U.S. Department of Homeland Security

SCIENCE AND TECHNOLOGY DIRECTORATE

Remote Identity Validation Rally (RIVR) Identity Document Validation (IDV) Results Webinar

Science and Technology

Arun Vemury
Senior Advisor
Biometric & Identity Technology Center

Yevgeniy Sirotin
Technical Director
The Maryland Test Facility

February 2026

Agenda

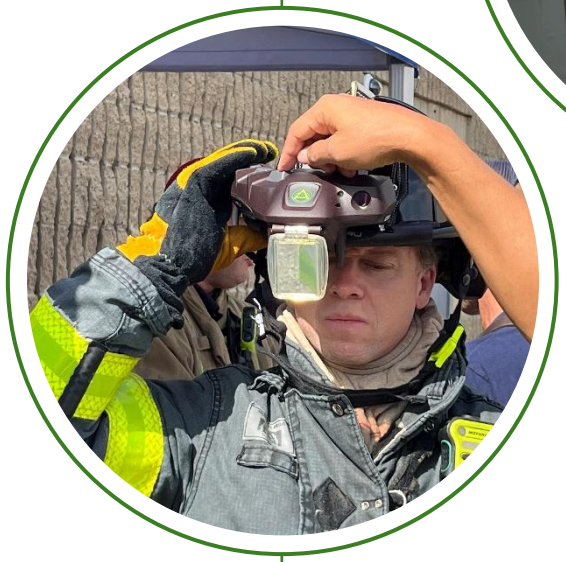
- Introduction
- Remote Identity Validation Rally (RIVR)
- RIVR: Identity Document Validation Evaluation
  - Data Used
  - System Requirements
  - Metrics and Benchmarks
- RIVR: Identity Document Validation Results
  - Evaluation Criteria
  - System Error Rates
  - Document Validation Error Rates
  - Detection Error Tradeoffs
- Summary & Conclusions

[SCIENCE AND TECHNOLOGY DIRECTORATE]

Operationalizing science and technology.

The Science and Technology Directorate (S&T) researches, develops, tests, and evaluates solutions needed to meet the growing demands of our nation's homeland security officials.

- We capture specific mission needs.
- We deliver impactful technology solutions.
- We conduct independent test and evaluation.



Science and Technology

Biometric & Identity Technology Center

The Science & Technology Directorate (S&T) conducts foundational research to ensure advancements in science and technology are harnessed in the development of cutting-edge solutions to new and emerging operational challenges.

- Drive biometric and identity innovation at the Department of Homeland Security (DHS) through Research, Development, Test, and Evaluation (RDT&E) capabilities
- Facilitate and accelerate understanding of biometrics and identity technologies for new, DHS use cases
- Drive efficiencies by supporting cross-cutting methods, best practices and solutions across programs
- Deliver subject matter expertise across the DHS enterprise
- Engage industry and provide feedback
- Encourage innovation across industry and academia

Remote Identity Validation

- Remote Identity Validation (RIV) technology is a tool to authenticate documents and verify the identity of users remotely
- These systems are complex, with multiple subsystems, and are increasing in popularity and adoption
- Industry performance benchmarks are not well defined, making it is difficult for organizations to test the effectiveness of these systems
- S&T is studying the current performance of RIV to help industry develop more secure, accurate, and robust technologies:
  - Remote Identity Validation Technology Demonstration (RIVTD) from 2023 to 2024
    - Comprehensively demonstrated performance of commercial RIV subsystems
    - Informed NIST digital identity guidelines
    - Identified metrics, performance gaps, and achievable performance benchmarks
  - Remote Identity Validation Technology Rally – currently ongoing

Remote Identity Validation Rally (RIVR)

- **Building on RIVTD Insights**: RIVTD identified key areas where RIV vendors should focus improvements, shaping the next phase of evaluation
- **Establishing Achievable Benchmarks**: RIVR sets industry-informed performance benchmarks based on RIVTD results, providing clear targets for improvement
- **Encouraging Innovation & Retesting**: Vendors can refine their technologies and participate in re-evaluation
- **Confidential & Industry-Driven**: Vendor names are aliased, allowing companies to self-attest participation while fostering industry-wide progress

Science and Technology

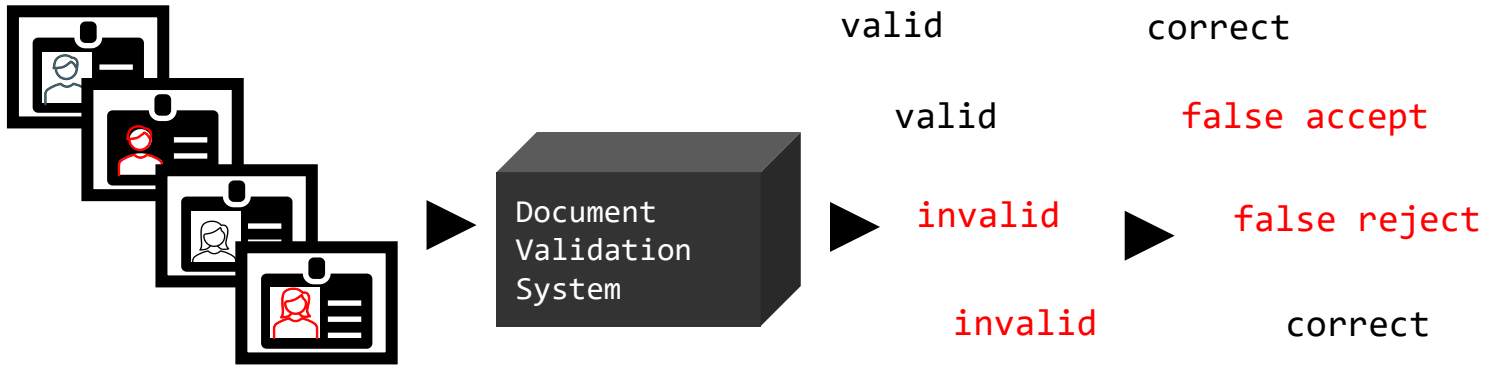
Identity Document Validation Track Process & Requirements

Science and Technology

Data Used

- Leveraged a large and growing collection of:
  - Images of genuine U.S. State-issued ID cards (e.g., driver's licenses)
  - Fraudulent IDs collected from the DHS HSI Forensic Laboratory
- RIV Document Validation Subsystems (DVSs) were evaluated based on their ability to determine if a **U.S. State-issued ID card** is genuine or fraudulent.

- System Combination:



Science and Technology

Genuine Document Dataset Composition

- Images of genuine documents were collected from volunteers under informed consent
- A total of 2,032 genuine documents were collected between 2023 and 2025
- Test staff collected both front and back images of fraudulent documents using each of three smartphones
  - All documents were validated using a multi-spectral government ID validation tool

State	N	State	N	State	N	State	N
AZ	7	GA	2	MI	2	PA	8
CA*	1,298	IL	1	MO	2	TX	6
CO	2	KS	1	MT	2	UT	2
CT	1	KY	1	NV	1	VA*	97
DE	1	MA	3	NY	5	WA	3
FL	2	MD*	526	OR	2	DC*	55
						Total:	2,032

Science and Technology

Sample Images Across Collection Methods and Devices

- Genuine document images were acquired on each of three smartphones. Data collection included images with and without mild perspective distortion:

*Volunteers consented to have images used in government presentations. ID documents redacted to protect privacy.

Science and Technology

Fraudulent Document Dataset Composition

- Images of fraudulent documents were collected in partnership with the DHS HSI Laboratory
- A total of 1,938 fraudulent documents were included in this evaluation
- Test staff collected images of both the front and back of each fraudulent document using the same smartphones used for genuine documents
- Fraudulent document states of issue included all 50 states as well as the District of Columbia

Participating Systems

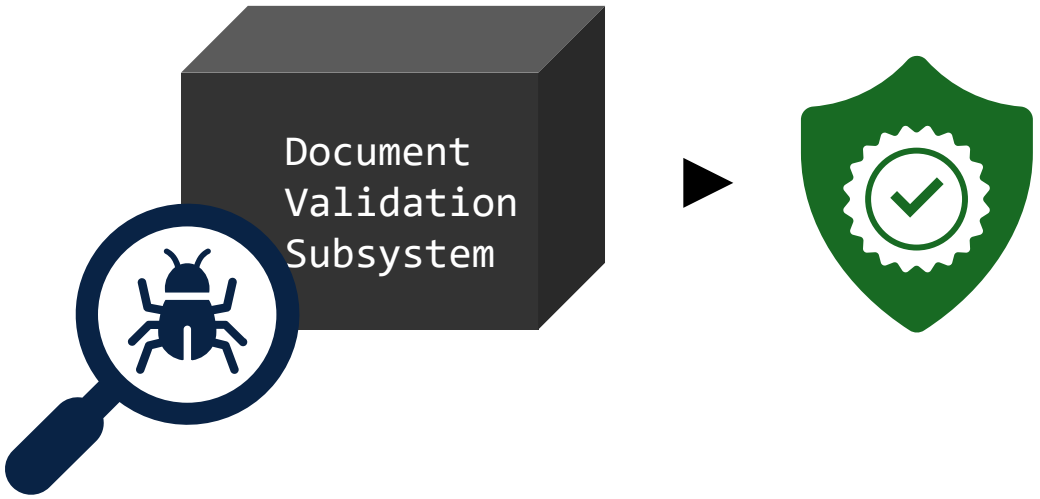
- 7 commercial Document Validation Subsystems (DVSs) participated in RIVR
  - Announced in July 2025
  - Applications due in July 2025
  - Submissions due in August 2025

- Representative of the state of the industry

- Each system was given a unique alias (DVS 1, DVS 2, etc.)

System Requirements

- Implement the MdTF Document Validation Application Programming Interface (API)
- A single Linux-based docker container
  - HTTP server on port 8080
  - Less than 5 GB in size
- Pass strict security assessments for deployment to government systems
- No outside functionality and no access to the internet
- Systems were evaluated on both government and non-government systems



Identity Document Validation Metrics

- **System Error Rate (SER)** – Proportion of documents for which the document validation subsystem returned an error message indicating the document failed to process
  - **Threshold: 0.10, Goal: 0.01**
- **Document False Reject Rate (DFRR)** – Proportion of genuine documents which the document validation subsystem either returned a system error or determined to be invalid
  - **Threshold: 0.10, Goal: 0.01**
- **Document False Accept Rate (DFAR)** – Proportion of fraudulent documents which the document validation subsystem determined to be valid
  - **Threshold: 0.10, Goal: 0.01**
- **Disaggregated to examine robustness for:**
  - Document state of issue
  - Smartphone type

RIVR set performance benchmarks for each metric:
Threshold – maximum high-performance error rate
Goal – target high-performance error rate

Science and Technology

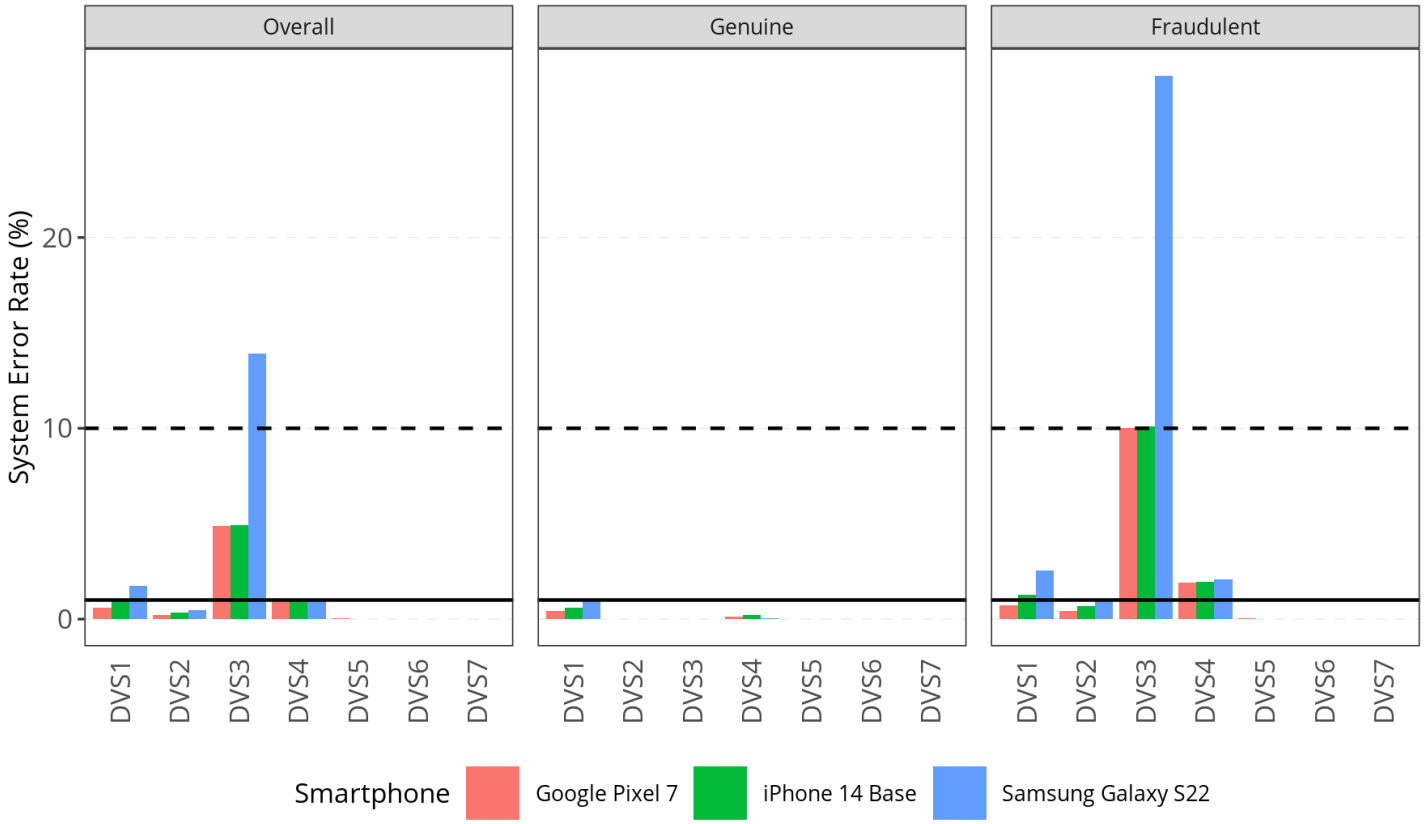
Identity Document Validation Track Results

Science and Technology

Methodology

- Document validation subsystems were evaluated in combination with different smartphones and disaggregated for different genuine document states of issue
  - The sample size was sufficient for disaggregation for CA, DC, MD, and VA
- Evaluation metrics were computed separately for each DVS smartphone pair, referred to as a system combination
- Overall performance for each tested system was assessed based on the maximum error rate value observed for each metric across the three smartphones
- A “failure is suspicious” policy was used to calculate document validation errors
  - A system error on a genuine document is considered a false rejection
  - A system error on a fraudulent document is **NOT** considered a false acceptance

System Error Rates

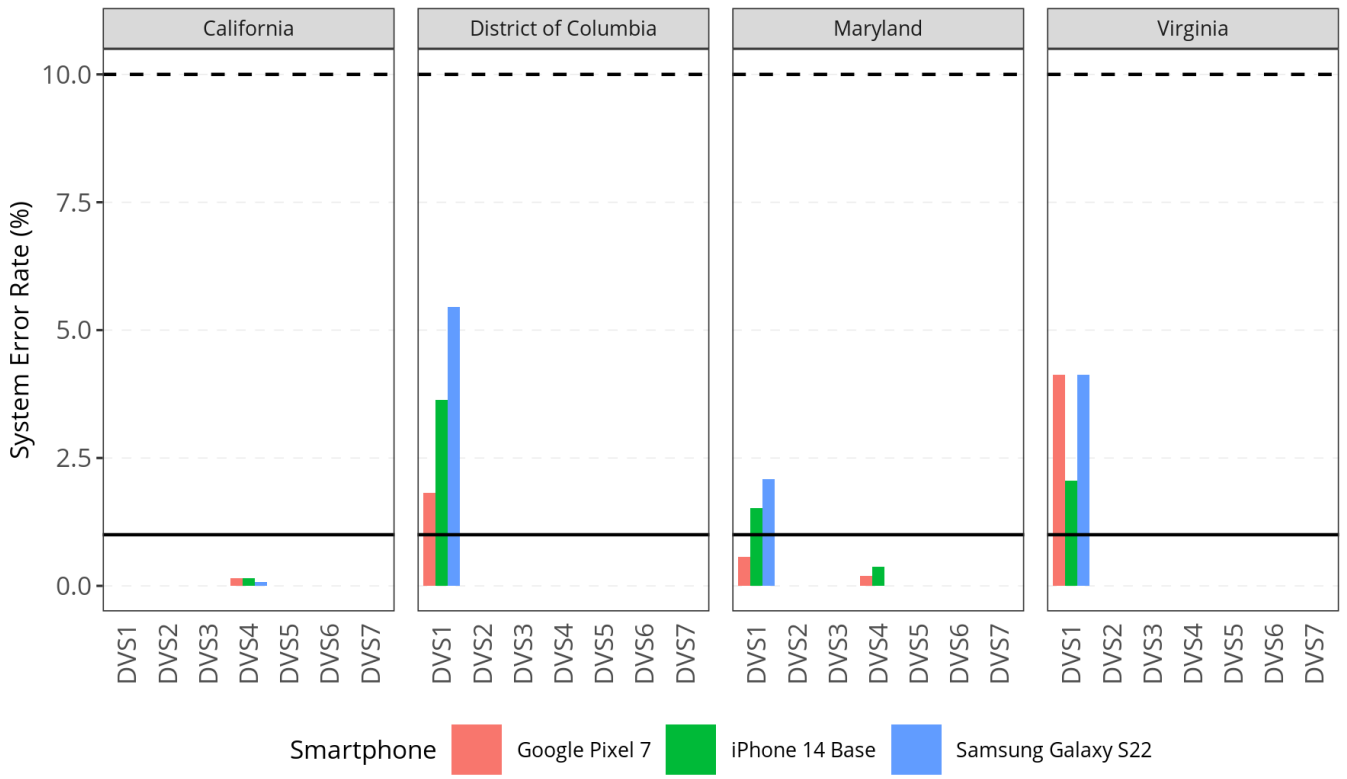


Bars correspond to performance of combinations of smartphone and DVS. Area below the black line indicates 1% or lower SER. Area below the dashed black line indicates 10% or lower SER.

- **Most systems performed well**
  - 4 of 7 DVS maintained overall SER below 1%
  - 6 of 7 DVS maintained overall SER below 10%
- **Error rates on fraudulent documents were higher**
- **Smartphone devices can affect performance**
  - DVS 3's SER was three times higher on the Samsung device as compared to the iPhone and Pixel

Science and Technology

State of Issue: Genuine System Error Rates

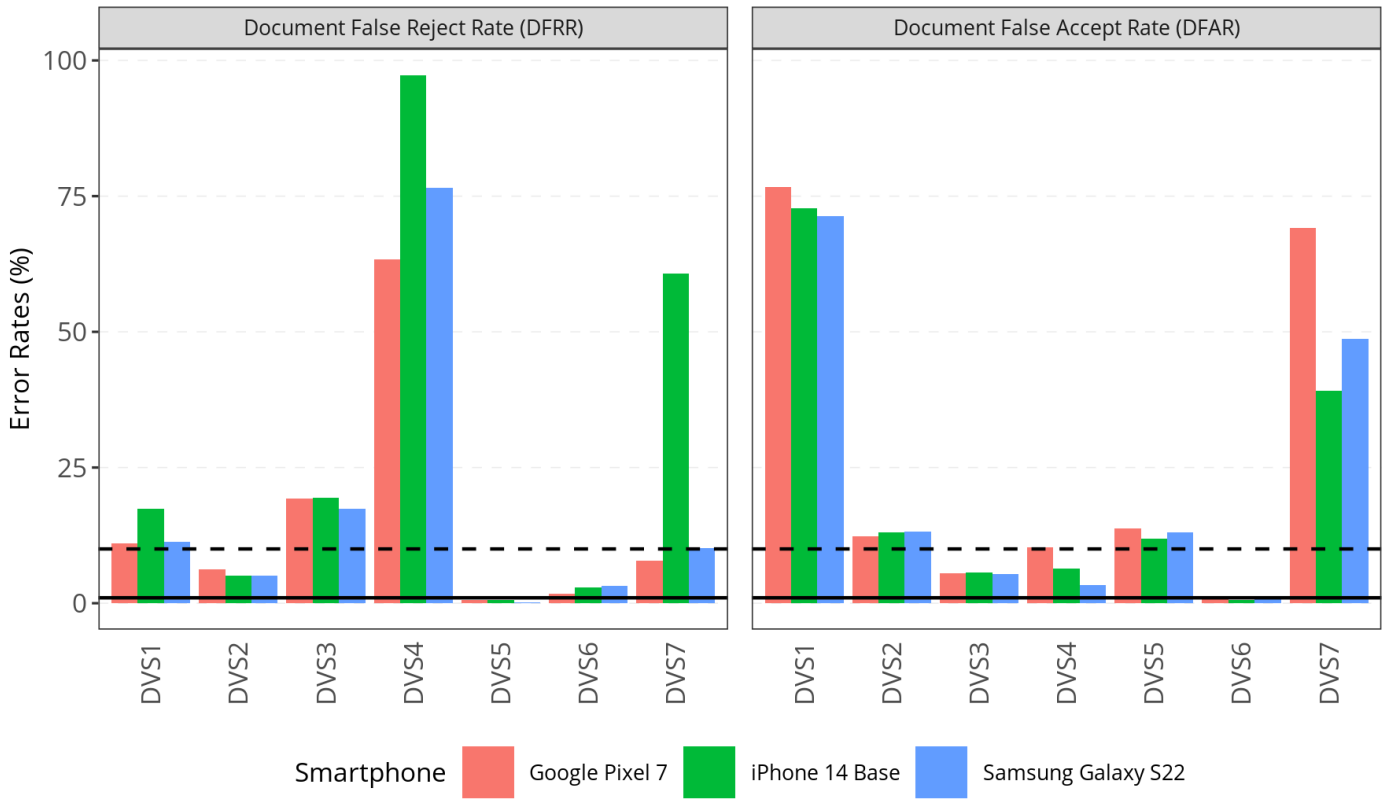


Bars correspond to performance of combinations of smartphone, states of issue, and DVS. Area below the black line indicates 1% or lower genuine SER. Area below the dashed black line indicates 10% or lower genuine SER.

- **Most DVS maintained low SER for different states of issue**
  - 6 of 7 DVS maintained genuine SER below 1%
- **Document state of issue can affect DVS performance**
  - DVS 1 had elevated SER for DC, MD, and VA documents
  - DVS 4 had elevated SER for CA and MD documents

Science and Technology

Document Validation Error Rates

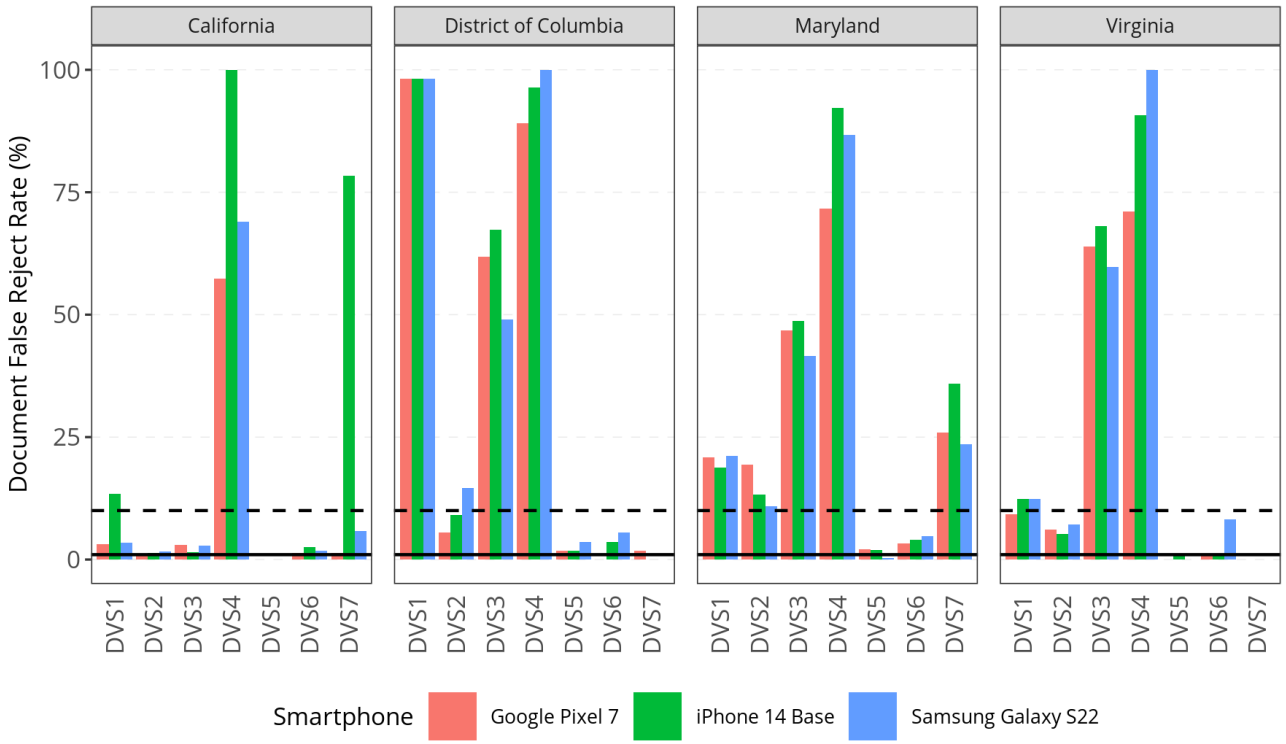


Bars correspond to performance of combinations of smartphone and DVS. Area below the black line indicates 1% or lower SER. Area below the dashed black line indicates 10% or lower SER.

- **Discriminating genuine and fraudulent documents is difficult**
  - DVS 6 was the only subsystem that maintained both DFRR and DFAR below 10%
  - DVS 5 and DVS 3 came close
- **Some systems had very high error rates**
  - DVS 4 tended to reject all documents, while DVS 1 tended to accept all documents regardless of if they were fraudulent or genuine
- **Smartphone devices can affect performance**
  - DVS 7 rejected six times more genuine documents on the iPhone as compared to the Samsung or Pixel

Science and Technology

State of Issue: Document False Reject Rates



Bars correspond to performance of combinations of smartphone, states of issue, and DVS. Area below the black line indicates 1% or lower genuine SER. Area below the dashed black line indicates 10% or lower genuine SER.

- **Document state of issue affected DVS performance**
  - Overall Performance was worst for MD documents and best for CA documents

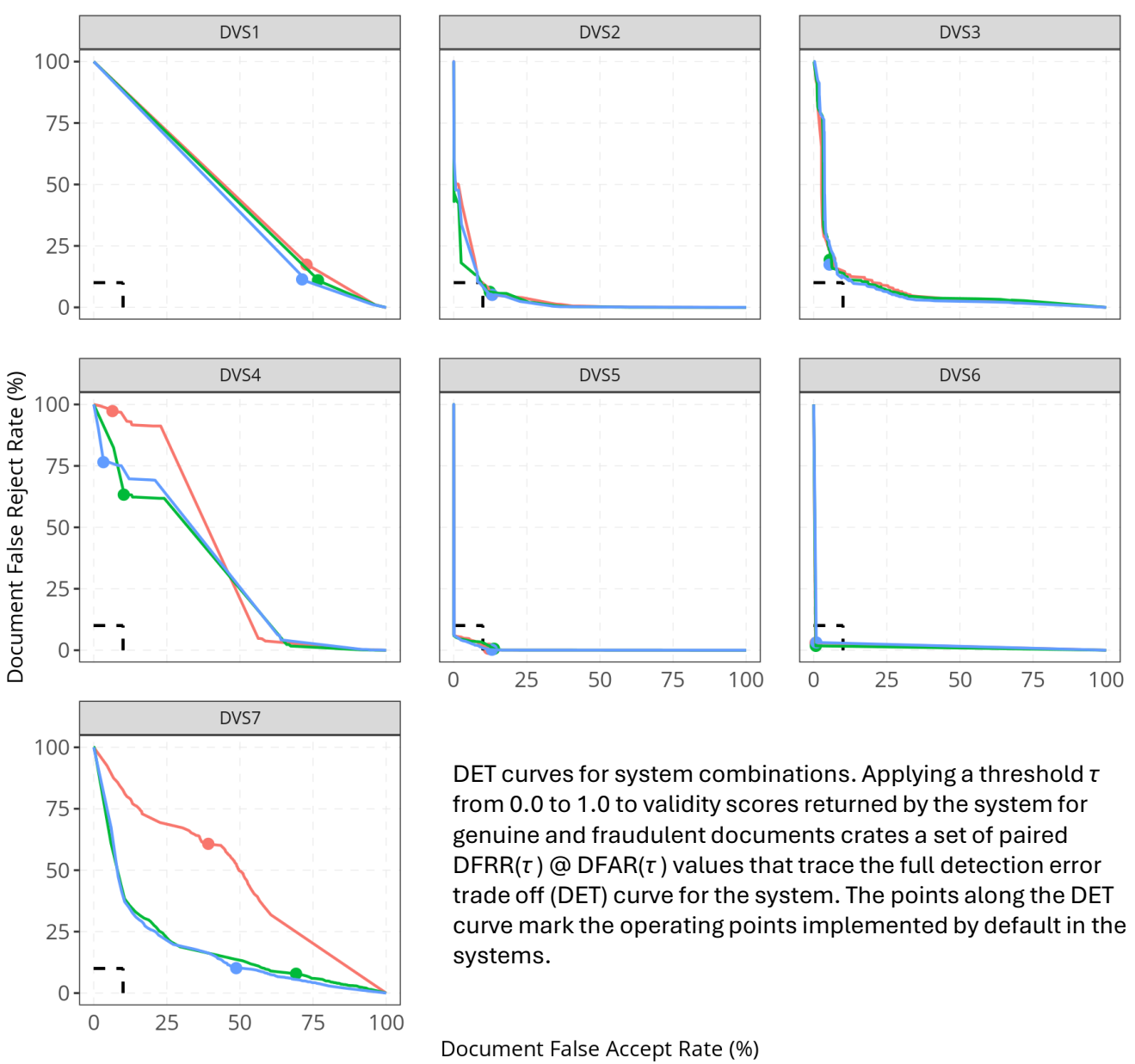
System combinations that met the threshold or goal		
Document State of Issue	DFRR < 10% (threshold)	DFRR < 1% (goal)
California	16	4
District of Columbia	11	3
Maryland	6	1
Virginia	13	5

- **Major impacts of state of issue**
  - DVS 3 met the DFRR threshold for CA but failed to do so for DC, MD, and VA

Science and Technology

Detection Error Trade Off

- DVS 6 had the best ability to discriminate between genuine and fraudulent documents
- DVS 5 had comparable discriminative power - given an alternative threshold DVS 5 could have met RIVR IDV performance benchmarks



DET curves for system combinations. Applying a threshold τ from 0.0 to 1.0 to validity scores returned by the system for genuine and fraudulent documents crates a set of paired DFRR(τ) @ DFAR(τ) values that trace the full detection error trade off (DET) curve for the system. The points along the DET curve mark the operating points implemented by default in the systems.

Science and Technology

Summary & Conclusions

Results Summary

DVS	1	2	3	4	5	6	7
SER	1.76%	0.48%	13.9%	1.06%	0.03%	0.00%	0.00%
DFRR	17.4%	6.3%	19.4%	97.3%	0.59%	3.15%	60.7%
DFAR	76.7%	13.2%	5.68%	10.3%	13.8%	0.88%	69.2%

Legend	
X	Met Goal
X	Met Threshold
X	Did not meet Threshold

- No tested systems met RIVR goals for all metrics
- **DVS6** met RIVR performance thresholds for all metrics
- **6 of 7** tested systems were unable to meet the threshold for at least one metric
- **3 of 7** tested systems performed worse than a coin flip on at least one metric

Conclusions

- Discriminating between genuine and fraudulent documents based only on visible light images is an extremely challenging task
  - Industry error rates can be high – few DVS can meet 10% error rate benchmarks
  - For individual DVS, performance can vary by:
    - Imaging device
    - Document state of issue
- Although the task is challenging, some DVS can reliably discriminate between genuine and fraudulent documents included in this evaluation
- Maintaining both DFRR and DFAR below 10% continues to be a challenging benchmark for most technology providers

Questions & Answers

- Contact information
  - peoplescreening@hq.dhs.gov
  - rivr@mdtf.org

- Visit our websites for additional information
  - To see additional work DHS S&T supports, visit www.dhs.gov/science-and-technology.
  - For information about this and other DHS S&T technology evaluations, visit https://mdtf.org.

Science and Technology