your Face Recognition Vendor

- Whose face recognition algorithm do you use? Is it your own, or a third party? If the latter, whose is it?
- Has that algorithm been benchmarked on NIST FRTE 1:1?
- Has that algorithm been benchmarked for demographic performance within that test?
- What is your worst-case FMR Max and on what population?
- How do you advocate for inclusion in Face Recognition internally?

Comparison of Demographic Performance, NIST FRTE 1:1 FMR Max

Demographic Variations of Top-25 NIST FRTE 1:1 Developers
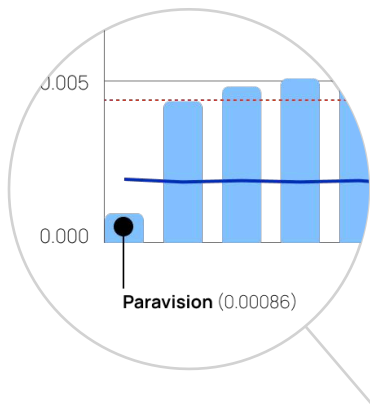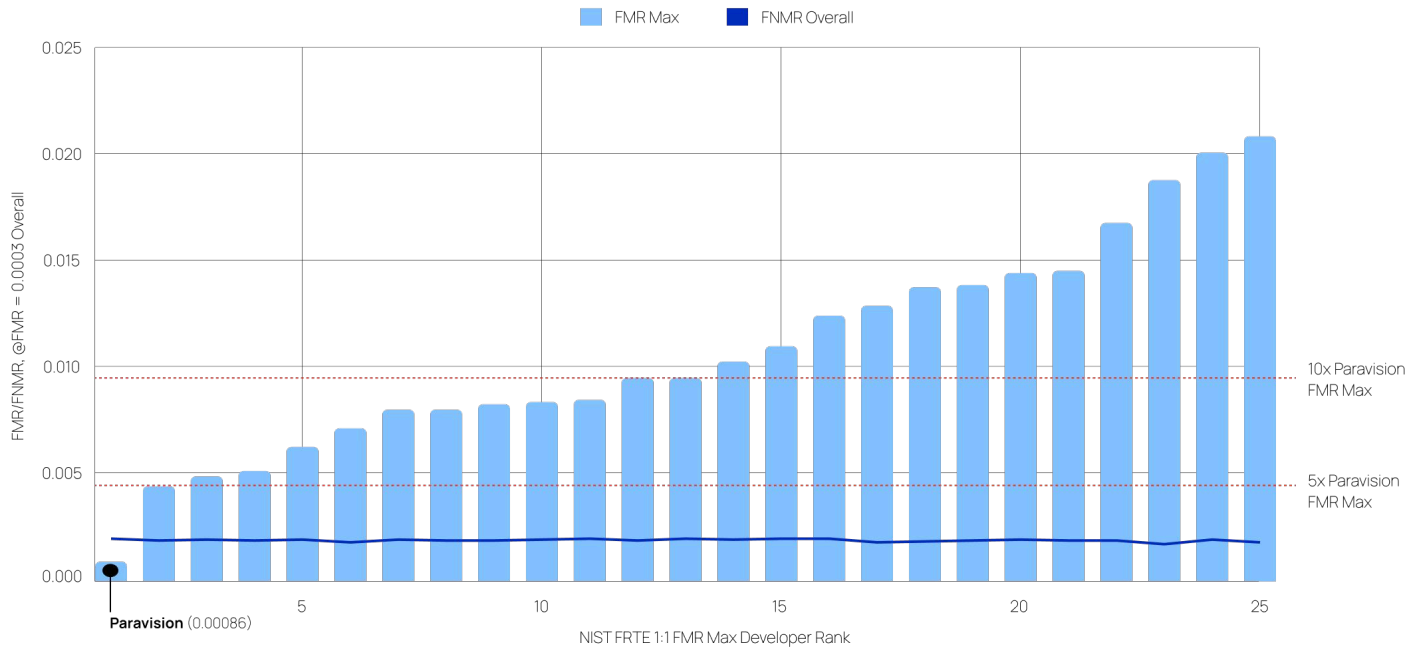
The above graph and the following table represent the top-25 vendors in the NIST 1:1 Verification leaderboard, ranked by their FMR Max from lowest to highest. When compared to the top-25 in the NIST FRTE 1:1 Verification leaderboard, Paravision's results are significantly better than any other vendor:

- The next best vendor's FMR Max is approximately **five times higher** than Paravision's.
- Compared to the worst-performing vendor in the top 25, Paravision's FMR Max is about **24 times lower**.
- Paravision's technology has its highest FMR (False Match Rate) in Eastern European males aged 12-20. **In contrast, the worst-case demographic group for all other vendors is Western African females aged 65-99.**

Demographic Variations of Top-25 NIST FRTE 1:1 Developers

| Rank	Algorithm	FMR Max	FMR Max vs Paravision	FMR Max Group	FNMR Overall
1	paravision_013	0.00086	1.0x	E.Europe M (12-20]	0.0019
2	toshiba_008	0.00439	5.1x	W.Africa F (65-99]	0.0017
3	rebs_001	0.00486	5.7x	W.Africa F (65-99]	0.0018
4	stcon_003	0.00508	5.9x	W.Africa F (65-99]	0.0017
5	kakao_009	0.00623	7.2x	W.Africa F (65-99]	0.0018
6	cloudwalk_mt_007	0.0071	8.3x	W.Africa F (65-99]	0.0015
7	samsungsds-002	0.00797	9.3x	W.Africa F (65-99]	0.0018
8	psl_012	0.00802	9.3x	W.Africa F (65-99]	0.0017
9	surrey_003	0.00826	9.6x	W.Africa F (65-99]	0.0017
10	roc_016	0.00831	9.7x	W.Africa F (65-99]	0.0018
11	clearviewai_001	0.00845	9.8x	W.Africa F (65-99]	0.0019
12	cybercore_003	0.00947	11.0x	W.Africa F (65-99]	0.0017
13	nhn_005	0.00951	11.1x	W.Africa F (65-99]	0.0019
14	tnitech_000	0.01023	11.9x	W.Africa F (65-99]	0.0018
15	neurotechnology_018	0.01089	12.7x	W.Africa F (65-99]	0.0019
16	vocord_010	0.01232	14.3x	W.Africa F (65-99]	0.0019
17	qazsmartvisionai_000	0.01282	14.9x	W.Africa F (65-99]	0.0015
18	viante_000	0.01366	15.9x	W.Africa F (65-99]	0.0016
19	cmcuni_001	0.01381	16.1x	W.Africa F (65-99]	0.0017
20	megvii_009	0.01441	16.8x	W.Africa F (65-99]	0.0018
21	idemia_010	0.01449	16.8x	W.Africa F (65-99]	0.0017
22	sensetime_008	0.01709	19.9x	W.Africa F (65-99]	0.0017
23	recognito_001	0.01878	21.8x	W.Africa F (65-99]	0.0013
24	adera_004	0.02001	23.3x	W.Africa F (65-99]	0.0018
25	intema_001	0.02071	24.1x	W.Africa F (65-99]	0.0015

Why FMR Matters

While many vendors in the competitive landscape of face recognition technology perform similarly on metrics like the False Non-Match Rate (FNMR) and the Minimum False Match Rate (FMR Min), significant disparities are evident in the Maximum False Match Rate (FMR Max). This metric, which measures the highest false match rate across all demographic groups, is a key indicator of a system's ability to effectively manage the worst-case scenarios of demographic performance.

What is FMR?	Imagine a bank application using face recognition system to identify each person correctly. If the system mistakenly says two different people are the same person, that's a false match. The False Match Rate (FMR) tells us how frequently these mistakes happen. A high FMR in a specific demographic group means that identity fraud and account takeovers are more likely for that specific group of people. A low FMR means that a system is better at accurately telling people apart, which is especially important to ensure fairness and accuracy for all demographic groups.

The specific FMR Max Demographic group is another important aspect to consider within the NIST FRTE 1:1 test results. It's crucial for face recognition technologies to prevent harm to already marginalized groups. **According to the NIST FRTE 1:1 Demographic Variations Leaderboard, Paravision's technology has its highest FMR (False Match Rate) in Eastern European males aged 12-20. In contrast, the worst-case demographic group for all other vendors in the top-25 is Western African females aged 65-99.**

Worst-case demographic group for all vendors, NIST FRTE 1:1 Demographic Variations

PARAVISION

0.5%
Eastern European Males, Aged 12-20

98%
Western African Females, Aged 65-99

1.4%
Eastern Asian Females, Aged 65-99

0.5%
Central American Females, Aged 65-99

0.5%
Western African Females, Aged 12-20
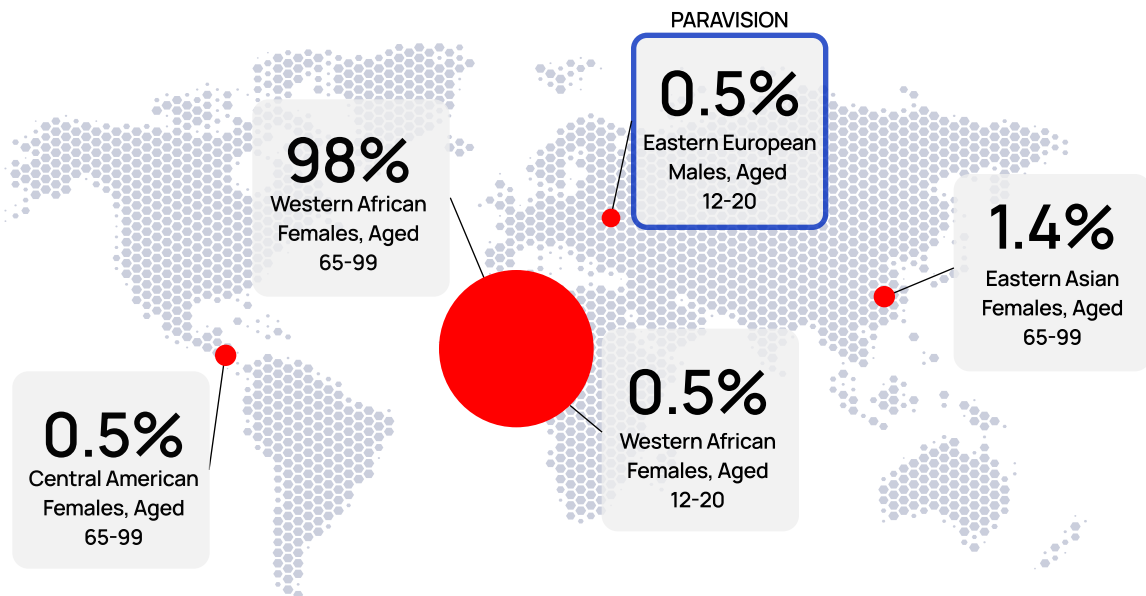
Paravision's Approach to Inclusion

While reducing bias is a crucial goal to which Paravision is deeply committed, we believe the ultimate goal should be inclusion through minimizing overall error rates. By striving to eliminate errors altogether, we aim to create a world where our AI solutions empower individuals universally, enhancing access and opportunities without compromising accuracy or fairness. The fence metaphor, often used in Diversity, Equity, and Inclusion discussions, can help illustrate different approaches to inclusion in face recognition systems:

The fence represents face recognition systems with high demographic variations, meaning they perform better for some groups than others. This approach is biased and creates significant risks for companies.
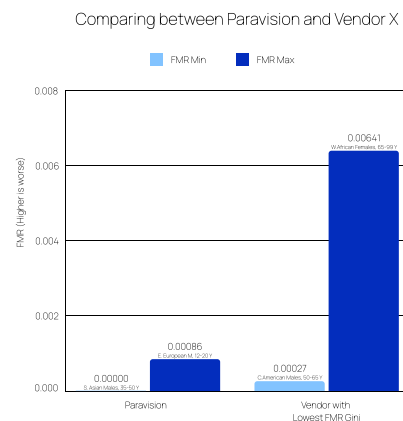
This image shows a system minimizing demographic variations with equal performance across groups but with high overall error rates. While fair and equal, these systems create barriers for users and operational risks for companies.

Paravision's approach: removing barriers by reducing error rates across all groups to an absolute minimum. This approach enables excellent performance for all groups, reducing friction and organizational risks.

What about FMR Gini?

In NIST's 1:1 Demographic Variations FMR leaderboard, the vendor with the lowest FMR Gini has a Gini of 0.38. While this indicates low differences between demographic groups, it's important to note the vendor's FMR Max and FMR Min to get a full picture. That same vendor's overall error rates are significantly higher than Paravision's as seen in the graph to the right. Although reducing differences between groups is vital for the industry, Paravision believes the goal should be to remove barriers completely and guarantee the lowest possible FMR Max, ensuring *good performance for all demographic groups.*

Comparing between Paravision and Vendor X

■ FMR Min ■ FMR Max

FMR (Higher is worse)

Paravision	Vendor with Lowest FMR Gini

0.00000 S. Asian Males, 35-50 Y
0.00086 E. European M, 12-20 Y
0.00027 C. American Males, 50-65 Y
0.00641 W. African Females, 65-99 Y

Turning Metrics into Real-World Language for Better Understanding

Understanding the technical metrics of face recognition technology can be challenging. To make these metrics more accessible, let's translate Paravision's False Match Rate (FMR) Max of 0.00086 into real-world scenarios.

False Match Rate (FMR) Max Explained

FMR Max measures the Maximum False Match Rate for a single demographic group at a set threshold of an overall False Match Rate of 0.0003. *This means that NIST sets the overall False Match Rate at 3 in 10,000, measures how the algorithms perform within each demographic group, and reports the worst case scenario.*

Paravision's FMR Max of 0.00086 means that at the threshold that delivers an overall FMR of 0.0003 (0.03%), the worst performing demographic group will see only 0.086% of identity verification attempts resulting in a false match. To put this into perspective:

- **Everyday Scenario: 10,000 Verifications**
  - Overall false match threshold of 0.0003 would mean 3 false matches in 10,000 verification attempts.
  - For every 10,000 verification attempts, a system using Paravision will get approximately 8.6 false matches for the worst-performing demographic group, and an even fewer false matches for all other demographic groups.
  - In contrast, the next best vendor's FMR Max of 0.00439 means they will see about 44 false matches per 10,000 verifications.
- **Larger Scale Scenario: 1 Million Verifications**
  - Overall false match threshold of 0.0003 would mean 300 false matches in 1 million verification attempts.
  - For every 1 million verification attempts, a system using Paravision will get approximately 860 false matches for the worst-performing demographic group.
  - In contrast, the next best vendor would see 4,390 false matches per 1 million verification attempts.
- **Massive Scale Scenario: 10 Million Verifications**
  - Overall false match threshold of 0.0003 would mean 3000 false matches in 10 million verification attempts.
  - For 10 million verification attempts, a system using Paravision will get approximately 8,600 false matches for the worst-performing demographic group.
  - In contrast, the next best vendor would experience 43,900 false matches.

Real-World Implications:

| Enhanced Security	Improved User Experience	Operational Efficiency
Lower false match rates translate to enhanced security, reducing the risk of unauthorized access and identity fraud.	With fewer false matches, users will experience smoother and more reliable identity verification processes, leading to higher satisfaction and trust.	Organizations can achieve higher operational efficiency with fewer interruptions caused by false matches, saving time and resources.

PARAVISION

Geographic Analysis of FMR

While NIST doesn't provide an apples-to-apples comparison of vendors' FMR for each demographic group, the NIST FRTE 1:1 Report Cards for vendors give us useful data for comparing the demographic performance across groups. In the graph and table below, you can find the cross-demographic FMR estimates for the top-25 vendors (as ranked by overall FNMR), when the enrollment image is compared to a probe image of the same country of origin, sex, and age.

The countries of origin have been clustered into 7 geographic regions as per NIST's method. Country-specific FMRs have been averaged into regional values to simplify the graph. This assumes equal weighting of the NIST 1:1 dataset for each country, and so actual regional performance might be slightly different (NIST doesn't publish country-specific dataset sizes). The regions are:

- Eastern Europe: Poland, Russia, Ukraine
- Central America: Mexico, El Salvador, Nicaragua
- West Africa: Nigeria, Liberia, Ghana
- Caribbean: Haiti and Jamaica
- Eastern Africa: Kenya
- South Asia: Iraq, Iran, Pakistan, India
- East Asia: Vietnam, Philippines, Korea, Japan, Thailand, China

Paravision singularly demonstrates consistent performance across all geographic areas, and outperforms all other vendors in Central America, Western Africa, the Caribbean, Eastern Africa, South Asia, and East Asia. This broad-spectrum accuracy is crucial because it ensures that the technology is equitable and effective for a diverse or global user base.

Demographic Analysis: False Match Rates of Top 25 NIST FRTE 1:1 Developers
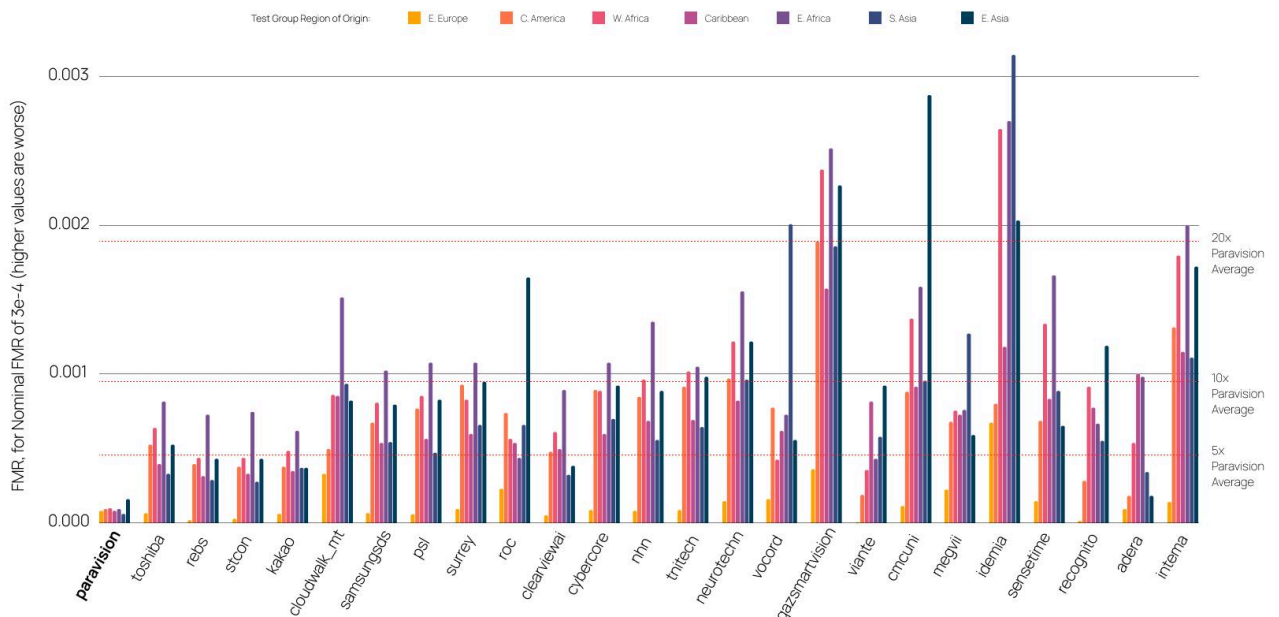
## FMR Demographic Performance of Top-25 NIST FRTE 1:1 Developers

Performance at FMR Nominal at 3e-5, with blue meaning better performance, and red meaning worse performance.

Algorithm	E. Europe	C. America	W. Africa	Carribbean	E. Africa	S. Asia	E. Asia
paravision_013	0.00008	0.00009	0.00010	0.00008	0.00009	0.00006	0.00016
toshiba_008	0.00007	0.00052	0.00064	0.00039	0.00081	0.00033	0.00052
rebs_001	0.00002	0.00040	0.00044	0.00031	0.00072	0.00029	0.00043
stcon_003	0.00003	0.00038	0.00044	0.00033	0.00074	0.00027	0.00043
kakao_009	0.00006	0.00038	0.00048	0.00035	0.00062	0.00037	0.00037
cloudwalk_mt_007	0.00033	0.00049	0.00086	0.00085	0.00151	0.00094	0.00082
samsungsds-002	0.00007	0.00067	0.00081	0.00054	0.00102	0.00054	0.00079
psl_012	0.00006	0.00076	0.00085	0.00056	0.00107	0.00047	0.00082
surrey_003	0.00009	0.00092	0.00082	0.00059	0.00107	0.00066	0.00094
roc_016	0.00022	0.00074	0.00056	0.00054	0.00044	0.00066	0.00165
clearviewai_001	0.00005	0.00048	0.00061	0.00049	0.00089	0.00032	0.00038
cybercore_003	0.00009	0.00089	0.00088	0.00060	0.00107	0.00070	0.00092
nhn_005	0.00008	0.00084	0.00096	0.00069	0.00135	0.00056	0.00089
tnitech_000	0.00008	0.00091	0.00101	0.00069	0.00105	0.00064	0.00098
neurotechnology_018	0.00014	0.00096	0.00122	0.00082	0.00155	0.00096	0.00121
vocord_010	0.00016	0.00077	0.00042	0.00062	0.00072	0.00200	0.00055
qazsmartvisionai_000	0.00036	0.00189	0.00237	0.00157	0.00251	0.00186	0.00226
viante_000	0.00000	0.00019	0.00036	0.00081	0.00043	0.00057	0.00092
cmcuni_001	0.00012	0.00088	0.00137	0.00091	0.00158	0.00095	0.00287
megvii_009	0.00022	0.00068	0.00075	0.00072	0.00076	0.00127	0.00059
idemia_010	0.00067	0.00080	0.00264	0.00118	0.00269	0.00314	0.00203
sensetime_008	0.00015	0.00068	0.00133	0.00083	0.00166	0.00088	0.00065
recognito_001	0.00001	0.00028	0.00091	0.00077	0.00066	0.00055	0.00119
adera_004	0.00009	0.00018	0.00054	0.00100	0.00098	0.00034	0.00018
intema_001	0.00014	0.00131	0.00179	0.00115	0.00200	0.00110	0.00172

FMR average

Low   High

Why Does Paravision Deliver Such Low Error Rates Across Demographics?

LEADING TECHNOLOGY

Paravision's face recognition systems are designed to be accurate and reliable across all demographic groups. Our advanced algorithms minimize demographic variations and deliver precise results, even in challenging scenarios, as evidenced by our leading performance in the NIST FRTE tests.

SOLID AI PRINCIPLES

Our strong AI principles, emphasizing ethical AI development, guide every aspect of our business and technology, ensuring fairness, transparency, and accountability from design to deployment.

DIVERSE TEAM

The diverse team at Paravision is crucial to developing inclusive technology. Incorporating various backgrounds and perspectives enables us to create innovative solutions that consider a wide range of demographic factors, enhancing the fairness and inclusivity of our systems.

DATA & BENCHMARKING

We believe in the importance of benchmarking to achieve fairness, and regularly participate in evaluations like the NIST FRTE 1:1 to ensure our technology meets high standards and continuously improves. Benchmarking provides an objective measure of performance, helping us identify and address any potential demographic variations.

GENUINE CARE

Paravision prioritizes the ethical implications of face recognition technology, striving to create solutions that respect and protect all users. This commitment to fairness and inclusivity ensures our technology benefits everyone, delivering equitable outcomes in all applications.
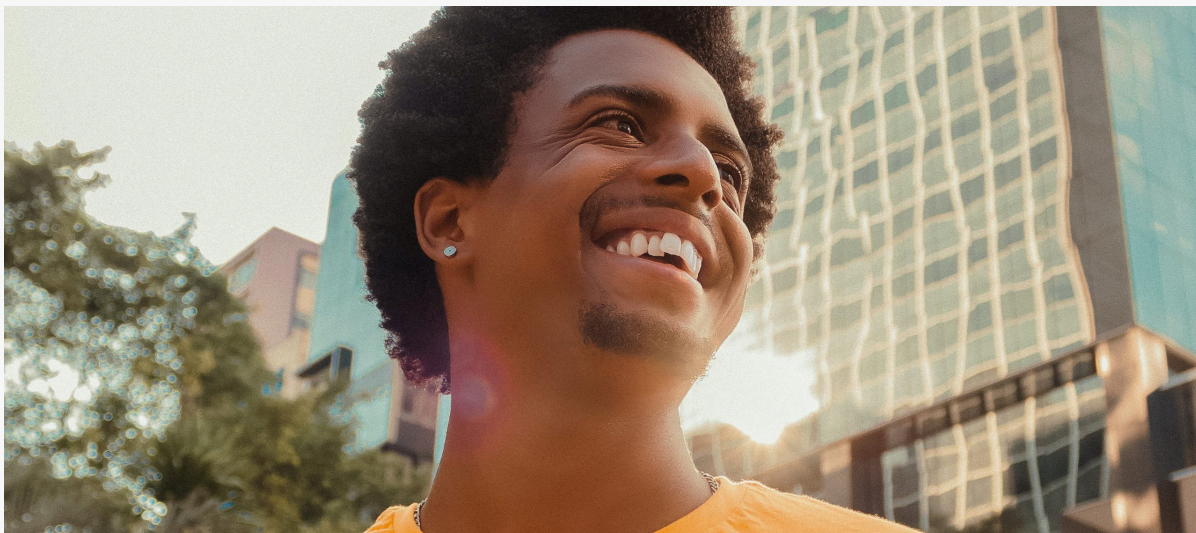
#1 in the U.S., UK & EU **NIST FRTE 1:1**

#1 in the U.S., UK & EU **NIST FRTE 1:N**

Top matching system **DHS Biometric Rally**

Best Face Recognition **AI Breakthrough**

PARAVISION

TRUSTED **IDENTITY AI**

paravision.ai/contact